Doctoral Dissertation

博士論文

Automatic 3-D Localization of Utility Pipes and Void from Ground Penetrating Radar Signals by Deep Learning and Digital Signal Processing

（地中レーダー信号の深層学習と解析学的信号処理による埋設管・空洞の三次元位置推定）

Takahiro Yamaguchi

山口　貴浩

Department of Civil Engineering, Graduate School of Engineering, the University of Tokyo

September 2020

Abstract

In recent years, aging of infrastructures is becoming an important issue. In terms of subsurface pipes, total distance of sewage pipes reaches up to 470,000 km in Japan. 30 % of the sewage pipes is going to be over 50 years old in 20 years. Void caused by deteriorated pipes brings 3,000 road collapse per year, which draws a large social concern. Moreover, old unknown pipes cause construction delays. In some cases, pipes are damaged to bring an economic loss to society.

In the research, we are focusing on Ground Penetrating Radar (GPR) method for noncontact, highspeed and dense 3D monitoring of conditions under the road. GPR method utilizes a vehicle-installed radar system to scan about 80 km/h, not requiring a traffic control. However, in GPR method, skilled inspectors detect anomalies manually from radar images. Furthermore, data volume is up to several Tera-Bytes for several tens km section, resulting in several months and million-yen (ten-thousand-dollar) inspection time and cost. Therefore, a highspeed and high accuracy automatic detection algorithm of subsurface pipes and void from radar data enables us to drastically reduce inspection time and cost.

The objective of the research is, “highspeed and accurate automatic 3D mapping of subsurface pipes and void from radar images by the combination of AI and signal processing.” Target objects are various including manhole, joint, subsurface pipes and void. The proposed algorithm is accurate and only needs 5 minutes for 1 km section.

In the application of pattern recognition approaches such as AI, signals behind manhole and joint caused by multiple reflections between antennas and objects were falsely detected as pipes and void. Therefore, manhole and joint were detected from surface reflections by nonlinear SVM utilizing a feature value. Classification accuracy was about 98 %, almost 100 %. Detected areas were removed from target regions of pipes and void.

A methodology to reproduce 3D reflection patterns by 2D-FDTD method was proposed. In the research, the number of void data was not enough for the training of deep learning models. Therefore, training data was produced by electromagnetic simulations based on FDTD method. To learn 3D spatial features of reflection patterns, 3D data should be reproduced. However, 3D-FDTD method is not feasible because of enormous calculation cost. Therefore, cross-sections of a 3D reflection pattern were approximated by 2D models. Then, regions were linearly interpolated in cylindrical coordinates to reproduce the 3D reflection pattern.

A high accuracy detection algorithm for pipes and void was developed utilizing 3D reflection patterns by deep 3D Convolutional Neural Network (3D-CNN) models. In terms of void, the classification accuracy of 3D-CNN was about 87 %, comparable to the detection accuracy of skilled inspectors. In terms of pipes, by three-category classification of transverse, longitudinal pipes and no pipe section, not only the existence but also the directions of pipes were estimated with 91 % classification accuracy. The classification accuracy of pipes and void are around 90 %, 10 - 20 % improvement compared to previous research. On the other hand, in a practical point of view, positions of pipes should be estimated in the regions detected by 3D-CNN. Therefore, Kirchhoff migration, one of the most accurate and simple migration methods, was applied to cross-section images to estimate positions and even inclinations of pipes.

Various objects including manhole, joint, subsurface pipes and void were visualized in a 3D map completely automatic, fast and accurately by taking the advantages of simulations, AI and inverse analysis. The proposed algorithm is expected to have an large impact on decision makings of road administrators and practice of subsurface sensing of infrastructures.

Contents

Chapter 1: Introduction

1.1 Background

In recent years, aging of subsurface infrastructures is becoming an important issue. In Japan, as shown in Fig. 1.1, total distance of sewage pipes reaches up to 470,000 km. About 3 % of sewage pipes, 14,000 km is over 50 years old [1), 2)]. In 10 years about 12 % and in 20 years about 30 % are going to be 50 years old. Vast amounts of pipes developed in 1970s are rapidly aging. Furthermore, damaged old pipes cause subsurface void, leading to a road collapse. Road collapses caused by deteriorated pipes are about 3,000 per year, becoming a critical social problem [3) - 5)]. Road collapses are frequently observed in urban areas such as Tokyo district and other designated cities. In Tokyo district, as shown in Fig. 1.2, road properties such as water, sewage, gas and communication pipes account for about 40 % of all the causes. About 90 % of the accidents caused by road properties is due to sewage pipes. In previous research, sewage pipes over 30 years were pointed out to tend to cause road collapses [4)]. Proper understanding of the existence of old sewage pipes and void is important to prevent severe road collapses.

In terms of subsurface pipes, unknown pipes cause construction delays. Sometimes damaged pipes bring influences on surrounding areas. Accidents of buried objects by Japanese major construction companies were reported over 100 cases per year [6), 7)]. Accidents on roads including bridges account for about 22 % of all the accidents. Especially, subsurface sensing before road construction is important. Water pipes consist of about 52 % of all the accidents as shown in Fig. 1.3 (Fig. 1.4) [8)]. On the other hand, electricity and communication cables have a large economical influence on surrounding areas though the number of the accidents is relatively small. Damaged gas pipes threaten the safety. The main reason of the accidents is a human error in construction work. In not a few cases, actual locations of pipes are different from design drawings or not written. Accurate positions of subsurface pipes are vital information for road administrators.

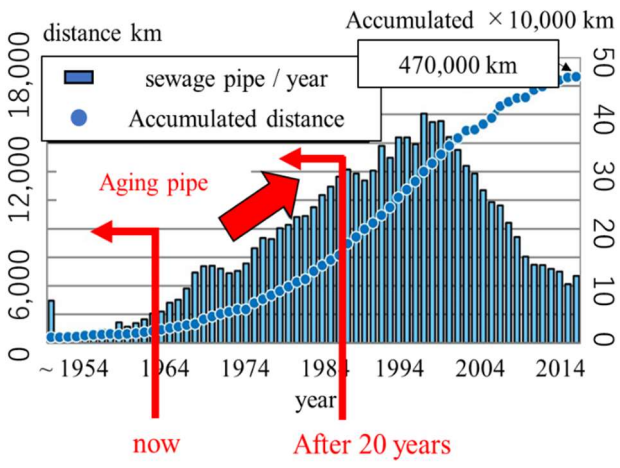


Fig. 1.1 Sewage pipe construction year [1)]

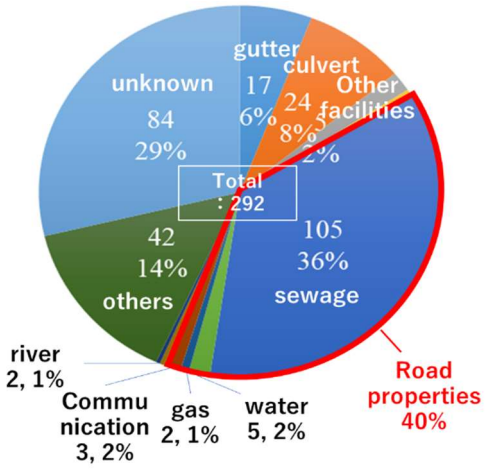


Fig. 1.2 Causes of road collapse [3)]

In terms of void, early detection before large a road collapse occur enables road administrators to conduct traffic regulation and take countermeasures to reduce an economic loss. Small void near subsurface pipes and substructures grows to several to several tens m size. Fig. 1.5 shows devastating incidents of road collapses [9) - 11)]. A road collapse in Fukuoka prefecture in 2016 was caused by a construction of subway (Fig. 1.5 (a)). A cave-in area was about 30 m by 30 m. Depth is about 15 m. There was no injury. Recovery work was completed in one week. The accident drew a large social concern. A road collapse in Tokyo district in 2008 was caused by sewage pipes under construction (Fig. 1.5 (b)). Soil came in the pipes. A cave-in area was about 5 m by 4 m. Depth is about 7 m. Motorcycle fell in to injure one person. Residents near the site evacuated because of the damaged water and gas pipes.

Non-destructive testing techniques for subsurface sensing such as radar are drawing a social interest. For example, cross-ministerial Strategic Innovation Promotion Program (SIP) spent 9.4 hundred million yen (about 9 million dollar) for the development of maintenance techniques and reduction of inspection cost in 2016. In terms of subsurface sensing, mainly about airport runway and pavement of port, void detection systems by radar were developed [12), 13)]. Subsurface sensing by radar is practiced by around 10 consultant companies in Japan. Guideline is being established [14), 15)]. Newly developed radar systems and Artificial Intelligence (AI) models for radar images are being introduced [16), 17)].

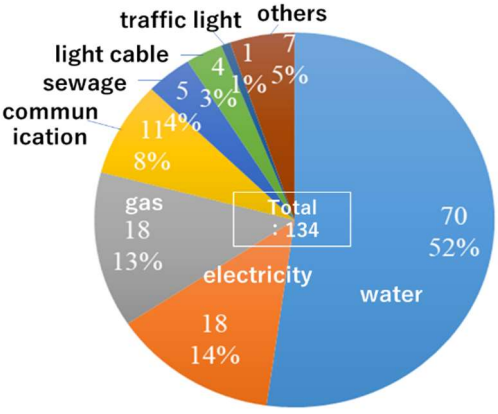


Fig. 1.3 Buried objects accident [3)]

(a) (b)

Fig. 1.4 Damaged subsurface pipes (a) Saitama prefecture(b) Nagasaki prefecture [8)]

(a) (b) (c)

Fig. 1.5 Road collapses (a) Fukuoka prefecture (b) Tokyo prefecture (c) Kyoto prefecture [9) - 11)]

1.2 Existing techniques

In terms of subsurface condition assessment, depending on target objects and anomalies, many non-destructive testing techniques are being developed. Most conventional method is visual inspection. Then, for the detection of buried objects and anomalies under the road, Falling Weight Deflectometry (FWD) and radar method are utilized. Each method has advantages and disadvantages. Taking the points into consideration, sometimes several methods are conducted together.

Visual inspection is conducted by inspectors on vehicle (Fig. 1.6) [18]. Visual inspection is targeted for anomalies on the surface of pavement. The advantage of visual inspection is the easiness to conduct without any special instruments. However, anomalies under the road cannot be inspected. Condition under the road can be inferred when anomalies such as void grow to deflect the road surface.

FWD is conducted by falling weight on the plate placed on the road surface [19) - 24)]. FWD measures deflection caused by impact force of the weight. The stiffness of the pavement above void decreases. Therefore, the existence of void can be inferred by the ratio of the deflection near the loading point and deflection distant from the loading point. The geometry of void can also be inferred. As shown in Fig. 1.7, there are portable type and vehicle-installed type. The advantage of FWD is the easiness to install without any special instruments. Data is quantitative and easily interpreted. However, as shown in Fig. 1.7, the method needs the installation of deflection sensors, resulting in traffic regulation. Furthermore, the time to measure the deflection is needed. A dense monitoring of the road surface is difficult.

Infrared camera method utilizes an infrared camera installed on the vehicle as shown in Fig. 1.8 (a) [25) - 27)]. The infrared camera measures the temperature of the road surface. Fig. 1.8 (b) shows a thermography image obtained by the camera. When delamination occurs just below the road surface, the temperature above the delaminated area is different from the surrounding area because of the insulation effect of air. Infrared camera method is simple and easy to interpret. The disadvantage is, the sufficient difference of the temperature of day and night is needed. Shady area cannot be evaluated. Moreover, detection accuracy depends on the size and depth of void. The effects of parameters were investigated in previous research [25), 26)].

Seismic method monitors inside soil, rocks and concrete by ultrasonic or elastic wave as shown in Fig. 1.9 [28) - 31)]. Waves reflect at the boundaries between soil layers, soil layer and buried objects, soil layer and void. A reflection amplitude depends on the difference of stiffness. The logic is similar to radar method. Propagating waves and medium are different. Typical operating frequency is lower than 1 MHz. The disadvantage of seismic method is, the results are difficult to interpret. The main target of seismic method is the estimation of a soil layer structure and properties. The assessment of integrality of concrete structures is also the target. There are several research about seismic method [28) - 30)]. The inactive detection of a crack inside concrete structures by emitted elastic waves is called Acoustic Emission (AE) [32)].

Electric method utilizes a current to measure the resistance of the medium between two electric poles [33)]. Electromagnetic (EM) method utilizes the inductive electromagnetic waves caused by coils. EM method estimates conductivity and permeability. There are few applications to civil infrastructures. In some cases, a

borehole is necessary to conduct the measurement.

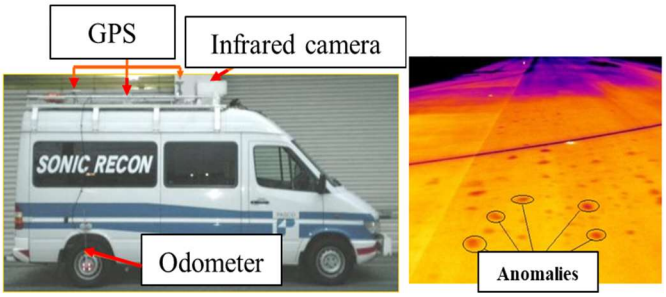
Most of the techniques above are campaign-type, which need sensor installation. On the other hand, this research focusses on Ground Penetrating Radar (GPR) method. In GPR method, antennas installed on vehicle transmit and receive electromagnetic waves to monitor the condition under the road [34) - 39)]. The target depth is about 1.5 m below the road surface, where most of subsurface pipes and void are considered to exist. The system is highspeed, up to 80 km/h. Therefore, GPR method does not need traffic regulation. The system scans under the road with dense measurement points. The target objects and anomalies include the both subsurface pipes and void. Especially, GPR method is recommended for the detection of subsurface pipes [14), 15)]. In the following sections, the disadvantages of GPR method were discussed in detail with up-to-date previous research.

Summarizing the discussions, the easiness to conduct and interpret, inspection time and cost, reliability of a method are the important points. GPR method has an advantage in the efficiency of measurement and high versatility of target objects. However, radar images are difficult to interpret, causing much inspection time and cost. Furthermore, detection accuracy is not considered to be enough. Recently, in the research field of GPR method, owing to the development of hardware and signal processing algorithms, these disadvantages are expected to be solved. The authors were focusing on GPR method for the detection of anomalies inside

Fig. 1.6 Visual inspection [18)]

(a) (b)

Fig. 1.7 Void detection by FWD (a) portable type (b) vehicle-installed type [23), 24)]



(a) (b)

Fig. 1.8 Infrared camera (a) camera-installed vehicle (b) thermography image [27)]

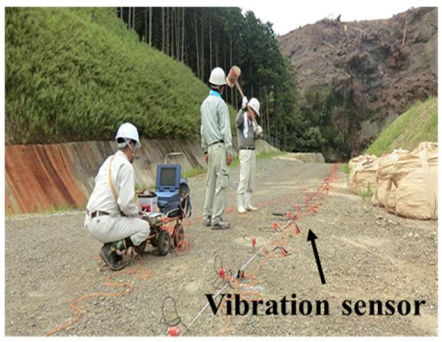


Fig. 1.9 Seismic method [31)]

the concrete decks of bridges [40), 41)]. GPR method has a potential to assess the condition inside the various infrastructures ranging from soil, asphalt pavement to concrete structures which is not enabled by conventional inspection methods.

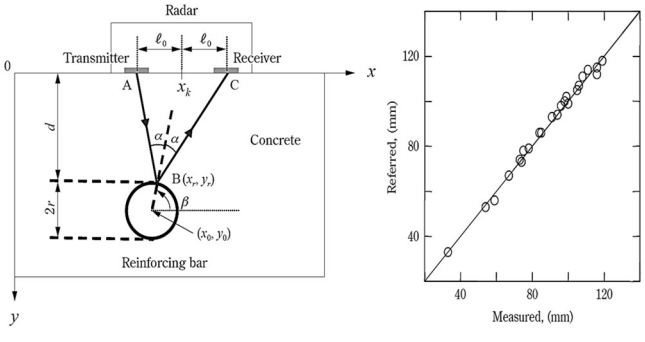
1.3 Principle and problems

1.3.1 Principle

GPR method was developed for subsurface sensing, whose history is old [34), 35), 42)]. First GPR was considered to be developed in 1910s. During World War II, GPR was utilized for geophysical purposes such as the condition of ground surface and soil. In 1970s, GPR for the exploration of the moon was developed. After 1970s, owing to the development of computer processor, GPR method prevailed in many research fields. Recently, GPR is utilized in archaeological exploration, geotechnical purposes, health monitoring of infrastructures, geophysical exploration, forensic medicine and planetary exploration. In terms of civil engineering, already in 1980s GPR was applied to the assessment of the integrality of a bridge deck [42)].

In terms of GPR method for civil engineering purposes other than subsurface sensing and integrality assessment of bridge decks, the estimation of concrete cover and asphalt layer thickness is an important application (Fig. 1.10 and Fig. 1.11) [43) - 48)]. In the applications, practically several mm accuracy is needed. In terms of concrete cover thickness, Tanaka and Okamoto (2010) focusses on Reinforced Concrete (RC) structures utilizing a portable type radar. Thickness was estimated with several % error considering the geometry of antennas and rebar related to the phase delay of measured signals. Loizos and Plati (2007) estimated the asphalt layer thickness utilizing a vehicle-installed type radar with several % error by inverse analysis.

In GPR method, common ground-penetrating radar is utilized. A measurement system consists of a radar transmitter/receiver, antennas, indicator, data logger and display (Fig. 1.12) [49), 50)]. Fig. 1.13 shows the principle of GPR method. Fig. 1.14 shows a conventional commercial system of GPR [51), 52)]. In GPR method,



(a) (b)

Fig. 1.10 Concrete cover thickness estimation (a) methodology (b) measured vs estimated [43)]

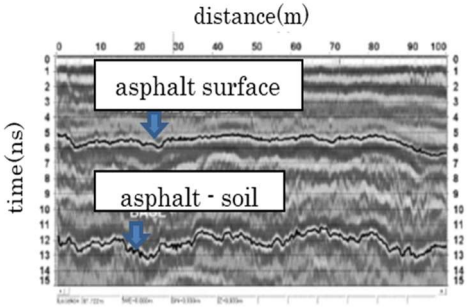


Fig. 1.11 Pavement thickness estimation by radar image [47)]

electromagnetic waves transmitted from a transmitter antenna reflect back at the boundaries such as soil and pipes or void. A receiver antenna measures the reflected waves. Reflection amplitude depends on the difference of the relative permittivity of two adjacent media. A signal at each measurement point is called A scan. A scan aligned in a certain direction is called B scan or a radar image. In the research, a multi-channel system was adopted. Obtained data was 3D. Considering the amplitude of each point, the dimension of the data can be regarded as 4D. A horizontal section at certain depth may have beneficial information.

1.3.2 Target objects and anomalies

A wide variety of objects are detectable by GPR method because any objects with the different relative permittivity from surrounding medium reflect waves. In the research, subsurface pipes and void are the targets. Other possible targets of the system may be a soil layer structure, soil condition such as water ratio, archaeological objects, landmine and mining. Landmine is several tens cm depth and size [53), 54)]. Archaeological objects and mining are several m to several tens m depth and size [36)]. Target depth and size are related to the operating frequency range and postprocessing algorithms.

Subsurface pipes are built under the road for the purpose of water, electricity, gas and communication (Fig. 1.15) [55)]. Subsurface pipes are categorized by materials. Typical materials are metal, plastic and concrete

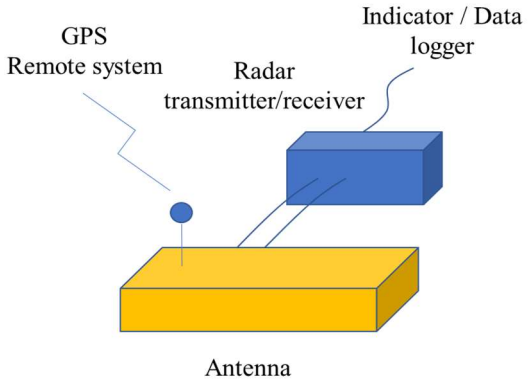


Fig. 1.12 Measurement system

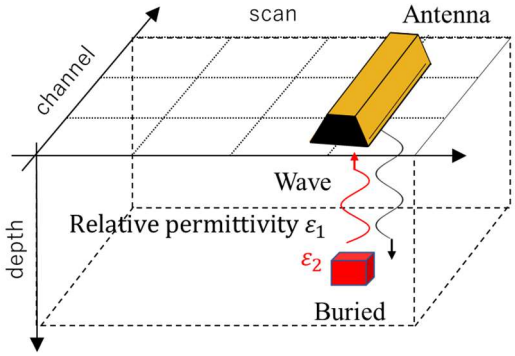
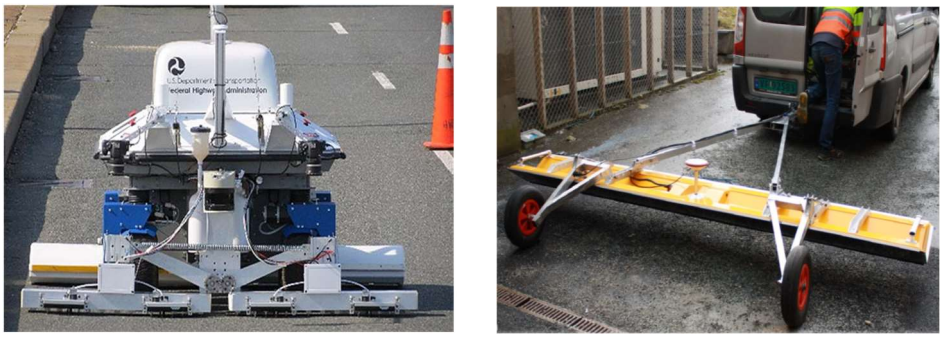


Fig. 1.13 Principle of GPR method



(a) (b)

Fig. 1.14 Commercial GPR system (a) RABIT (b) 3D-RADAR [51), 52)]

(hume pipes) [2]. In recent years, new types of materials such as alloy are adopted. Some conventional pipes are made of ceramic. In terms of electromagnetics, pipes can be classified into three categories, metal pipes, whose relative permittivity is infinitely large ($\approx 1e^7$), concrete pipes, whose relative permittivity is relatively large (5 - 10), ceramic and plastic pipes, whose relative permittivity is relatively small (1 - 2) [34) - 36), 39), 56)]. The relative permittivity of soil depends on water ratio and composition. Dried soil is about 2 - 5. Soil with water is above 10. Generally, relative permittivity increases at deeper region. Considering the target depth, the relative permittivity of soil is assumed 3 in void models and 2 - 20 in pipe models in the research. The relative permittivity of air is 1 and water is 80. Consequently, various types of waveforms of signals appear depending on the material of a pipe, condition inside the pipe, air or filled with water and surrounding soil. Diameter of pipes ranges from several cm to several tens cm. Large concrete pipes reach up to several m. In the research, the assessment of the integrality of pipes is out of focus. Some research focusses on the integrality assessment utilizing Pipe Penetrating Radar (PPR) [57), 58)].

Void is an air region inside soil, which frequently occurs near subsurface pipes and structures (Fig. 1.16) [4), 59), 60)]. Void consists of an air area and surrounding loosen soil area. Void is caused by soil coming into damaged pipes or deflection of subsurface structures. Void grows by the cycle of water supply and drainage. The relative permittivity of void is the same as air, 1 in dried condition. It may increase depending on the water content. The geometry of void is affected by the locations of pipes and properties of soil. Typically, void is a several tens cm to several m diameter dome-like shape. Large void reaches up to several tens m. Sometimes soil just below the pavement layer falls off forming an air layer. Void should be detected before a road collapse occurs. Several cm void is not important in terms of the management of road surface and difficult to be detected by inspectors. Therefore, larger than 0.5 m is the target of the research considering the definition of subsurface void [3), 4)].

1.3.3 Problems

In this section, three problems of GPR method is pointed out. The first problem is, radar images are usually difficult to interpret (Fig. 1.17). For example, in subsurface sensing by GPR method, a radar image consists of a direct-coupling wave, reflection from road surface, asphalt-soil-layer and soil-layer boundaries. The

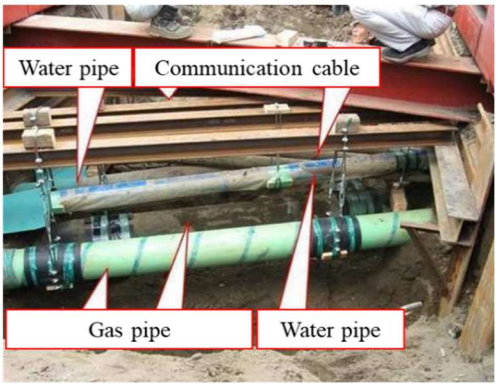


Fig. 1.15 Subsurface pipe[55)]

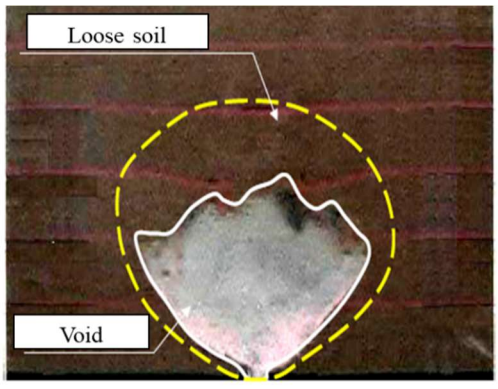


Fig. 1.16 Experimental void model [60)]

radar image is also affected by objects such as rocks and other artificial objects, which is called clutters. Because of these undesired signals, the detection of subsurface pipes and void is not an easy task. A radar image is interpreted by skilled inspectors with special preprocessing algorithms and image display programs. Usually, one measurement takes several months and million-yen (tens-thousand-dollar) inspection time and cost. Therefore, an automatic detection algorithm may reduce significant inspection time and cost.

The second problem is, a large amount of data is collected in one measurement. Total distance is ranging from several hundreds m to several tens km. One section of measured data is several hundreds MB to several hundreds GB. Ordinarily, a multichannel measurement system is utilized. If a 29 channel system is adopted, the total volume of all the channels is up to several TB (Fig. 1.18). When UHF band is utilized, assuming the relative permittivity of soil is 3, the resolution in a depth direction is several cm. Larger the relative permittivity, higher the resolution is. The resolutions in scan and channel directions are also several cm. Measured data is dense cm-order resolution 3D data. Inspectors are trying to detect buried objects and anomalies from the whole amount of data image by image. Therefore, a highspeed algorithm is necessary for the detection of target objects and anomalies by GPR method.

The third problem is, in terms of void detection, the number of measurement data by GPR method is few. Accordingly, machine learning and deep learning approaches are difficult because the accuracy depends on the number of training data. As stated in Section 1.1, road collapses by deteriorated pipes are about 3,000 per year. GPR method is not common in practice. It is not realistic to collect void data before a road collapse occurs. Therefore, in previous research an experimental field is constructed by embedding Styrofoam in soil. However, the number of data is not enough even by experimental field data. For example, the number of training data needed for deep learning is several hundred to several thousand. It is not feasible to train models only by measurement and experimental field data. Moreover, the reflection patterns of actual void are various. The effect of the number of training data on the accuracy of pipes detection is also not clear.

Summarizing the above discussions, for the detection of subsurface pipes and void by GPR method, an automatic, highspeed and accurate algorithm is needed which can be applied to a large amount of

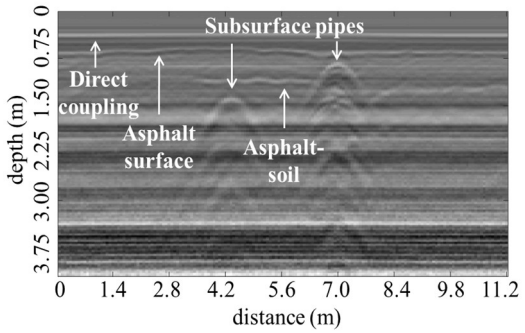


Fig. 1.17 Interpretation of radar image

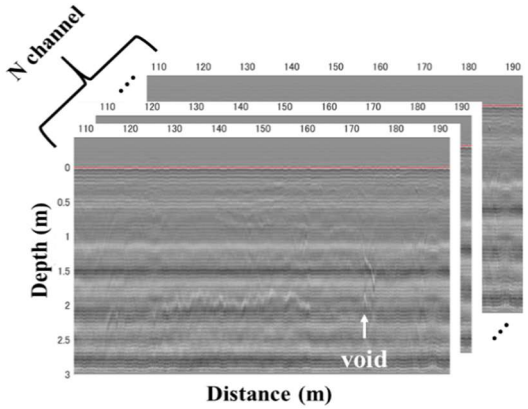


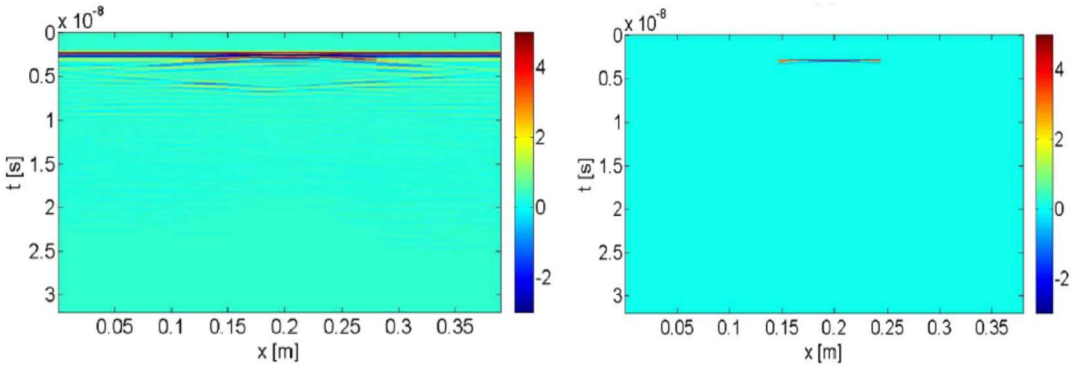
Fig. 1.18 Measured data

measurement data. Inspectors manually detect buried objects and anomalies form each radar image. Therefore, inspection time and cost are enormous. Furthermore, the algorithm should be reliable enough to replace the manual inspection. In the development of the automatic detection algorithm, training data is needed. Especially, the number of void data is not enough. Therefore, training data is produced by electromagnetic simulations. Then, an AI model is developed utilizing the produced data.

1.4 Previous research

1.4.1 GPR signal processing

Signal processing applied to radar signals for subsurface sensing are in common with other radar systems and seismic wave method [49), 50), 61)]. In terms of seismic wave method, Yilmaz (1987) discussed in detail the conventional algorithms [62), 63)]. Signal processing is divided into two parts, preprocessing and postprocessing. Postprocessing is categorized into common filtering process (filtering), object detection and localization (detection) and parameter estimation. Detection is in some cases manual and in other cases automatic. Estimated parameters are often related to physical parameters such as soil relative permittivity and diameter



(a) before clutter removal (b) after clutter removal

Fig. 1.19 Detection of steel bar inside soil by clutter removal [64)]

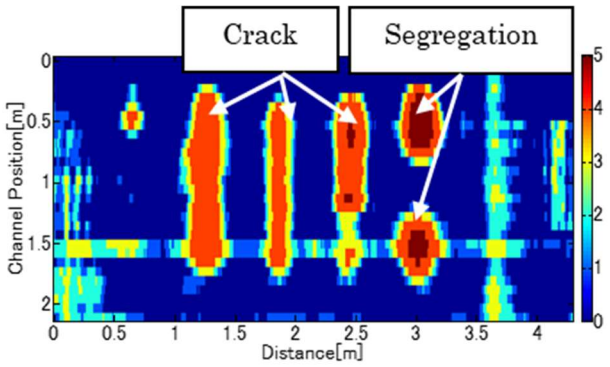


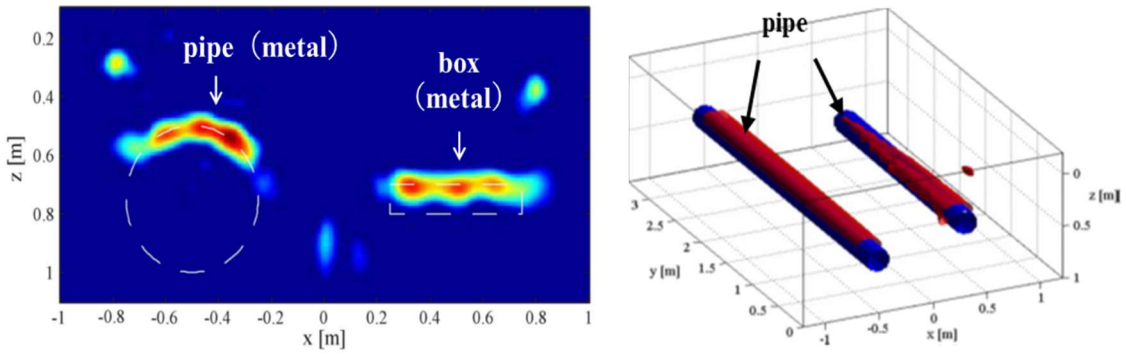
Fig. 1.20 Deterioration map of bridge deck model [33)]

of pipes. In some cases, detection and parameter estimation are conducted simultaneously. Appropriate algorithms are different depending on targets.

Preprocessing is fundamental algorithms necessary for the display of a radar image. When a radar system adopts frequency modulated method, obtained frequency domain data should be converted to time domain data by inverse Fourier transform. Additionally, the extraction of target sections, upsampling and frequency filtering may be applied. To convert from a time range to a distance range, the relative permittivity of a propagating medium should be estimated to calculate propagation velocity. Frequently the accurate estimation of permittivity is difficult. Targets whose depth is known is utilized or a certain value is assumed. Preprocessing algorithms are essential and installed on most of commercial systems.

Postprocessing is intended for amplifying desired signals, detecting target objects and estimating the parameters of the objects. Sometimes the combination of algorithms is adopted. Especially, in automatic detection and parameter estimation many research is conducted. Clutter removal is to remove undesired signals caused by the reflections from ground surface and media boundaries (Fig. 1.19) [64), 65)]. BackGround Removal (BGR) is one of the typical clutter removal algorithms. Gain correction is to compensate for the attenuation of input waves in a lossy medium. Deconvolution is to increase the clarity of signals. In measurement process, transmitted impulse waves are significantly altered from pulsed waves. The small reflections from target objects are often covered by other reflections. Deconvolution enhances the peaks of desired signals by suppressing sidelobes. Migration is to increase the clarity in horizontal directions by considering the characteristics of spreading waves, medium and antenna directivity.

In automatic detection, the existence and positions of target objects are automatically estimated. Sometimes the distribution of probability is mapped. In the following section, previous research about the automatic detection of subsurface pipes and void is introduced. Other than pipes and void, bridge deck deterioration was mapped in Fig. 1.20 [41)]. After parameter estimation, the quantitative evaluation of target objects is possible. Previous research about parameter estimation is also introduced in the following section.



(a) (b)

Fig. 1.21 Reconstruction of buried objects (a) 2D section (b) 3D geometry [66)]

Other than subsurface pipes and void, the estimation of the distribution of electromagnetic parameters of soil, positions and diameter of rebars in RC structures, asphalt layer thickness are also considered as parameter estimation [43), 44), 47), 48)].

1.4.2 Previous research

In this section, previous research about the automatic detection and parameter estimation of subsurface pipes and void is discussed. In terms of subsurface pipes, 4 topics are introduced. The first point is, pipe shape reconstruction by migration and other inverse analysis methods such as tomographic reconstruction and deconvolution [66) - 69)]. Migration methods are compared in Section 7.2. In migration methods, integral operation is applied on an image assuming a hyperbolic reflection pattern. Soldovieri *et al*. (2008) derived the theoretical formulas of tomographic reconstruction applying Kirchhoff approximation and Truncated Singular Value Decomposition (TSVD). The algorithm was validated by experimental field data as shown in Fig. 1.21. Li (2014) proposed the algorithm to estimate reflection coefficients by blind deconvolution, which assumes an input signal is unknown. However, previous research focused on several m to several tens m experimental field data, implying migration methods and inverse analysis need high calculation cost. Therefore, it is not feasible to simply apply these methods on measurement data whose distance is several hundred m to several tens km.

The second point is, the detection of hyperbolic reflection patterns by Hough transform, Neural Network (NN) and other pattern recognition approaches [70) - 76)]. Simi *et al.* (2008) developed Hyperbola Automatic Detection Algorithm (HADA) based on image processing and Hough transform (Fig. 1.22). Gamba and Lossani (2000) detected the hyperbolic reflection patterns of subsurface pipes by the combination of image processing and simple 1 layer NN model. Al-Nuaimy *et al.* (2000) detected the regions of target reflection patterns by the combination of window function and NN (Fig. 1.23). Pasolli *et al.* (2009) proposed the algorithm to extract the hyperbolic reflection patterns by the combination of pattern matching and Genetic Algorithm (GA). The algorithm was validated by electromagnetic simulations. Support Vector Machine (SVM) was adopted to estimate the materials of detected objects. Shihab and Al-Nuaimy (2005) proposed

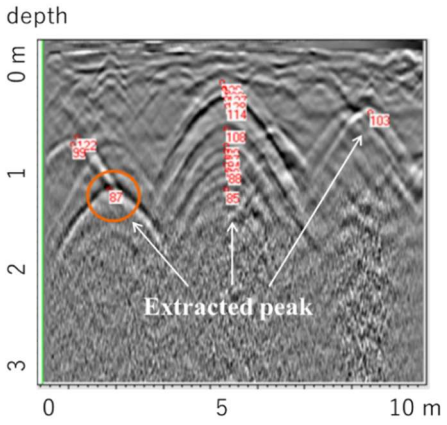


Fig. 1.22 Pipe peak extraction by HADA [70)]

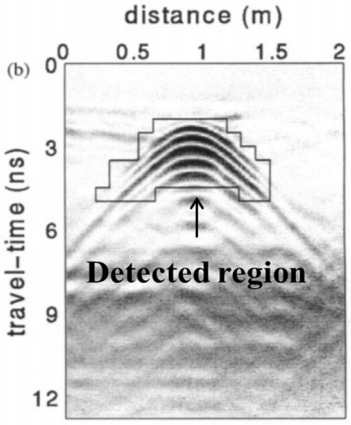
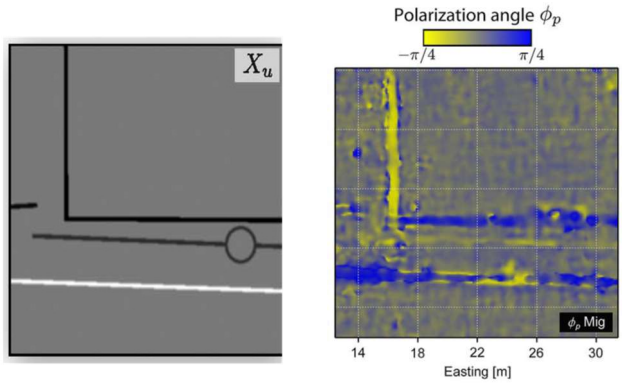


Fig. 1.23 Reflection region by NN [72)]

the algorithm to estimate the diameter of pipes by pattern matching. Other research utilized optimization methods such as Multi Agent (MA) and Bayes approaches [77), 78)]. The detection accuracy of these algorithms was about 70 - 80 %. Therefore, the detection accuracy of previous research is not considered to be enough. The advancement of a detection algorithm by a deep Convolutional Neural Network (CNN) model and utilization of 3D reflection patterns considered in manual inspection may increase the accuracy. Though the estimation of the materials and diameter of pipes is out of scope, the previous algorithms may be applicable after the detection of the pipes reflection patterns.

The third point is, the estimation of the direction of pipes by utilizing antenna polarization [79), 80)]. Boniger and Tronicke (2012) adopted the antennas polarized in two orthogonal directions. The research proposed a new index based on Principal Component Analysis (PCA) to map the polarization corresponding to the directions of pipes as shown in Fig. 1.24. Shaari *et al.* (2010) proposed a system model representing antenna polarization to simulate signals obtained by electromagnetic simulations. In the research, a common commercial system which can detect both the transverse and longitudinal pipes was adopted. Therefore, the



(a) (b)

Fig. 1.24 Mapping of material by polarization (a) design plan (b) estimated map [79)]

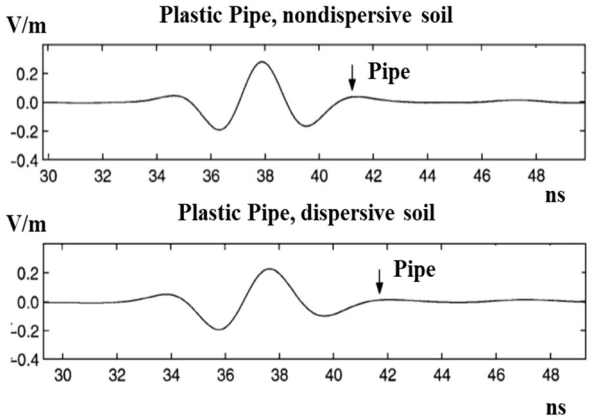


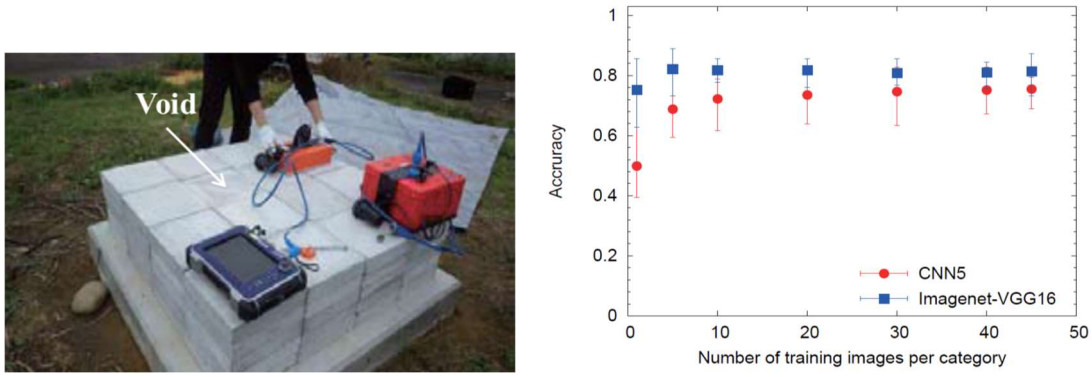
Fig. 1.25 Attenuation in dispersive attenuation [82)]

estimation of not only the existence but also the directions of pipes was targeted in this research. In the research, the directions of pipes were estimated more simply with less calculation cost by the three-category classification of 3D-CNN.

The fourth point is, the reproduction of reflection patterns by electromagnetic simulations [81), 82)]. Warren *et al.* (2016) developed an electromagnetic simulation platform specialized in GPR environment by Finite Difference Time Domain (FDTD) method, which is called gprMax. The pattern of a 3D bowtie antenna model in inhomogeneous soil was estimated. Wang and Oristaglio (2000) estimated the reflection patterns of pipes by 3D-FDTD. The attenuation of soil layer had frequency dependence (Fig. 1.25). Some previous research also provided platforms for electromagnetic simulations by FDTD method. In the research, the training of pipes was conducted by measurement data. On the other hand, because there was few void data, the training data of void was produced by electromagnetic simulations by FDTD method, gprMax.

In terms of void, Sonoda *et al.* developed a CNN model for the classification of void radar images [83) - 86)]. Sonoda and Kimoto (2018) proposed an FDTD void model to train a CNN model by the produced data. The classification accuracy of some conventional CNN architectures such as VGG16 was compared. The comparison of conventional CNN architectures is discussed in Section 6.5.2. The classification accuracy of experimental field data was about 80 % (Fig. 1.26). Park *et al.* (1996) developed an inverse analysis method by considering the geometries of antennas and void to reconstruct a void shape. However, there is no research about the classification accuracy of real measurement void data. Again, classification accuracy may increase by utilizing the 3D reflection patterns of void. In recently years, owing to the advancement of hardware and development of Graphics Processing Unit (GPU), 3D-FDTD becomes possible in reasonable calculation time and cost. However, even by GPU, simply producing several hundred training data by 3D-FDTD is not feasible.

Summarizing the above discussions, in terms of subsurface pipes, the calculation cost of inverse analysis is enormous. Therefore, in the research, after pipes and void regions were localized by AI model, migration



(a) (b)

Fig. 1.26 void image classification by CNN (a) experimental field (b) accuracy [83)]

method was applied to the detected regions. The utilization of a deep CNN model and 3D reflection patterns may increase the accuracy. In terms of void, the accuracy of real measurement data should be evaluated. Also in void detection, 3D reflection patterns may increase the accuracy. However, the application of 3D-FDTD is not feasible considering the calculation cost.

1.5 Objective and overview

1.5.1 Objective

From the discussions of previous sections, the research objective is "development of 3D automatic subsurface pipes and void detection algorithm by deep learning and signal processing from radar images." The methodologies adopted in the research are briefly summarized below.

The target of the research is subsurface pipes and void under the road. In the research, a large number of pipe data was collected by inspectors, though the number of void data was too small for the training of a deep learning model. Therefore, in terms of void, training data was produced by electromagnetic simulations by FDTD method. In the development of a 2D-CNN model, training data can be reproduced by 2D-FDTD. However, the reproduction of 3D training data by 3D-FDTD method is not feasible because of calculation cost. Therefore, by the combination of 2D-FDTD models and interpolation method, the methodology to reproduce 3D void data by 2D-FDTD was proposed.

In the application of pattern recognition approaches such as AI models, false signals behind manhole and joint caused by the multiple reflections between antennas and objects were falsely detected as pipes and void. In the research, an automatic manhole and joint detection algorithm from surface reflections was developed by SVM. Detected regions were removed from the target regions of pipes and void.

The reflection patterns of pipes and void was accurately detected by a deep CNN model. The 3D reflection patterns were utilized by 3D-CNN to increase the classification accuracy. The developed 3D-CNN model can detect not only the existence but also the directions of transverse and longitudinal pipes by three-category classification. After the detection, for practical mapping, especially the positions of pipes should be estimated in the detected regions. Box-by-box search of 3D-CNN was applied to the whole data to localize the possible regions. Then, migration method was applied to the detected regions to estimate the pipes' positions. Kirchhoff migration, one of the simplest and most accurate migration methods, was applied to the cross-section images of the detected pipes.

In the research, by the combination of electromagnetic simulations, AI and inverse analysis, a highspeed and high-accuracy automatic detection algorithm for manhole, joint, pipes and void was proposed. A final output is a 3D map of each target, which is expected to have a large impact on GPR practice for civil engineering purposes.

1.5.2 Overview

The thesis consists of 8 chapters. Fig. 1.27 shows the flow chart of the contents.

"Chapter 1: Introduction" stated general background, existing techniques and definition of target objects and anomalies. Then, the problems of GPR method was introduced. The objective and overview of the

research were also introduced.

"Chapter 2: Measurement system and data" stated the configurations of the measurement system and measurement conditions. Then, the programs and hardware adopted in the research were introduced. The general characteristics of pipes and void data collected by inspectors were qualitatively analyzed. The experimental field site was also introduced.

"Chapter 3: Electromagnetic characteristics" introduced the fundamental electromagnetic characteristics of subsurface sensing environment. The main factors affecting the detection accuracy of pipes and void were discussed in terms of electromagnetics and signal processing. The flow chart of the basic preprocessing algorithms applied on the training data was explained.

"Chapter 4: Manhole and joint detection by SVM" proposed the manhole and joint detection algorithm by SVM. The methodology to produce training data and indexes to evaluate a trained model was discussed. The classification accuracy of the optimized SVM model was compared to a CNN model.

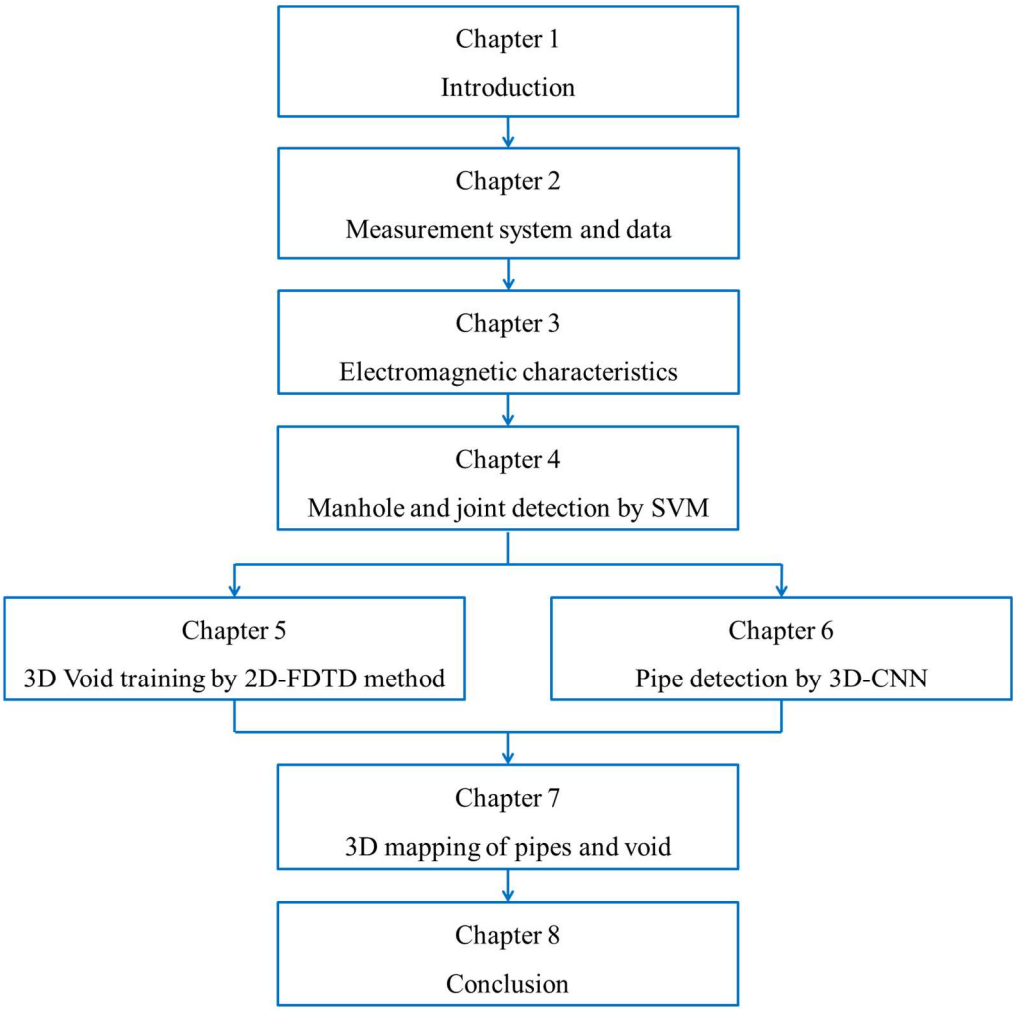


Fig. 1.27 Flow chart of the research

“Chapter 5: 3D Void training by 2D-FDTD method” proposed the 2D-FDTD void model. The effects of the characteristics of antennas, void and measurement environment on reflection patterns were evaluated. The methodology to reproduce 3D void data by 2D-FDTD method was proposed. The proposed methodology was validated by measurement data. Based on the proposed method, a 3D-CNN model was developed. The classification accuracy of the 3D-CNN model was compared to a 2D-CNN model.

“Chapter 6: Pipes detection by 3D-CNN” proposed a deep 3D-CNN model to accurately classify the 3D reflection patterns of pipes. The training method and effect of the number of training data were discussed. Preprocessing algorithms and model architecture were optimized by a 2D-CNN model. Then, the sizes of convolution filters and input data of 3D-CNN were optimized for each direction pipe. The classification accuracy of 3D-CNN was compared to each direction pipe 2D-CNN model. Finally, a three-category classification model for two-direction pipes was developed.

“Chapter 7: 3D mapping of pipes and void” proposed an estimation method of the positions of each target. Manhole and joint detected by SVM were visualized by Hough transform. The 3D-CNN model of pipes and void was applied on the whole region except for the regions of manhole and joint. Migration method was applied to the cross-section images of the detected 3D regions. The positions of pipes and void were estimated and visualized in a 3D map.

“Chapter 8: Conclusion” summarized the conclusions and impacts of the research, flow chart of the proposed algorithm and future work.

Chapter 2: Measurement system and data

In this chapter, a measurement system, data and computing environment are stated. In the research, a detection algorithm was developed by the measurement data collected by inspectors. The general characteristics of subsurface pipes and void were qualitatively analyzed by the measurement data. In the research, an experimental field was newly constructed by the authors. In the research, two types of data were utilized. “measurement data” was used for the training of AI models and “experimental field data” for the validation of a migration method.

2.1 Measurement system

In the research, a vehicle-installed multi-channel system manufactured by 3D-Radar was utilized (Fig. 2.1) [51), 87)]. Fig. 2.2 shows a transmitting/receiving antenna and radar transmitter. 3D-Radar is targeted for subsurface sensing such as detecting buried objects and estimating asphalt layer thickness. Typical to GPR method, the system consists of an antenna, radar transmitter, indicator, data logger and display. The system records a position by a GPS and odometer. Antennas are usually installed on the front or rear of a vehicle. The antennas are covered with an electromagnetic shield to prevent electromagnetic waves leak out of the system to meet restrictions of measurement environment.

The antenna consists of transmitting/receiving antenna units aligned orthogonal to a scan direction as shown in Fig. 2.3 (multi-channel). Combination patterns of transmitting/receiving antennas are switched to measure electromagnetic waves at each measurement point (bistatic method) [87)]. In the research, 29 measurement points (29 channels) were adopted. Channel distance is 0.075 m. Accordingly, a width in a channel direction is about 2 m, covering a half of a lane at one measurement. The antennas are placed with 20 cm distance from the ground surface (air-coupled method), which are inclined to make V-shape. The

Fig. 2.1 Measurement system

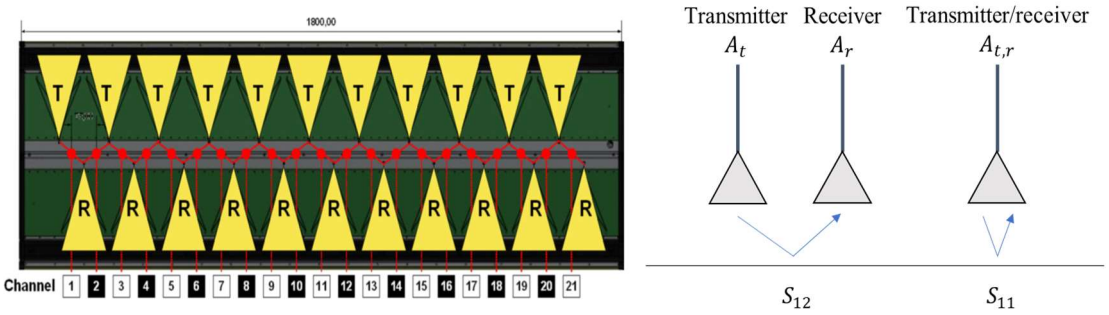
(a)

(b)

Fig. 2.2 Radar system (a) transmitter/receiver antenna (b) radar transmitter [87)]

details of antenna configurations are not published. Therefore, the type and geometry of the antennas are not known. In the research, the effect of antenna directivity was discussed in Section 5.3.2 by electromagnetic simulations with an array antenna model.

3D-Radar adopts Stepped Frequency Continuous Wave (SF-CW) method [88), 89)]. In SF-CW method, an operating frequency range and step width are the important parameters. In the research, 50 MHz - 3030 MHz (Very High Frequency, VHF and Ultra High Frequency, UHF band, Fig. 2.4), which is utilized for FM radio, mobile phone and telecommunications, was adopted. The step width was 20 MHz [90)]. Data with a smaller step width was downsampled to unify the step width of measurement data. Wider an operating frequency range is, higher the resolution in a depth direction is. Furthermore, smaller a step width is, larger the maximum depth becomes. However, because the time to sweep the frequency range increases, scan pitch and vehicle speed should be limited. Scan pitch is 0.07 m. Vehicle speed is 80 km/h. In a practical sense, the operating frequency range and step width stated above was reasonable considering the desired resolution, maximum depth and vehicle speed. In the research, the bandpass filter of 200 MHz - 3 GHz was applied to the measured data because the nominal frequency range of antennas was that range. The raw text data shown



(a) (b)

Fig.2.3 Measurement pattern (a) antennas (b) bistatic and monostatic method [87)]

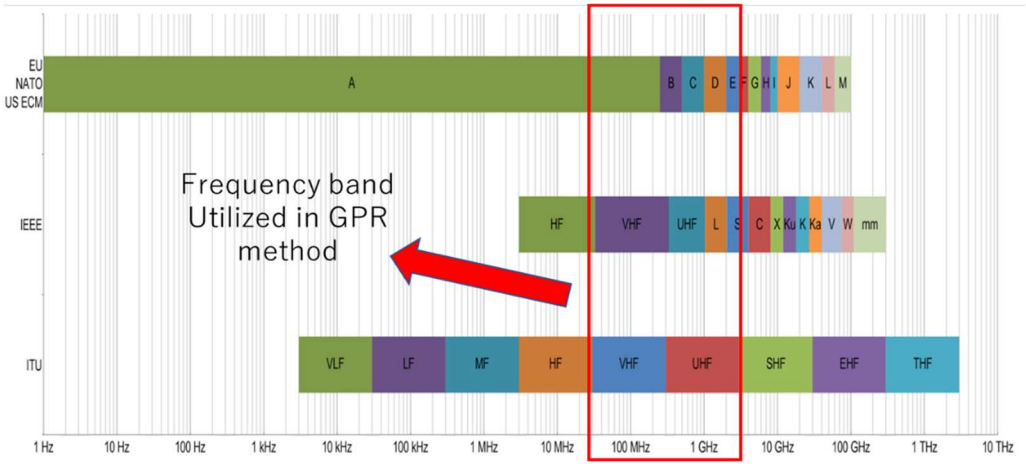


Fig. 2.4 Utilized frequency band, IEEE, comparison to other codes [90)]

in Fig 2.5 consists of a real and imaginary part of a Fourier coefficient at each frequency, channel and scan number. The text data was converted to 3D frequency domain data. Then, after preprocessing, the data was converted to 3D time domain data and displayed as radar images.

3dr Examiner is an image display program of 3D-Radar shown in Fig. 2.6 [91]. Fig. 2.6 (a) displays a measurement course by a GPS and odometer. Fig. 2.6 (b) displays each cross-section image of the obtained data. 3dr Examiner shows transverse, longitudinal and horizontal cross-section images in real time after fundamental preprocessing. 3dr Examiner has postprocessing algorithms such as automatic tracing of layer boundaries and estimation of medium permittivity by Common Mid-Point (CMP) method. In the research, the raw text data shown in Fig. 2.5 was extracted to evaluate the effects of preprocessing algorithms.

2.2 Measurement data

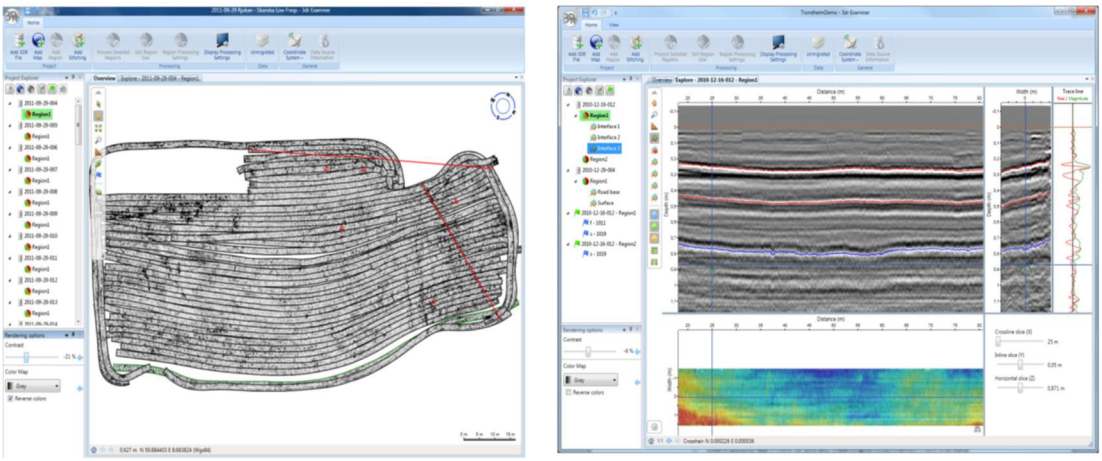
2.2.1 Configurations

In the research, pipes data measured in Nagano prefecture, void data in Aichi and Nagano prefectures and

```
1 X(Sample ID),Y(Channel No.),Frequency(MHz),Spectrum(Real),Spectrum(Imaginary)
2 0,0,50,0.01497118,0.008478344
3 0,0,60,0.01762769,0.006624783
4 0,0,70,0.0143639,0.003750249
5 0,0,80,0.01299973,0.002758577
6 0,0,90,0.01459013,0.005739014
7 0,0,100,0.03181823,0.01288504
8 0,0,110,0.02612373,0.001131147
9 0,0,120,0.02233676,0.005903714
10 0,0,130,0.03079997,0.006471218
11 0,0,140,0.03615966,0.000364047
12 0,0,150,0.05467868,-0.009314926
13 0,0,160,0.04507872,-0.02309084
14 0,0,170,0.03152865,-0.02865927
15 0,0,180,0.04425203,-0.02225877
16 0,0,190,0.04406048,-0.04726332
17 0,0,200,0.0421629,-0.04768755
18 0,0,210,0.02309299,-0.06855573
19 0,0,220,0.001210167,-0.05361544
20 0,0,230,0.004350244,-0.04512416
```

Imaginary part of Fourier coefficient
Real part of Fourier coefficient
Frequency
Channel No.
Scan No.

Fig. 2.5 Measured raw text data

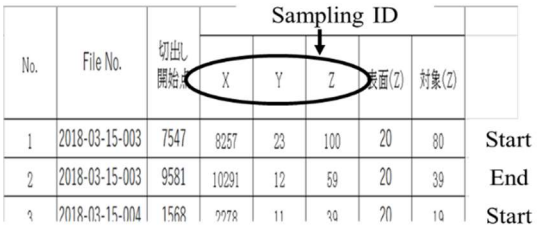


(a) (b)

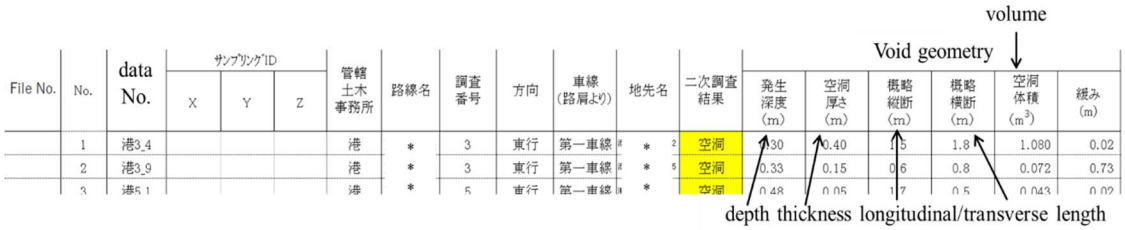
Fig. 2.6 3dr Examiner (a) GPS data (b) radar images [92]

manhole and joint data in Nagano prefecture was used. Fig. 2.7 shows the results of the inspection. Inspectors manually extracted hyperbolic reflection patterns of pipes and void from radar images. The points of apexes of hyperbolics and file numbers were recorded. In terms of transverse pipes, scan and depth numbers were recorded. Transverse pipes were assumed to lie in all the channels. In terms of longitudinal pipes, start and end points, channel and depth numbers were recorded. Longitudinal pipes were assumed to lie linearly between start and end points. In terms of void, assuming an ellipsoid or cuboid body, the points of apexes of hyperboloid reflection patterns was recorded. Scan, channel and depth numbers were the parameters. In terms of void, an excavation survey was conducted to obtain the actual depth and geometry of void. Thickness, transverse, longitudinal lengths and volume of void were measured. Manhole and joint were detected from surface reflections. The scan and channel numbers of centers of manhole were recorded. The scan number of joint was recorded assuming joint passes through all the channels.

Pipes, manhole and joint data utilized in the research was collected by inspectors. Void data was collected by inspectors and confirmed by the excavation survey. Therefore, in terms of pipes, manhole and joint detection, there is a possibility that inspectors miss target objects (type I error) and falsely detect (type II error). In terms of void, only type I error can occur. In the research, the inspection results are assumed to be true. The accuracy of the developed classification algorithms was evaluated by the inspection results. The



(a)



(b)

No.	dataID	x(scan)	y(ch)	z(depth)	manhole	joint
1	2017080214-N36E137-01-001000	146	18	50		
2	2017080214-N36E137-01-001000	155	4	50		1
3	2017080214-N36E137-01-001000	175	14	50		1

(c)

Fig. 2.7 Examples of detection results by inspectors (a) longitudinal pipes (b) void (c) manhole and joint

objective of the research is to develop a reliable algorithm which can replace the decision of skilled inspectors. To be precise, the existence of pipes and no pipe sections in measurement data cannot be confirmed. The detection accuracy of actual pipes and void can be evaluated by the experimental field data, which was shown in Section 7.4.

Table 2.1 - Table 2.4 show the numbers of pipes, void, manhole and joint divided for each measurement date. Table 2.1 and Table 2.2 show the numbers of pipes and 2D training images. The number of transverse pipes is 2,641. Manhole and joint are 6,296 and 2,164 respectively. The numbers of transverse pipes, manhole and joint are considered to be enough even for the training of deep learning models. The number of longitudinal pipes 730 is relatively small. However, the number of 2D training images is comparable with transverse pipes because the lengths of longitudinal pipes were relatively large, several m to several tens m. The classification accuracy of multiple-category classification models is affected by the number of test data of each category. Therefore, in this thesis, detection and false detection rates were shown for each category by a confusion matrix. On the other hand, the number of void is 88. Two thirds were used for training. Therefore, training data was around 50, which is not enough for deep learning. An excavation survey was conducted for 35 data among 88 data collected in 2017/4/27 - 2017/4/28. The total distance of pipes data was about 230 km. Void was about 10 km. Manhole and joint were 350 km. The inspectors detected 1 pipe per 100 m, 1 void per 150 m and 1 manhole and joint per 50 m to 150 m. Pipes and void measurement may

Table 2.1 Transverse pipe

date	Pipe (image num.)
2018/3/15	1,040（30,160）
2018/3/19	709（20,561）
2017/8/4	204（5,916）
2017/8/5	688（19,952）
total	2,641（76,589）

Table 2.2 Longitudinal pipe

date	Pipe (image num.)
2017/8/2	730（33,759）

Table 2.3 Void

date	Void (image num.*)
2017/4/27	18（310*）
2017/4/28	17（284*）
2018/3/9 - 2018/3/13	44 (1276*)
2018/5/16 - 2018/5/17	9 (261*)
total	88（2131*）

*channel direction

Table 2.4 Manhole and joint

date/class	Manhole	Joint
2017/8/2	443	91
2017/8/3	1,141	301
2017/8/4	1,219	411
2017/8/5	1,723	664
2017/8/6	579	237
2017/8/7	1,191	460
total	6,296	2,164

be conducted in the areas where target objects exist with high possibility. The numbers of manhole and joint were considered to correspond to the real situation.

2.2.2 Analysis results

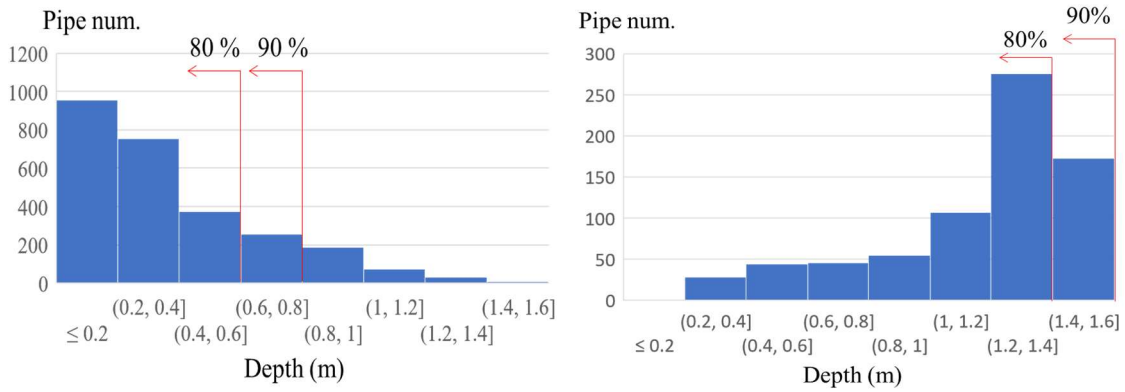
To discuss the general characteristics of pipes and void, the inspection results were analyzed. Fig. 2.8 shows the histograms of the depth of the detected transverse and longitudinal pipes assuming the relative permittivity of soil is 20. The depth of longitudinal pipes was represented by the average of the depth of start and end points. Depth d was derived by a sampling point from the road surface s, the relative permittivity of soil ε_s, light velocity c, maximum frequency f_{max} considering the round trip of propagating waves.

$$d = sc/\sqrt{\varepsilon_s}\,/f_{max}/2 \tag{2-1}$$

In the research, $c = 3 \times 10^8$ m/s and $f_{max} = 3$ GHz. ε_s of actual soil is ranging from 2 to 20. d at $\varepsilon_s = 2$ is three times larger than $\varepsilon_s = 20$. ε_s varies at each measurement point and is not uniform in a depth direction. Generally, relative permittivity in deeper regions is larger because of water content. The histograms shown in Fig. 2.8 may be affected by these factors.

From Fig. 2.8, the depth of most pipes was about 0.2 m for transverse pipes and 1.2 m - 1.4 m for longitudinal pipes. The maximum depth of the measurement system was about 1.5 m. Considering the target depth range, the results were consistent [39), 87)]. The depth widely ranged from 0.2 m - 1.5 m. The histograms shown in Fig. 2.8 was affected by the actual depth of pipes and characteristics of the measurement system. Because of the attenuation of electromagnetic waves, it is difficult to detect pipes in deeper regions. Therefore, there may be more pipes in deeper regions than detected.

In terms of longitudinal pipes, Fig. 2.9 shows the length, channel number and inclination angles in horizontal and depth directions. Lengths of pipes and inclination angles were calculated from the extracted start and end points. From Fig. 2.9, the lengths of most longitudinal pipes was about 0.50 m. However, the lengths of some pipes were up to 10 m. The maximum length was about 40 m. There was no tendency in the



(a) (b)

Fig. 2.8 Pipes depth (a) transverse pipes (b) longitudinal pipes

histogram of channel number. The inclination angles of most pipes were within ±2° in both directions, indicating the pipes are precisely placed parallel to lane and horizontally. However, the angles of some pipes were over ±20°. The maximum angle was about 47°. At 20°, a pipe shifts about 40 m per 100 m section. The inclinations of pipes are not negligible in the case of longitudinal pipes.

Fig. 2.10 shows the histograms of depth, channel number, transverse and longitudinal lengths, thickness and volume of void. From Fig. 2.10, the depth of most void was around 0.5 m. The depth uniformly ranged from 0.2 m to 0.65 m. In a channel direction, void lay around 13 to 17 channels, corresponding to the middle of the lane. The lengths in transverse and longitudinal directions was around 1 m. Thickness was around 0.5 m. Lengths widely ranged from 0.5 m to 1.5 m. Thickness was around 0.2 m to 0.5 m. Volume ranged from 0.1 m^3 to 0.3 m^3. Large void is about 1 m^3. Assuming an ellipsoid body with transverse and longitudinal lengths 1 m and thickness 0.3 m, volume will be about 0.16 m^3. The shape of most void is considered to be close to an ellipsoid body. However, some void is up to 1 m^3. In a geotechnical point of view, a cuboid body can be assumed which occurs when the soil just below the road surface falls off. From Fig. 2.10 (e), the

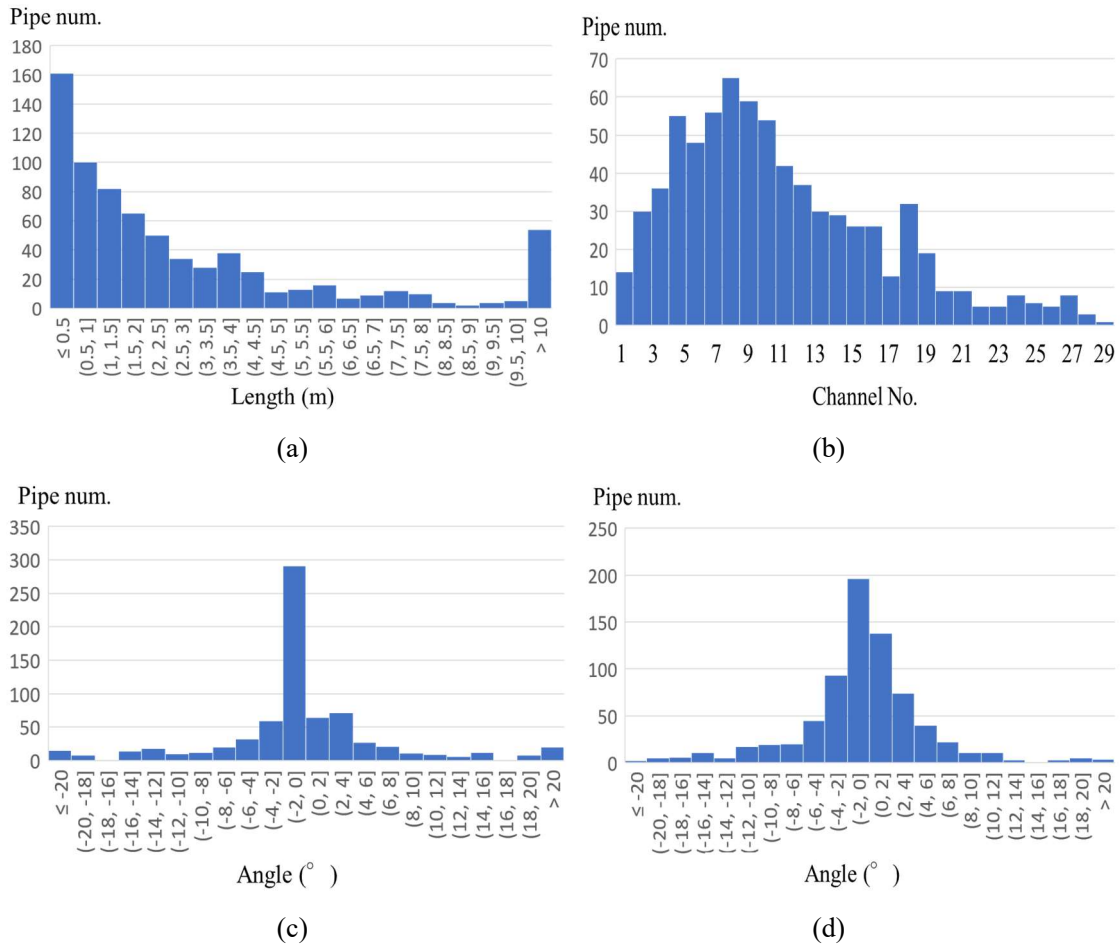


Fig. 2.9 Analysis of transverse pipes (a) length (b) channel number (c) inclination angle in channel direction (d) inclination angle in depth direction

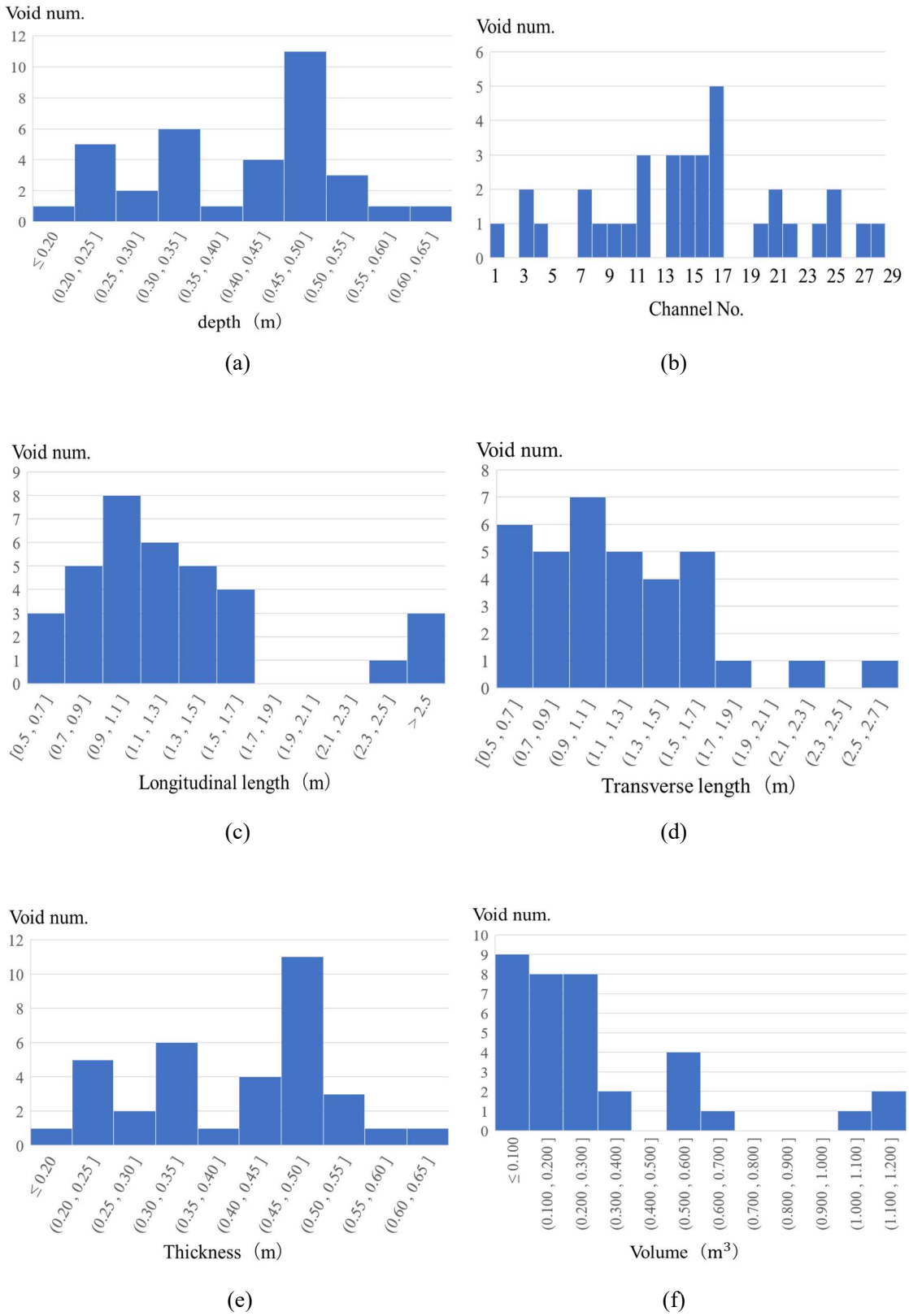


Fig. 2.10　Analysis of void (a) depth (b) channel number (c) transverse length (d) longitudinal length (e) thickness (f) volume

volume of some void was under $0.1\ m^3$. The smallest volume was about $0.04\ m^3$. Actually, small void is not important for the management of the road surface. Too small void cannot be detected because of the limitation of the measurement system. The largest volume was about $1.2\ m^3$. Because too large void will cause the deflection of the road surface leading to a road collapse, road administrators should have taken countermeasures. Therefore, the target volume of void in the research is considered to be around $0.05\ m^3$ - $1\ m^3$.

From the above discussions, in electromagnetic simulations of void, an ellipsoid body whose depth is 0.5 m below the road surface with 1m length and 0.3 m thickness is assumed as a standard model. The randomness in depth, channel number, lengths and thickness was introduced. A cuboid body with the same randomness was also introduced in simulations.

2.3 Experimental field site

In this section, the experimental field sites of pipes and void are introduced. The objective is to quantitatively evaluate the estimation accuracy of the positions of pipes and void. The void data was also utilized to validate the simulation model.

2.3.1 Pipes

The experimental field site of pipes was newly constructed by the authors in Gunma prefecture. The details of the site are shown in Appendix A. Fig. 2.11 (a) shows the site. Fig. 2.12 shows the construction process. The site is about 6 m width and 40 m distance. The site consists of 3 transverse and 3 longitudinal pipes. 1 manhole is connected to one of the longitudinal pipes. The pipes are made of metal, plastic and concrete. The diameter of the pipes ranges from 10 cm to 30 cm. The pipes are placed with 1 m to 10 m intervals. The depth of the pipes ranges from 0.5 m to 1.5 m. The transverse and longitudinal pipes were placed alternately in different distance and depth as shown in Appendix A to reproduce the real 3D

(a) (b)

Fig. 2.11 Experimental field sites (a) pipe (b) void

arrangement of pipes. Styrofoam was embedded as target objects to obtain reference points in radar images. The measurement was conducted on 3 courses as shown in Appendix A. All the 3 courses passed through 3 transverse pipes. The 2 courses passed along a longitudinal pipe. In Section 7.4.1, only the results of course 2 were introduced for the brevity of space. The results of other courses were similar.

2.3.2 Void

The experimental field site of void was owned by Ministry of Land, Infrastructure, Transport and Tourism in Kinki district. The objective of the measurement was to demonstrate the performance of the system and inspection accuracy. Fig 2.11 (b) shows the experiment yard. The site is about 4 m width and 30 m distance. In the 4 measurement courses total 20 void were buried. All the void was modeled by ellipsoid or cuboid Styrofoam. The depth of the void ranged from 0.3 m to 1.5 m. The length in a horizontal direction is from 0.1 m to 1 m. Some metal and plastic pipes were buried in the experiment site, though only the void data was utilized in the research.

2.4 Computing environment

In the research, a GPU computing environment was developed to train a deep learning model and conduct electromagnetic simulations with reasonable calculation time. Fig. 2.13 shows a commercial GPU and concept of Central Processing Unit (CPU) and GPU. GPU is constituted with several hundred to several thousand cores. GPU is the hardware for high performance video output by parallel computing. Though GPU is directly connected to a monitor screen, it can be utilized in various processing tasks coworking with CPU (General Purpose Computing on Graphics Processing Unit, GPGPU). By utilizing GPU, 5 times to at most 20 times faster calculation speed is obtained compared to CPU [92]. In the research, GeForce GTX 1080 Ti manufactured by NVIDIA was used [93]. The performance of GPU is limited by the performance of a battery and cooling system installed in PC. In the research, G-Tune MASTERPIECE i1640 series manufactured by Mouse Computer was used [94]. GPU was utilized for the simulations of 2D and 3D-FDTD

(a) (b)

Fig. 2.12 Construction of pipe site (a) metal pipe (b) concrete pipe

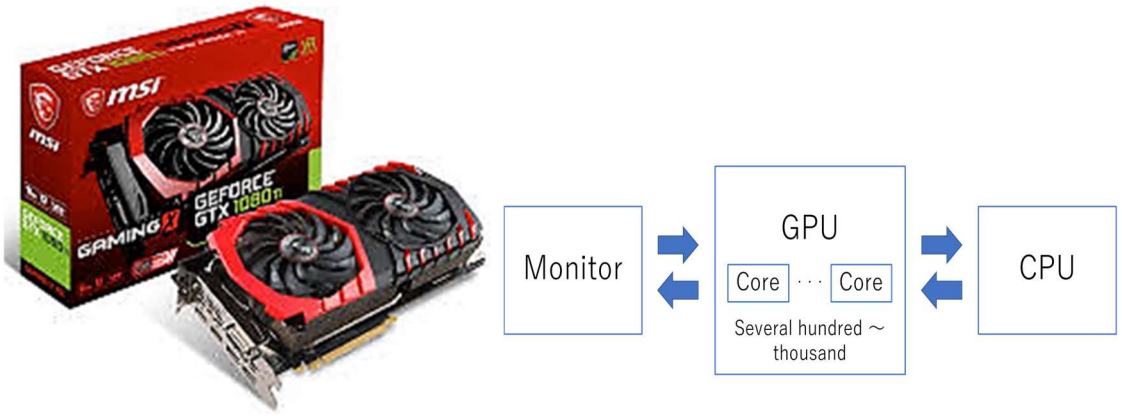
and training of deep learning models. An ordinal laptop PC, Panasonic CF-SZ5 series was used for simple tasks such as displaying detection results and training SVM models [95].

In the research, MATrix LABoratory (MATLAB) was used to produce training data and plot detection results (Fig. 2.14 (a)) [96), 97)]. MATLAB was a general computing platform especially targeted for matrix operation developed by MathWorks. MATLAB is known for the simple and convenient interactive platform. MATLAB R2019b was used in the research. gprMax was used for electromagnetic simulations by FDTD method (Fig. 2.14 (b)) [81), 98)]. gprMax is an open source program written by Python focusing on GPR method [99), 100)]. Any types of media can be introduced. Antenna analysis can easily be conducted by designated programs. 1D, 2D and 3D-FDTD are offered. In the research, gprMax Ver. 3 was used. For the training of CNN models, TensorFlow as shown in Fig. 2.14 (c) was adopted [101) - 103)]. TensorFlow is a machine learning library written by C language and Python released by Google. TensorFlow can deal with 3D-CNN. The newest version of MATLAB can also deal with 3D-CNN. The difference is, TensorFlow includes many training options and trained deep learning models. TensorFlow was utilized for training 3D-CNN models and MATLAB for applying the trained models in the research. By utilizing TensorBoard, displaying learning curves and visualization of deep learning model architectures are easily conducted [104)]. In the research, one of the libraries of TensorFlow for designing deep learning models, Keras (Ver. 2) was adopted. Other Python libraries, scikit-learn (Ver. 1) and matplotlib (Ver. 1) were used for showing training results.

Output data was ".mat" format. ".mat" data was converted to ".npy" format to read in Python programs by an open source program run in MATLAB language [105)]. Other data formats ".json" and ".hdf5," which are list and structure-like formats respectively, were utilized to save simulation results by FDTD method and trained CNN models. The both formats can read in MATLAB.

2.5 Summary

In the research, a vehicle-installed type multi-channel measurement system was utilized. 3D-Radar adopts



(a) (b)

Fig. 2.13 GPU (a) NVIDIA GTX 1080Ti (b) architecture of GPU [94)]

SF-CW method. Therefore, not only scan pitch, but also an operating frequency range and step width were the important parameters. These parameters were optimized to obtain the highest resolutions in depth and horizontal directions with practical vehicle speed.

Pipes, void, manhole and joint data were collected in Nagano and Aichi prefecture by the inspectors. The points of each object in radar images were recorded. The number of pipes, manhole and joint data were enough to train a deep learning model. The void training data was produced by electromagnetic simulations. The excavation survey was conducted to obtain the geometry of void. From the analysis of the detection results, the lengths and depth of pipes were various. In terms of longitudinal pipes, the inclinations of pipes

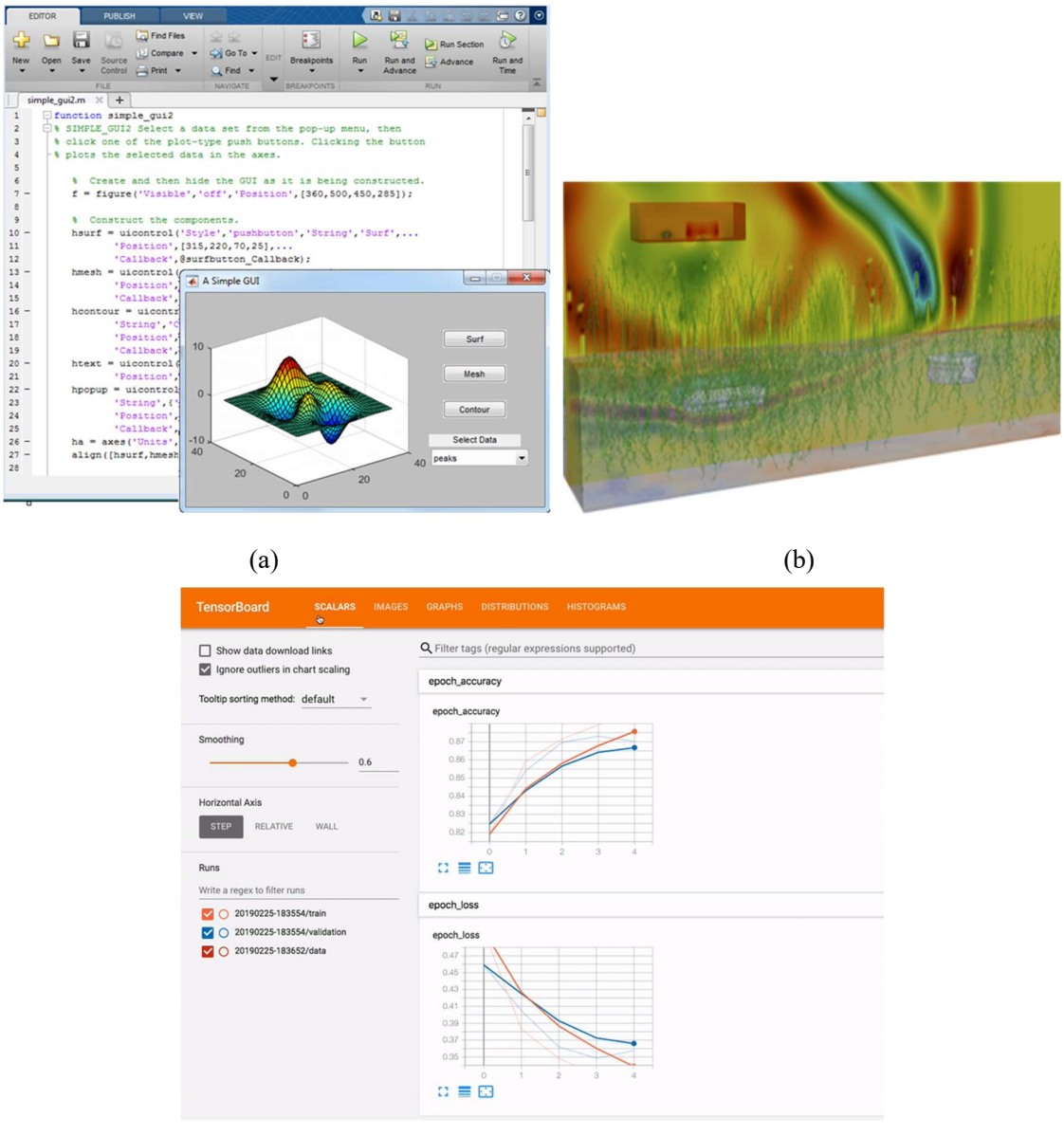


Fig. 2.14 Programs (a) MATLAB (b) gprMax (c) TensorFlow (TensorBoard) [96), 98), 104)]

were not negligible. From the discussions, in electromagnetic simulations, void was modeled by ellipsoid and cuboid bodies with random depth, lengths and thickness.

In the research, the experimental field site of pipes was newly constructed by the authors. The experimental field site of void was also utilized. The objective is to quantitatively evaluate the estimation accuracy of the positions of pipes and void and to validate electromagnetic simulation models. The site consists of transverse and longitudinal pipes with various types of materials and diameter in various depth. Void was reproduced by ellipsoid or cuboid Styrofoam buried in soil with various sizes in various depth. All the data used in this research was introduced in this chapter.

In terms of a computational environment, MATLAB was used to process data and plot the detection results. gprMax was used for the electromagnetic simulations by FDTD method. TensorFlow was used for the training of deep learning models. GPU was utilized for tasks with high calculation cost, such as electromagnetic simulations and training of deep learning models.

Chapter 3: Electromagnetic Characteristics

In this chapter, the fundamental electromagnetic characteristics of subsurface sensing environment was introduced. The maximum depth of SF-CW method was discussed. The effect of radar arrangement on signals was also discussed. Considering these parameters, basic preprocessing algorithms were studied. The flow chart of preprocessing algorithms was shown. In the previous chapter, the optimal parameters were based on the limitation of hardware and measurement environment. In this chapter, the appropriate preprocessing process was derived from digital signal processing theories.

3.1 Electromagnetic properties

3.1.1 Free space loss

Free space loss is one of the fundamental characteristics of propagating electromagnetic waves. In free space, the transmission formula by Friis (1946) is well known [106]. From the Friis formula, the relationship between transmitting power P_t and receiving power P_r is written by absolute gain G_t of an isotropic antenna which transmits the same power to every direction, an effective area of the antenna A_r and distance between transmitting and receiving antennas r as shown in Eq. (3-1). From Eq. (3-1), P_r is inversely proportional to the square of r (Eq. (3-2)).

$$\frac{P_r}{P_t} = \frac{G_t A_r}{4\pi r^2} \tag{3-1}$$

$$\propto \frac{1}{r^2} \tag{3-2}$$

Considering impedance in free space Z_0, substituting the relationship between power and electric field intensity Eq. (3-3) in Eq. (3-1), the relationship between the electric field of a transmitting antenna E_t and receiving antenna E_r can be derived as shown in Eq. (3-4). From Eq. (3-4), E_r is inversely proportional to r irrespective of transmitting wave length λ.

$$\frac{P_{r,t}}{A_{r,t}} = \frac{{E_{r,t}}^2}{Z_0} \tag{3-3}$$

$$\frac{E_r}{E_t} \propto \frac{1}{r} \tag{3-4}$$

Eq. (3-1) holds true only at far field, where r is large enough to fulfill Eq. (3-5) with the maximum size of transmitting and receiving antennas a. The size of antennas utilized in the research is about several tens cm. To satisfy Eq. (3-5), r should be larger than 1 m at $f = 3$ GHz and 3 m at $f = 1$ GHz. Therefore, at a low frequency range in shallow depth, Eq. (3-5) may not seem true. Considering the relative permittivity of soil and round trip distance of waves, at most frequency ranges and depth Eq. (3-5) holds.

$$r \geqq \frac{2a^2}{\lambda} \tag{3-5}$$

3.1.2 Permittivity

In many research fields of GPR applications, the electromagnetic characteristics of natural materials and artificial objects were experimentally estimated [34) ~ 36), 39), 56), 107), 108)]. For example, many laboratory testing methods were proposed to estimate the electromagnetic characteristics of soil. The relative permittivity and conductivity of materials at 100 MHz are shown in Table 3.1 by Daniels (2004) [34)]. From Table 3.1, the parameters have large variations depending on water content. Topp's formula shows the relationship between a water volume ratio θ_v and relative permittivity ε_r (Eq. (3-6)).

$$\theta_v = -0.0503 + 0.0292\varepsilon_r - 5.5 \times 10^{-4}{\varepsilon_r}^2 + 4.3 \times 10^{-6}{\varepsilon_r}^3 \tag{3-6}$$

Because of the inhomogeneous and time-dependent natures of materials, it is further difficult to appropriately estimate electromagnetic parameters on site.

Dielectric loss tangent $\tan\delta$ at each angular frequency ω is defined by Eq. (3-7) with permittivity ε and conductivity σ.

$$\tan\delta = \frac{\sigma}{\omega\varepsilon} \tag{3-7}$$

From Maxwell's equations, an attenuation factor $F_\alpha(f)$ and attenuation coefficient $\alpha(f)$ are derived below with permeability μ.

$$F_\alpha(f) = e^{-\alpha(f)} \tag{3-8}$$

$$\alpha(f) \cong 2\pi f\sqrt{\mu\varepsilon}\sqrt{\frac{1}{2}\left(\sqrt{1+\tan\delta^2}-1\right)} \tag{3-9}$$

Conductivity and an attenuation coefficient depend on frequency while relative permittivity does not depend on frequency.

3.2 Factors affecting maximum depth

Table 3.1 Electromagnetic parameters of materials at 100 MHz [34)]

Material	Relative permittivity ε	Conductivity σ (Sm^{-1})
Air	1	0
asphalt (dry)	2 - 4	$10^{-2} : 10^{-1}$
asphalt (wet)	6 - 12	$10^{-3} : 10^{-1}$
water	81	$10^{-6} : 10^{-2}$
sand (dry)	2 - 6	$10^{-7} : 10^{-3}$
sand (wet)	10 - 30	$10^{-3} : 10^{-2}$
soil (dry)	4 - 10	$10^{-2} : 10^{-1}$
soil (wet)	10 - 30	$10^{-3} : 10^{-0}$

3.2.1 Maximum depth

In the research, SF-CW method was adopted (Fig. 3.1) [88), 89)]. SF-CW method transmits a frequency modulated wave which is the same as a pulsed wave from a mathematical point of view. Because of the limitation of hardware, SF-CW method can transmit a stronger signal than transmitting a pulsed wave. GPR method transmits a signal with constant frequency $f_n = f_0 + n\Delta f \quad (n = 0,1,\dots,N)$ from $t_n = n\Delta t$ to $t_{n+1} = (n+1)\Delta t$. Frequency is swept from minimum frequency $f_{min} = f_0$ to maximum frequency $f_{max} = f_0 + N\Delta f$ (Fig. 3.2). A received signal has time duration $T = (N+1)\Delta t$, which is obtained by inverse Fourier transform of a measured spectrum. Therefore, the maximum depth D_{max} is derived by the round-trip of propagation waves with velocity v, relative permittivity ε, and light velocity c.

$$D_{max} = \frac{vT}{2} \tag{3-10}$$

$$v = \frac{c}{\sqrt{\varepsilon}} \tag{3-11}$$

Replacing the time duration T of Eq. (3-10) by the reciprocal of frequency pitch Δf, D_{max} is written below. Eq. (3-12) is the maximum depth derived from mathematical conditions.

$$D_{max} = \frac{v}{2\Delta f} \tag{3-12}$$

In the research, $\Delta f = 20$ MHz was adopted. Assuming $c = 3.0 \cdot 10^8$ m/s and $\varepsilon = 1$, $D_{max} = 7.5$ m from Eq. (3-12). In GPR method, target depth is from several tens cm to 1.5 m below the road surface. Δf can be doubled or 3 times larger in shallower regions with low relative permittivity. $\Delta f = 20$ MHz is considered to be a safer side assumption.

3.2.2 Signal Noise Ratio (SNR)

Even if target depth is smaller than D_{max}, a received signal should be large enough to detect target objects. In radar images, other than reflections from target objects, undesired signals called clutter and system noise

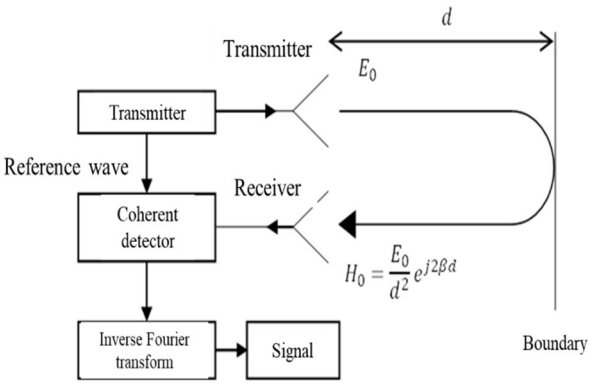


Fig. 3.1 Measurement system

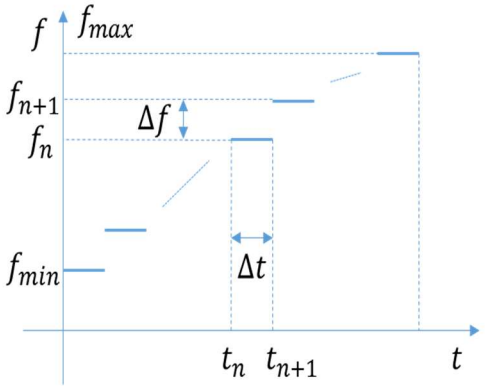


Fig. 3.2 SF-CF method

(called noise in this thesis) are superposed. Target signals should be at least larger than clutter and noise [49), 50)]. Assuming the power of an undesired signal P_N and target signal P_S, Signal Noise Ratio (SNR) is defined as below.

$$SNR = \frac{P_S}{P_N} \tag{3-13}$$

SNR can be improved by suppressing noise, clutter and/or amplifying transmitted waves. Noise is affected by utilized instruments, which is difficult to control. Clutter depends on measurement environment. In some cases, by applying a bandpass filter, noise and clutter can be suppressed to a certain level.

The power of a transmitted wave P is an important parameter. Larger P is, larger SNR becomes. However, P is limited by hardware and social restrictions of measurement environment. In the research, a commercial system is utilized, which is considered to transmit enough P within restrictions. A transmitted frequency band F is also an important parameter. Electromagnetic waves attenuate from several tens dB to several hundreds dB in a lossy medium such as wet soil. Loss depends on F. Lower F is, smaller loss is. However, a resolution, which corresponds to the detectable size of target objects, is larger when F is lower. Strictly speaking, the optimal F may also be affected by the characteristics of soil on site. In the research, the minimum frequency $f_{min} = 50$ MHz and maximum frequency $f_{max} = 3030$ MHz were adopted from a practical point of view.

3.3 Radar arrangement

3.3.1 Effect on measured waves

In the research, Bistatic method was adopted. As shown in Fig. 3.3, in Bistatic method, transmitting and receiving antennas are placed with certain distance. Transmitted waves by a transmitting antenna are received by a receiving antenna. The received signal $x(t)$ is a function of a transmitted impulse wave $s(t)$, factor representing the effects of propagating media and medium boundaries between antennas and target objects $u(t)$, reflection coefficients including target objects and clutter $e(t)$ [62) - 64)]. The symbol $*$ represents the

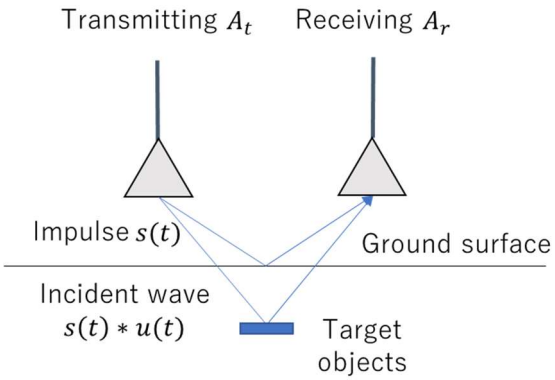


Fig. 3.3 Components of received wave

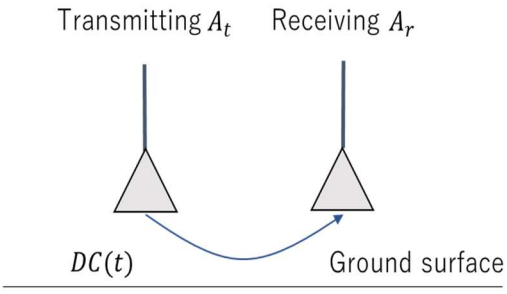


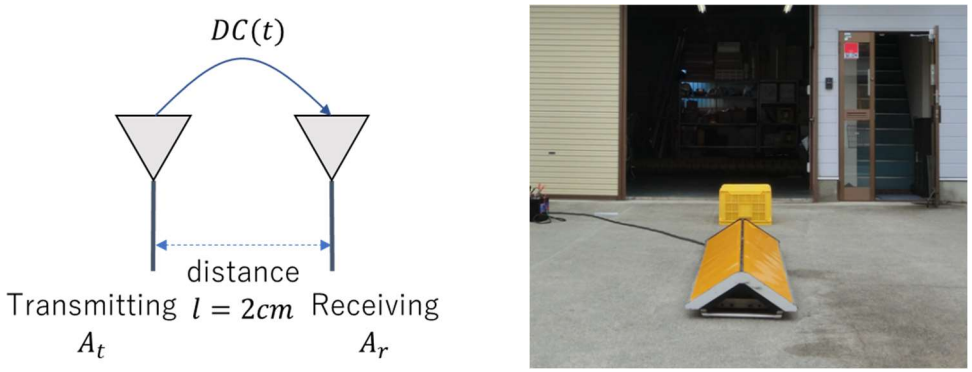
Fig. 3.4 Direct coupling

convolution of two functions.

$$x(\mathrm{t}) = s(t) * u(t) * e(t) \tag{3-14}$$

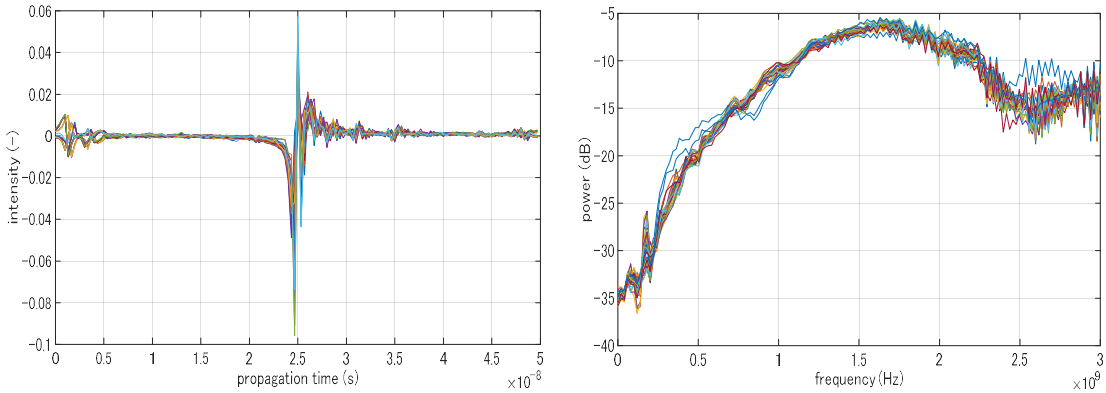
$u(t)$ includes many factors such as the interaction between antennas and ground surface, attenuation inside a medium and effect of a measurement system. The model represented by Eq. (3-14) holds true not only in GPR method but also in many fields of geophysical surveys. In GPR method, the reflection from ground surface and attenuation inside lossy soil are important. The problem is, Eq. (3-14) cannot consider the effect of direct coupling $DC(t)$. $DC(t)$ is a wave directly propagated from a transmitting antenna to a receiving antenna in the air. $DC(t)$ is affected not only by $s(t)$ but also antenna geometries and arrangement. $DC(t)$ is added to the right side of Eq. (3-14), which is independent from $e(t)$.

To estimate $e(t)$ from Eq. (3-14), the effects of $DC(t)$ and $s(t)$ should be removed. In the research, $DC(t)$ was measured beforehand to subtract from measured waves. From Wiener-Khinchin's theorem, convolution in the time domain is converted to product in the frequency domain [109) - 113)]. Eq. (3-15) is the Fourier transform of Eq. (3-14).



(a) (b)

Fig. 3.5 Direct coupling (a) schematic view (b) photo



(a) (b)

Fig. 3.6 Measured direct coupling (a) signal (b) power spectrum

$$x(f) = s(f) \times u(f) \times e(f) \tag{3-15}$$

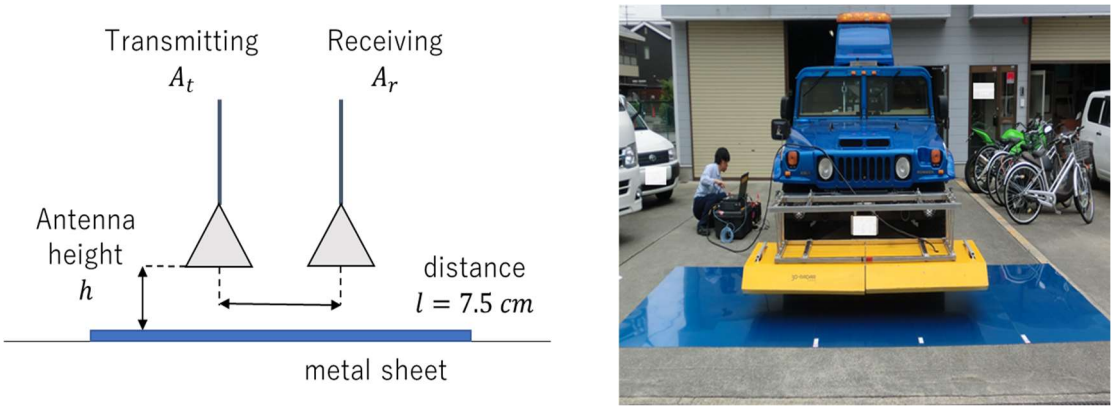
Measured waves after removal of direct coupling was divided by $s(f)$ in the frequency domain, which was also measured in advance.

3.3.2 Direct coupling

Direct coupling was obtained as shown in Fig. 3.5 placing antennas upward. Fig. 3.6 shows a measured signal and power spectrum. The signals and power spectra of all the channels (29 channels) were plotted. In this thesis, the unit of the vertical axis of a power spectrum of electric field intensity was dB using Eq. (3-16) below.

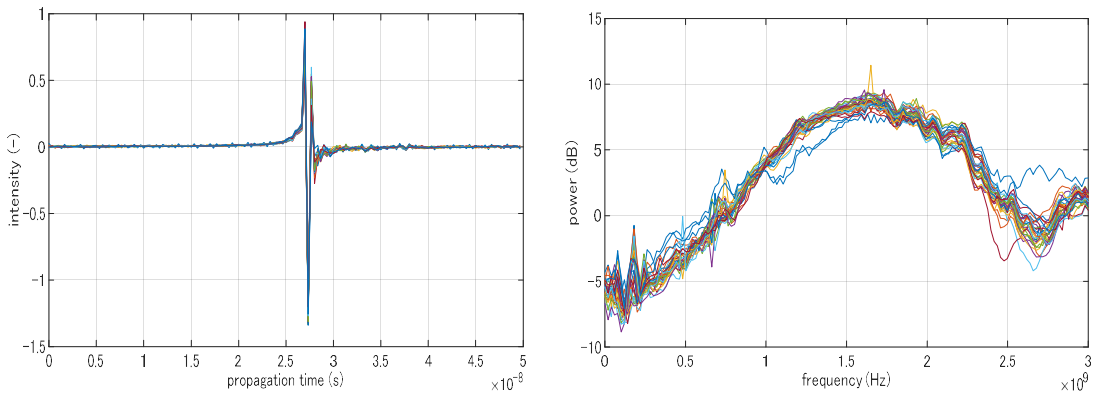
$$P_r = 20 \log_{10} E_r \tag{3-16}$$

As shown in Fig. 3.6, the positions of peaks of 29 channels were almost the same. However, because of channel characteristics, the signals have variations. The peaks of direct coupling were smaller by about -15 dB than impulse waves, indicating the effect of direct coupling is small at shallower depth. However, at



(a) (b)

Fig. 3.7 Impulse wave (a) schematic view (b) photo



(a) (b)

Fig. 3.8 Measured impulse wave (a) signal (b) power spectrum

deeper around 1.5 m, because of the attenuation of transmitted waves, direct coupling is not negligible.

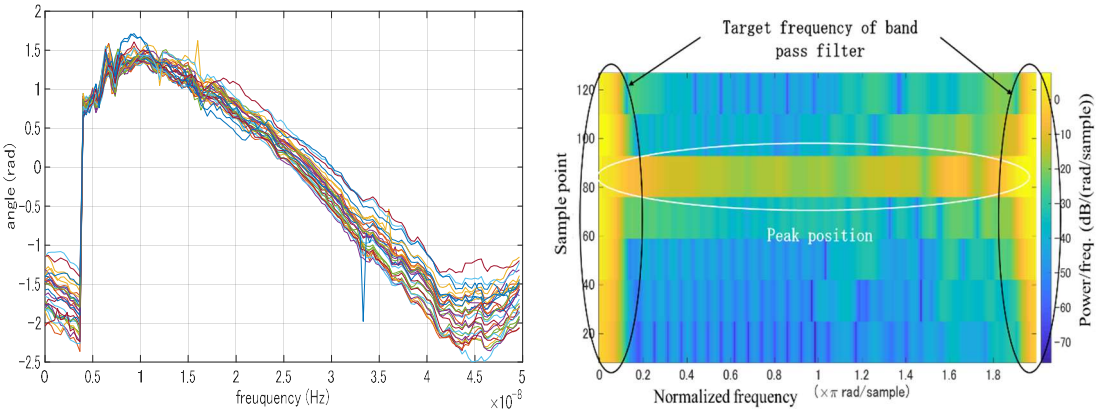
3.3.3 Impulse wave

A transmitted impulse wave was approximately obtained from the reflection from a metal sheet as shown in Fig. 3.7. Fig. 3.8 shows the signal and power spectrum. In the measurement of the impulse wave, antenna height h is an important parameter. $h = 20$ cm was the same as the measurement condition. The plotted impulse waves were after the removal of direct coupling.

The reflection from a metal sheet can be considered as a transmitted impulse wave because the reflection coefficient of metal is close to 1. The size of the metal sheet should be larger enough than the antennas. The system transmits constant power at every frequency. However, the received impulse wave was different from a pulsed signal as shown in Fig. 3.8 because of the effects of antennas and interactions between antennas and target objects. The power spectrum was attenuated in lower and higher frequency ranges. Therefore, by dividing the impulse wave in the frequency domain, the power spectrum is flattened to increase the quality of radar images, which is also called "whitening." The impulse wave characterizes the measurement system. Fig. 3.9 (a) shows a phase spectrum. Fig. 3.9 (b) shows a spectrogram. From Fig. 3.9 (a), the impulse wave has ideal characteristics of a linear phase and localized signal intensity in a time range [62), 63), 109) - 113)]. From Fig. 3.9 (b) at certain low and high frequency ranges the signal intensity was not localized because the frequency range is outside the nominal frequency range as stated in Section 2.1.

3.4 Flow chart of preprocessing

Fig. 3.10 shows the flow chart of preprocessing. Preprocessing algorithms adopted in the research are in common with an ordinary GPR system. Measured data is real and imaginary parts of a Fourier coefficient at each scan and frequency. After preprocessing in the frequency and time domains, 3D time domain data is processed as a final output. In the research, all the signal processing algorithms were applied in MATLAB. Inverse Fourier transform and other functions are built-in functions of MATLAB. The algorithms are not



(a) (b)

Fig. 3.9 Impulse wave (a) phase spectrum (b) spectrogram

special for this purpose.

Conversion to 3D frequency domain data is to process 3D frequency domain data from text data of Fourier coefficients. As stated in Section 2.1, the measured data was sorted by scan, channel number and frequency. Therefore, after extracting a Fourier coefficient at each point, data was aligned by frequency in first dimension, channel number in second dimension and scan number in third dimension.

Removal of direct coupling subtracts direct coupling from measured data. Fig. 3.11 shows a raw image and one example of signals. Fig. 3.12 shows an image after removal of direct coupling. The first and largest peak in Fig. 3.12 is direct coupling. From Fig. 3.12, after removal of direct coupling, the peak was suppressed, indicating the effect of direct coupling was reduced. From a mathematical point of view, removal of direct coupling can also be conducted in the time domain. However, to apply a calibration filter, direct coupling should be removed beforehand. For simplicity of the programs, removal of direct coupling was conducted firstly in the frequency domain. All the signals shown below were after removal of direct coupling.

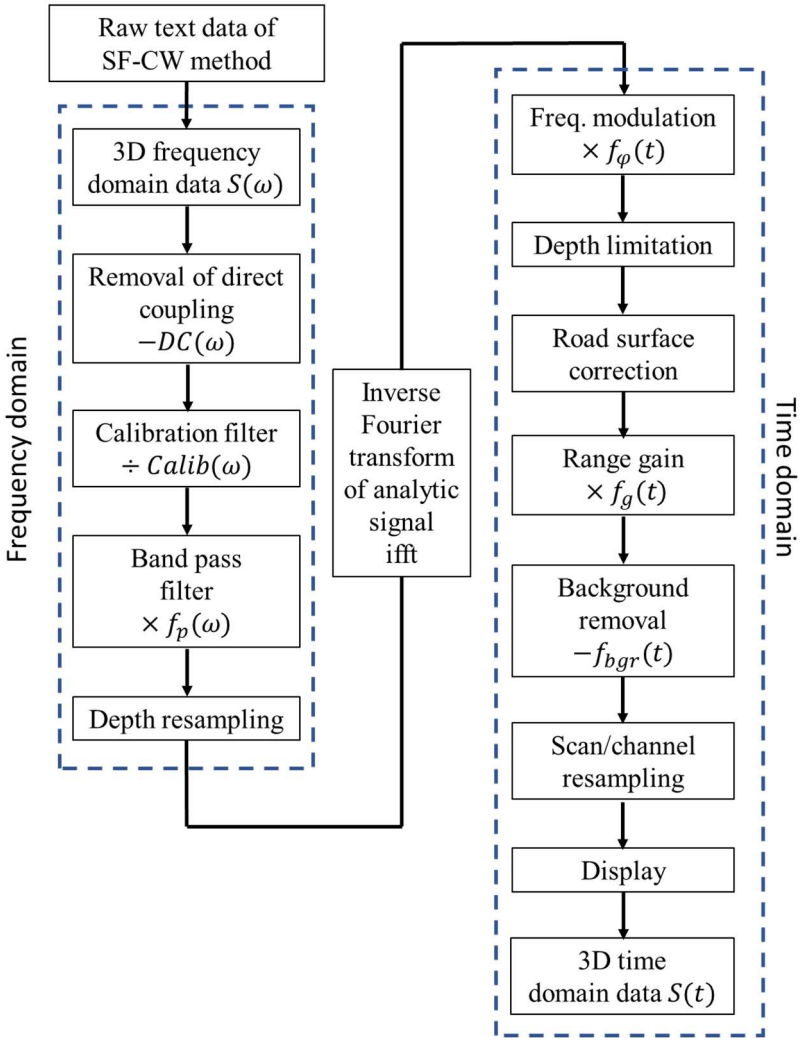


Fig. 3.10 Flow chart of preprocessing

Application of calibration filter is to divide measured waves by a transmitted impulse wave in the frequency domain. Fig. 3.13 shows an image after calibration filter. From Fig. 3.13, a clear peak was observed at the depth of asphalt pavement surface. The measured waves are the superposition of reflected waves at each medium boundary. A distribution of reflection coefficients after calibration does not depend on the waveform of an impulse wave. Division in the frequency domain corresponds to deconvolution in the time domain. Deconvolution needs more calculation cost than division. Therefore, a calibration filter was applied in the frequency domain.

Application of bandpass filter is to apply bandpass filter in the frequency domain. The frequency range corresponds to the nominal frequency range of antennas utilized in the research, 200 MHz - 3 GHz. A rectangular window with lower frequency limit f_d and upper limit f_u was adopted.

$$f_p(f) = 1 \quad (f_d \leq f \leq f_u) \tag{3-17}$$

$$= 0 \quad (f \leq f_d, f_u \leq f) \tag{3-18}$$

Other bandpass filters can be utilized which shows a tapering effect in frequency band boundaries. For example, Hamming and Gauss windows show smaller sidelobe peaks and larger mainlobe widths than a rectangular window [109) - 112)]. Radar images are affected by the types of window functions. However, in terms of machine learning and AI approaches, the effect of a window function is considered to be small.

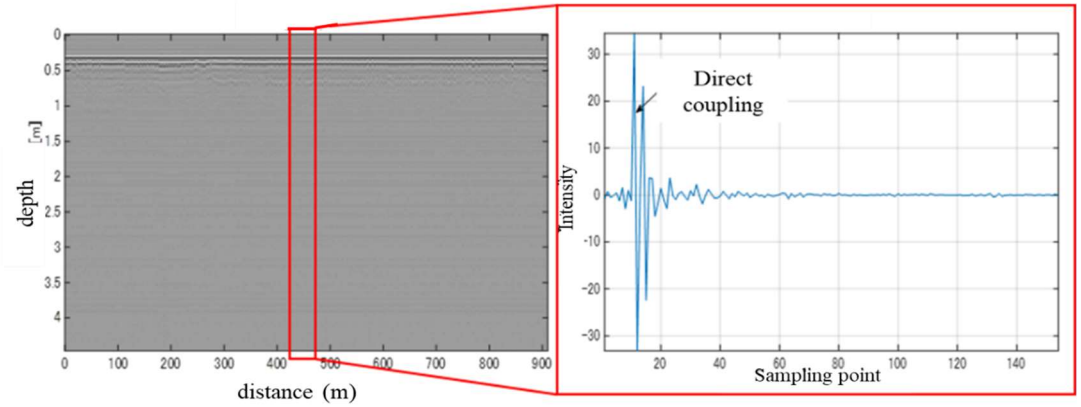
Depth resampling resamples data in a depth direction. Ordinarily, the quality of radar images improves after upsampling though the effect is considered to be small in terms of machine learning and AI approaches. To downsample the data, a certain frequency range is extracted in the frequency domain. In the time domain, it corresponds to thinning out data after applying an antialiasing filter. To upsample the data, zeros were added to a higher frequency range in the frequency domain, which is called zero padding. Zero padding corresponds to interpolate between sampling points by sinc function in the time domain [109) - 112)].

$$f_{sinc}(t) = \frac{\sin(t)}{t} \tag{3-19}$$

Typically, the number of data is the power of 2 for IFFT. In the research, the number of the data was increased to the double of the closest power of 2 after zero padding. The inverse Fourier transform of an upsampled signal is an analytic signal, which has power only in a frequency range lower than Nyquist frequency. A target signal is a real part of an analytic signal.

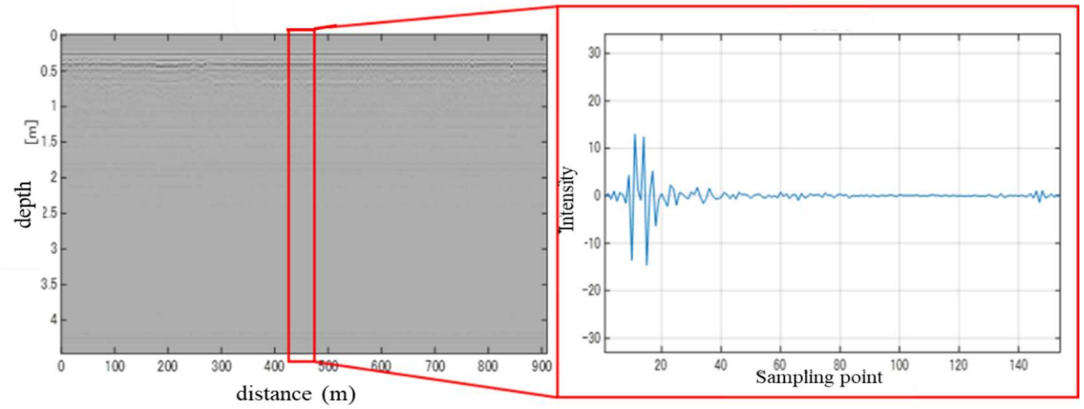
Fast Fourier transform (FFT) is a fast version of Fourier Transform (FT), which converts a time domain signal to a frequency domain signal [109) - 112)] with low calculation cost. Inverse Fast Fourier Transform (IFFT) is an inverse function of FFT. FT converts time domain data $x(n)$ to frequency domain data $X(k)$ by a function below.

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-i\frac{2\pi}{N}kn} \tag{3-20}$$



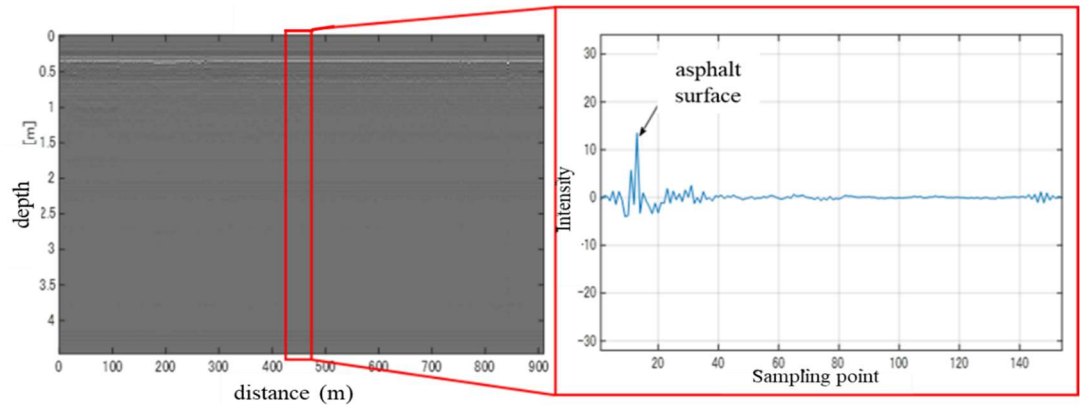
(a) (b)

Fig. 3.11 Raw image (a) image (b) signal



(a) (b)

Fig. 3.12 After removal of direct coupling (a) image (b) signal



(a) (b)

Fig. 3.13 After calibration (a) image (b) signal

On the other hand, operation $W(k) = e^{-i\frac{2\pi}{N}k}$ shows periodicity below.

$$W(k) = W(k+N) \tag{3-21}$$

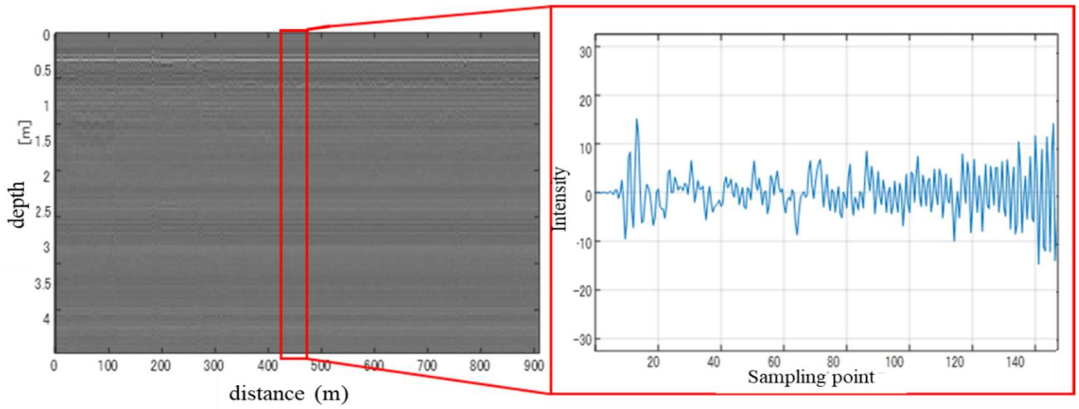
From a Cooley-Tukey's FFT algorithm, when the number of data is the power of 2, Fourier transform of $x(n)$ is defined by odd and even data $x_{even}(n)$, $x_{odd}(n)$ from Eq. (3-21). FFT accelerates calculation by recursively calculating smaller data $x_{even}(n)$, $x_{odd}(n)$. IFFT is derived from Inverse Fourier Transform (IFT) in the same way. In the research, IFFT was applied to the data after upsampling.

Frequency modulation is to shift a spectrum by multiplying a phase rotator function in the time domain [109) - 112)]. The operation to shift a spectrum $P(f)$ by Δf in the frequency domain corresponds to multiply a phase rotator function f_φ below in the time domain.

$$f_\varphi(t) = e^{-2\pi i \Delta f t} \tag{3-22}$$

In the research, frequency pitch was 20 MHz. The frequency range was 50 MHz - 3030 MHz. Therefore, the total number of data was 150 points. After simple Fourier transform by the built-in function of MATLAB "ifft," the frequency range of converted data was 0 MHz - 2980 MHz. Therefore, to appropriately process the data, frequency modulation with shift $\Delta f = 50$ MHz should be applied after IFFT.

Depth limitation is to extract a certain depth range in the time domain to reduce an amount of data. Depth limitation depends on target objects. From the measurement conditions, the maximum depth was about 7.5 m. Considering the relative permittivity of soil, the maximum depth became around 2 m. From the discussion of Section 2.2.2, depth limitation was not adopted considering the depth range of target objects. Road surface correction is to align the depth of road surface by extracting peaks. Misalignment occurs because of the vibration of the vehicle and road surface unevenness. Therefore, the information of road surface unevenness is lost after road surface correction. Sometimes at rough road surface, road surface correction causes the discontinuity of radar images. In the research, road surface correction was not adopted.



(a) (b)

Fig. 3.14 After range gain (a) image (b) signal

Range gain is to compensate for the attenuation and loss of electromagnetic waves (Fig. 3.16). Range gain is to multiply a gain function f_g in the time domain, which is the function of depth $z = ct/2$.

$$f_g(z) = z^{\alpha} 10^{\beta z} \tag{3-23}$$

Range gain is composed of the terms of the power of z and 10. The term of z represents free space loss. The term of 10 represents the attenuation in lossy medium. α and β are the parameters. From the discussion of Section 3.1.1, α is close to 1. The parameters are manually adjusted by inspectors data by data. In the research, the effect of range gain was evaluated in Section 6.5.1.

BGR is to remove the multiple reflections between antennas, road surface and medium boundaries. Multiple reflections show horizontally uniform reflection patterns. Signals after BGR $f_b(x,z)$ is obtained by subtracting moving averages in a scan direction from signals $f(x,z)$ at scan x.

$$f_b(x,z) = f(x,z) - \frac{1}{n+1} \sum_{x'=x-n\Delta x/2}^{x+n\Delta x/2} f(x',z) \tag{3-24}$$

A window length of a moving average n is the parameter. n depends on the scale of target reflection patterns. Too large n cannot remove multiple reflections while too small n removes target reflection patterns. The effect of BGR was evaluated in Section 6.5.1.

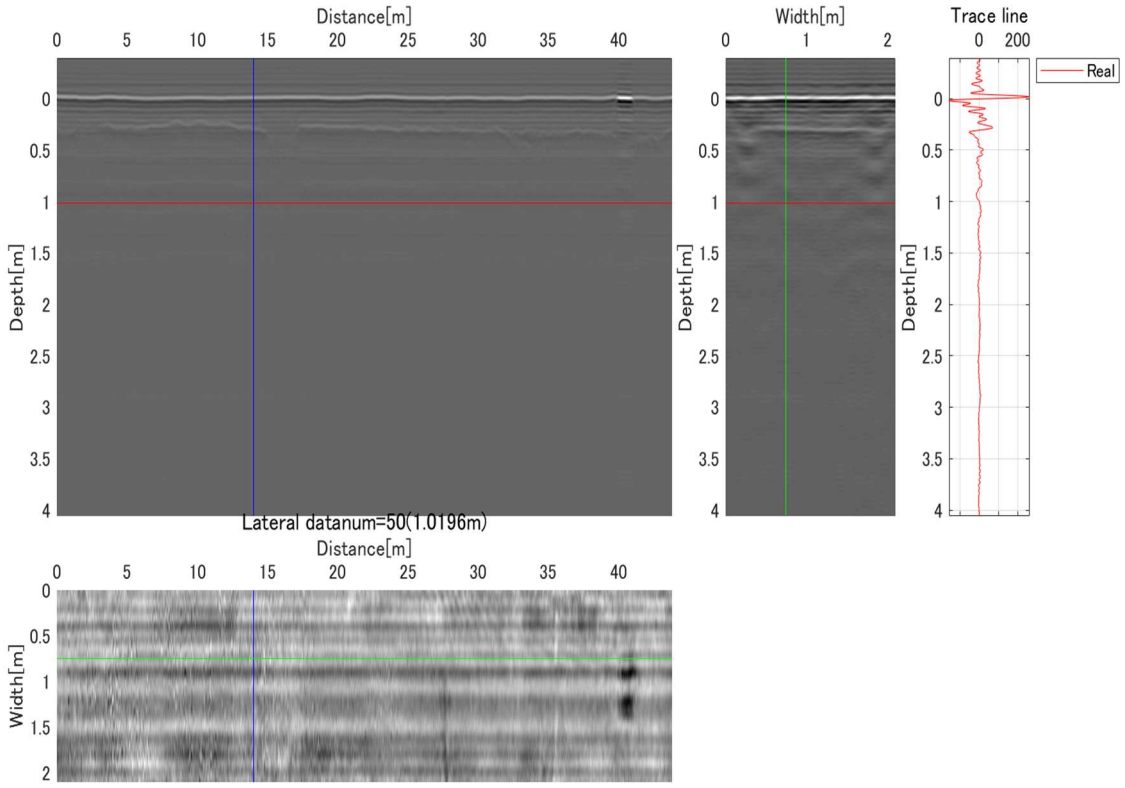


Fig. 3.15 Example of radar images

Scan/channel resampling is to resample data in a horizontal direction. As stated above, in terms of machine learning and AI approaches, the effect of upsampling is considered to be small. Radar images are usually displayed in scan, channel and depth direction sections at a certain position (Fig. 3.15). A signal of a certain point was also shown in Fig. 3.15. In the research, radar images were displayed by gray scale. Simulation results were shown by parula color. A volume of final output data was up to several TB. Therefore, an external hard disk drive is necessary.

3.5 Summary

The fundamental electromagnetics of subsurface sensing environment was introduced. Free space loss and permittivity of typical materials were discussed. The principle of SF-CW method was explained. The effect of a frequency range and step width on the maximum depth was discussed. The effect of antenna arrangement on measured waves was evaluated. In the research, bistatic and air-coupled method was adopted. Therefore, removal of direct coupling was necessary. A direct coupling wave was obtained by placing antennas upward. In the research, a calibration filter was applied on the measured data. A transmitted impulse wave was estimated by a reflection from a metal sheet.

The flow chart of preprocessing was shown. Input data was a Fourier coefficient at each point. The data was converted to 3D frequency domain data with frequency, channel and scan directions. After IFFT, the data was converted to 3D time domain data with depth, channel and scan directions. In the frequency domain, removal of direct coupling, calibration and bandpass filters were applied. In the time domain, frequency modulation, range gain and BGR were applied. The effects of range gain and BGR were evaluated in Section 6.5.1. The following chapters were about the training of machine learning and AI models by the data processed in this chapter.

Chapter 4: Manhole and joint detection by SVM

In this chapter, an SVM model for the detection of manhole and joint was developed. The principle of SVM was introduced. For the basis of the following discussions, the evaluation methods for classifiers and methodology to produce training data was discussed. The effects of the size of training images, preprocessing and feature values on classification accuracy were evaluated. A 3 category SVM model was developed considering the optimized parameters of each 2 category classifier. Output categories were manhole, joint and pavement section. Preparing for the discussions of the following chapters, the developed SVM model was compared to a deep CNN model.

4.1 Classification by SVM

4.1.1 Proposed flow chart

In the research, image classification by an SVM model is composed of 3 steps, preprocessing filter, feature extraction and application of an SVM model referring to previous research [114) - 120)]. Fig. 4.1 shows a flow chart of image classification by an SVM model. A Laplacian filter was applied to input images which have an optimized size. Then, Histogram of Oriented Gradients (HOG) features of input images were calculated. HOG feature will be an input vector of the SVM model. A nonlinear dividing plane of an SVM model was defined by a Radial Basis Function (RBF) kernel. The output of an SVM model was 3 categories, manhole,

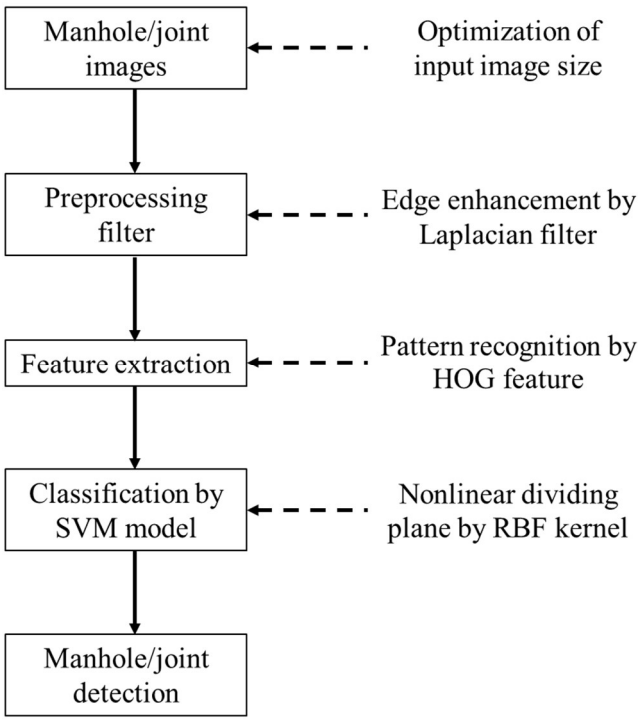


Fig. 4.1 Flow chart of classification by SVM

joint and pavement section. In this chapter, Section 4.4 corresponds to the optimization of an SVM model. Section 4.5 explains how to extend the category of the SVM model to multiple classes. Section 4.6 compares the optimized SVM model with a CNN model.

4.1.2 Principle of SVM

The function f of SVM is defined by an N dimension feature vector $x = (x_1, x_2, \dots, x_n)$, model parameter $w = (w_1, w_2, \dots, w_n)$ and bias b.

$$f(x) = w \cdot x + b \tag{4-1}$$

With certain w and b, a feature space is divided into 2 categories by the sign of f as shown in Fig. 4.2. f of an input x corresponds to the distance between x and $f = 0$. Therefore, considering an output class $y = 1$ or -1, $\gamma = y \cdot f$ is defined as a margin whose sign represents a category. With an allowable margin γ', a slack variable ξ is defined below.

$$\xi = \max(0, \gamma - \gamma') \tag{4-2}$$

ξ is 0 when x is correctly classified or γ is smaller than γ'. When γ is larger than γ', ξ is positive. γ' prevents an SVM model from overfitting. w and b which minimize the summation of ξ are considered to be the most appropriate SVM parameters.

Utilizing a nonlinear mapping function Φ for a nonlinear dividing plane, an SVM model can also be defined in a feature space $\Phi(x)$ as below.

$$f = w \cdot \Phi(x) + b \tag{4-3}$$

By solving the dual problem of Eq. (4-1), w can be represented by the linear combination of y and $\Phi(x)$ with a coefficient α.

$$w = \alpha \cdot y \cdot \Phi(x) \tag{4-4}$$

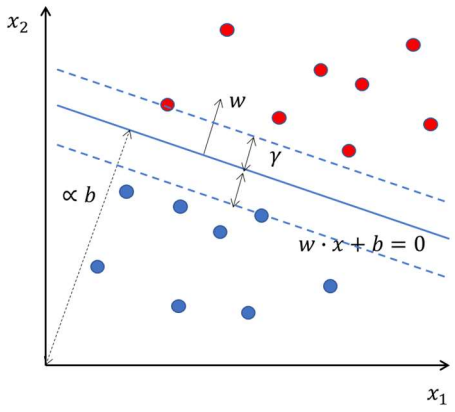


Fig. 4.2 Dividing plane and margin of SVM

Substituting Eq. (4-4) in Eq. (4-3), $\Phi(x)$ is replaced by a kernel function K, called kernel trick.

$$f = \alpha \cdot y \cdot K(x) + b \tag{4-5}$$

$$K(x) = \Phi(x) \cdot \Phi(x) \tag{4-6}$$

RBF and polynomial kernels below are suggested from the mathematical conditions of K [114) - 116)].

$$F(x_i, x_j) = \exp\left(-||x_i - x_j||^2\right) \tag{4-7}$$

$$F(x_i, x_j) = (1 + x_i x_j)^p \tag{4-8}$$

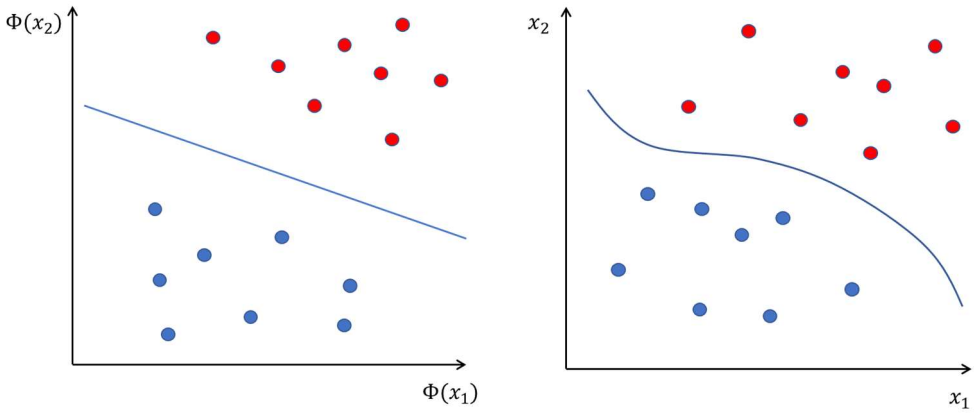
An SVM model is linear when an identity function is used as K. The complexity of a dividing plane is related to the nonlinearity of K as shown in Fig 4.3. Complex functions increase accuracy. However, too complicated functions cause overfitting. Appropriate K depends on the characteristics of data.

A cost (objective) function L to be minimized is defined below with ξ which corresponds to the distance from a dividing plane as defined in Eq. (4-2).

$$L(x) = N_u \sum_{i=1}^{n} \xi(x_i) + \frac{1}{2}||w||^2 \tag{4-9}$$

$||w||^2/2$ serves as a regularization term. N_u is a regularization factor. With large N_u, w is optimized to minimize the distance from the dividing plane. With small N_u, w is optimized to avoid too large w, which is related to overfitting. N_u depends on the kernels and data. From Appendix B, N_u did not affect the results. Therefore, $N_u = 0.5$ was adopted.

The parameters w and b are estimated by differentiating L by w and b. ξ should be positive. The inequality conditions are included by considering Karush-Kuhn-Tucker conditions [114) - 116)]. In the research,



(a) (b)

Fig. 4.3 Kernel trick (a) feature space after kernel trick (b) before kernel trick

the SVM model outputs not only the categories but also the probability of each class p. p is derived from the posterior probability of Bayes' theorem.

Because Eq. (4-9) has many local solutions, optimization methods were needed. Iterative Single Data Algorithm (ISDA) is based on gradient descent method. Sequential Minimal Optimization (SMO) accelerates calculation utilizing the analytical solution α of two random inputs x. The both methods asymptotically estimate the optimal solutions with a certain learning rate. In the research, from Appendix B, the optimization method does not affect calculation time and classification accuracy. Calculation time was at most 5 minutes because the training data set was relatively small.

An appropriate initial α may accelerate the training process. However, there is no research about manhole and joint detection from radar images by SVM. In the research, all the elements of the initial α was 0.5. Removal of outliers in the training process may assist the robust training preventing overfitting. The effect depends on the data. In the research, from Appendix B, an outlier rate r did not affect the results. Therefore, outlier removal was not conducted.

4.2 Evaluation of classifiers

The results of two category classification are categorized into 4 patterns by the combination of a classifier output and actual class (true class) as shown in Fig. 4.4 (a). A classifier with n category outputs is also evaluated by n^2 patterns. Fig. 4.4 (b) shows a 3 category case. The matrixes shown in Fig. 4.4 is called a confusion matrix. The classification accuracy of an n category classifier can be evaluated by the same procedure as a 2 category case diving n classes into the target class and other classes. Therefore, this section assumes a 2 category classifier.

True Positive (TP) or True Negative (TN) are the cases when positive or negative data is correctly classified as positive or negative. False Negative (FN) is the case when positive data is falsely classified as negative. FN is also called a Type I error or "missing". False Positive (FP) is vice versa, also called a Type II error or false detection. Therefore, classification accuracy (accuracy) in this thesis is the rate of correctly

predict \ actual	**positive**	**negative**
positive	True Positive (TP)	False Negative (FN)
negative	False Positive (FP)	True Negative (TN)

(a)

predict \ actual	**class 1**	**class 2**	**class 3**
class 1	True	False	False
class 2	False	True	False
class 3	False	False	True

(b)

Fig. 4.4 Results of classifier (a) 2 categories (b) 3 category confusion matrix

classified data among all the data as defined below.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{4-10}$$

Classification accuracy cannot consider the characteristics of classifiers such as the detection rate of positive data. Precision is the rate of true positive data among the data classified as positive. Recall (True Positive Rate, TPR, detection rate) is the rate of the data classified as positive among all the positive data.

$$Precision = \frac{TP}{TP + FP} \tag{4-11}$$

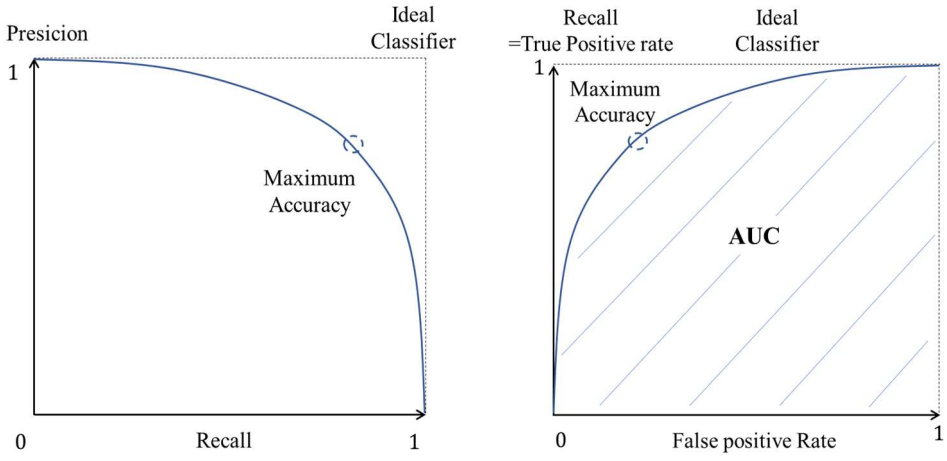
$$Recall = \frac{TP}{TP + FN} \tag{4-12}$$

F value is an integrated index of Precision and Recall.

$$F = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{4-13}$$

Some classifiers may detect more positive data from the same data set while others detect less. The precision of former classifiers is higher while the recall is smaller even though classification accuracy is the same. There is a trade-off relationship. In the research, classifiers output the probability of each category. By changing the probability threshold of each class, precision and recall can be adjusted with the same classification accuracy. Fig. 4.5 (a) shows a PR curve with recall in a horizontal axis and precision in a vertical axis. When the numbers of data of two categories are the same, the maximum classification accuracy is achieved at the point where a slope of a PR curve is -1.

When negative data is much more than positive data, TN and FP will be much larger than TP and FN. Recall shows a valid value while precision is almost zero because the denominator of precision becomes extremely large. The classifier which categorizes all the data as negative may show high classification



(a) (b)

Fig. 4.5 Indexes (a) PR curve (b) ROC curve

accuracy. Therefore, a new index is necessary. Recall is one of the appropriate indexes. A detection rate of negative data among all the negative data (False Positive Rate, FPR) is also a candidate.

$$FPR = \frac{FP}{FP + TN} \tag{4-14}$$

Fig. 4.5 (b) shows a Receiver Operator Characteristics (ROC) curve with FPR in a horizontal axis and recall in a vertical axis. When the numbers of data of two categories are the same, the maximum classification accuracy is achieved at the point where a slope of a ROC curve is 1. The areas below PR and ROC curves are Area Under the Curve (AUC). AUC corresponds to the overall performance of classifiers, which does not depend on a probability threshold. AUC of a ROC curve was adopted in the research.

In the research, the numbers of data of categories were adjusted to be almost the same. Therefore, classification accuracy has the same meaning as AUC. However, in the application of classifiers to the whole measurement data, regions of negative data are much larger than regions of target objects. Classifiers may provide much false detection. Therefore, in the research, not only classification accuracy, precision and recall, but also AUC were utilized to evaluate classifiers. AUC is considered to be the most accurate index to evaluate object detection performance. In the research, the probability threshold was relatively high to reduce FPR. The operating range was a left side of a ROC curve. Estimation accuracy of positions of target objects is an another topic as discussed in Chapter 7.

The data was divided into training (teacher) data and test (validation) data to train and evaluate classifiers. The number of test data should be large enough. In the research, manhole and joint data was divided by the ratio 5 : 1. Pipe data was divided by the ratio 4 : 1. Void data was divided by the ratio 2 : 1 considering the required data for transfer learning and model evaluation. When the number of data is large, test data can further be divided into the data for model evaluation and optimization of hyper parameters such as model architectures. However, in the research, the number of the data was not enough to tune hyper parameters.

In the extraction of healthy data, when the number of test data is small, the classification accuracy of each test data set may have variations. The most accurate evaluation method may be to divide the data into N sets to use $N - 1$ sets as training and remaining 1 set as test data, which is called cross validation. The average of classification accuracy of N combinations of data sets is utilized. The disadvantage is, cross validation needs much training time. In the research, cross validation was adopted in the evaluation of an SVM model and CNN model for void detection.

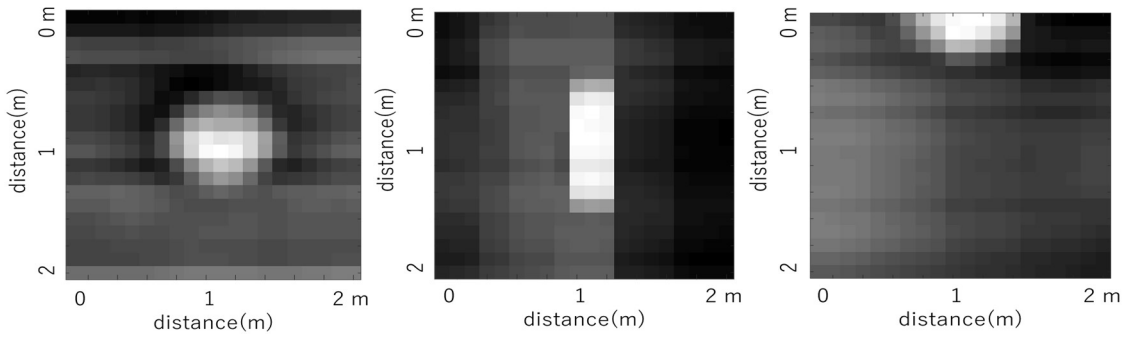
4.3 Production of training data

4.3.1 Methodology

Manhole and joint training data were extracted referring to the coordinate points recorded by inspectors. The center of the data corresponds to the recorded point. The effect of the size of data is discussed in Section 4.4.1. Fig. 4.6 shows the examples of manhole. Fig. 4.7 shows the examples of joint. From Fig. 4.6 (a), (b), the reflections of most manhole were strong, having geometric patterns such as circle and square. The diameter of manhole is several tens cm to 1 m (5 - 15 pixels). A circle shape was dominant though square

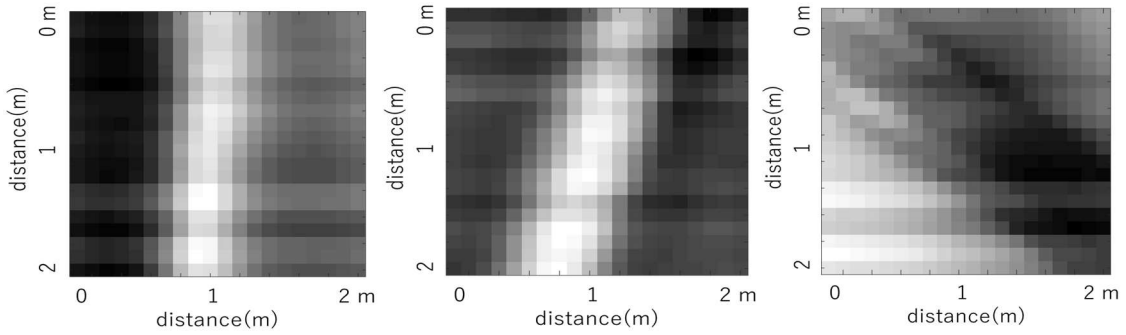
and other types were not negligible. Fig. 4.6 (c) includes only the part of manhole, where detection is considered to be difficult. Joint shows linear reflection patterns as shown in Fig. 4.7 (a), (b). The reflections of most joint were also strong because the majority of joint is made of metal. Some joint shows an opposite phase as shown in Fig. 4.7 (c), possibly covered by rubber.

Training data of pavement section without manhole and joint was randomly extracted as shown in Fig. 4.8. Pavement section was affected by the characteristics of the system and inhomogeneity of the properties



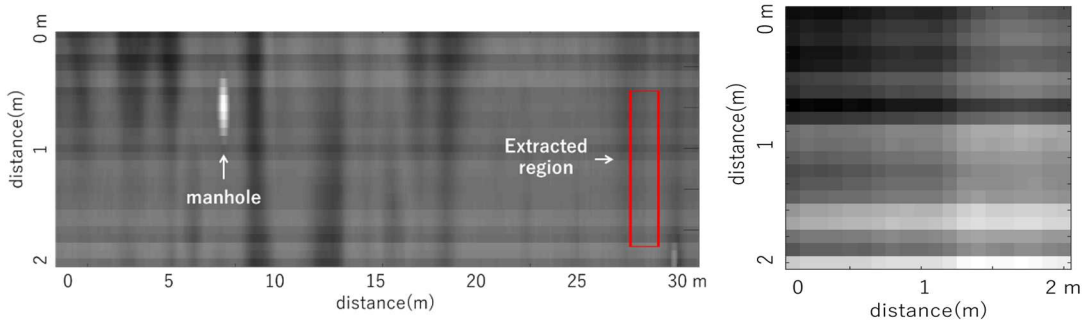
(a) (b) (c)

Fig. 4.6 Surface reflections of manhole.



(a) (b) (c)

Fig. 4.7 Surface reflections of joint.



(a) (b)

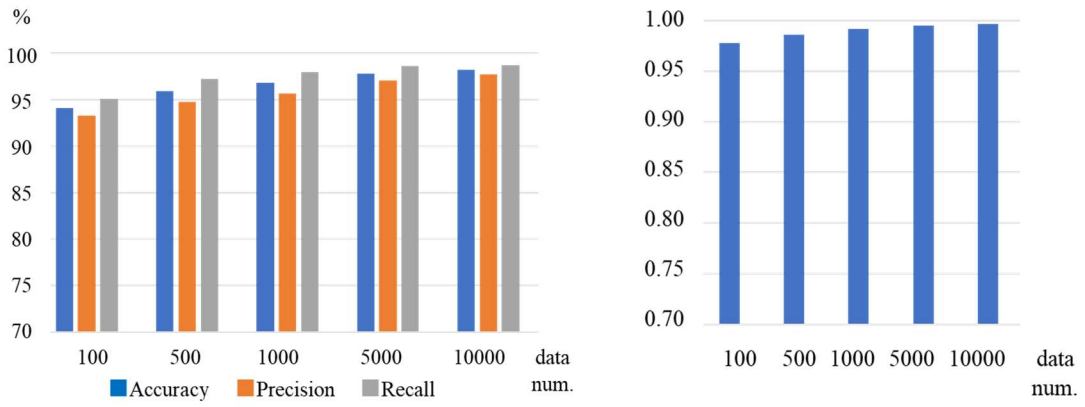
Fig. 4.8 Extraction of pavement section (a) extraction (b) extracted image.

of pavement surface such as permittivity and road unevenness. The effects of these variations were small comparing to the reflection amplitudes of manhole and joint. The reflection patterns of manhole and joint were apparent from the training data shown in Fig. 4.6 and Fig. 4.7.

The number of pavement section data was the same as manhole and joint in two category classification. In three category classification, the number of pavement section data was intermediate between manhole and joint. In the research, the number of manhole was larger than joint. Therefore, the detection rate of each category was also shown as a confusion matrix in three class classification. The sizes of manhole, joint and pavement section data were the same. The same preprocessing filters and feature extraction methods were applied. The data was not normalized because values themselves have a meaning. A data type was double precision floating point with positive and negative values.

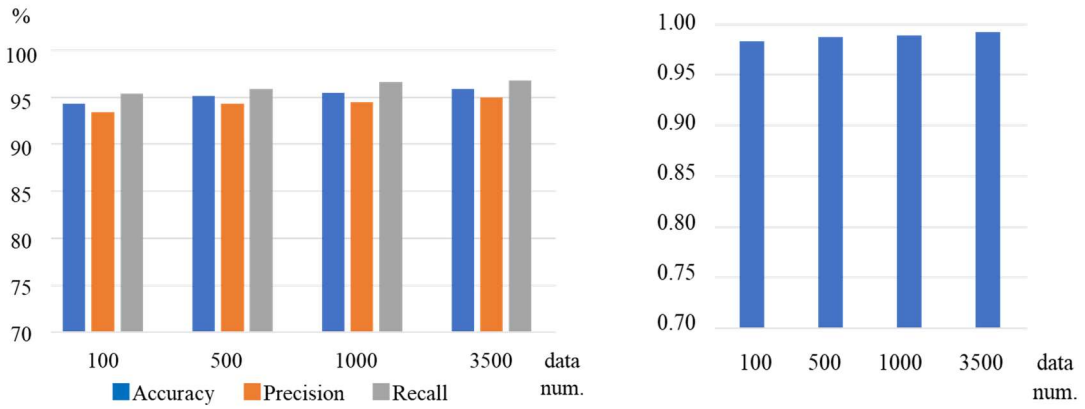
4.3.2 Effect of number of training data

The number of training data may affect classification accuracy. Fig. 4.9 and Fig. 4.10 show the effect of the number of training data N on classification accuracy and AUC of an SVM model shown in Fig. 4.1.



(a) (b)

Fig. 4.9 Effect of number of training data (manhole) (a) Accuracy, Precision, Recall (b) AUC



(a) (b)

Fig. 4.10 Effect of number of training data (joint) (a) Accuracy, Precision, Recall (b) AUC

Because the numbers of manhole, joint and pavement section data were the same, accuracy, precision and recall show similar values. Though AUC does not necessarily correspond to classification accuracy one by one, AUC shows a similar tendency. AUC shows a relatively high value in the research, which depends on the characteristics of test data.

From 4.9 and Fig. 4.10, the classification accuracy of manhole converged at $N = 1000$. Joint converged at $N = 100$. The results indicate the number of data was enough for training SVM models. The classification accuracy of manhole at $N = 5000$ was about 98 %. Even at $N = 100$, classification accuracy was about 94 %, only by 4 % decrease. The classification accuracy of joint at $N = 1000$ was about 96 % and at $N = 100$ was about 94 %, only by 2 % decrease in this case. The required number of data was much smaller than CNN models as shown in Section 6.2. The number of the parameters of an SVM model to be trained was much smaller than a CNN model. Furthermore, only some support vectors affect the results. Feature values may efficiently extract features. The generalization capability of an SVM model is high [1)-10)]. On the other hand, because of a simple model architecture, the classification accuracy of an SVM model is expected to be the same or smaller than a deep CNN model. In the case of void, because the number of data was about 88 and training data was about 50, even an SVM model was difficult to be applied. In SVM model training, validation accuracy was constant at each training. Furthermore, cross validation was conducted. Therefore, the variations of classification accuracy of each training and test data set were almost zero.

4.4 Optimization of algorithm

4.4.1 Size of input image

A square region of a surface reflection with (p, p) pixels was extracted as shown in Fig. 4.11. Then q

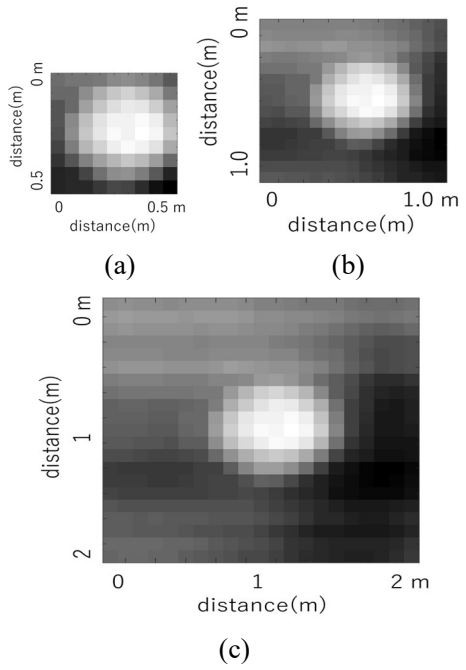


Fig. 4.11 Size of input image

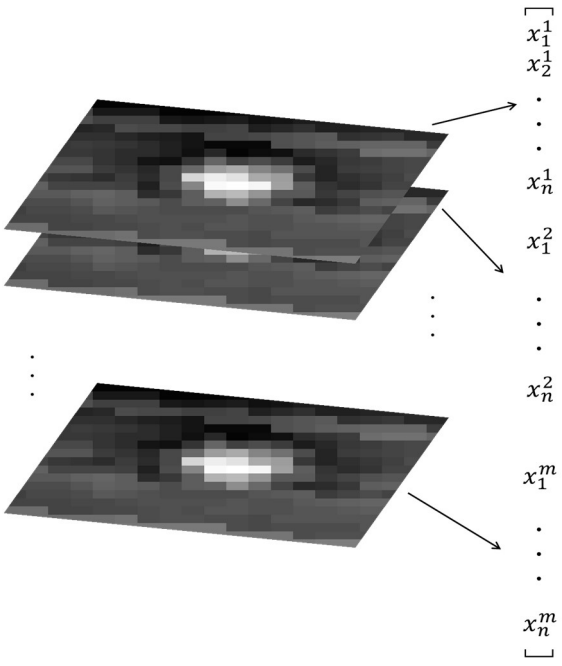
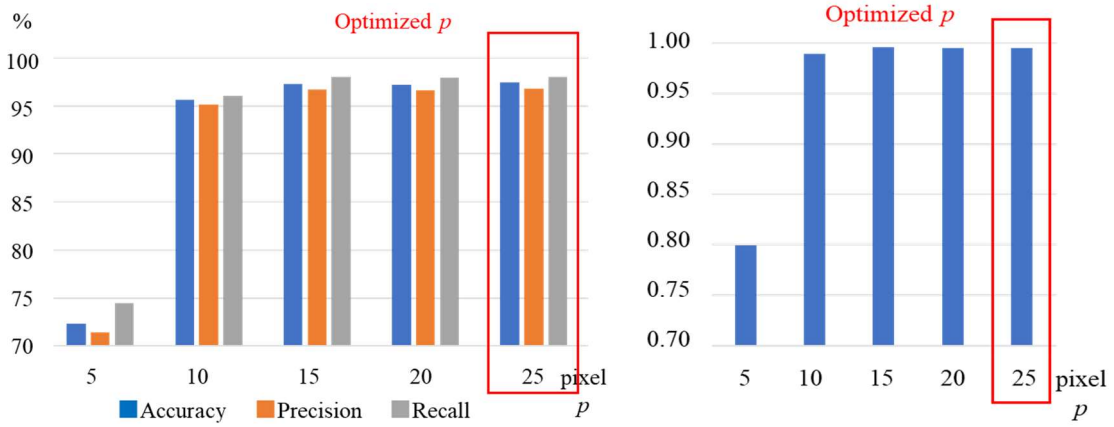


Fig. 4.12 3D data

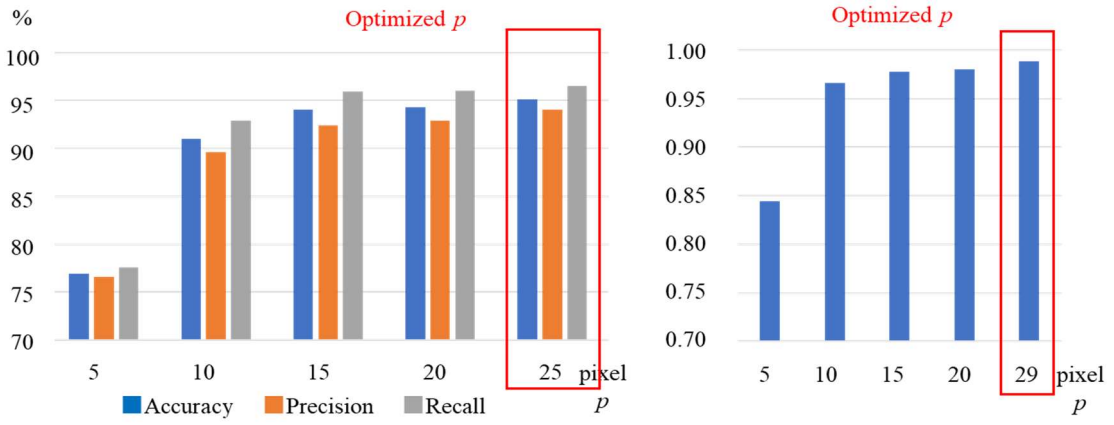
cross-sections in a depth direction was extracted to compose 3D data. An SVM model can only handle 1D data. Therefore, after extracting feature values in each cross-section, vectors were concatenated as shown in Fig. 4.12. p and q were the parameters to be considered.

The effect of p on the classification accuracy of the SVM model shown in Fig. 4.1 was shown in Fig. 4.13 and Fig. 4.14. $q = 1$, which corresponds to 2D data. In the case of manhole, classification accuracy was about 72 % at $p = 5$ from Fig. 4.13. Classification accuracy was about 97 % at $p = 15$, by about 25 % increase. Over $p = 15$, classification accuracy converged. Considering the diameter of manhole, $p = 5$ was not enough while $p = 15$ was enough. In the case of joint, classification accuracy was about 77 % at $p = 5$ from Fig. 4.14. Classification accuracy was about 95 % at $p = 25$, by about 18 % increase. Classification accuracy was maximum at $p = 25$ because the reflection patterns of joint lay in all the channels. Therefore, from the results of Fig. 4.13 and Fig. 4.14, to conduct three-category classification of manhole, joint and pavement section, $p = 25$ was the optimal parameter.



(a) (b)

Fig. 4.13 Effect of p on classification accuracy (manhole) (a) Accuracy, Precision, Recall (b) AUC



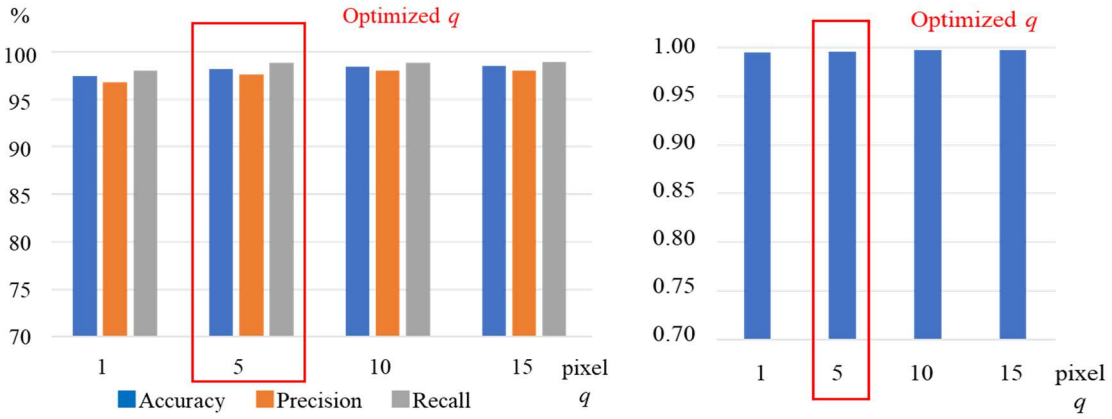
(a) (b)

Fig. 4.14 Effect of p on classification accuracy (joint) (a) Accuracy, Precision, Recall (b) AUC

The effect of q on classification accuracy at $p = 25$ is also shown in Fig. 4.15 and Fig. 4.16. From the results of Fig. 4.15 and Fig. 4.16, classification accuracy was about 98 % and 96 % at $q = 5$, by about 1 % increase compared to $q = 1$ in the both cases. Multiple reflections between antennas and manhole or joint appear below manhole and joint. By utilizing the information in a depth direction, the reflection patterns of manhole and joint were more accurately classified. Classification accuracy converged over $q = 5$. From the results of Fig. 4.13 - Fig. 4.16, $p = 25$ and $q = 5$ were considered to be the optimal parameters. These parameters were adopted hereafter.

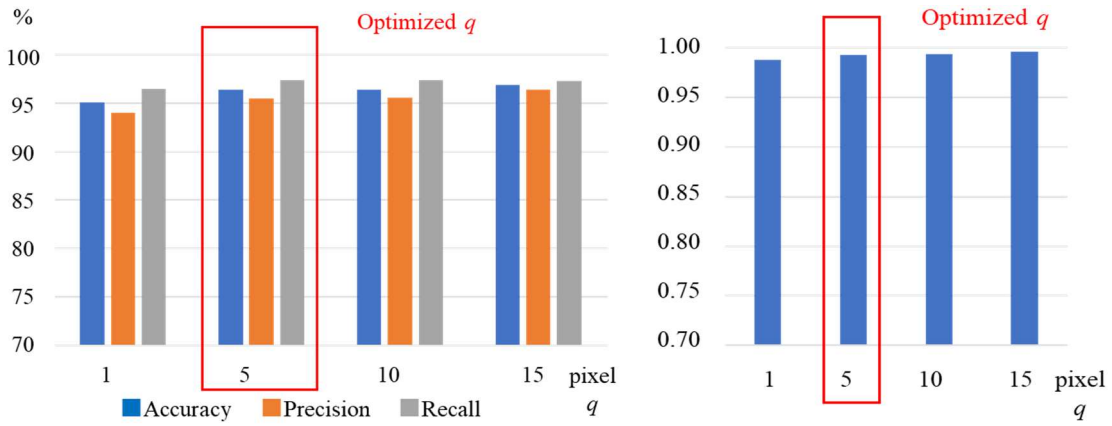
4.4.2 Preprocessing filter

In image processing, preprocessing filters are convoluted to images to enhance image quality and efficiently extract features [122) - 126)]. A second derivative filter is one of the most fundamental edge enhancement filters. A second derivative filter is defined in a certain direction. Considering the coordinate (x, y), Laplacian filter f_l was derived below by combining the second derivative filters in x and y



(a) (b)

Fig. 4.15 Effect of q on classification accuracy (manhole) (a) Accuracy, Precision, Recall (b) AUC



(a) (b)

Fig. 4.16 Effect of q on classification accuracy (joint) (a) Accuracy, Precision, Recall (b) AUC

directions. The raw image and after a Laplacian filter are shown in Fig. 4.17 (a) (b).

$$f_l = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix} \tag{4-15}$$

Usually, edge enhancement filters amplify noise. Therefore, smoothing filters are used together. Typical filters of the combination of edge enhancement and smoothing filters are a Laplacian of Gaussian (LoG) filter and Sobel filter [123), 124)]. A LoG filter is derived by the second derivative of a Gaussian filter, one of the smoothing filters. In terms of a Sobel filter, a filter f_{sx} in a x direction is the combination of a derivative filter in a x direction and Gaussian filter in a y direction as shown below.

$$f_{sx} = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \tag{4-16}$$

Two images after applying filters in a x direction F_{sx} and y direction F_{sy} were obtained from one image. The summation of the square of F_{sx} and F_{sy} is the final output of a Sobel filter.

$$F_s = \sqrt{{F_{sx}}^2 + {F_{sy}}^2} \tag{4-17}$$

The image after Sobel filter is shown in Fig. 4.17 (c).

Edge detection is enabled by setting thresholds after applying edge enhancement filters to images. After

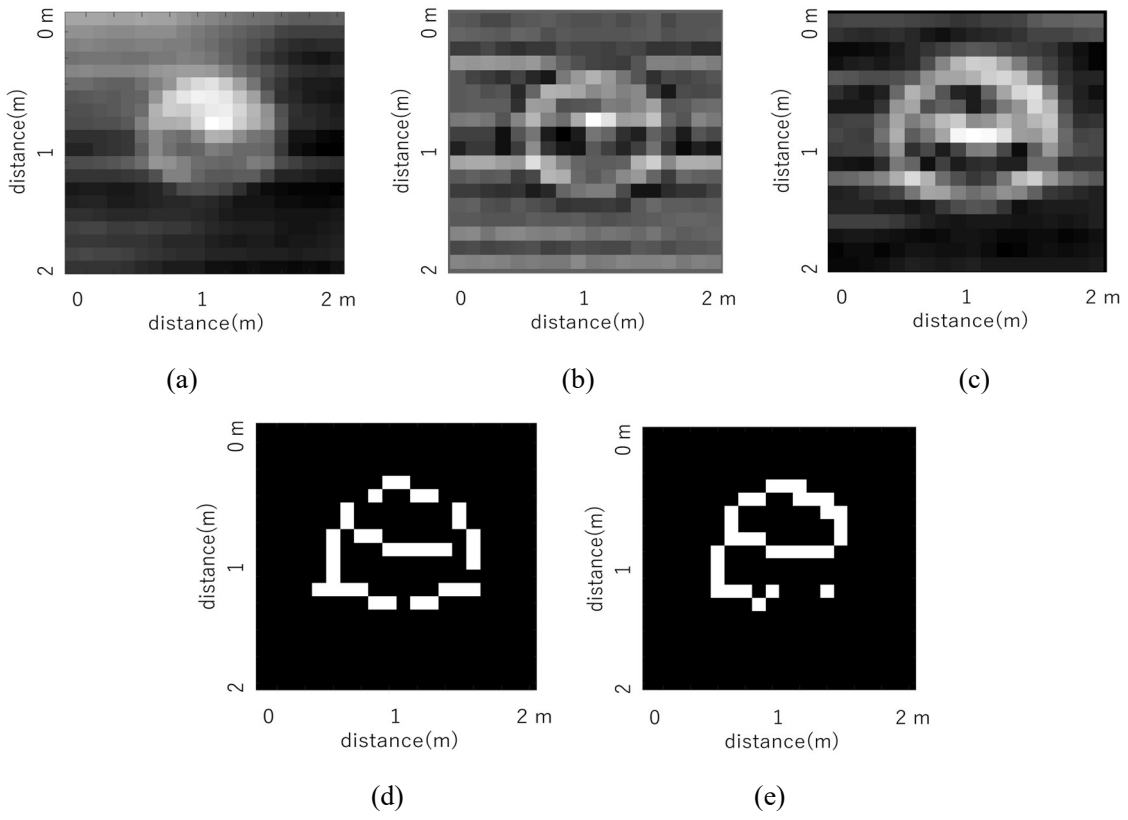
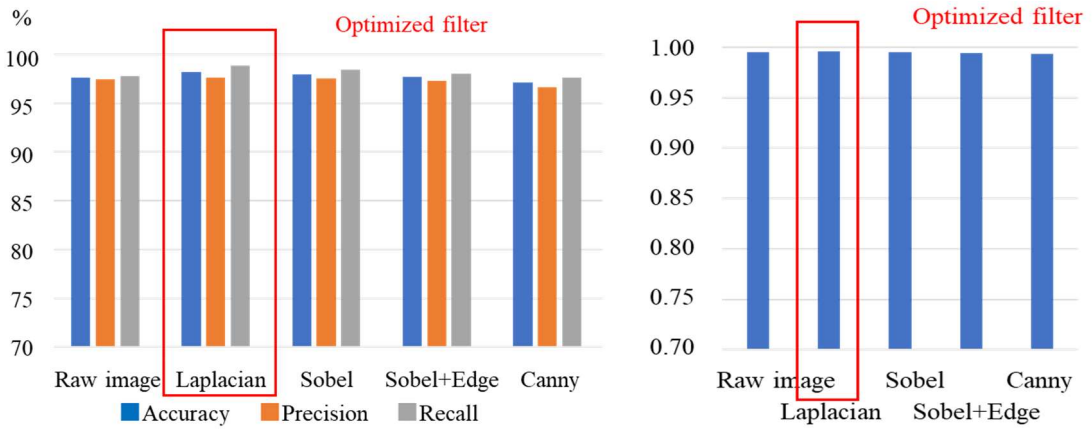


Fig. 4.17 Preprocessing filters (a) raw image (b) Raplacian filter (c) Sobel filter (d) Sobel filter + threshold (e) Canny method

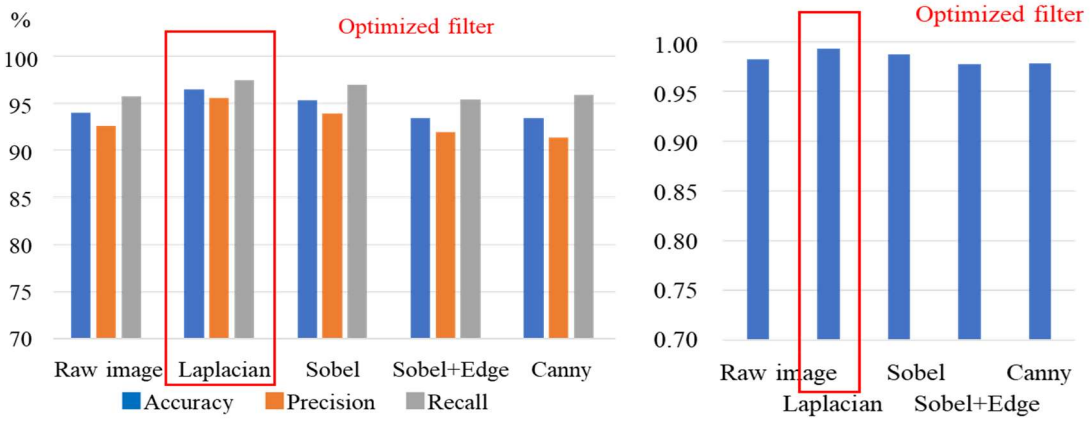
edge detection, images are binarized. Threshold method with a Sobel filter is shown in Fig. 4.17 (d). Other typical edge detection method is Canny method [124]. In Canny method, a gradient of each pixel is calculated after applying Gaussian and Sobel filters. The extremum points of gradients are extracted by comparing the gradients of adjacent pixels. Obtained points are categorized by two thresholds into edge, no edge and pending which will further be inspected by the categories of adjacent pixels. The image after Canny method is shown in Fig. 4.17 (e).

The effects of preprocessing filters on classification accuracy are shown in Fig. 4.18 and Fig. 4.19. Laplacian filter shows the highest classification accuracy from Fig. 4.18 and Fig. 4.19. Classification accuracy is about 98 % and 97 % in the case of manhole and joint respectively, by about 1 % and 2 % increase compared to a raw image. The most appropriate preprocessing filter depends on the characteristics of target features. After a Laplacian filter, the edges of manhole and joint were emphasized. After edge



(a) (b)

Fig. 4.18 Effect of preprocessing filters on classification accuracy (manhole) (a) Accuracy, Precision, Recall (b) AUC



(a) (b)

Fig. 4.19 Effect of preprocessing filters on classification accuracy (joint) (a) Accuracy, Precision, Recall (b) AUC

detection, classification accuracy decreases by about 1 - 2 %. The SVM model may utilize not only the edges of manhole and joint but also the whole feature of an image. From the above discussions, a Laplacian filter was adopted in the research.

4.4.3 Feature value and kernel function

Typical feature values utilized in object recognition are Scale-Invariant Feature Transform (SIFT), Speeded-Up Robust Features (SURF) and HOG features [123), 127), 128)]. The extraction of SIFT features consists of detecting feature points and calculation of feature values. Firstly, a raw image $I(x,y)$ is convoluted by a Gaussian filter f_G defined below.

$$f_G = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{4-18}$$

A scale parameter σ is changed to obtain images for each scale, which are called Difference of Gaussian (DoG) images. DoG images compose 3D data. Feature points are obtained by searching extremums in 3D space of DoG images. The extremums have scale-invariance.

The regions around the feature points whose size corresponds to σ are extracted as target images. Images I_x, I_y after a smoothing filter in each direction are substituted in Eq. (4-19) and Eq. (4-20) to calculate the histograms of angle θ with brightness m as a weight.

$$m(x,y) = \sqrt{\left(\frac{dI_x}{dx}\right)^2 + \left(\frac{dI_y}{dy}\right)^2} \tag{4-19}$$

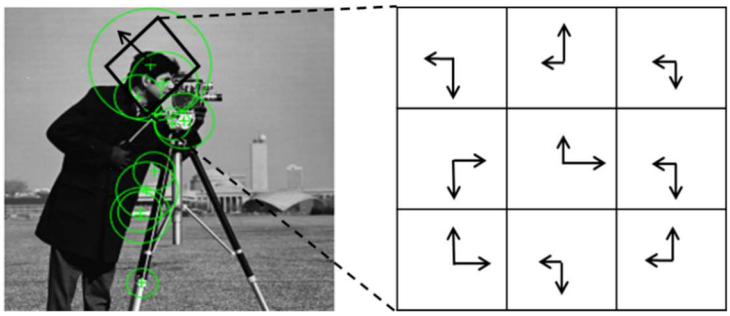


Fig. 4.20 SIFT feature and histograms

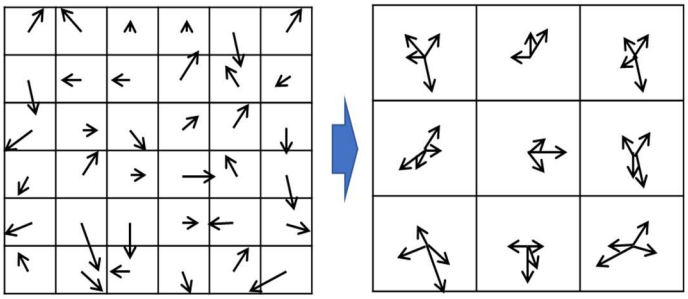


Fig. 4.21 Brightness in each pixel and conversion to histograms of cells.

$$\theta(x,y) = \arctan\{\frac{I_y(x,y)}{I_x(x,y)}\} \tag{4-20}$$

The regions are rotated by the mode angles of the histograms θ to accurately extract feature values as shown in Fig. 4.20. The regions are divided into $N_x \times N_y$ cells. The histograms of θ with a weight m at $N_x \times N_y$ cells are SIFT features. The dimension of SIFT features is $N = N_x \times N_y \times N_\theta$ with the number of bins of a histogram N_θ. SURF features are a simplified version of SIFT features. A Gaussian filter of SIFT features is approximated by a simple box filter. θ is roughly estimated by wavelets.

HOG features do not need the extraction of feature points. Brightness $m(x,y)$ and angle $\theta(x,y)$ at each pixel $I(x,y)$ is directly calculated. Then, neighboring pixels are integrated into a cell to plot the histograms of θ with a weight m as shown in Fig. 4.21. Then, neighboring cells are integrated into a block to normalize the histograms in blocks. HOG features are utilized in tracking objects from movies because

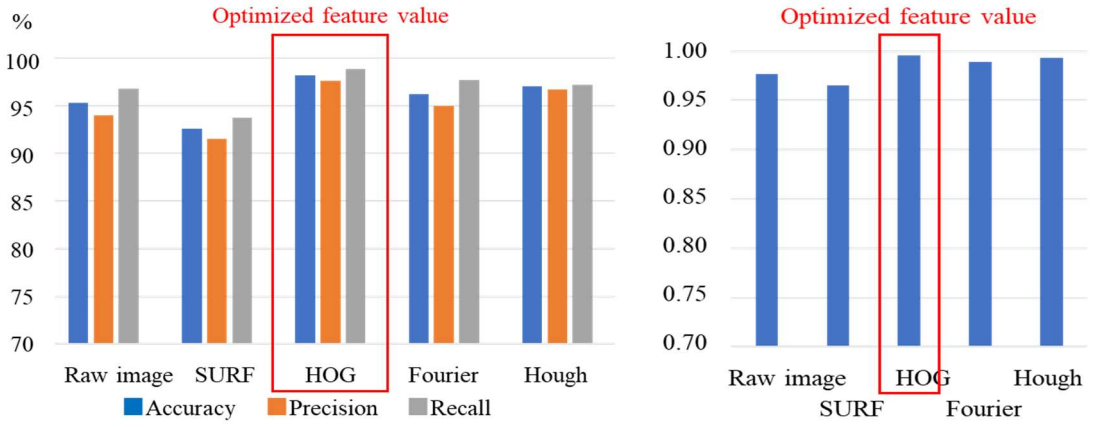


Fig. 4.22 Effect of feature value on classification accuracy (manhole) (a) Accuracy, Precision, Recall (b) AUC

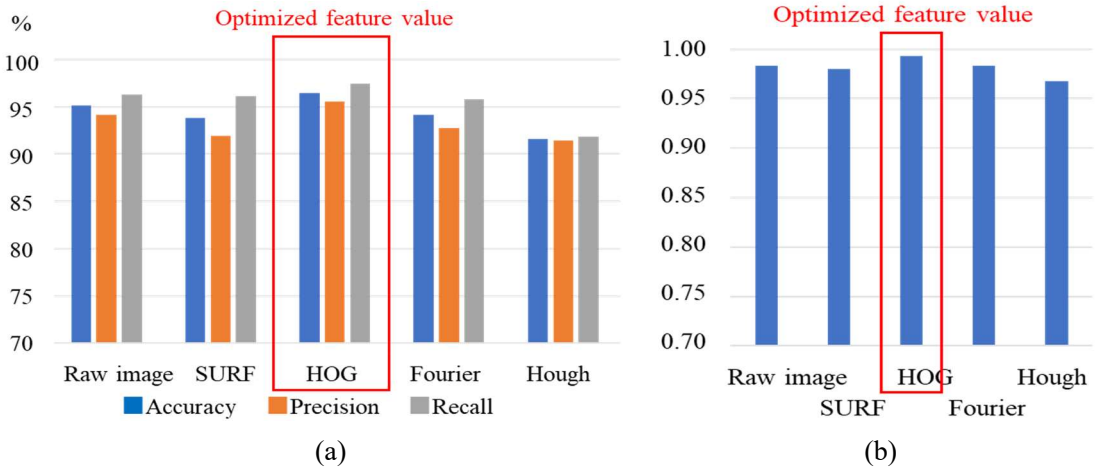


Fig. 4.23 Effect of feature value on classification accuracy (joint) (a) Accuracy, Precision, Recall (b) AUC

calculation cost is low [119), 120), 128)]. The disadvantage is, HOG features are not robust to geometric transformations such as scale change and rotation.

Appropriate feature values depend on the characteristics of target features. In the research, the classification accuracy of SURF and HOG features was compared to raw images. Other feature values for comparison were Fourier and Hough transform. In Hough transform, edges were detected after a Sobel filter. The theory of Hough transform is explained in Section 7.2.1. Fig. 4.22 and Fig. 4.23 show the effect of feature values on classification accuracy. From Fig. 4.22 and Fig. 4.23, HOG features show the highest classification accuracy. HOG features improved classification accuracy by about 3 % in the case of manhole and 1 % in the case of joint compared to raw images. HOG features efficiently extract the geometric features of the edges of manhole and joint. The classification accuracy of SURF features slightly decreased in all the cases. There was a difficulty in choosing an appropriate threshold. SURF features extract many extremums

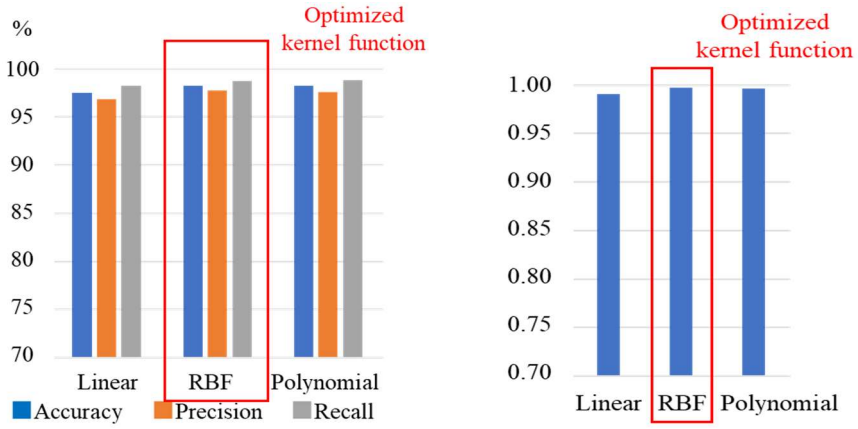


Fig. 4.24 Effect of kernel function on classification accuracy (manhole) (a) Accuracy, Precision, Recall (b) AUC

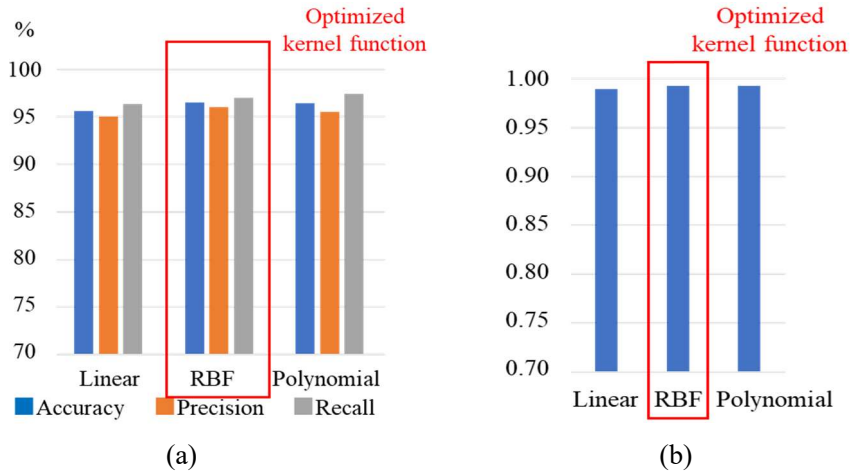
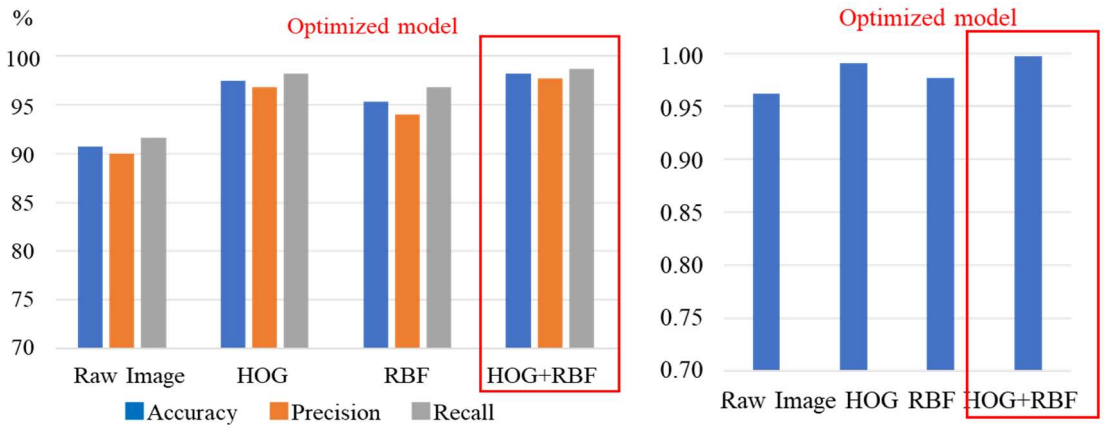


Fig. 4.25 Effect of kernel function on classification accuracy (joint) (a) Accuracy, Precision, Recall (b) AUC

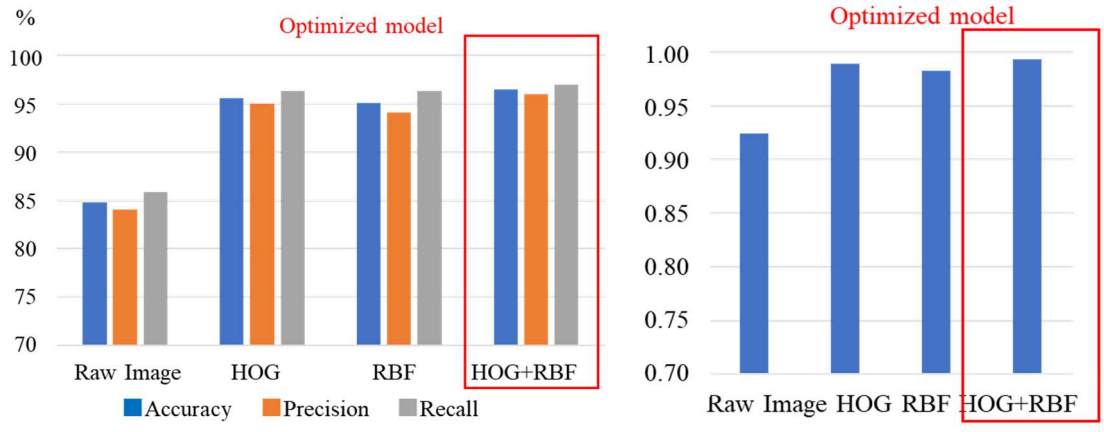
irrelevant to manhole and joint, possibly related to noise when a threshold is too small. No extremum is extracted when a threshold is too large. Fourier and Hough transform improved the classification accuracy of manhole by about 1 - 2 % while the classification accuracy of joint decreased by about 1 - 4 %. In the edge detection of Hough transform, most information of the images was lost. Fourier coefficients are equivalent to raw images though Fourier coefficients may not be an efficient feature value for detecting manhole and joint. From the results of Fig. 4.23 and Fig. 4.24, HOG features were adopted.

Another important hyper parameter except for a normalization factor and outlier rate are a kernel function. Fig. 4.24 and Fig. 4.25 shows the effect of a kernel function on classification accuracy. Compared to raw images, classification accuracy improved slightly by about 1 % in the case of an RBF kernel and Polynomial kernel of two degrees in the both manhole and joint cases. A nonlinear kernel may increase the performance of the SVM model. The calculation cost of an RBF and Polynomial kernel is comparable. In the research,



(a) (b)

Fig. 4.26 Comparison of the combination of feature value and kernel function (manhole) (a) Accuracy, Precision, Recall (b) AUC



(a) (b)

Fig. 4.27 Comparison of the combination of feature value and kernel function (joint) (a) Accuracy, Precision, Recall (b) AUC

RBF kernel was chosen as one of the most standard kernel functions.

Fig. 4.26 and Fig. 4.27 show the effect of the combination of feature values and kernel functions on classification accuracy compared to raw images. From Fig. 4.26 and Fig. 4.27, even only HOG features or an RBF kernel improved classification accuracy by about 5 - 7 % in the case of manhole and 10 - 11 % in the case of joint. The utilization of either feature values or nonlinear kernels drastically increased the performance. The combination of HOG features and an RBF kernel shows the highest classification accuracy, about 96 % in the case of manhole and 98 % in the case of joint. It was about 8 % increase in the case of manhole and 12 % increase in the case of joint. Fig. 4.28 shows the examples of detected manhole. The proposed algorithm detect manhole even if manhole is small and a part of manhole is observed as shown in Fig. 4.28 (c), (d). Manhole cannot be detected when reflection patterns are weak and noise is dominant as shown in Fig. 4.28 (e), (f).

4.5 Extension to multi-class SVM

A three-category, manhole, joint and pavement section SVM model was developed considering the optimal parameters of each category. The three-category model enables unified evaluation and high-speed detection of manhole and joint than applying each two-category SVM model. SVM is inherently a two-category model.

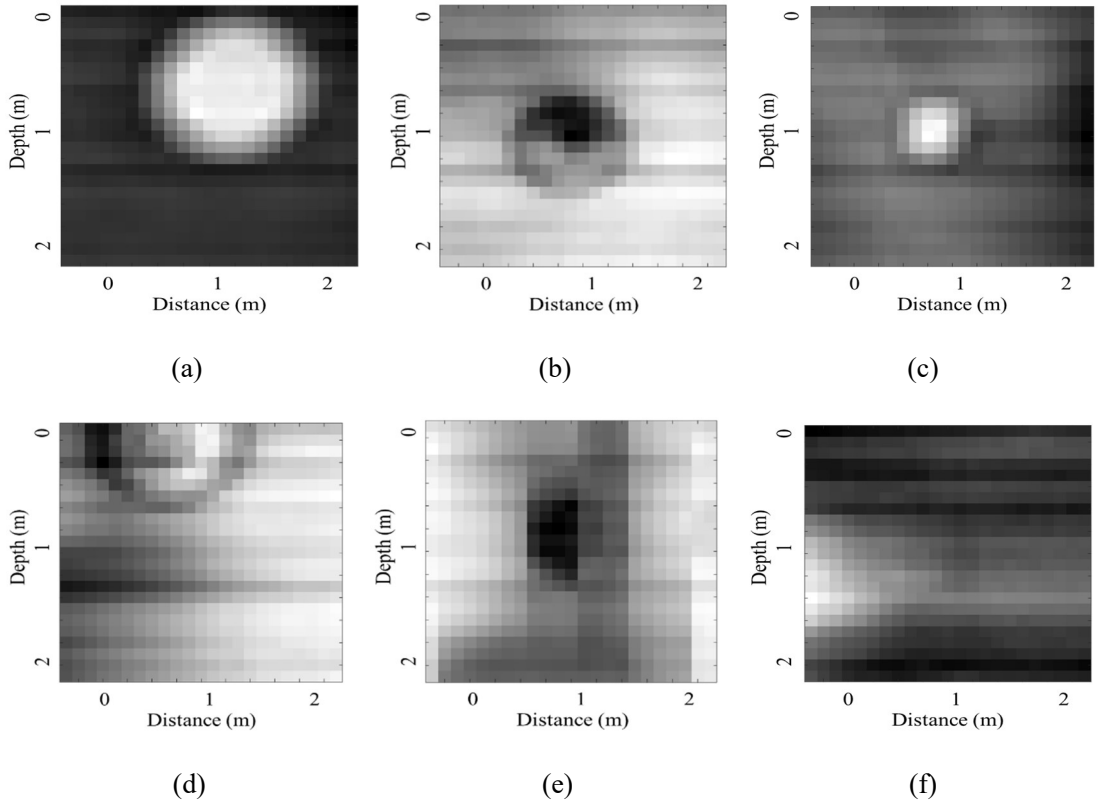
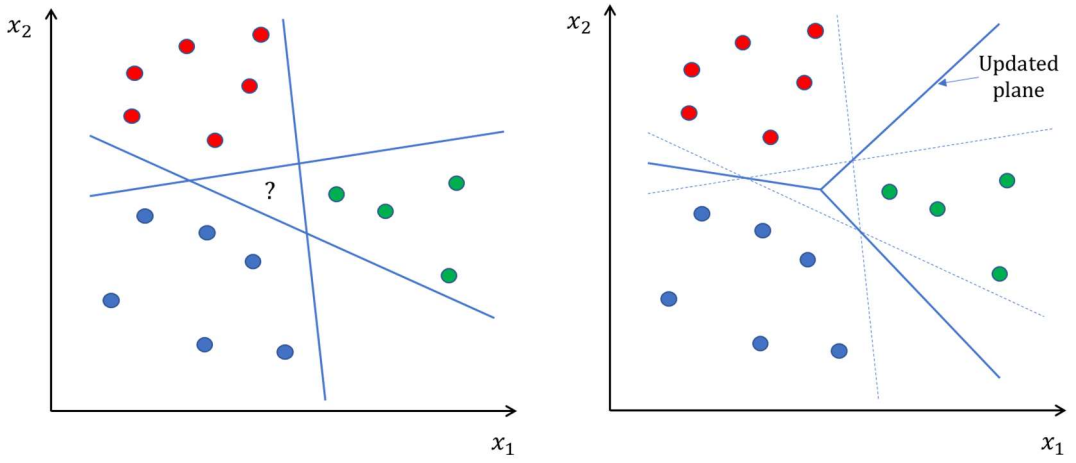


Fig. 4.28 Examples of detection (a) (b) detected from raw images (c) (d) detected only by HOG feature and RBF kernel (e) (f) not detected by HOG feature and RBF kernel

Multi-class classification can be conducted by one-to-one, one-to-others and Error Correcting Output Code (ECOC) method [114), 115)].

One-to-one method for N class classification utilizes $n = {}_N C_2$ two-category classifiers. n classifiers output n binary data for one input. The simplest idea is the principle of majority rule. The most voted class by n classifiers is the output category. The total training time of one-to-one method is relatively small because a data set for training each classifier is small. However, this method yields an undeterminable region as shown in Fig. 4.29 (a). Previous research proposed some solutions such as hierarchizing n classifiers or assigning categories by Bayesian estimation utilizing the values of dividing planes f (Fig. 4.29 (b)). In one-to-others method, the data is divided into a target class and other $N-1$ classes for the training of each classifier. The problem of an undeterminable region occurs also in one-to-others method. The category with the maximum f may be adopted as an output class. The methodology to assign categories may affect classification accuracy in the both methods.

ECOC method utilizes M classifiers for N ($< M$) category classification. Each classifier corresponds to each feature. M classifiers output M binary data. As shown in Fig. 4.30, referring to a code book, the class



(a) (b)

Fig. 4.29 Three-class SVM model (a) dividing plane (b) updated plane

Code book

category index	1	2	3
1	1	0	1
2	0	1	0
3	0	1	1

Category allocated by Hamming distance

Output

1
0
0

Fig. 4.30 Hamming distance

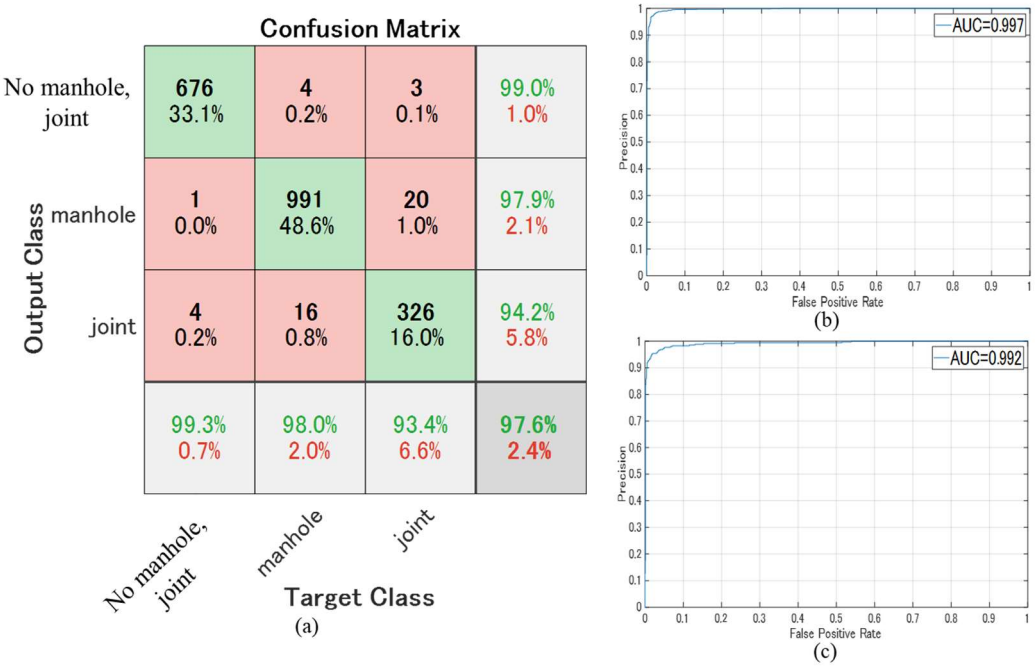


Fig. 4.31 Three-category CNN (a) confusion matrix (b) (c) ROC curves of manhole and joint

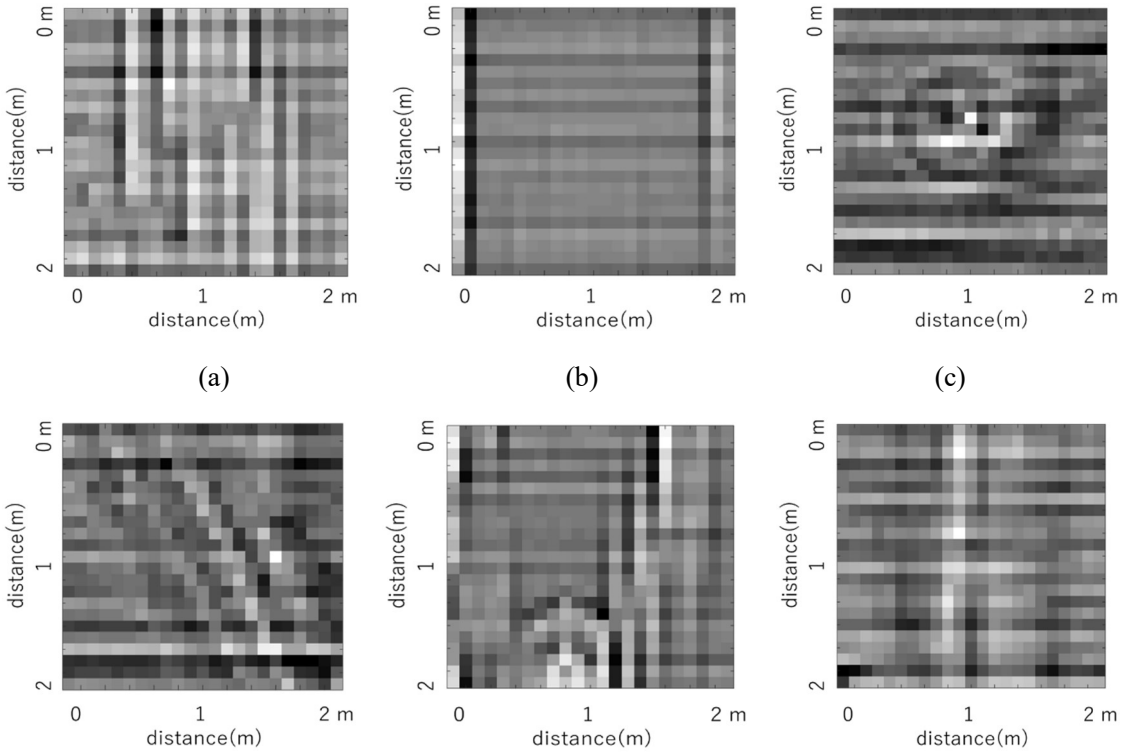


Fig. 4.32 Examples of false detection (true, classified) (a) (pavement section, manhole) (b) (pavement section, joint) (c) (manhole, pavement section) (d) (joint, pavement section) (e) (manhole, joint) (f) (joint, manhole)

of minimum binary loss is output. Binary loss $\hat{l}_k$ of a class k, cost function g, corresponding code m_{kj} and output code s_j is defined below.

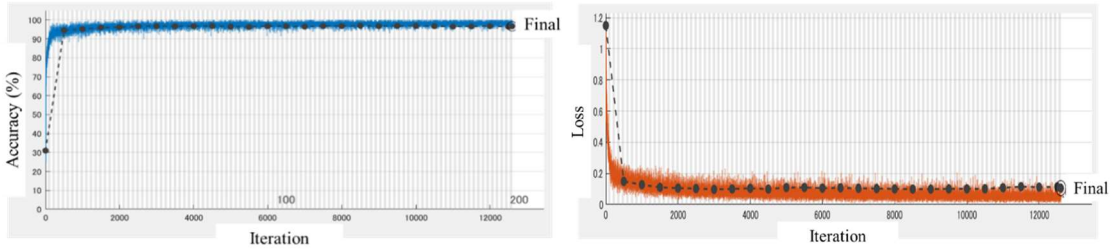
$$\hat{l} = \sum_j g(m_{kj}, s_j) \tag{4-21}$$

In the research, Hamming distance was adopted as g.

$$g = \{\, 1 - \mathrm{sign}(m_{kj} \times s_j) \,\} \tag{4-22}$$

ECOC method can develop accurate classifiers with less calculation cost by choosing Hamming distance a cost function. In the research, a three-category SVM model was developed by ECOC method.

Fig. 4.31 shows a confusion matrix of a three-category SVM model trained by ECOC method. ROC curves of each category are also shown. From Fig. 4.31 (a), the numbers of the false detection of manhole and joint as pavement section and vice versa were relatively small. Manhole falsely detected as joint and vice versa were dominant. The classification accuracy of manhole and other categories was about 99 %. The classification accuracy of joint was about 96 %. Classification accuracy was comparable with each two class SVM. The total classification accuracy of the three categories was about 98 %, intermediate between each two class SVM. From the results, the developed SVM model accurately classifies the manhole, joint and



(a) (b)

Fig. 4.33 Training of CNN model (a) learning curve (b) loss curve

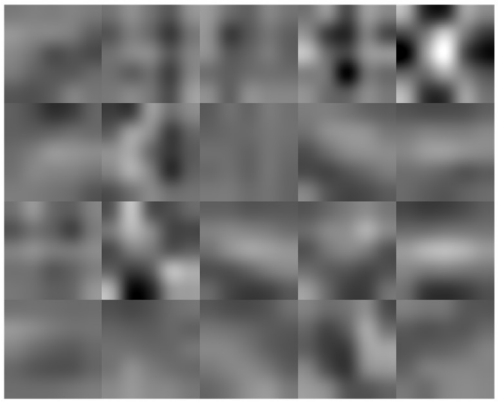


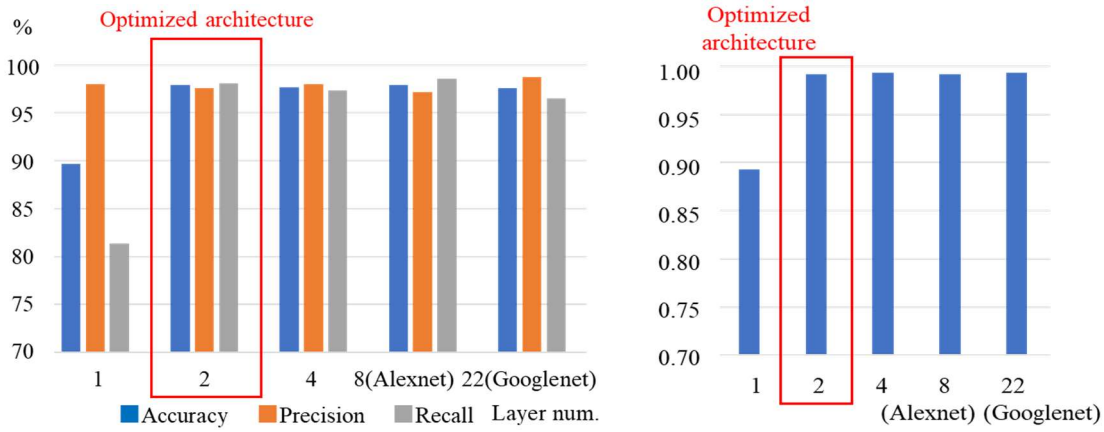
Fig. 4.34 Patterns of convolution filters

pavement section images.

Fig. 4.32 shows the examples of false detection. As shown in Fig. 4.31 (a), (b), pavement section falsely detected as manhole or joint show noise patterns. As shown in Fig. 4.31 (c), (d), in the case of the manhole and joint falsely classified as pavement section, the reflection patterns of manhole and joint were weak. When reflection patterns were covered by noise, images were falsely classified. In the case shown in Fig. 4.31 (f), the weak reflection pattern of joint was falsely detected as manhole. Furthermore, in the case of manhole falsely detected as joint as shown in Fig. 4.31 (e), only the part of manhole was observed in the image. By improving the performance of the system, classification accuracy may further increase.

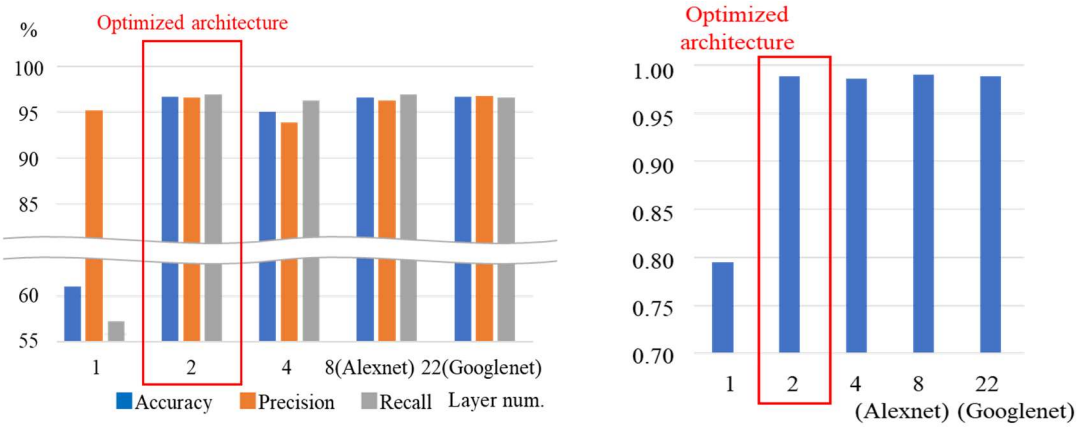
4.6 Comparison to CNN model

The optimized SVM model was compared to a CNN model to evaluate the effect of deep learning. Fig.



(a) (b)

Fig. 4.35 Effect of number of CNN layer on classification accuracy (manhole) (a) Accuracy, Precision, Recall (b) AUC



(a) (b)

Fig. 4.36 Effect of number of CNN layer on classification accuracy (joint) (a) Accuracy, Precision, Recall (b) AUC

4.33 shows learning, validation and loss curves. The output of a CNN model was three categories, manhole, joint and pavement section. Fig. 4.34 shows the examples of trained convolution filters. The convolution filters show certain patterns. Filters are considered to correspond to the features of manhole and joint. Fig. 4.35 and Fig. 4.36 show the effect of the number of layers on classification accuracy. In each case the size ad number of convolution filters were optimized. Compared to a 1 layer simple NN model, classification accuracy increased by about 8 % in the case of manhole and 21 % in the case of joint by adding one convolution layer. However, classification accuracy did not increase by utilizing 4 layers (2 convolution and 2 fully connected), conventional CNN models such as AlexNet (8 layers) and GoogleNet (22 layers). Generally, model performance increases by deep layer architectures. From the results, highly abstract features were not utilized in the classification of manhole and joint.

Fig. 4.37 shows the optimized 3D-CNN model obtained by increasing the dimension of input data. The

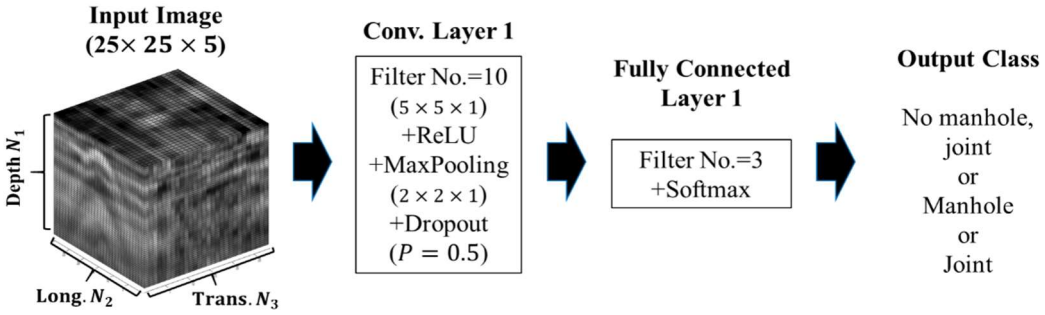


Fig. 4.37 Optimized CNN model

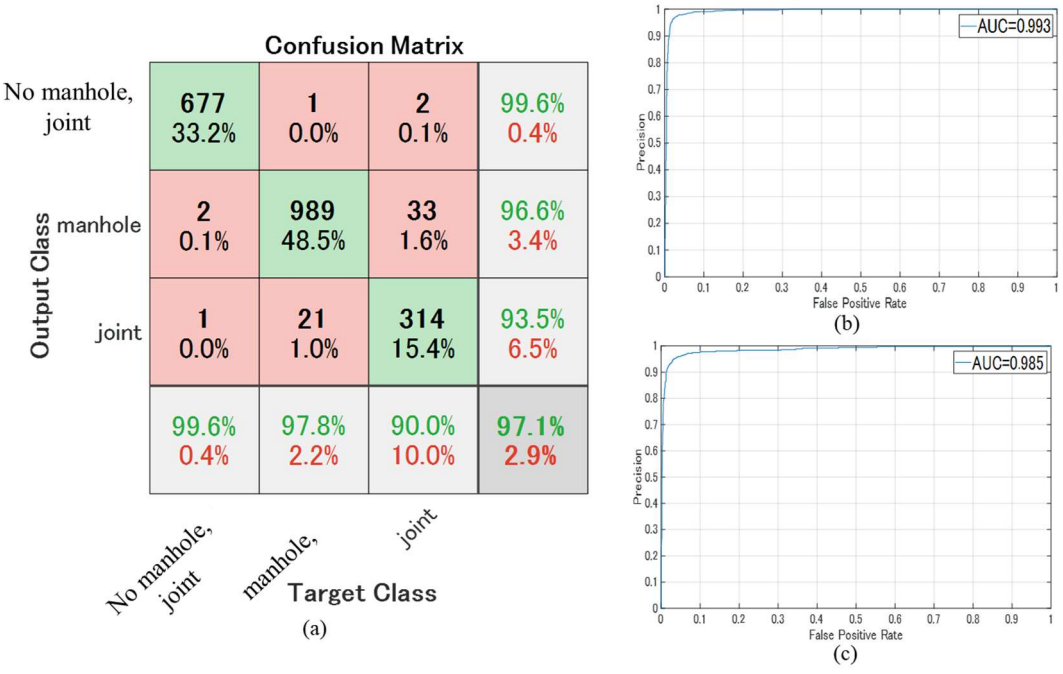


Fig. 4.38 Three category CNN (a) confusion matrix (b) (c) ROC curves of manhole and joint

size of input data is the same as the proposed SVM model. The sizes of convolution filters were also optimized. In this case, the classification accuracy of 2D and 3D convolution filters was the same. Fig. 4.38 shows the three-category confusion matrix of 3D-CNN and each category ROC curves. From Fig. 4.38 (a), total classification accuracy was about 97 %, close to the SVM model. The optimized SVM and 3D-CNN models showed the same level performance. However, the calculation cost of 3D-CNN was much higher than SVM because of complicated model structures. Therefore, in the research, in terms of manhole and joint detection, an SVM model was utilized. In Section 6.5.2, in terms of pipe detection, classification accuracy increased by CNN utilizing deep layer architectures compared to an SVM model.

4.7 Summary

An SVM model for the detection of manhole and joint was developed. The flow chart of image classification by SVM and theory of SVM were explained. The evaluation methods of classifiers were introduced. The discussion is applicable to the whole content of the thesis. In the research, the number of data of each category was adjusted to be the same. Classification accuracy, precision and recall show the similar values. In this thesis, classification accuracy and AUC were shown for each training case. AUC is more precise because AUC is not affected by the number of data of each category. The obtained data was divided into training and test data. Because of the problem of training time, cross validation was adopted only in the training of an SVM model and CNN model for void detection.

In terms of SVM model optimization, the effect of the size of input data was evaluated. Then, the combination of preprocessing filters and feature values was optimized. Even either feature values or nonlinear kernels drastically increased classification accuracy because the performance of classifiers was improved. By utilizing a Laplacian filter as a preprocessing filter, HOG features as feature values and an RBF kernel as a kernel function, the highest classification accuracy was obtained. The edges of manhole and joint were appropriately extracted by the proposed model.

The number of the output classes of an SVM model was augmented to three categories, manhole, joint and pavement section by ECOC method. Total classification accuracy was about 98 %, intermediate between each two class SVM model. The developed SVM model accurately classifies manhole, joint and pavement section. The classification accuracy of the SVM model was close to the optimized CNN model. The increase of the number of layers does not help improve classification accuracy of the CNN model because highly abstract features were not utilized. Considering calculation cost, 3D-CNN has no advantages in the detection of manhole and joint. The developed SVM model was utilized in Section 7.1 to estimate the positions of manhole and joint.

Chapter 5: 3D void training by 2D-FDTD method

In this chapter, the methodology to produce 3D void data by 2D-FDTD method was proposed. A 3D-CNN model was trained by the produced data. The theory of FDTD method was explained. The effects of electromagnetic parameters of soil and geometry of void on reflection patterns were analyzed. A 3D reflection pattern was simulated by 3D-FDTD method. Then, the 3D reflection patterns were reproduced by 2D-FDTD method. The simulated reflection patterns were compared to experimental field and measurement data. The classification accuracy of a 3D-CNN model was compared to a 2D-CNN model. The simulation results of pipes were introduced in Appendix C.

5.1 Theory of FDTD method

Other than FDTD method, Finite Element Method (FEM) and Method of Moments (MoM) are the popular electromagnetic simulation methods [129), 130)]. FDTD method divides a target region into grids. The difference equations of electric and magnetic fields are derived to sequentially calculate the fields of the next step. FDTD method is often applied to the analysis of propagation of waves in an open region. FEM divides a region into finite elements. The functionals of elements with initial and boundary conditions are solved. FEM is mainly applied to a closed region with complex surfaces. MoM is one of the Boundary Element Methods (BEM). MoM represents an electric current as the linear combination of fundamental functions. Maxwell equations are decomposed by fundamental functions and then integrated. The calculation cost of MoM depends on the choice of fundamental functions. In GPR method, the shapes of objects are usually not so complex. However, a target region is open and relatively wide. Therefore, FDTD method was adopted in the research.

5.1.1 Maxwell's equations, constitution laws and boundary conditions

In FDTD method, the differences of Maxwell's equations are solved in each step [129), 130)]. Maxwell's equations consist of Faraday's law (the change of magnetic flux density B produces an electric field E), Ampère's law (electric flux density D produces a magnetic field H) and electric and magnetic charge conservation laws [131) - 133)]. J_e, ρ_e represents an electric current and electric charge density. J_m, ρ_m represents a magnetic current and magnetic charge density.

$$\nabla \times E + \frac{\partial B}{\partial t} = -J_m \tag{5-1}$$

$$\nabla \times H - \frac{\partial D}{\partial t} = J_e \tag{5-2}$$

$$\nabla \cdot D = \rho_e \tag{5-3}$$

$$\nabla \cdot B = \rho_m \tag{5-4}$$

Among the 8 variables of Maxwell's equations, E and D, B and H, J_e and E, J_m and H are not independent. The variables are related by the electromagnetic characteristics of a propagating medium, permittivity ε, permeability μ, electric conductivity σ_e and magnetic conductivity σ_m. Eq. (5-5) - Eq. (5-8) is called constitution laws. ${J_e}^{ex}$, ${J_m}^{ex}$ is external electric and magnetic current density.

$$D = \varepsilon E \tag{5-5}$$

$$B = \mu H \tag{5-6}$$

$$J_e = \sigma_e E + {J_e}^{ex} \tag{5-7}$$

$$J_m = \sigma_m H + {J_m}^{ex} \tag{5-8}$$

ε, μ, σ_e, σ_m is time and space dependent. Space dependency is also called inhomogeneity. A region which consists of several media such as asphalt and soil layers and medium whose characteristics are spatially distributed are the examples of inhomogeneity. If parameters ε, μ, σ_e, σ_m depend on the direction of propagating waves, it is called anisotropy. Parameters which have anisotropy is represented by tensors. Some media show nonlinearity and frequency dispersion (frequency dependence). In terms of subsurface sensing by GPR method, generally propagating medium is assumed to be isotropic and linear. In lossy media such as wet soil, σ_e, σ_m are assumed to have frequency dispersion.

From Eq. (5-1) - Eq. (5-8), unknown variables are 8. Therefore, all the variables can be solved for provided initial and boundary conditions. Electromagnetic waves are caused by inducing a current or voltage in antennas. In terms of boundary (continuity) conditions, usually there is no electric and magnetic currents on medium boundaries. Therefore, the components of D, B parallel to boundaries are continuous (Fig. 5.1). The components perpendicular to boundaries are discontinuous by electric charge density ω_e and magnetic charge density ω_m on the boundaries.

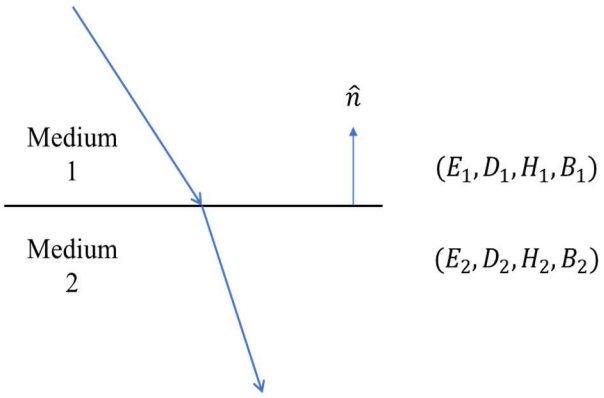


Fig. 5.1 Boundary conditions

$$\hat{n} \cdot (D_1 - D_2) = \omega_e \tag{5-9}$$

$$\hat{n} \cdot (B_1 - B_2) = \omega_m \tag{5-10}$$

$$\hat{n} \times (H_1 - H_2) = 0 \tag{5-11}$$

$$\hat{n} \times (E_1 - E_2) = 0 \tag{5-12}$$

In the case of a perfect electric and magnetic conductor, $E = H = 0$ inside the conductor. The parallel components of E and H of adjacent media should also be 0. Metal pipes were assumed as perfect electric and magnetic conductors in the research. Electromagnetic waves completely reflect at the surface of metal pipes.

5.1.2 Yee's algorithm

In FDTD method, the differences of Eq. (5-1) - Eq. (5-4) are calculated in time and space directions (Yee's algorithm) [134]. In Fig. 5.2 (a), in a time direction, an integer-order electric field E^n and non-integer-order magnetic field $H^{n+\frac{1}{2}}$ are assigned. The differential terms are replaced by the central differences below.

$$\frac{\partial \mathrm{E}}{\partial \mathrm{x}}\Big|_{t=(n-\frac{1}{2})\Delta t} = \frac{E^n - E^{n-1}}{\Delta t} \tag{5-13}$$

$$\frac{\partial \mathrm{H}}{\partial \mathrm{x}}\Big|_{t=n\Delta t} = \frac{H^{n+\frac{1}{2}} - H^{n-\frac{1}{2}}}{\Delta t} \tag{5-14}$$

Substituting Eq. (5-13) and Eq. (5-14) in Eq. (5-1) and Eq. (5-2) provides the time-differential equations below.

$$E^n = a_e E^{n-1} + b_e[\nabla \times H^{n-\frac{1}{2}} - {J_e}^{ex\,n-\frac{1}{2}}] \tag{5-15}$$

$$H^{n+\frac{1}{2}} = a_m H^{n-\frac{1}{2}} - b_m[\nabla \times E^n - {J_m}^{ex\,n}] \tag{5-16}$$

The coefficients a_e, b_e, a_m, b_m are functions of the electromagnetic parameters ε, μ, σ_e and σ_m of medium and time step Δt.

$$a_e = \frac{1 - \frac{\sigma_e \Delta t}{2\varepsilon}}{1 - \frac{\sigma_e \Delta t}{2\varepsilon}} \tag{5-17}$$

$$b_e = \frac{\frac{\Delta t}{\varepsilon}}{1 + \frac{\sigma_e \Delta t}{2\varepsilon}} \tag{5-18}$$

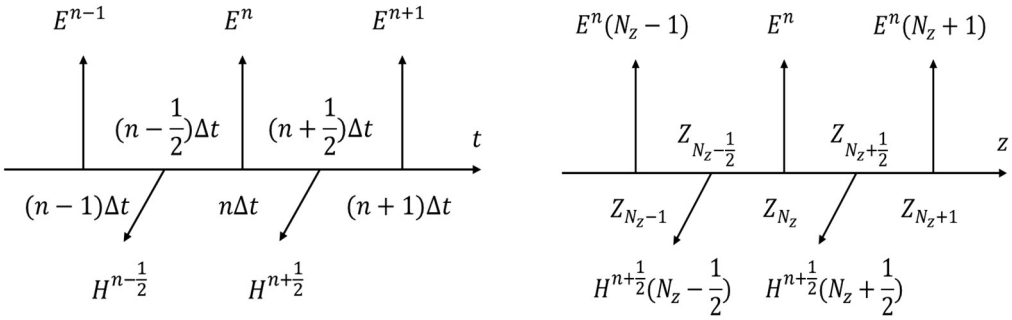
$$a_m = \frac{1 - \frac{\sigma_m \Delta t}{2\mu}}{1 - \frac{\sigma_m \Delta t}{2\mu}} \tag{5-19}$$

$$b_m = \frac{\frac{\Delta t}{\mu}}{1 + \frac{\sigma_m \Delta t}{2\mu}} \tag{5-20}$$

From Eq. (5-15) and Eq. (5-16), E^n and $H^{n+\frac{1}{2}}$ are calculated by E^{n-1} and $H^{n-\frac{1}{2}}$. E^n and H^n follow the law of causality. E^n brings $H^{n+\frac{1}{2}}$ and $H^{n+\frac{1}{2}}$ brings E^{n+1}.

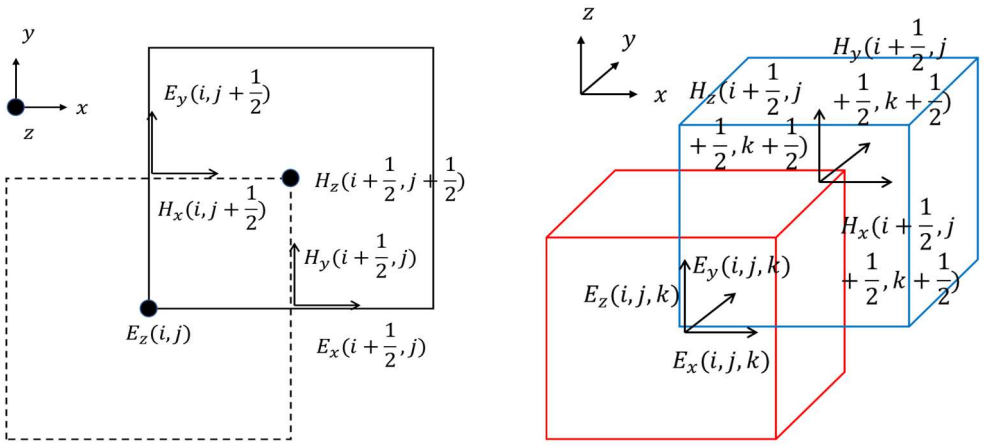
Eq. (5-15) and Eq. (5-16) include space differentials. As shown in Fig. 5.2 (b), in a space direction (z), an integer-order electric field $E^n(z = N_z)$ and non-integer-order magnetic field $H^{n+\frac{1}{2}}(z = N_z + \frac{1}{2})$ are assigned. The central differences provide the space-differential equations below.

$$\nabla \times E^n|_{z=N_z+\frac{1}{2}} = \frac{E^n(N_z + 1) - E^{n-1}(N_z)}{\Delta z} \tag{5-21}$$



(a) (b)

Fig. 5.2 Difference of FDTD method (a) time direction (b) space direction



(a) (b)

Fig. 5.3 Yee cell (a) two dimension (b) three dimension

$$\nabla \times H^{n+\frac{1}{2}}|_{z=N_z} = \frac{H^{n-\frac{1}{2}}(N_z+\frac{1}{2}) - H^{n-\frac{1}{2}}(N_z-\frac{1}{2})}{\Delta z} \tag{5-22}$$

Substituting Eq. (5-21) and Eq. (5-22) in Eq. (5-15) and Eq. (5-16) provides the recursive formulas of E^{n-1} and $H^{n-\frac{1}{2}}$, E^n and $H^{n+\frac{1}{2}}$. In the programs, the sets of Eq. (5-15) and Eq. (5-16), Eq. (5-21) and Eq. (5-22) were alternatively solved.

A two-dimensional wave is represented by electric fields E_x, E_y and a magnetic field H_z (TE mode), magnetic fields H_x, H_y and an electric field E_z (TM mode). Any plane waves are decomposed into TE and TM modes. The number of differential equations of each mode is 3. The total number of differential equations is 6. The number of terms of each equation also increases. Fig. 5.3 (a) shows the example of a two-dimensional Yee cell. $E_z(i,j)$ is calculated by $H_x(i,j\pm\frac{1}{2})$ and $H_x(i\pm\frac{1}{2},j)$. $H_z(i,j)$ is calculated by $E_x(i,j\pm\frac{1}{2})$ and $E_y(i\pm\frac{1}{2},j)$. In the case of a three-dimensional wave, the numbers of differential equations and terms of each equation further increase. Fig. 5.3 (b) shows a three-dimensional Yee cell. E is calculated by surrounding H and H is calculated by surrounding E by the same procedure.

Δx should be small enough to obtain reasonable results. Δx is usually one tenth of the wave length of the maximum frequency considering the permittivity of propagating medium [129), 130)]. In the research, the wave length of the maximum frequency was about 10 cm. Considering the relative permittivity of soil, Δx should be less than 10 mm. However, smaller Δx is, larger the number of cells is, resulting in large calculation cost. Δt should follow the equation below with propagating velocity v considering the stability of the recursive equations Eq. (5-15) and Eq. (5-16), called Courant-Friedrichs-Lewy (CFL) condition.

$$\Delta t < \frac{1}{v\sqrt{(\frac{1}{\Delta x})^2+(\frac{1}{\Delta y})^2+(\frac{1}{\Delta z})^2}} \tag{5-23}$$

When CFL condition does not hold, numerical error diverges to make solutions unstable. Δt depends on Δx. For example, at $v = 3\times10^8\ \mathrm{m/s^2}$ and $\Delta x = \Delta y = \Delta z = 5\ \mathrm{mm}$, Δt should be smaller than 6.7×10^{-1} s. Smaller Δx is, smaller Δt should be. Small Δt increases the number of steps, resulting in large calculation cost. The following section discusses the optimal cell size Δx.

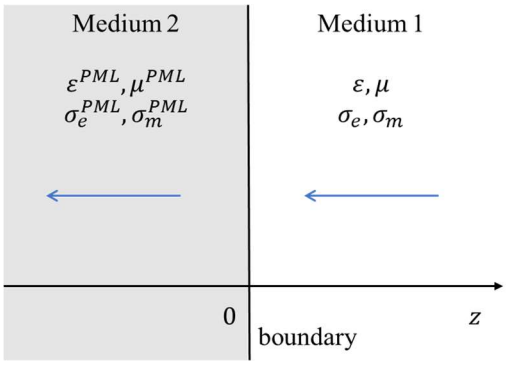


Fig. 5.4 PML boundary

5.1.3 Absorbing boundary conditions

In subsurface sensing environment by GPR method, electromagnetic waves propagate infinitely distant. In electromagnetic simulations, an analysis area should be limited to conduct simulations with reasonable calculation cost. Therefore, the boundary of the region should be Absorbing Boundary Condition (ABC). Mur's boundary condition and Berenger's Perfectly Matched Layer (PML) are the typical ABC conditions. Considering numerical error, PML is prevailing in FDTD method [135].

A reflection coefficient R and transmission coefficient T of medium boundaries are represented by impedance Z_1, Z_2 of two adjacent media.

$$R = \frac{Z_2 - Z_1}{Z_2 + Z_1} \tag{5-24}$$

$$T = \frac{2Z_2}{Z_1 + Z_2} \tag{5-25}$$

$$Z_{1,2} = \sqrt{\frac{\mu(1 + \frac{\sigma_m^{1,2}}{j\omega\mu})}{\varepsilon(1 + \frac{\sigma_e^{1,2}}{j\omega\varepsilon})}} \tag{5-26}$$

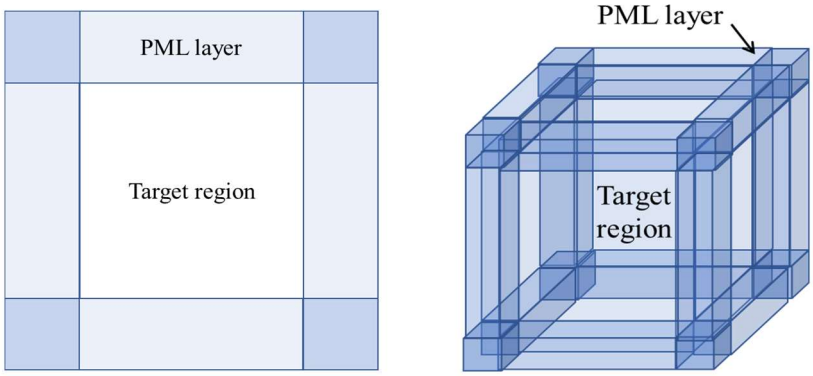
If σ_e, σ_m is small enough, for example in lossless medium, R, T is rewritten below.

$$R = \frac{\sqrt{\varepsilon_1} - \sqrt{\varepsilon_2}}{\sqrt{\varepsilon_1} + \sqrt{\varepsilon_2}} \tag{5-27}$$

$$T = \frac{2\sqrt{\varepsilon_1}}{\sqrt{\varepsilon_1} + \sqrt{\varepsilon_2}} \tag{5-28}$$

In the thesis, Eq. (5-27) and Eq. (5-28) are utilized to discuss the intensity of reflection patterns.

Fig. 5.4 shows one-dimensional PML conditions. Electric conductor σ_e^{PML} and magnetic conductor

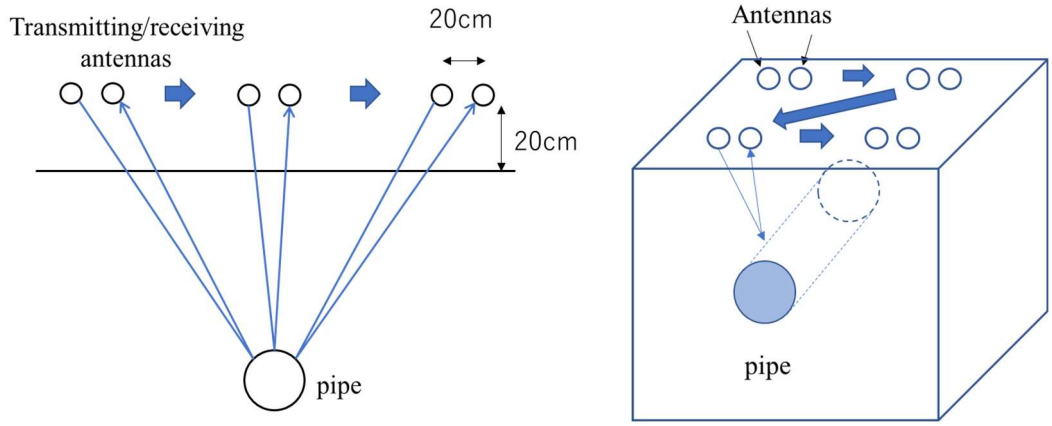


(a) (b)

Fig. 5.5 PML layer settings (a) two dimension (b) three dimension

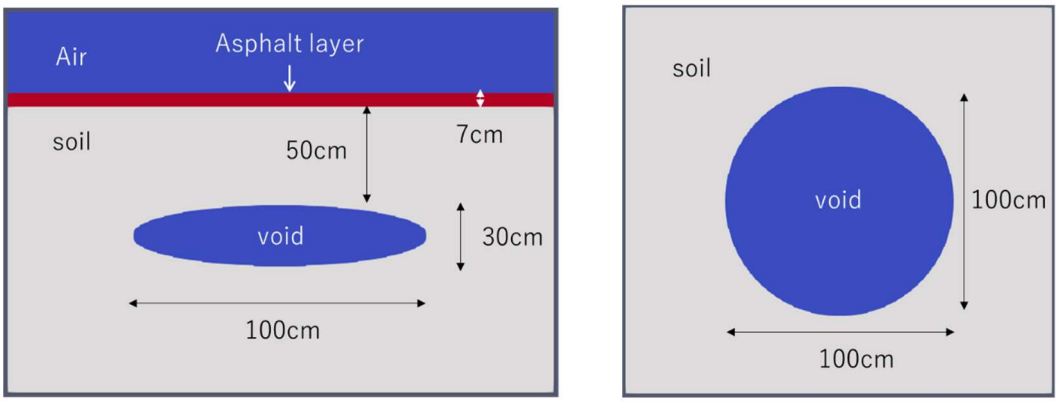
σ_m^{PML} are the parameters of PML. Impedance matching condition $Z_{PML} = Z$ at the boundaries is derived from Eq. (5-24) considering $R = 0$. σ_e^{PML} and σ_m^{PML} are related by Z. Electromagnetic waves attenuate by the factor $e^{-Z\sigma_e^{PML}}$ in PML. In gprMax, considering the electromagnetic characteristics of adjacent media, σ_e^{PML} and σ_m^{PML} are automatically assigned to minimize numerical error at the boundaries. To sufficiently attenuate electromagnetic waves, the number (thickness) of PML should be large enough. On the other hand, considering calculation cost, the number of PML cannot be too large.

In a two-dimensional plane, the number of equations of impedance matching condition does not increase because only the components of fields perpendicular to the boundaries follow the conditions. The number of PML increase depending on the dimension of a simulation model. Fig. 5.5 (a) shows two-dimensional PML. To cover the region, total 8 PML is needed. Fig. 5.5 (b) shows three-dimensional PML. Total 26 PML is needed. The calculation cost of one PML also increases according to the increase of the dimension. In the research, the effect of the number of layers on reflection patterns was discussed.



(a) (b)

Fig. 5.6 Antennas arrangement (a) two dimension (b) three dimension



(a) (b)

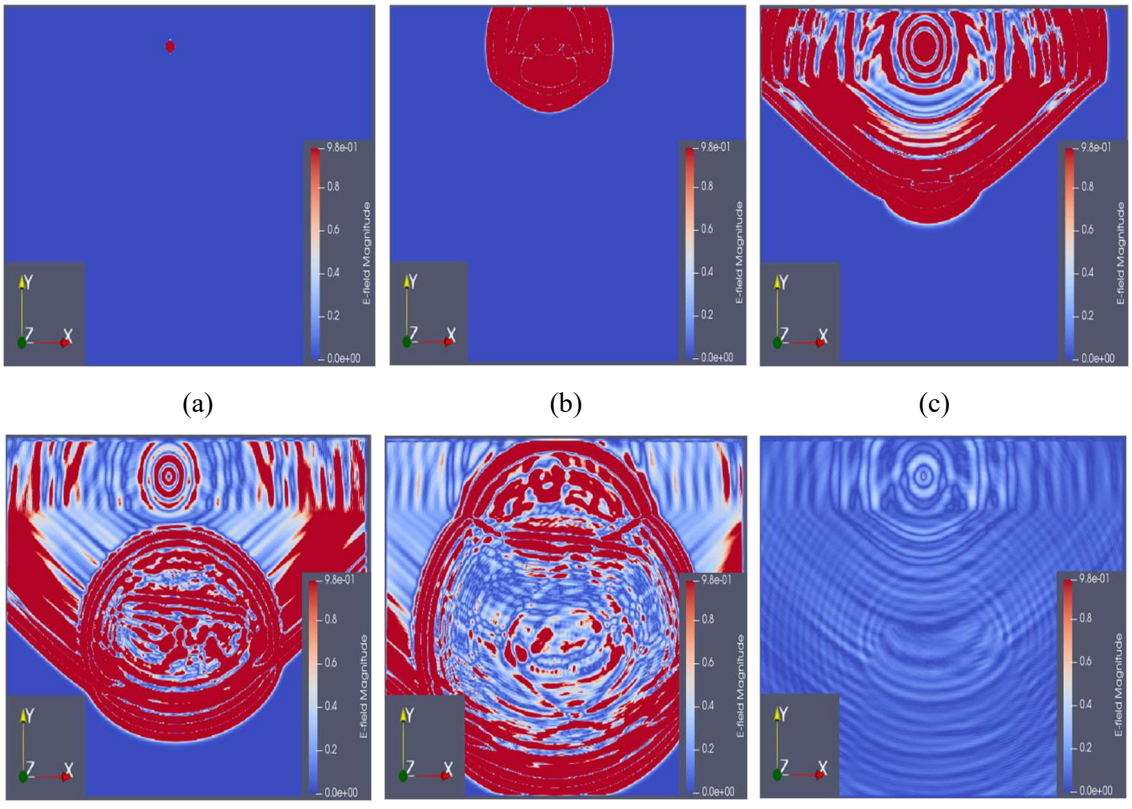
Fig. 5.7 Three-dimensional void model (a) vertical section (b) horizontal section

5.2 Simulation conditions

5.2.1 Production of radar data

One simulation provides a time history of electric field intensity at a certain observation point. This is called A scan. To obtain a 2D radar image, which is also called B scan, the set of transmitting and receiving antennas is moved in a scan (longitudinal) direction with certain san pitch (Fig. 5.6 (a)). In the research, a 2.5 m by 2.5 m square region is considered as a target area. Scan pitch is 7 cm. Simulations were conducted at total 32 observation points. Calculation cost is high. The calculation time to obtain one image is about 2 minutes even by GPU.

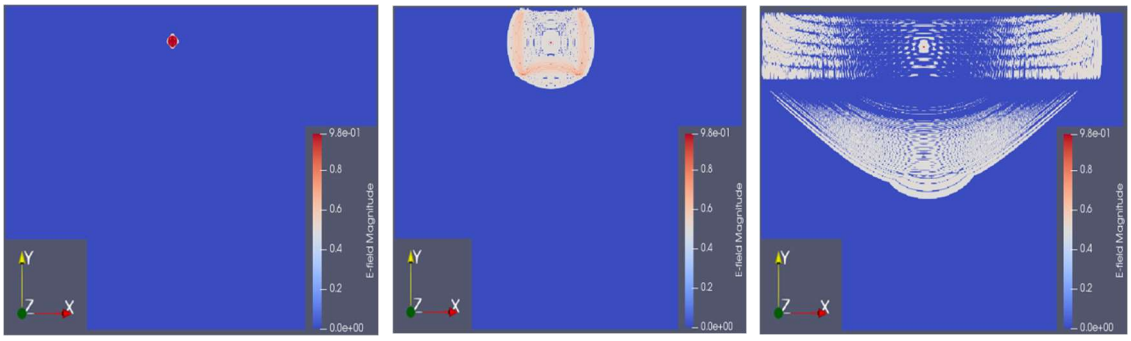
To obtain 3D data, the set of antennas is moved in an observation area (Fig. 5.6 (b)). A time history is obtained at each scan and channel point. A simulated area is a 2.5 m by 2.5 m by 2.5 m 3D region. Channel distance is 7.5 cm. In the research, the calculation time of 3D-FDTD method to obtain a time history is about 30 times larger than 2D-FDTD method. Simulations were conducted at 32 points in a scan direction and 29 points in a channel direction, total 928 observation points. Considering a channel direction, the number of simulation cases is also about 30 times larger. The calculation time to produce one data will be about 900 times larger. It will take one day to produce 1 data and 500 days for 500 data. Therefore, producing 3D



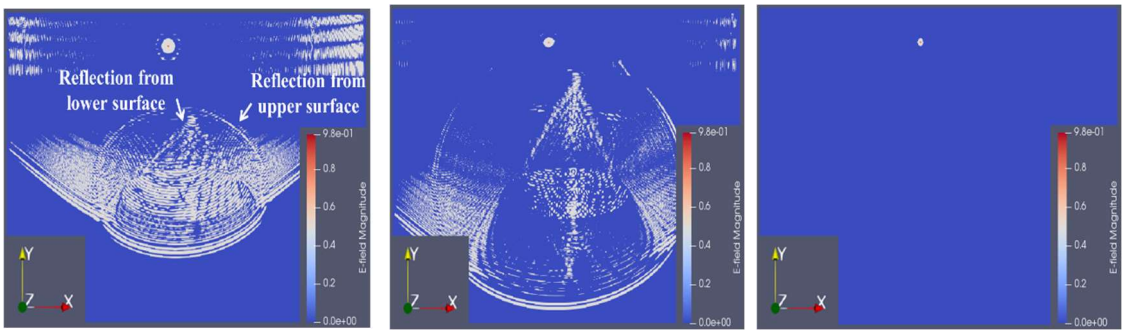
(a) (b) (c)

(d) (e) (f)

Fig. 5.8 Snapshots of 2D void reflection pattern (a) 0 s (b) 2.5×10^{-9} s (c) 5.0×10^{-9} s (d) 7.5×10^{-9} s (e) 10×10^{-9} s (f) 20×10^{-9} s

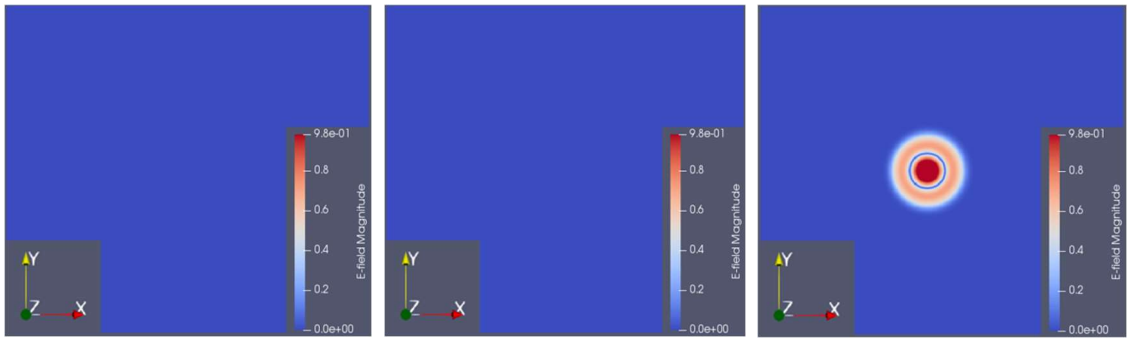


(a) (b) (c)

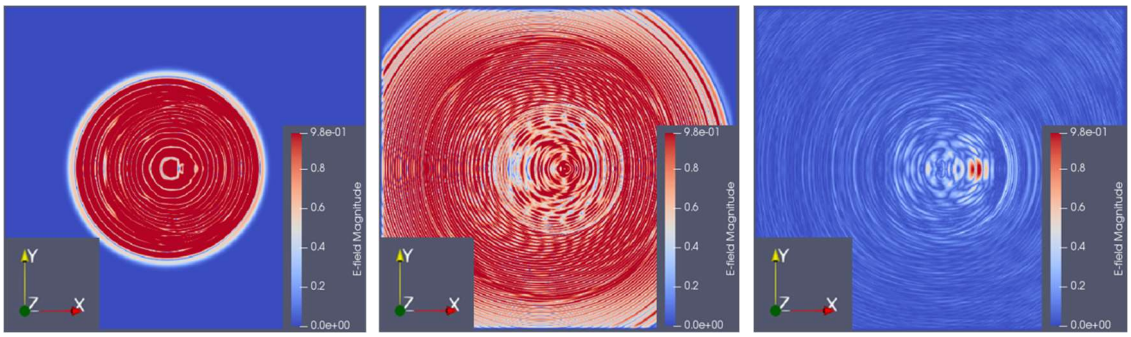


(d) (e) (f)

Fig. 5.9 Snapshots of 3D void reflection pattern (vertical sections) (a) 0 s (b) 2.5×10^{-9} s (c) 5.0×10^{-9} s (d) 7.5×10^{-9} s (e) 10×10^{-9} s (f) 20×10^{-9} s



(a) (b) (c)



(d) (e) (f)

Fig. 5.10 Snapshots of 3D void reflection pattern (horizontal sections) (a) 0 s (b) 2.5×10^{-9} s (c) 5.0×10^{-9} s (d) 7.5×10^{-9} s (e) 10×10^{-9} s (f) 20×10^{-9} s

training data by 3D-FDTD is not feasible.

Fig. 5.7 shows a typical 3D void model of 3D-FDTD method adopted in the research. A 2D void model utilizes the cross-section of the 3D void model shown in Fig. 5.7 (a). The void model consists of air, asphalt, soil layers and void. The relative permittivity of air and void is 1. Asphalt is 4. Soil is 3. A lossless homogeneous medium was assumed. The thickness of an air layer is 40 cm. An asphalt layer is 7 cm. A soil layer is 50 cm. The shape of the typical void model shown in Fig. 5.7 is ellipsoid. Vertical and horizontal sections are ellipse. The lengths in depth and horizontal directions are the parameters. The thickness of void shown in Fig. 5.7 is 30 cm. The lengths in scan and channel directions are 1 m. PML is adopted. Considering the actual measurement conditions, the distance between antennas and asphalt surface is 20 cm. The distance between two antennas is 20 cm.

An input is a voltage time history at a transmitting point polarized in a channel (transverse) direction, which is called a hard source. A negative peak can be observed at asphalt surface. A positive peak appears at the upper surface of void and negative peak at the lower surface. The phase of simulated data was reversed considering the consistency with measurement data.

Fig. 5.8 shows the distribution of electric field intensity (snapshot) of a 2D-FDTD void model shown in

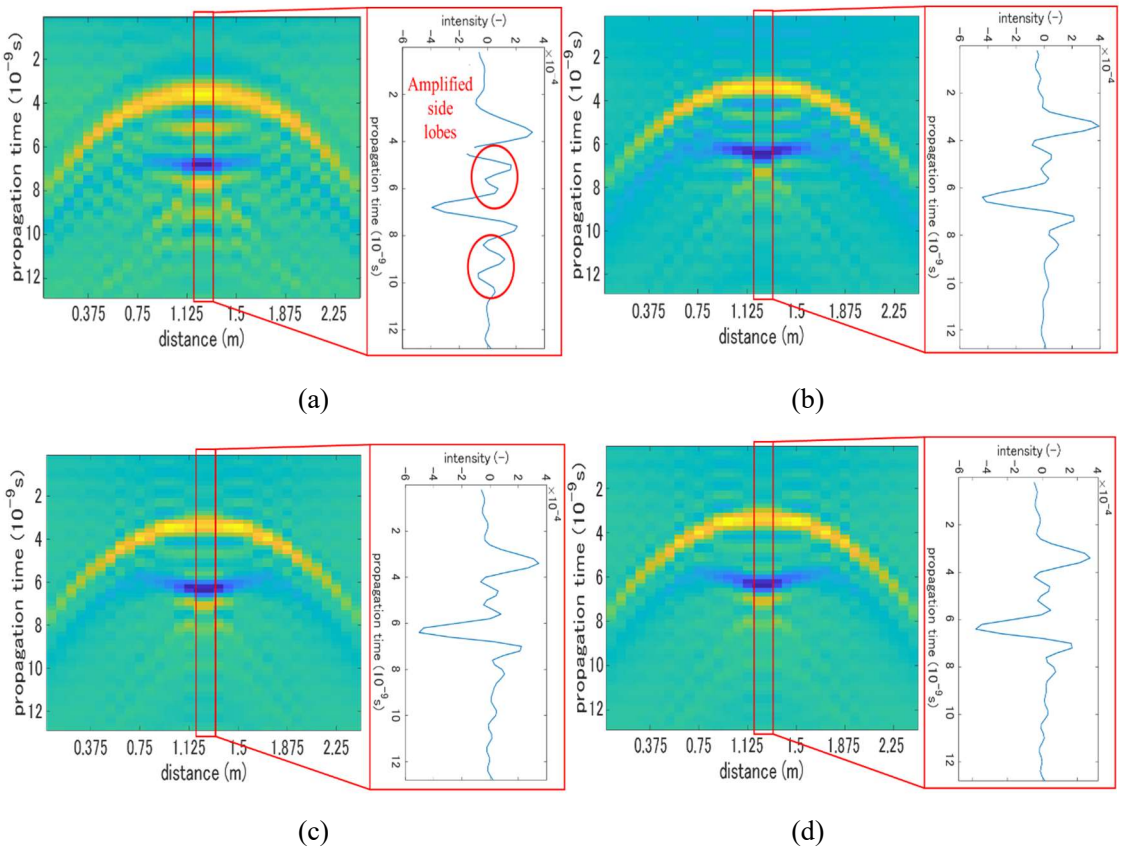


Fig. 5.11 Effect of cell size on reflection patterns and 1/2 point signal (a) 20 mm (b) 10 mm (c) 5 mm (d) 1 mm

Fig. 5.7 with time duration $t_d = 3 \times 10^{-8}$ s. In t_d, electromagnetic waves travel about 9 m in air. Considering the round-trip distance and relative permittivity of soil, the maximum depth will be about 2 m - 3 m, indicating t_d is large enough. The propagation paths of waves in the void model are complicated. A Wave reach the asphalt surface in Fig. 5.8 (b) and the upper surface of the void in Fig. 5.8 (c). The reflections from the upper and lower surface of the void reach the asphalt layer in Fig. 5.8 (d), (e) respectively. After sufficient time, waves attenuate as shown in Fig. 5.8 (f).

Fig. 5.9 and Fig. 5.10 show the vertical and horizontal sections of the snapshots of the 3D void model. Each section passes through the center of the void. In the vertical section, a wave reaches the asphalt surface in Fig. 5.9 (b) and inside the void in Fig. 5.9 (c). The reflection from the upper surface of the void reaches the asphalt layer in Fig. 5.9 (d). In Fig. 5.9 (d), (e), the reflection from the lower surface of the void focusses on the observation point. After sufficient time, waves attenuate as shown in Fig. 5.9 (f). In the horizontal section, a wave reaches the center of the void in Fig. 5.10 (c) and diffuses in Fig. 5.10 (f).

5.2.2 Cell size and number of PML

Fig. 5.11 shows the effect of a cell size Δx on the reflection patterns and signals. In the thesis, a signal at the 1/2 point (1.25 m) in a scan direction is shown. If necessary, a signal at the 1/4 point (0.625 m or 1.875

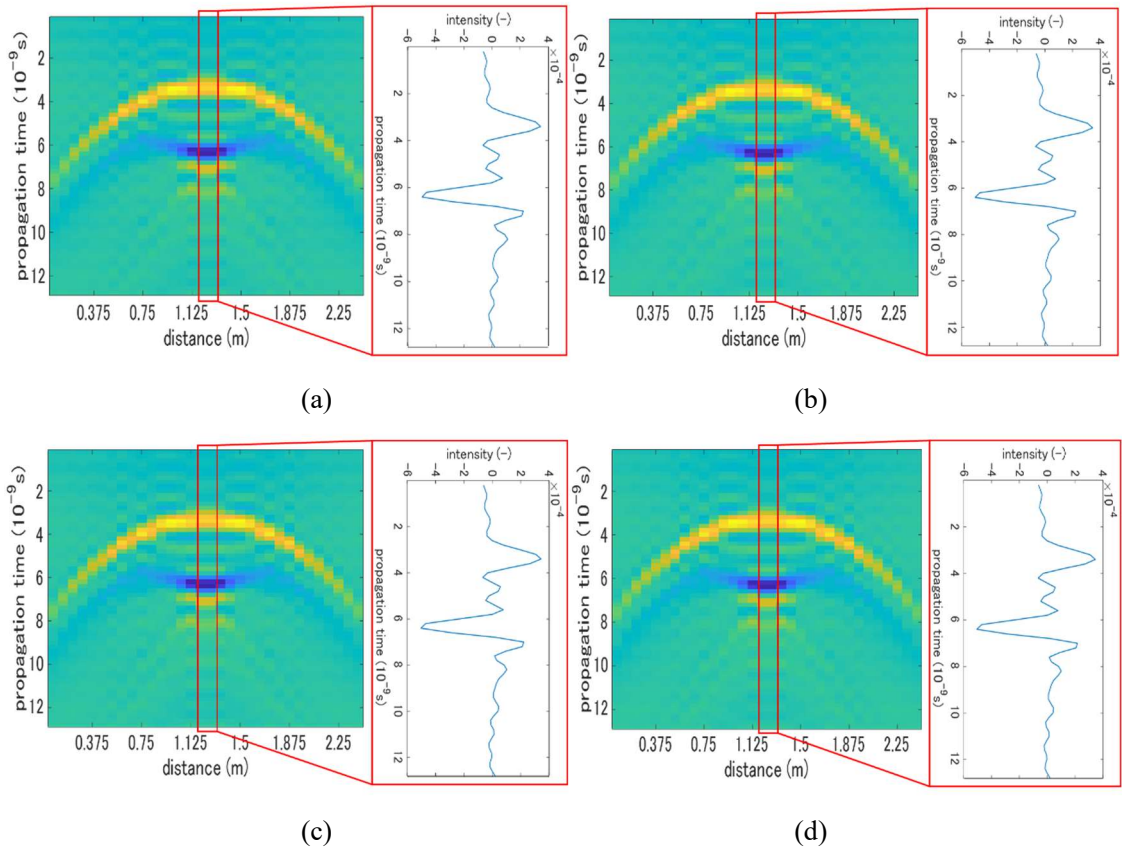


Fig. 5.12 Effect of the number of PML on reflection patterns and 1/2 point signal (a) 5 layers (b) 10 layers (c) 20 layers (d) 50 layers

m, the same signals because images were symmetrical) is also shown. From Fig. 5.11 (a), sidelobes were amplified at $\Delta x = 20$ mm. On the other hand, in the case $\Delta x = 10, 5, 1$ mm shown in Fig. 5.11 (b) - (d), the sidelobes decreased and converged. Therefore, $\Delta x = 10$ is the optimal size in this case. Propagation velocity is proportional to the reciprocal of the square root of relative permittivity. The relative permittivity of wet soil is over 10. Dried soil is around 3. Therefore, Δx should be one half in wet soil. On the other hand, calculation cost increases by the square or cube of a cell size. The calculation time of an image is several tens second to several minutes at $\Delta x = 10, 5$ mm while several hours to 1 day at $\Delta x = 1$ mm. In the research, considering accuracy and calculation cost, $\Delta x = 5$ mm was adopted.

Fig. 5.12 shows the effect of the number of PML n on the reflection patterns. If $n = 0$, the reflection patterns would be totally different because all the waves reflect back at the boundaries. However, even at $n = 5$ as shown in Fig. 5.12 (a), the clear reflection pattern was observed. Reflection patterns converged over $n = 5$. Therefore, $n = 5$ was adopted in the research.

5.3 Factors affecting reflection patterns

5.3.1 Input wave

Fig. 5.13 shows the examples of input waves. The maximum positive peaks of input waves are normalized in this section. The simplest case is a 1 wavelength sine wave shown in Fig. 5.13 (a). A Gauss wave shown in Fig. 5.13 (b) has an amplitude only in a limited frequency range. A Gauss wave is written below with a

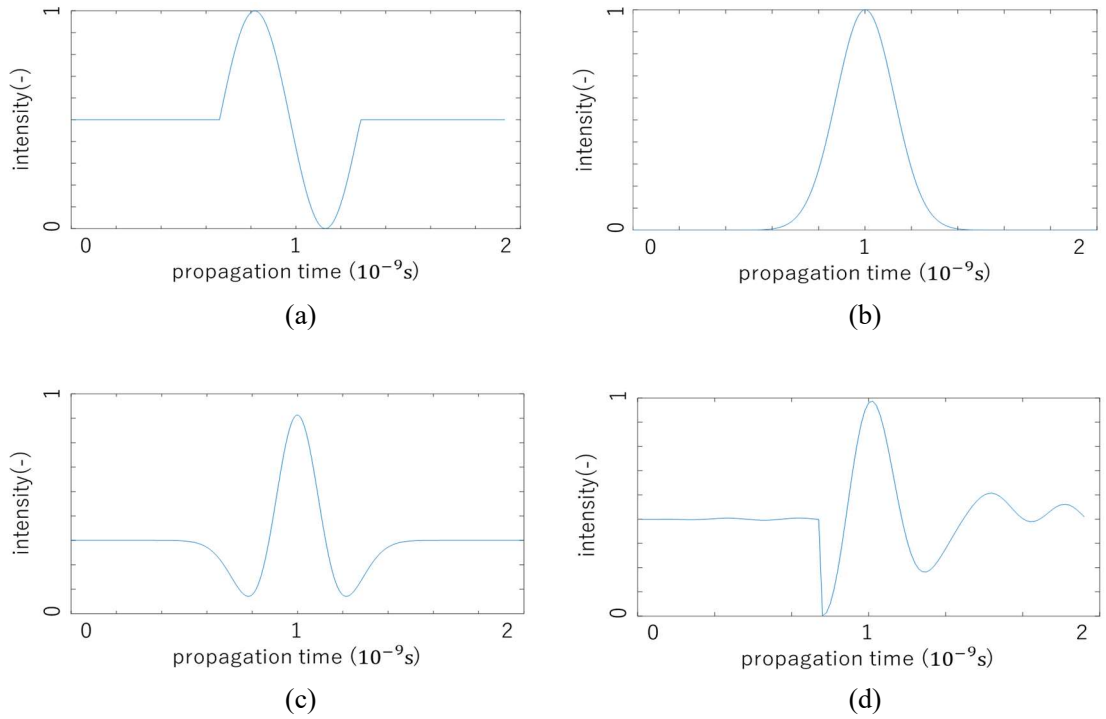


Fig. 5.13 Input waves (a) sine wave (b) Gaussian wave (c) Ricker (Mexican Hat) wave (d) Measured impulse wave

parameter σ related to a frequency range.

$$g(x) = \exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{5-29}$$

A Gauss wave is written by the same function both in the time and frequency domains, which is easy to design. A wave which is closest to the transmitted impulse wave is a Ricker wave shown in Fig. 5.13 (c). A Ricker wave is derived by the second derivative of a Gauss wave.

$$r(x) = \left(1 - \frac{x^2}{\sigma^2}\right)\exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{5-30}$$

A Ricker wave has two negative peaks around the positive peak. In the frequency domain, amplitude in a high frequency range is attenuated. The spectrum is wide in right side. A Ricker wave is often utilized as an impulse wave in FDTD simulations [82), 129), 130)]. An actual measured impulse wave is a pulsed wave distorted in a low and high frequency range. As shown in Fig. 5.13 (d), the wave has positive and negative peaks. The wave has causality. The electric field intensity before certain time is zero. In the following discussions, to apply the same preprocessing flow as measured signals, a 500 MHz - 3 GHz bandpass filter was applied to all the simulation results. Then, all the signals were calibrated by the impulse wave.

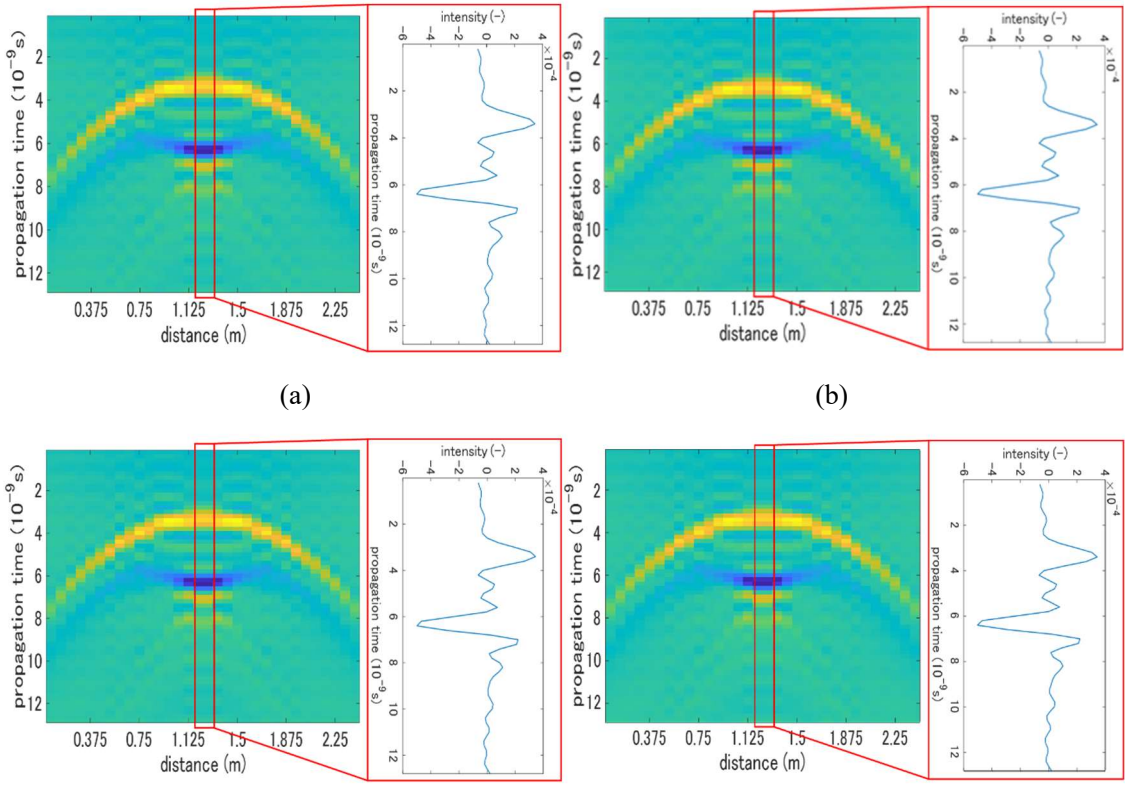


Fig. 5.14 Effect of input wave (a) sine wave (b) Gaussian wave (c) Ricker wave (d) measured impulse wave

Fig. 5.14 shows the effect of an input wave on reflection patterns. All the cases show the same reflection patterns. As stated above, all the signals were calibrated by the impulse wave. Therefore, the effect of the input wave was removed. In the research, for the simplicity of the programs, a Ricker wave was adopted, which did not affect the results.

5.3.2 Antenna directivity

The details of antenna characteristics are not published. There are some references about the measured antenna patterns of each frequency as shown in Fig. 5.15 [88] and impulse wave. Antenna patterns have some variations depending on frequency. On the whole, the antennas are designed to transmit and receive the maximum intensity just below the two antennas. However, directivity is relatively small possibly to cover all the area below the antennas. The types of antennas and geometry are not known. In this section, the effect of antenna directivity was evaluated by an array antenna model shown in Fig. 5.16.

In an array antenna model, electric field intensity $E_i(r,\theta)$ in polar coordinates caused by a source i at distance r_i is written by an intensity factor w_i, element pattern $g_i(\theta)$, wave number k_0 [49), 132), 133)].

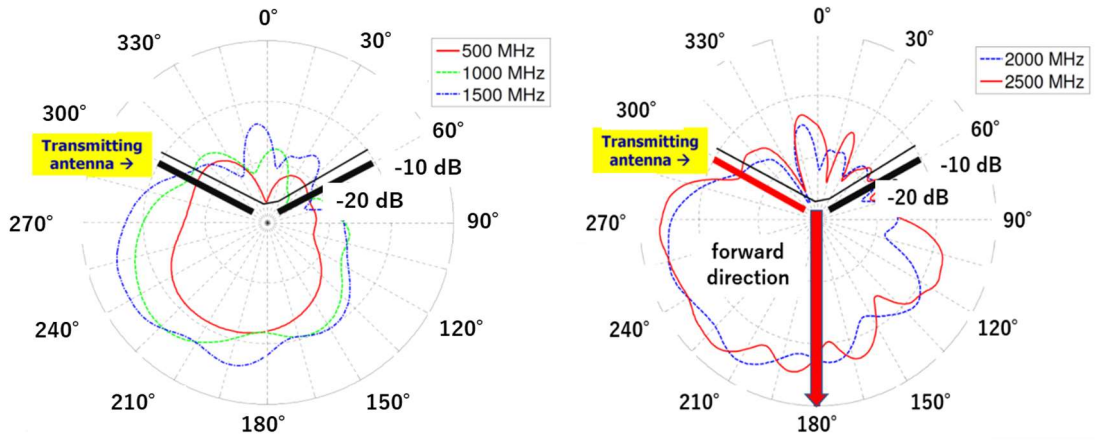
$$E_i(r,\theta) = w_i g_i(\theta) \frac{\exp\left(-jk_0\left(r - \frac{r}{|r|}\cdot r_i\right)\right)}{|r|} \tag{5-31}$$

The summation of the electric field intensity of all the sources E is written below.

$$E(r,\theta) = g_i(\theta) AF(\theta) \frac{\exp\left(-jk_0 r\right)}{|r|} \tag{5-32}$$

$$AF(\theta) = \sum_{i=1}^{N} w_i \exp\left(jk_0 \frac{r}{|r|}\cdot r_i\right) \tag{5-33}$$

AF represents the effect of the arrangement of sources, which is called an array factor. In the research, phase delay of each source is not considered. g_i and w_i is the same for all the sources. Therefore, the antenna



(a) (b)

Fig. 5.15 Antenna pattern (a) 500, 1000, 1500 MHz (b) 2000, 2500 MHz

pattern only depends on AF. AF of $2N+1$ point sources aligned in a x axis with a constant interval d as shown in Fig. 5.16 is written below.

$$AF(\theta) = \sum_{i=-N}^{N} \exp\left(j2dik_0\sin(\theta)\right) \tag{5-34}$$

AF depends on frequency $f = vk_0/2\pi$, d and an antenna length $L = (2N+1)d$.

Fig. 5.17 (a) shows the effect of f on antenna patterns at $L = 12$ cm, $d = 10$ mm. At $f = 3$ GHz, the power of electric field intensity at $\theta = \pi/2$ compared to $\theta = 0$ is about -7.0 dB. At $f = 1.5$ GHz, the

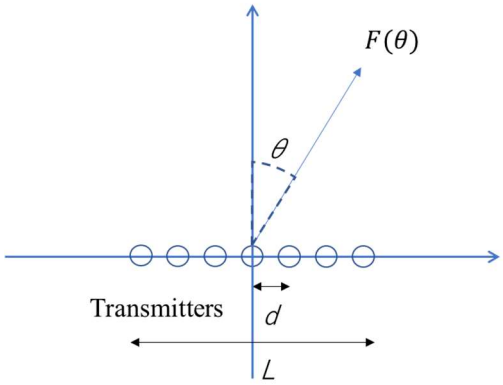


Fig. 5.16 Array antenna model

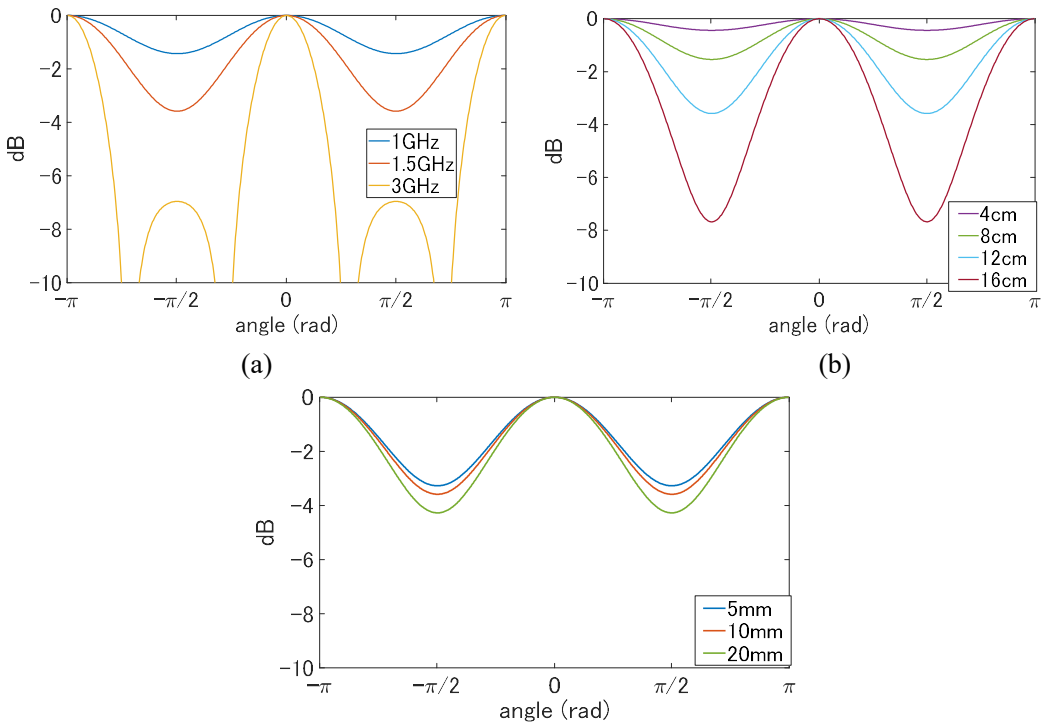


Fig. 5.17 Array factor (a) frequency (b) antenna length (c) antenna distance

power is about -3.6 dB. At $f = 1$ GHz, the power is about -1.4 dB. If $\lambda = v/f$ is much larger than L, the antenna pattern is not affected by f. At $L = 12$ cm and $f = 3$ GHz ($\lambda = 10$ cm), the effect of f is not negligible. The center frequency of the system is $f_c = 1.5$ GHz. Therefore, the discussions below assume $f = 1.5$ GHz.

Fig. 5.17 (b) shows the effect of L on the antenna patterns at $d = 10$ mm. Compared to the case $L = 16$ cm when the power of $\theta = \pi/2$ is about -7.7 dB, the power is about -3.6 dB at $L = 12$ cm, about -1.5 dB at $L = 8$ cm and about -0.4 dB at $L = 4$ cm. From Fig. 5.15 (a), the change of the power at different angles at $f = 1.5$ GHz is about -2 dB to -5 dB. Therefore, the equivalent antenna length is around $L = 8$ cm to 12 cm. The actual length of the antennas is expected to be around 5 cm - 15 cm from their sizes, indicating the estimated antenna length is reasonable. Fig. 5.17 (c) shows the effect of d on the antenna patterns at $L = 12$ cm. The difference of the power between $d = 5$ mm and 20 mm was at most 1 dB. d is relatively small compared to L and λ. Therefore, the effect of d was small. In the research, considering calculation cost, $d = 20$ mm was adopted. Fig. 5.18 shows the maximum power of the received waves of all the frequency range. At $L = 8$ cm and 12 cm, the power is -5 dB to -10 dB. The results were consistent with Fig. 5.15 (a).

Table 5.1 Effect of directivity on peak value of 1/4 point

	Peak value (-) $\times 10^{-4}$	Difference to no directivity (%)
No directivity	3.46	—
Antenna length 8 cm	3.36	−2.9
Antenna length 12 cm	3.29	−4.9

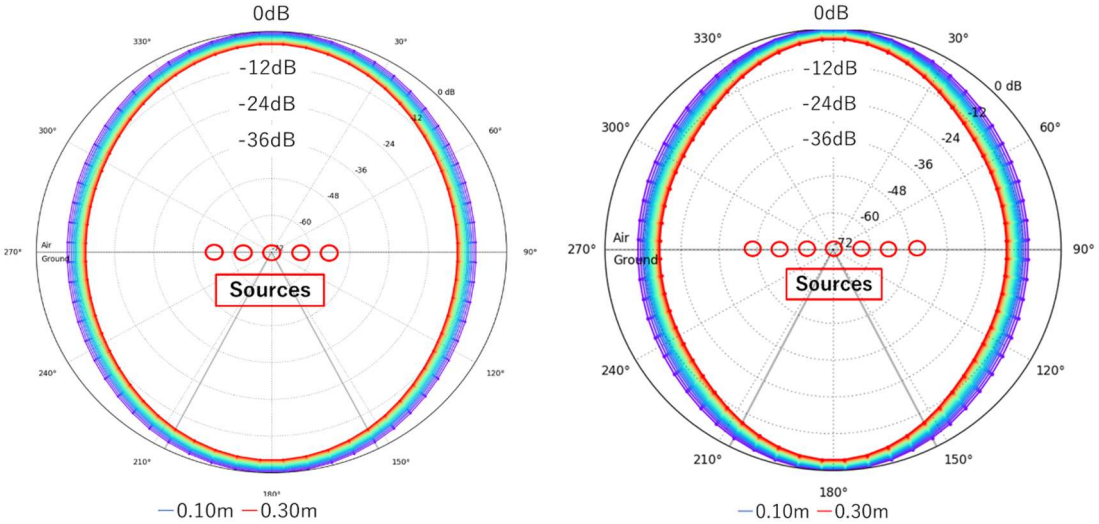


Fig. 5.18 Antenna patterns (a) antenna length $L = 8$ cm (b) $L = 12$ cm

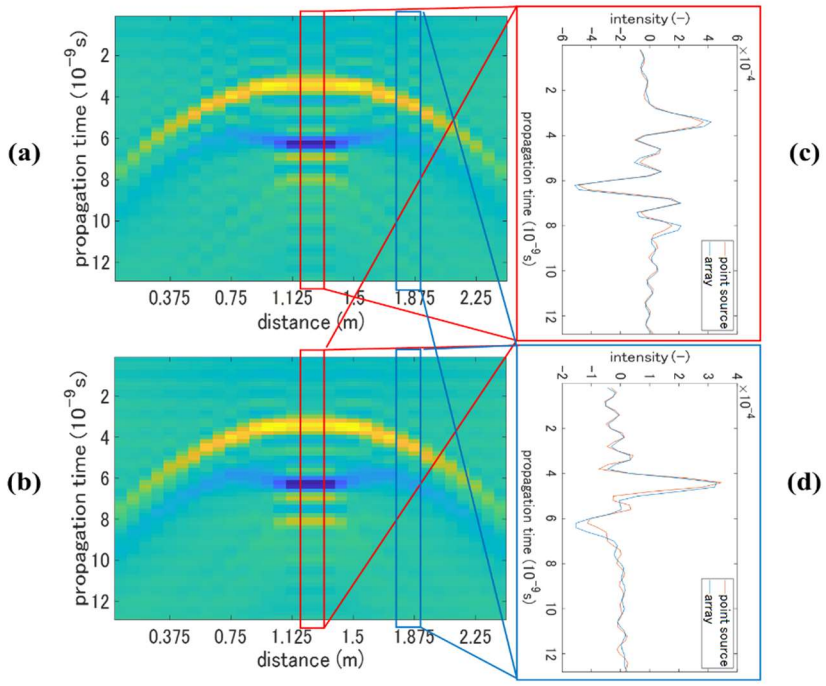


Fig. 5.19 Effect of antenna length (a) (b) no directivity (point source) and 8 cm array antenna (c) (d) 1/2 and 1/4 point signals

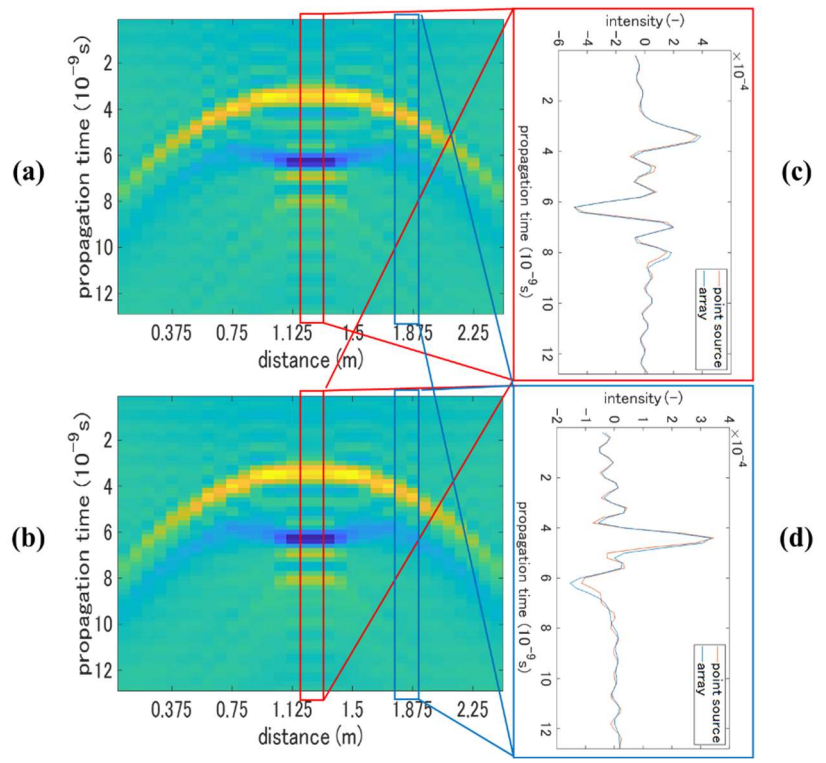


Fig. 5.20 Effect of antenna length (a) (b) no directivity (point source) and 8 cm array antenna (c) (d) 1/2 and 1/4 point signals

The effect of antenna patterns is expected to be large at the deeper region of radar images because the reflections from a wider area affect the results. Fig. 5.19 and Fig. 5.20 shows the effect of L on the reflection patterns and signals at 1/2 and 1/4 points. Table 5.1 shows the difference of the peak values of 1/4 point signals compared to a no directivity case. From Fig. 5.19 and Fig. 5.20, the reflection patterns almost did not change irrespective of the antenna length. From Table 5.1, the differences compared to the no directivity case were at most 5 %, suggesting the effect is negligible. The size of the antenna was small enough compared to the void. Therefore, the point source was accurate enough to approximate the antenna model.

5.3.3 Asphalt layer

The important parameters of an asphalt layer are layer thickness, relative permittivity and randomness of the parameters. Fig. 5.21 shows an example of a 7 cm asphalt layer with ± 0.5 cm randomness. Because the cell size of the model was 5 mm as stated in Section 5.2.2, realizable randomness was at least 0.5 cm. In gprMax, the randomness of layer thickness and inhomogeneity of medium is introduced by fractals. Fractals are composed of small similar figures [99), 136)]. The fractal dimension D of a figure consisted of m $1/n$-size smaller similar figures is defined below.

$$D = \log_n m \tag{5-35}$$

Sherpinski's gasket is shown in Fig. 5.22 (a). The figure is the case $n = 2$, $m = 3$. Therefore, $D = 1.59$. Fig. 5.21 is the case $D = 1.5$. Fractal dimension is related to the complexity of geometry. In gprMax, fractal dimension and upper and lower limits of parameters are assigned. As shown in Fig. 5.22 (b), provided D

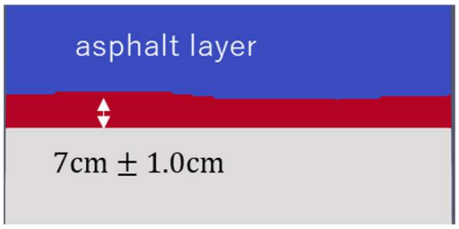
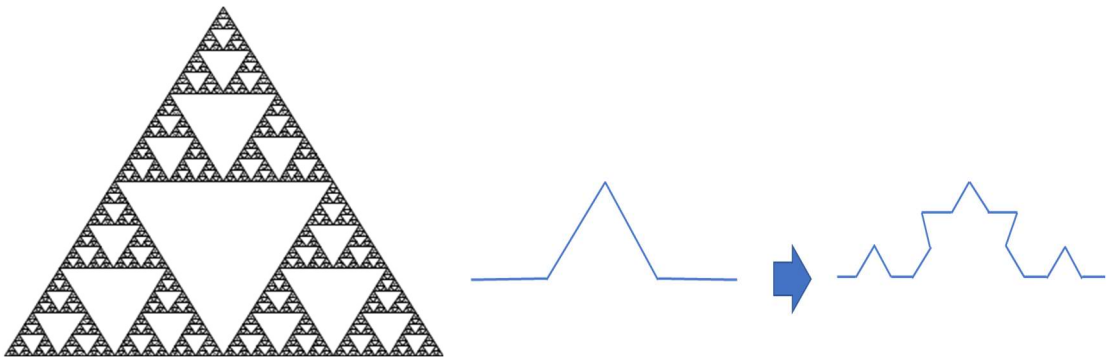


Fig. 5.21 Randomness of asphalt layer thickness



(a) (b)

Fig. 5.22 Fractals (a) Sherpinski's gasket (b) production of a fractal

and upper and lower limits, the parameters of $1/n$-th points is randomly assigned. The same procedure is applied to the divided $n+1$ small regions.

Fig. 5.23 (a) (b) show the effect of the thickness of asphalt layer h_a on the reflection patterns. At $h_a = 6\ \mathrm{cm}, 8\ \mathrm{cm}$ the reflection patterns did not change regardless of h_a. Fig. 5.23 (c) (d) show the effect of relative permittivity ε_a. The reflection patterns did not change regardless of ε_a. The maximum wavelength of the transmitted wave is 10 cm. Furthermore, the size of void is large enough compared to the changes of h_a and ε_a. Therefore, the changes did not affect the reflection patterns. From Eq. (5-28), a transmission

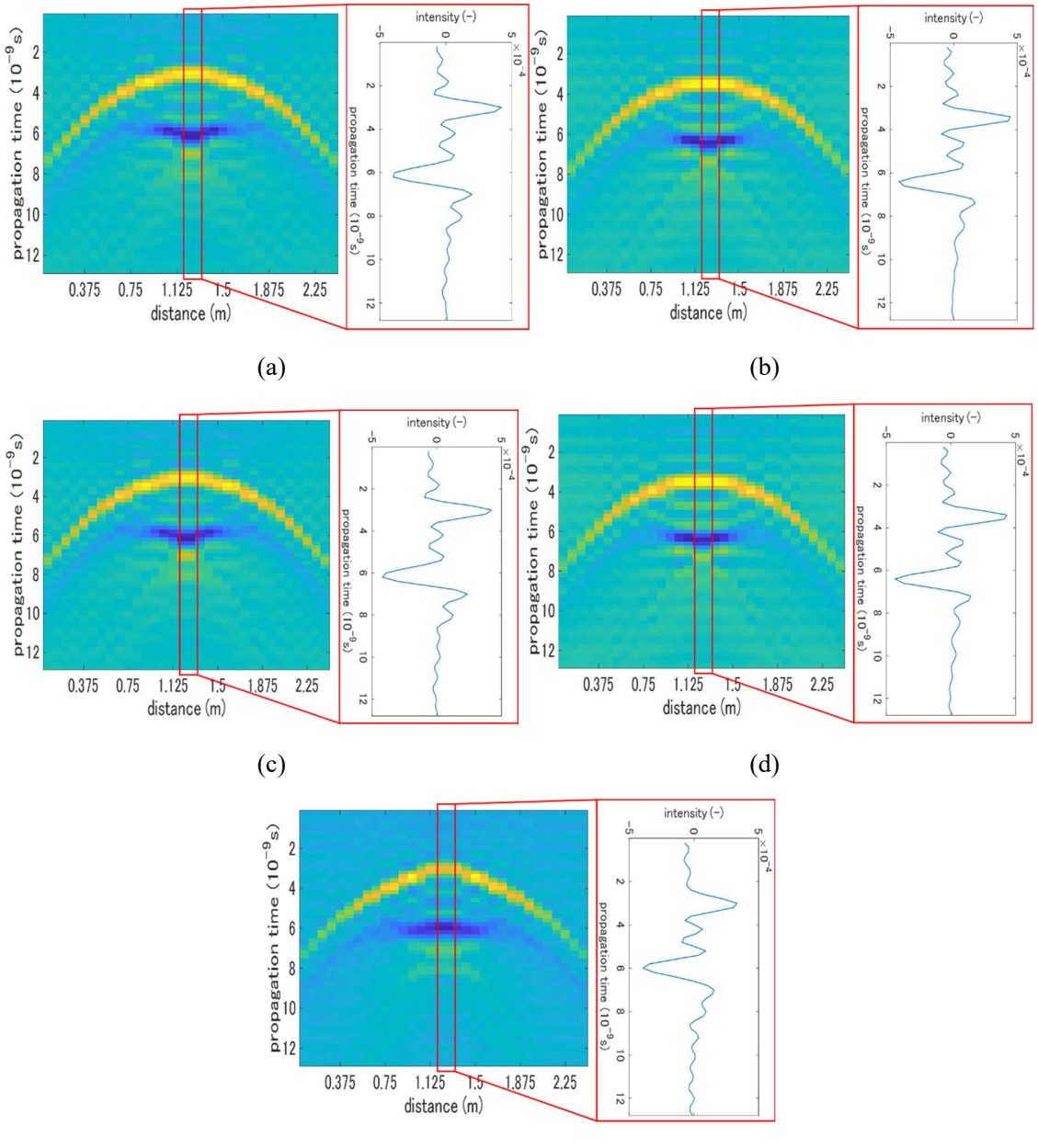


Fig. 5.23 Effect of asphalt layer on reflection patterns (a) (b) thickness 6 cm, 8 cm (c) (d) relative permittivity 3, 5 (e) randomness of thickness 7 ± 0.5 cm

coefficient $T = 0.73$ at $\varepsilon_a = 3$ and $T = 0.62$ at $\varepsilon_a = 5$, only 15 % difference. Therefore, the effect of ε_a on peak values is also small From Fig. 5.23 (a) - (d). Fig. 5.23 (e) shows the reflection pattern of the asphalt layer shown in Fig. 5.21. Because the effect of h_a is small, randomness also did not affect the results. From a practical point of view, the randomness of the thickness of asphalt layer is at most several mm. Therefore, the effect is further small. In the research, $h_a = 7$ cm, $\varepsilon_a = 4$ were assumed. The randomness of the thickness of asphalt layer was not introduced.

5.3.4 Soil layer

The important parameters of soil layer are relative permittivity ε_s, conductivity σ_s and randomness of the parameters. ε_s, σ_s are expected to affect reflection patterns. Fig. 5.24 (a) (b) show the effect of ε_s on the reflection patterns. Larger ε_s is, the propagation velocity of electromagnetic waves decreases. Compared to $\varepsilon_s = 2$, the reflection pattern was delayed in the case $\varepsilon_s = 10$. Considering the optical distance between the antennas and void, higher relative permittivity is, narrower hyperbolics are. Peak values depend on the difference of the relative permittivity of void and surrounding soil. The peak value of the 1/2 point signal became larger in the case $\varepsilon_s = 10$. From Eq. (5-29), a reflection coefficient $R = 0.172$ at $\varepsilon_s = 2$ while $R = 0.519$ at $\varepsilon_s = 10$, three times larger. Fig. 5.24 (c) (d) show the effect of σ_s on the reflection patterns. Compared to $\sigma_s = 0.001$ s/m, the signal was weakened at 0.01 s/m, indicating the waves heavily

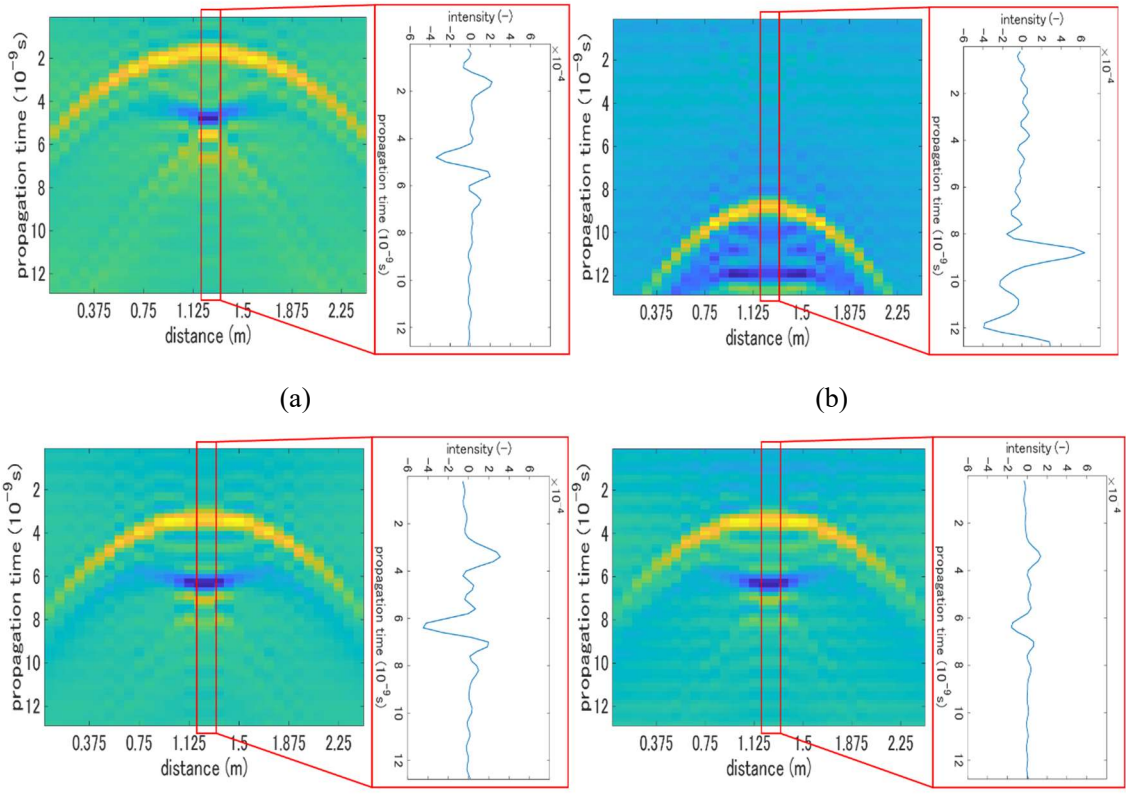
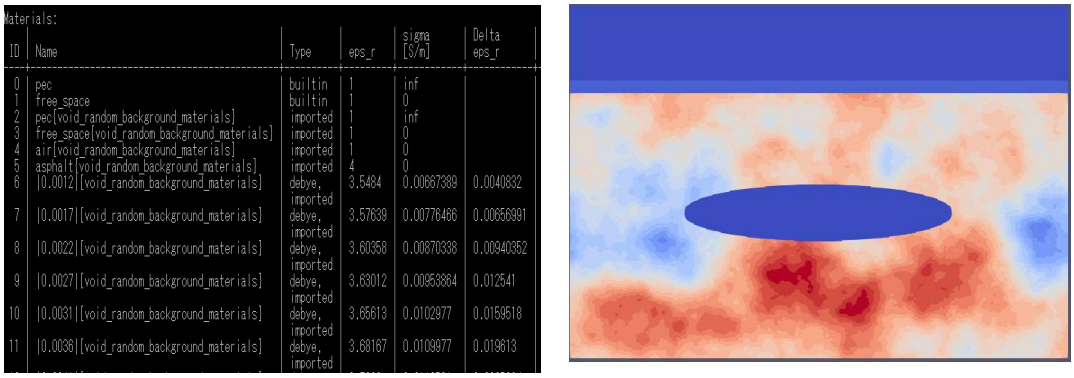


Fig. 5.24 Effect of soil layer on reflection patterns (a) (b) relative permittivity 2, 10 (c) (d) conductivity 0.001, 0.01s/m

attenuated in the soil layer. The electromagnetic characteristics of actual soil depends on the water content and type of soil. In the case of dried sand, ε_s is around 3 and σ_s is almost 0 s/m. In the case of wet soil, ε_s can be up to 20 and σ_s can be up to 0.01 s/m [34) - 36)].

In gprMax, Peplinski soil model is assumed [56)]. The ratio of sand and clay and density of soil are given. Water content is randomly assigned to produce multiple medium patterns. The media are placed on soil layer

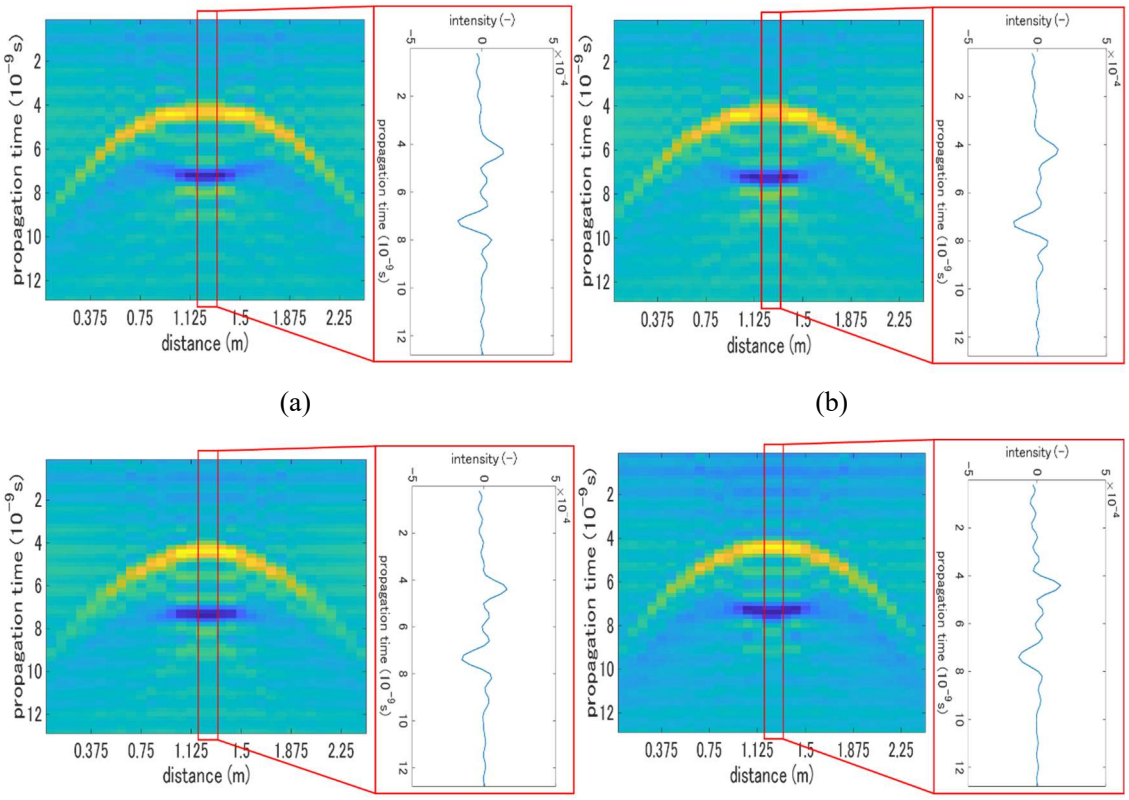


Materials:

ID	Name	Type	eps_r	sigma [S/m]	Delta eps_r
0	pec	builtin	1	inf	
1	free_space	builtin	1	0	
2	pec[void_random_background_materials]	imported	1	inf	
3	free_space[void_random_background_materials]	imported	1	0	
4	air[void_random_background_materials]	imported	1	0	
5	asphalt[void_random_background_materials]	imported	4	0	
6	[0.0012][void_random_background_materials]	debye, imported	3.5484	0.00667389	0.0040832
7	[0.0017][void_random_background_materials]	debye, imported	3.57639	0.00776466	0.00656991
8	[0.0022][void_random_background_materials]	debye, imported	3.60358	0.00870338	0.00540352
9	[0.0027][void_random_background_materials]	debye, imported	3.63012	0.00953864	0.012541
10	[0.0031][void_random_background_materials]	debye, imported	3.65613	0.0102977	0.0159518
11	[0.0036][void_random_background_materials]	debye, imported	3.68167	0.0109977	0.019613

(a) (b)

Fig. 5.25 Randomness of soil layer (a) relative permittivity (b) medium distribution



(a) (b)

(c) (d)

Fig. 5.26 Effect of randomness of soil layer on reflection patterns (a) (b) water content 0.001 - 0.01, 0.002 - 0.005 (c) (d) fractal dimension 1.0, 2.0

based on fractal structures. Fig. 5.25 shows the distribution of the electromagnetic characteristics. In the research, the ratio of sand and clay is 0.5. The weight of soil is $2\ \mathrm{g/cm^3}$, sand is $2.66\ \mathrm{g/cm^3}$. The number of media is 20. Fig. 5.26 (a) shows the case of the range of water content $w = 0.002$ - 0.005 at fractal dimension $D = 1.5$. Fig. 5.26 (b) shows the case $w = 0.001$ - 0.01. $w = 0.002$ - 0.005 corresponds to $\varepsilon_s = 2.5$ - 3.5 and $\sigma_s = 0.002\ \mathrm{s/m}$ - $0.005\ \mathrm{s/m}$. $w = 0.001$ - 0.010 corresponds to $\varepsilon_s = 2$ - 5 and $\sigma_s = 0.001\ \mathrm{s/m}$ - $0.01\ \mathrm{s/m}$. Fig. 5.26 (c) shows the case $D = 1.0$ at the same range of water content

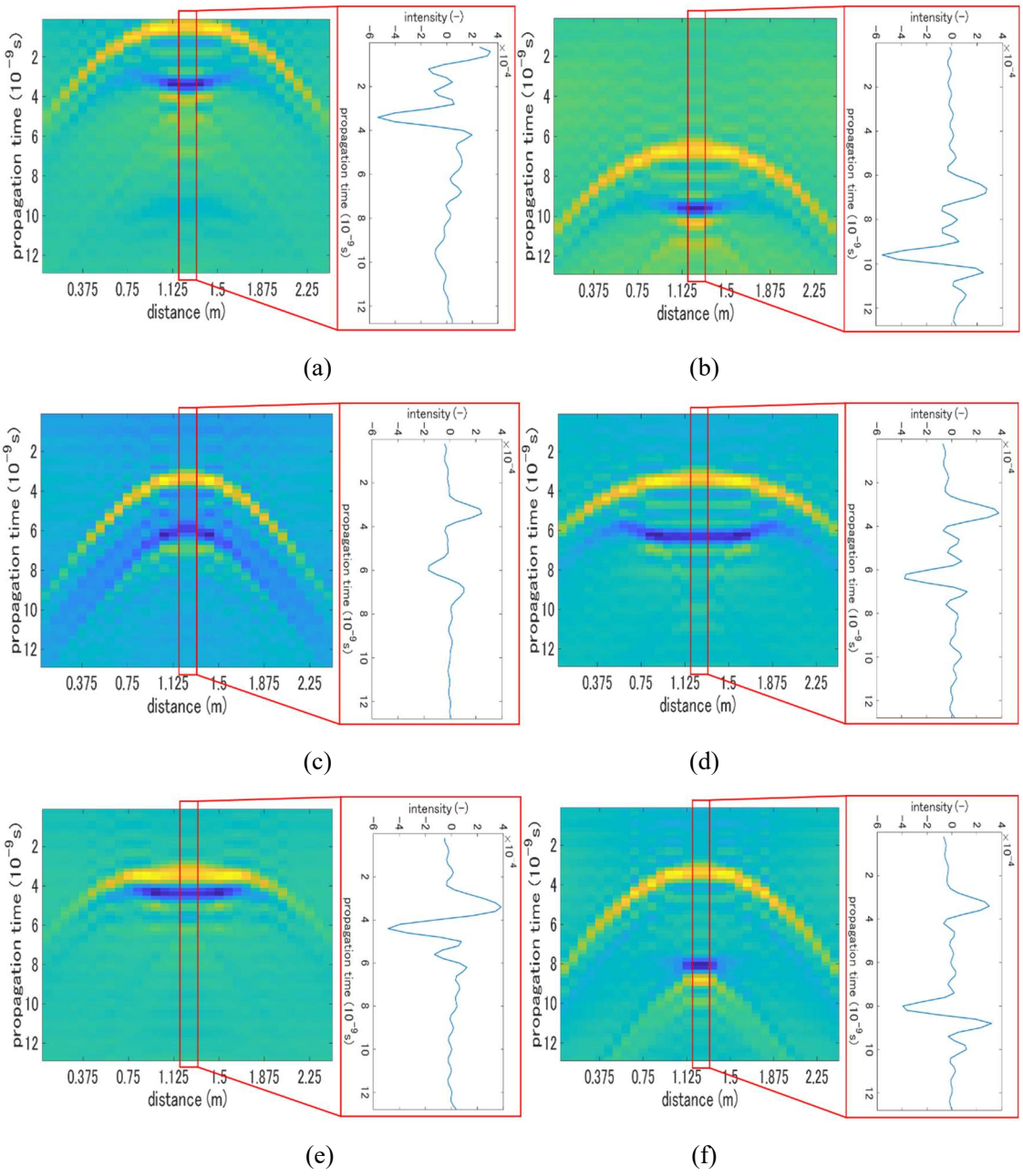


Fig. 5.27 Effect of geometry of void on reflection patterns (a) (b) depth 0.3, 0.7 m (c) (d) scan direction length 0.5, 2 m (e) (f) thickness 0.1, 0.5 m

$w = 0.001$ - 0.01. Fig. 5.26 (d) shows the case $D_s = 2.0$. All the signals were weakened by introducing water content. However, comparing Fig. 5.26 (a), (b) and Fig. 5.26 (c), (d), the reflection patterns did not change regardless of the change of D and range of w, indicating the inhomogeneity of soil layer does not affect the results. Considering the wavelength of transmitted waves and size of void, the scale of the randomness of soil layer was relatively small. Therefore, in the research, $\sigma_s = 0.001\ \mathrm{s/m}$ - $0.01\ \mathrm{s/m}$ is assumed not considering the inhomogeneity of soil layer.

5.3.5 Geometry of void

The important parameters of void are depth d, thickness (the length in a depth direction) t and the length in a scan direction h. Fig. 5.27 (a), (b) show the effect of d on reflection patterns. Compared to $d = 0.3$ m, in the case $d = 0.7$ m the reflection pattern delayed because the depth increased. Fig. 5.27 (c), (d) show the effect of h. Compared to $h = 0.5$ m, the hyperbolic becomes wider in the case $h = 1.5$ m. Fig. 5.27 (e), (f) show the effect of t. Compared to $t = 0.5$ m, in the case $t = 0.1$ m the hyperbolic was wider. In these two cases, the void was flatter. On the other hand, from Fig. 5.27 (c) - (f), the peak values did not change. Reflection coefficient is determined by the difference between the relative permittivity of void and surrounding soil. When the size of void is larger than the spatial resolution in depth and horizontal directions, the peak values are the same irrespective of the geometry of void. In the research, d, t and h are randomly assigned referring to the results of the excavation survey stated in Section 2.3.2.

5.4 Methodology to produce 3D data

5.4.1 Approximation by 2D model

There is some research about the comparison of 2D and 3D-FDTD method focusing on the delamination inside bridge deck [137), 138)]. However, there is no research about void under the road. In 2D-FDTD method, a model is assumed to be uniform in a direction perpendicular to a target plane. An input wave is also uniform in a perpendicular direction, which is called a plane wave. On the other hand, in 3D-FDTD method, a model has geometry also in a perpendicular direction. An input wave is a spherical wave propagating in three-dimensional space.

Fig. 5.28 shows the reflection patterns of the 2D and 3D-FDTD void models where the lengths of the void in scan and channel directions were 1.5 m and 0.5 m respectively. When the difference of the lengths in scan and channel directions was largest, the difference between the reflection patterns of 2D and 3D models became largest. The most severe case is discussed in this section. Table 5.2 shows the comparison of the peak values. From Fig. 5.28 (a) (b), the reflection patterns of the 2D and 3D void models were similar. From Table 5.2, the maximum error was at most 20 %, indicating a 3D model can be approximated by a 2D model.

There are two possible reasons. The first reason is, the void existed 50 cm below the road surface. Considering the relative permittivity of asphalt and soil layers, the void was far enough from the antennas. Waves transmitted from the antennas spread into a propagating medium transforming to a plane wave. Therefore, the geometry of void in a perpendicular direction did not affect the results. The second reason is, the maximum wavelength of transmitted waves is 10 cm. The shape of void is considered to be flat compared

to the wavelength of transmitted waves. Therefore, from the above discussions, a 3D model can be approximated by a 2D model. In 3D-FDTD method, the 3D positions of two antennas may affect the results. In the research, because the depth and size of void are much larger than the distance between the two antennas, the effect of the 3D positions of two antennas were ignored [137), 138)].

From Fig. 5.28 (c), the 2D void model slightly underestimates the negative peak of the 1/2 point signal. In terms of free space loss, a spherical wave attenuates by the square of distance while a plane wave attenuates proportional to distance. Therefore, the 2D void model is expected to overestimate signals in a deeper region. On the other hand, the 2D void model cannot consider the focusing of waves in a channel direction. Fig. 5.29 explains the propagation of waves in the 2D and 3D void models in a channel direction. The 2D void is modeled as an air layer in a perpendicular direction. Therefore, the waves with the same intensity appear at

Table 5.2 Comparison of peak values of 2D and 3D-FDTD method

	3D-FDTD Peak value $\times 10^{-4}$	2D-FDTD Error (%)
1/2 positive peak	3.93	−6.9 %
1/2 negative peak	−5.88	+16.7 %
1/4 peak	3.31	+4.5 %

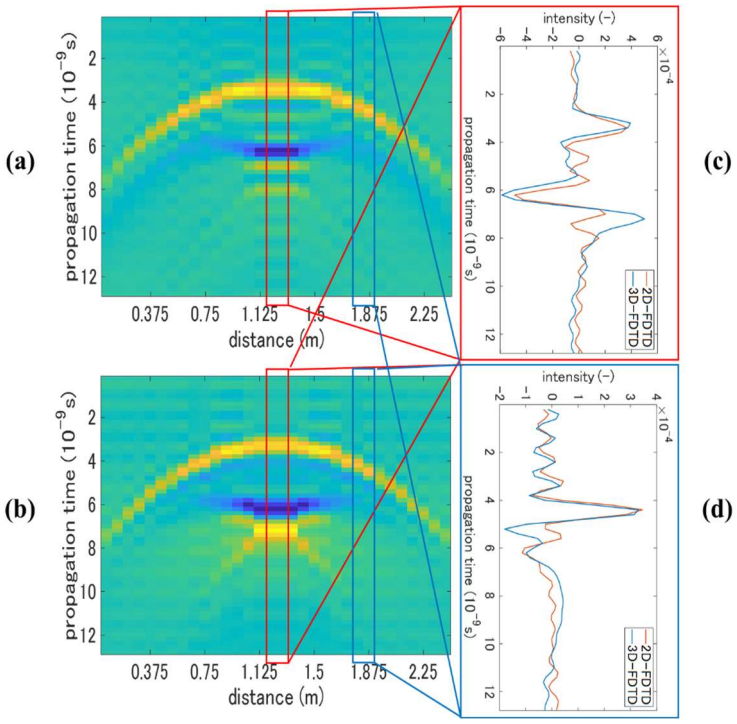


Fig. 5.28 Comparison of 2D and 3D-FDTD method (a) 3D-FDTD (b) 2D-FDTD (c) 1/2 point signal (d) 1/4 point signal

the upper and lower surface of void. On the other hand, in the 3D void model, depending on the depth and curvature of the lower surface of void, reflected waves focus at the receiving antenna. This phenomenon is dominant as shown in the snapshot of Fig. 5.9 (d). However, because the error is within 20 %, the 2D void model is considered to approximate the 3D void model considering the variation of measured signals.

The antennas have directivity also in a channel direction. However, the details of the antennas are not known. Furthermore, a detailed 3D antenna model needs large calculation cost. Therefore, a point source was adopted in the research. The error caused by ignoring the directivity in a channel direction was evaluated in Section 5.4.3 comparing to the measurement data. In an actual measurement system, there is a variation of channel characteristics. In the research, because the measured waves were calibrated by the impulse wave for each channel as stated in Section 3.4, the variation of channels was ignored.

5.4.2 Reconstruction of 3D reflection patterns

A void model has symmetricity around a center axis. Considering a horizontal section with a center axis as origin, the signals in first quadrant are the same as second, third and fourth quadrant. Therefore, calculation cost can be reduced to one fourth considering the symmetricity of reflection patterns. Also, in 2D-FDTD method, calculation cost can be one half. This characteristic was utilized in producing training data.

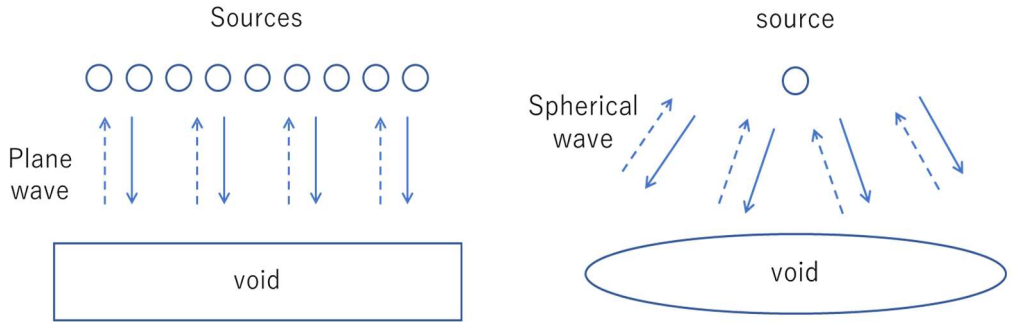


Fig. 5.29 Propagation of electromagnetic waves (a) 2D plane wave (b) 3D spherical wave

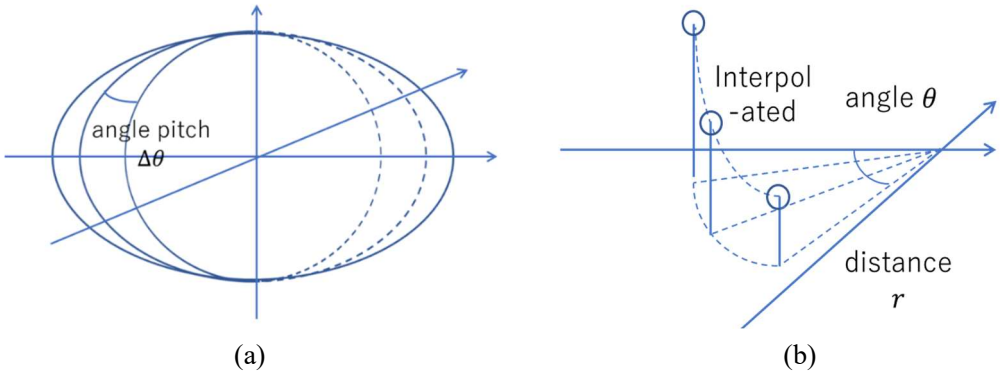


Fig. 5.30 Reproduction of 3D reflection pattern (a) 2D cross-sections (b) interpolation in cylindrical coordinates

Reflection patterns are a hyperboloid along the upper surface of void. Therefore, a cylindrical coordinates system is more convenient. A reflection pattern is represented by $I(r,\theta)$ at distance r, an angle from a scan direction θ. From the discussions of the previous section, a 3D void model can be approximated by a 2D void model. Therefore, considering $\theta_n = n\Delta\theta$ with angle pitch $\Delta\theta$, a 3D void reflection pattern can be reconstructed by the combination of the sections $I(r,\theta_n)$ at θ_n. $\Delta\theta$ should be small enough to obtain an accurate 3D reflection pattern. However, smaller $\Delta\theta$ is, the number of sections N increases, resulting in large calculation cost. $N = 18$ at $\Delta\theta = 5°$ and $N = 90$ at $\Delta\theta = 1°$. Accuracy depends on the interpolation method. In the research, any section $I_i(r,\theta)$ is linearly interpolated by the closest two simulated sections $I(r,\theta_n)$, $I(r,\theta_{n+1})$ below.

$$I_i(r,\theta) = \frac{\theta_{n+1}-\theta}{\theta_{n+1}-\theta_n} I(r,\theta_n) + \frac{\theta-\theta_n}{\theta_{n+1}-\theta_n} I(r,\theta_{n+1}) \tag{5-35}$$

From Eq. (5-35), at $\theta = \theta_n$, $I_i(r,\theta) = I(r,\theta_n)$ and at $\theta = \theta_{n+1}$, $I_i(r,\theta) = I(r,\theta_{n+1})$. Fig. 5.30 summarizes the procedure of approximating a 3D reflection pattern by a 2D void model. As shown in Fig. 5.30 (a), 2D simulations are conducted by 2D-FDTD method at each section $\theta_n = n\Delta\theta$ passing through the center of the void. Then, as shown in Fig. 5.30 (b), linear interpolation is conducted. The following discussion assumed $I(r,\theta)$ is a true reflection pattern. $I_i(r,\theta)$ and $I(r,\theta)$ at each $\Delta\theta$ were compared.

Assuming the lengths of the void in scan and channel directions a, b respectively, Fig. 5.31 shows the reflection patterns at $\theta = 22.5°$ in the case $a = b = 1.0$ m and $\Delta\theta = 45°$. From Fig. 5.31 (a) (b), the

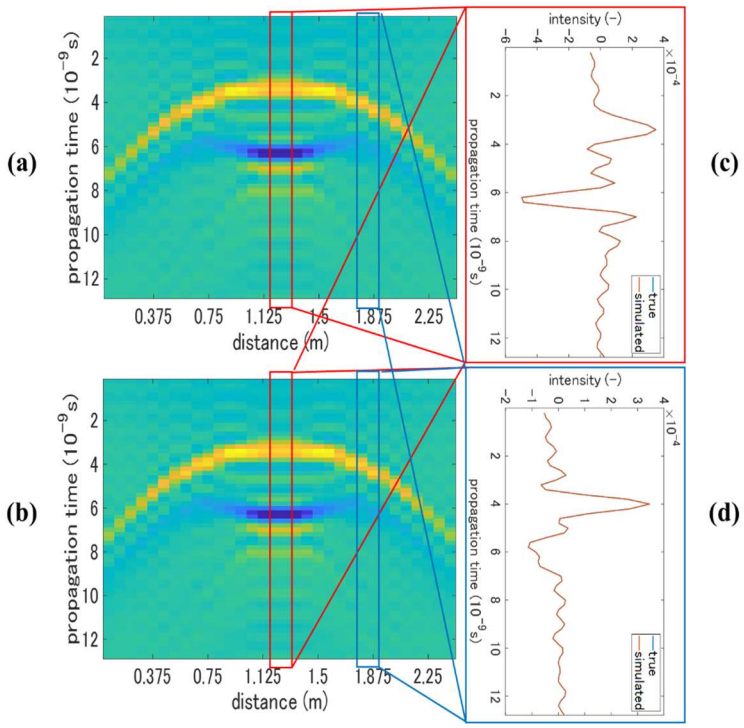


Fig. 5.31 Comparison of true and interpolated images ($a = b = 1.0$ m, $\Delta\theta = 45°$, $\theta = 22.5°$)
(a) (b) true and interpolated reflection patterns (c) (d) 1/2 and 1/4 signals

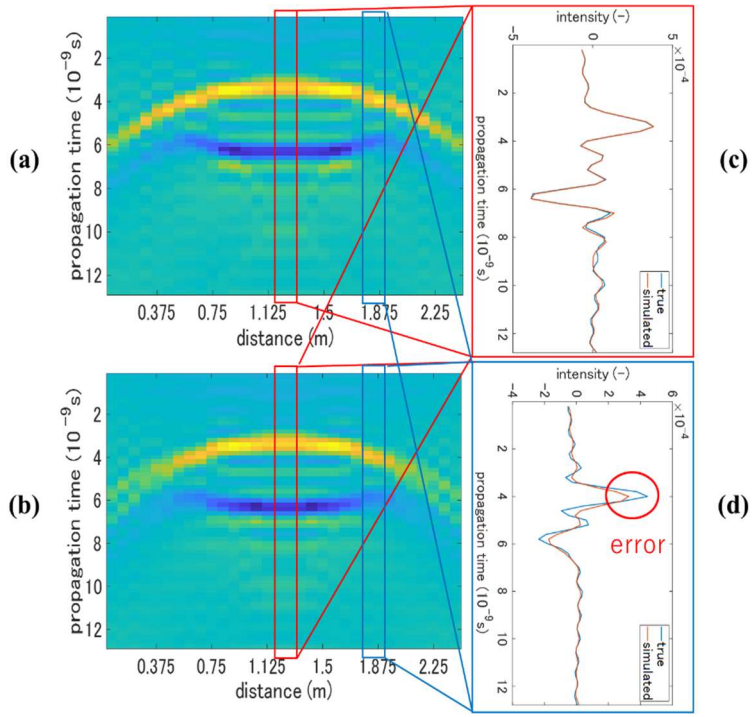


Fig. 5.32 Comparison of true and interpolated images ($a = 1.5$ m, $b = 0.5$ m, $\Delta\theta = 45°$, $\theta = 22.5°$) (a) (b) true and interpolated reflection patterns (c) (d) 1/2 and 1/4 signals

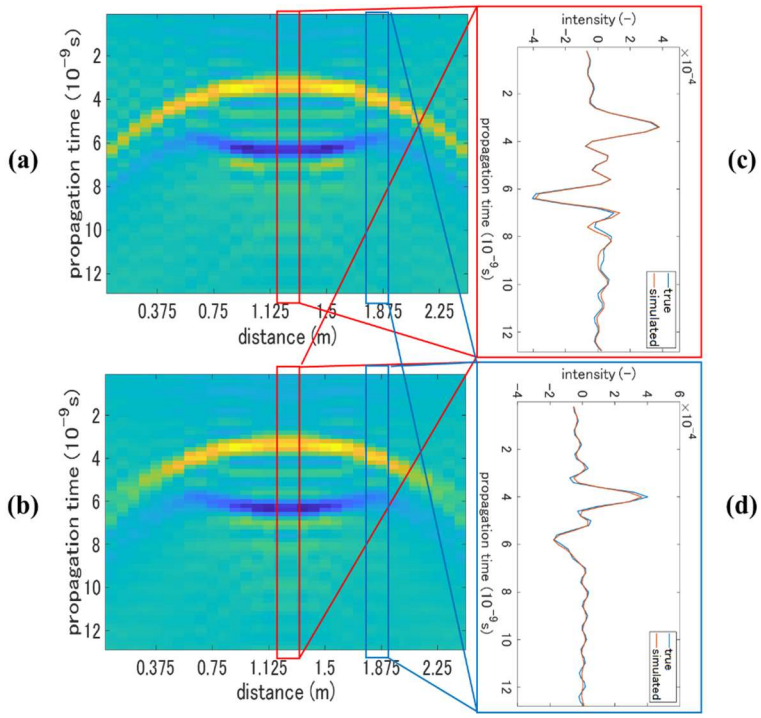


Fig. 5.33 Comparison of true and interpolated images ($a = 1.5$ m, $b = 0.5$ m, $\Delta\theta = 22.5°$, $\theta = 33.8°$) (a) (b) true and interpolated reflection patterns (c) (d) 1/2 and 1/4 signals

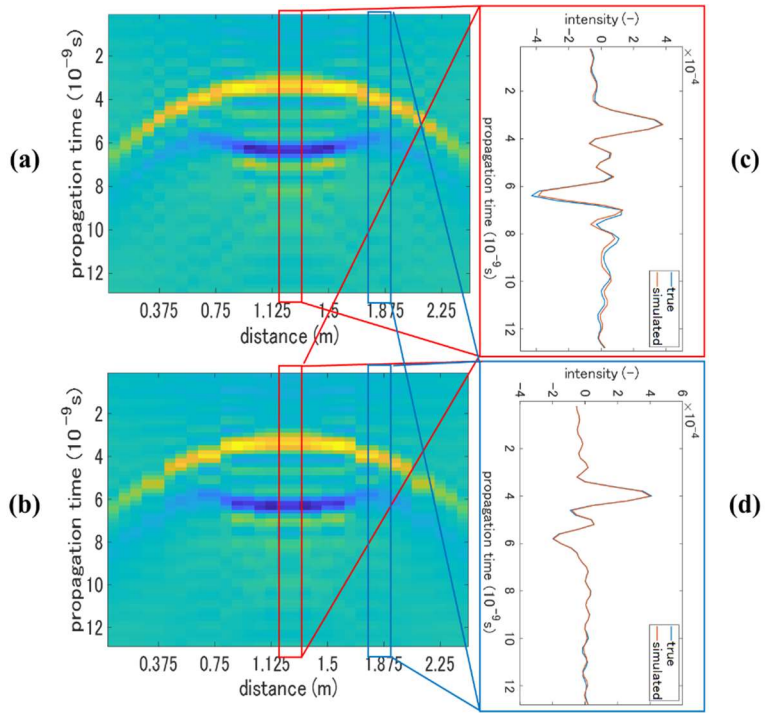


Fig. 5.34　Comparison of true and interpolated images ($a = 1.5$ m, $b = 0.5$ m, $\Delta\theta = 11.3°$, $\theta = 39.4°$) (a) (b) true and interpolated reflection patterns (c) (d) 1/2 and 1/4 signals

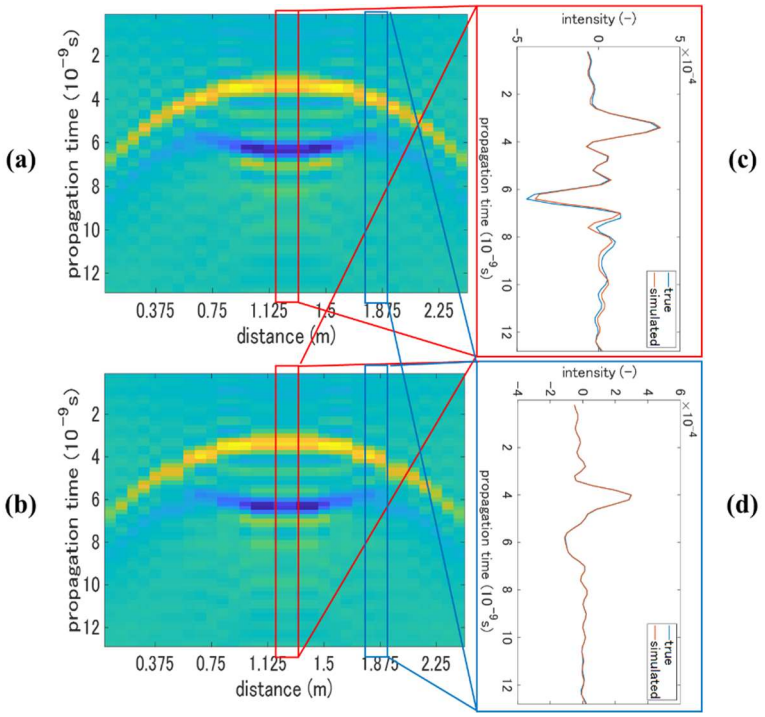


Fig. 5.35　Comparison of true and interpolated images ($a = 1.5$ m, $b = 0.5$ m, $\Delta\theta = 5.6°$, $\theta = 42.2°$) (a) (b) true and interpolated reflection patterns (c) (d) 1/2 and 1/4 signals

reflection patterns were almost the same. From Fig. 5.31 (c) (d), the 1/2 and 1/4 points signals are also the same. At $a = b$, the void is a sphere. Therefore, $I(r, \theta)$ does not depend on θ. At any θ, $I_i(r, \theta)$ equals to $I(r, \theta)$.

On the other hand, when the difference between a and b is large, the difference of the horizontal lengths of void at section θ_n and θ_{n+1} in Eq. (5-35) becomes large, resulting in the large error of $I_i(r, \theta)$. Fig. 5.32 - Fig. 5.35 show the reflection patterns assuming $a = 1.5$ m, $b = 0.5$ m when the difference of the lengths in scan and channel directions becomes largest. Larger $\Delta\theta$ is, larger the error is. Fig. 5.32 shows the sections at $\theta = 22.5°$ in the case $\Delta\theta = 45°$. From Fig. 5.32 (c), (d), the signal at a 1/2 point is similar while the signal at a 1/4 point have large error. Table 5.3 shows the error of the peak values of a 1/4 point. At

Table 5.3 Error of true and interpolated reflection patterns

	$\Delta\theta = 45°$ Error (%)	$\Delta\theta = 22.5°$ Error (%)	$\Delta\theta = 12.3°$ Error (%)	$\Delta\theta = 6.1°$ Error (%)
1/4 point peak	−63.8 %	−9.2 %	−2.2 %	+0.8 %

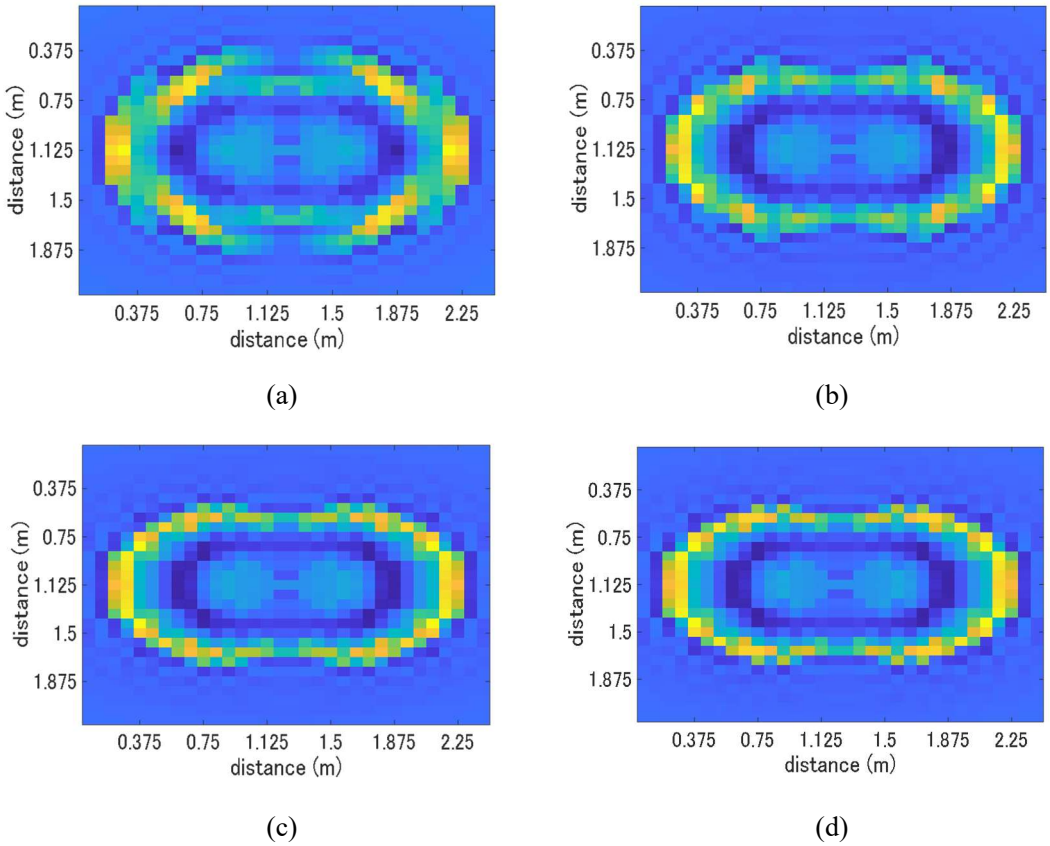


Fig. 5.36 Horizontal sections of interpolated reflection patterns (a) (b) (c) (d) $\Delta\theta = 45°,\ 22.5°,\ 11.3°,\ 5.6°$ respectively

$\Delta\theta = 45°$, the peak value is underestimated by about 64 %. Fig. 5.33, Fig. 5.34 and Fig. 5.35 show the reflection patterns at $\theta = 33.8°, 39.4°, 42.2°$ in the cases $\Delta\theta = 22.5°, 11.3°, 5.6°$ respectively. The reflection patterns are almost the same. The 1/4 signals are also reproduced accurately. The length of the void in a scan direction in the interpolated section L can be formulated as below.

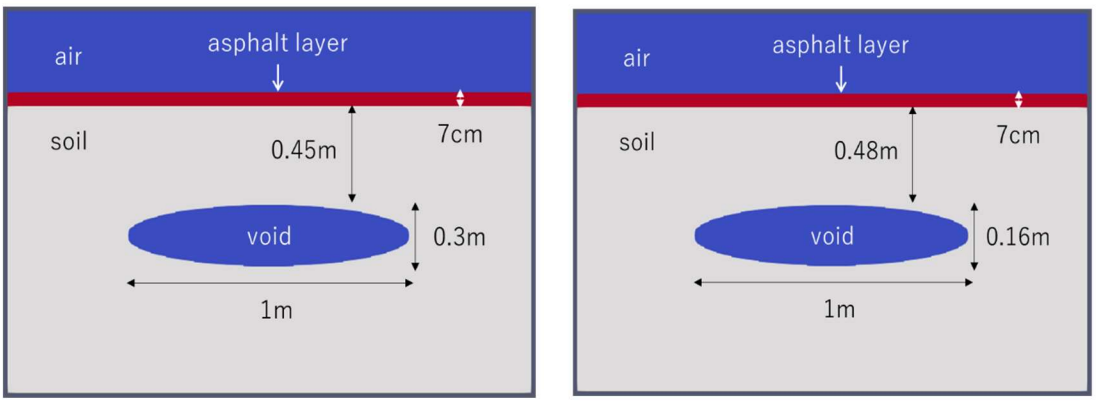
$$L = \sqrt{(a\cos\theta)^2 + (b\sin\theta)^2} \tag{5-36}$$

In the case $a = 1.5$ m and $b = 0.5$ m, $L = 1.50$ m at the section $\theta = 0°$. $L = 1.40$ m at $\theta = 22.5°$ and $L = 1.12$ m at $\theta = 45°$. Therefore, in the case $\Delta\theta = 22.5°$, the maximum difference of L at different angles is 28 cm. The difference is much larger than scan pitch 7 cm, causing large error. In the case $\Delta\theta = 11.3°$, the differences of L among $\theta = 22.5°, 33.8°, 45°$ are at most 16 cm. The difference is relatively small. The proposed interpolation method can reproduce the reflection patterns as shown in Table 5.3.

Fig. 5.36 shows the comparison of the horizontal section of the reflection patterns passing through the center of the void. From Fig. 5.36 (a), in the case $\Delta\theta = 45°$ discontinuity occurred while in the case $\Delta\theta = 22.5°$ the ellipsoid shape of the void was reconstructed. The reflection patterns converged over $\Delta\theta = 22.5°$. Therefore, in the research $\Delta\theta = 22.5°$ was adopted. The calculation time of the proposed algorithm was 10 minutes for 1 data, reduced to 0.7 % compared to 3D-FDTD method.

5.4.3 Comparison to measurement data

Reproduced 3D reflection patterns were compared to the experimental field data and measurement data. Fig. 5.37 shows the geometry of void. The experimental field model shown in Fig. 5.37 (a) was 0.3 m thickness at 0.45 m depth. The measurement model shown in Fig. 5.37 (b) was 0.16 m thickness at 0.48 m depth. The lengths in a scan direction of the two cases were 1 m and 0.5 m respectively. The lengths in a channel direction were the same as the lengths in a scan direction in the both cases. Fig. 5.38 and Table 5.4 compare the results of the experimental field data. Fig. 5.39 and Table 5.5 compare the results of the measurement data. From Fig. 5.38 and Table 5.4, the reflection patterns and peak values of 1/2 and 1/4 points



(a) (b)

Fig. 5.37 Geometry of model for validation (a) experimental field (b) measurement

signals were close. The reflection pattern of the experimental field model was accurately reproduced within 20 % error. On the other hand, from Fig. 5.39 (a), (b), the reflection pattern of the measurement model was different from the simulated pattern. From Fig. 5.39 (c), the peaks of a 1/2 point were close. However, from Fig. 5.39 (d), the 1/4 point signal was totally different. From Table 5.5, the difference of the peak values was up to 100 %. In the case of the measurement data, because of the effects of soil layer boundaries above the void, clutters of other buried objects and complex geometry of void, the model itself was different from the actual situation. Furthermore, the electromagnetic characteristics of the model may be different because of water content and loosen soil around the void. The conclusion is, it is difficult to exactly reproduce the reflection patterns of measurement data. These characteristics affect the validation accuracy of a trained CNN model in the following section.

Table 5.4　Comparison of peak values of experimental field data and simulation

	Experimental field Peak value× 10^{-4}	Simulation Error (%)
1/2 positive peak	5.03	+10.9 %
1/2 negative peak	−4.29	−18.7 %
1/4 peak	4.49	−16.5 %

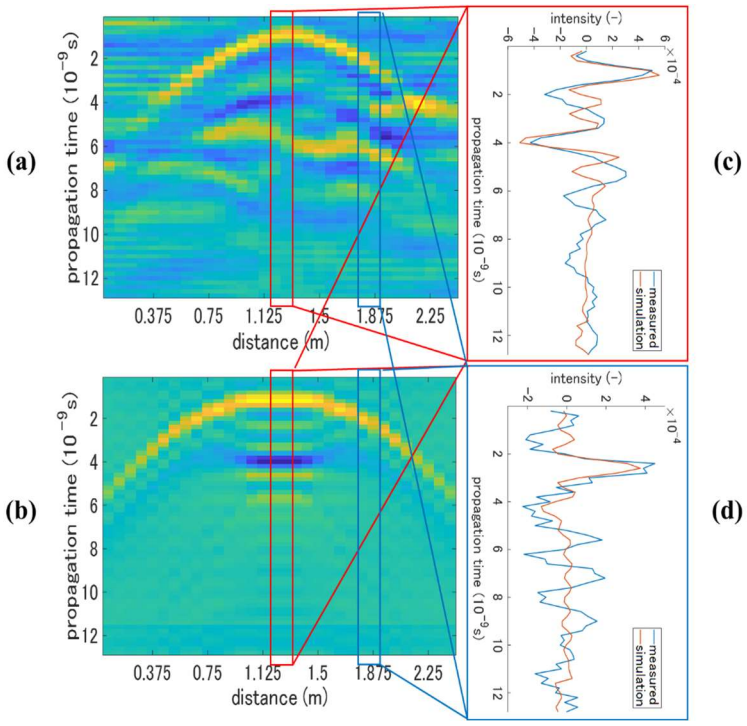


Fig. 5.38　Reproduction of experimental field data (a) (b) measured and simulated reflection patterns (c) (d) 1/2 and 1/4 point signals

5.5 Training by 3D-CNN

5.5.1 2D and 3D-CNN models

Referring to the discussions of the previous sections, training data was produced by FDTD method. From a geotechnical engineering point of view, the shapes of actual void are not only an ellipsoid body but also a cuboid body. In terms of 2D and 3D void, the two types of models shown in Fig. 5.40 and Fig 5.41 were assumed. The thickness of asphalt layer is 7 cm. The relative permittivity of asphalt layer is 4. The ranges of the relative permittivity and conductivity of soil layer are 2 - 10 and $0.001\ \mathrm{s/m}$ - $0.01\ \mathrm{s/m}$ respectively. From the discussions of Section 2.3.2, the ranges of the depth, length in a horizontal direction and thickness of void are 0.2 m - 0.5 m, 0.5 m - 1.5 m, 0.1 m - 0.5 m respectively. In terms of a 3D void model, the range of the length in a channel direction is also 0.5 m - 1.5 m, which was independently determined from the

Table 5.5 Comparison of peak values of measurement data and simulation

	Measurement Peak value$\times 10^{-4}$	Simulation Error (%)
1/2 positive peak	4.94	+13.2 %
1/2 negative peak	−6.58	+28.0 %
1/4 peak	1.50	+98.7 %

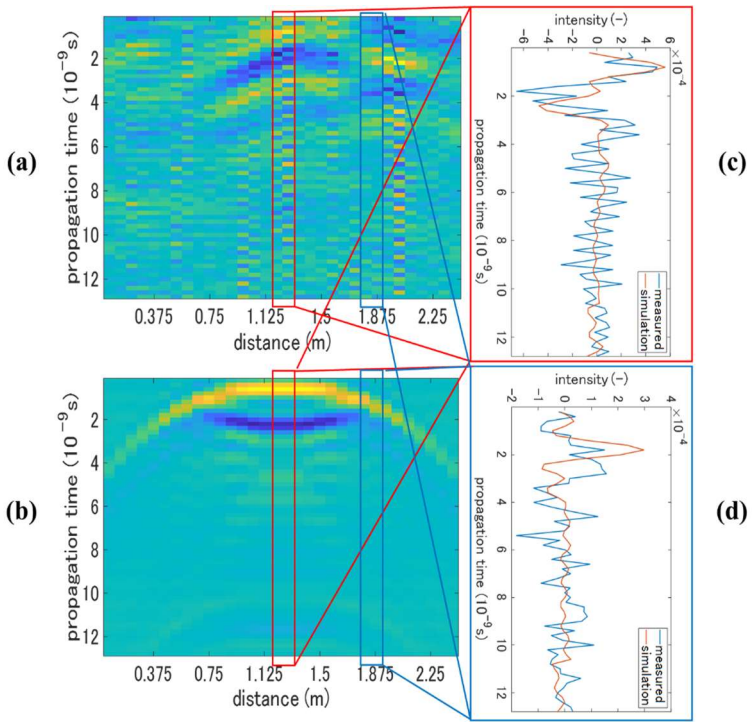
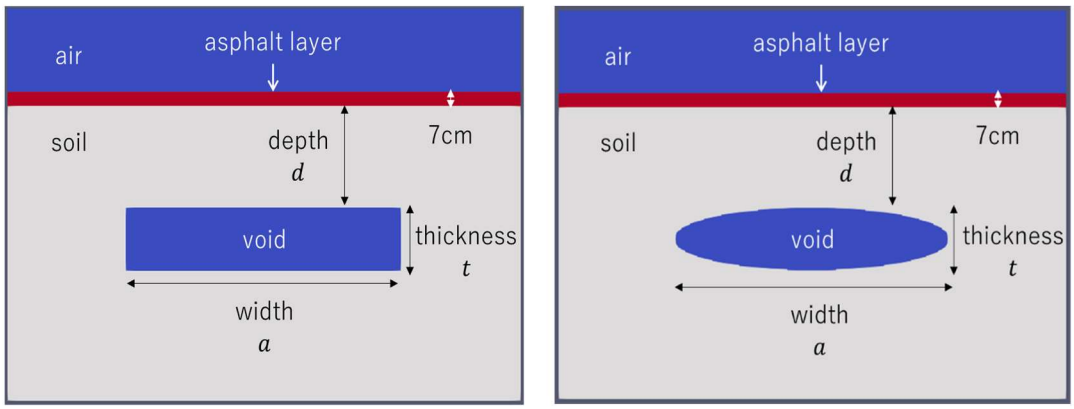
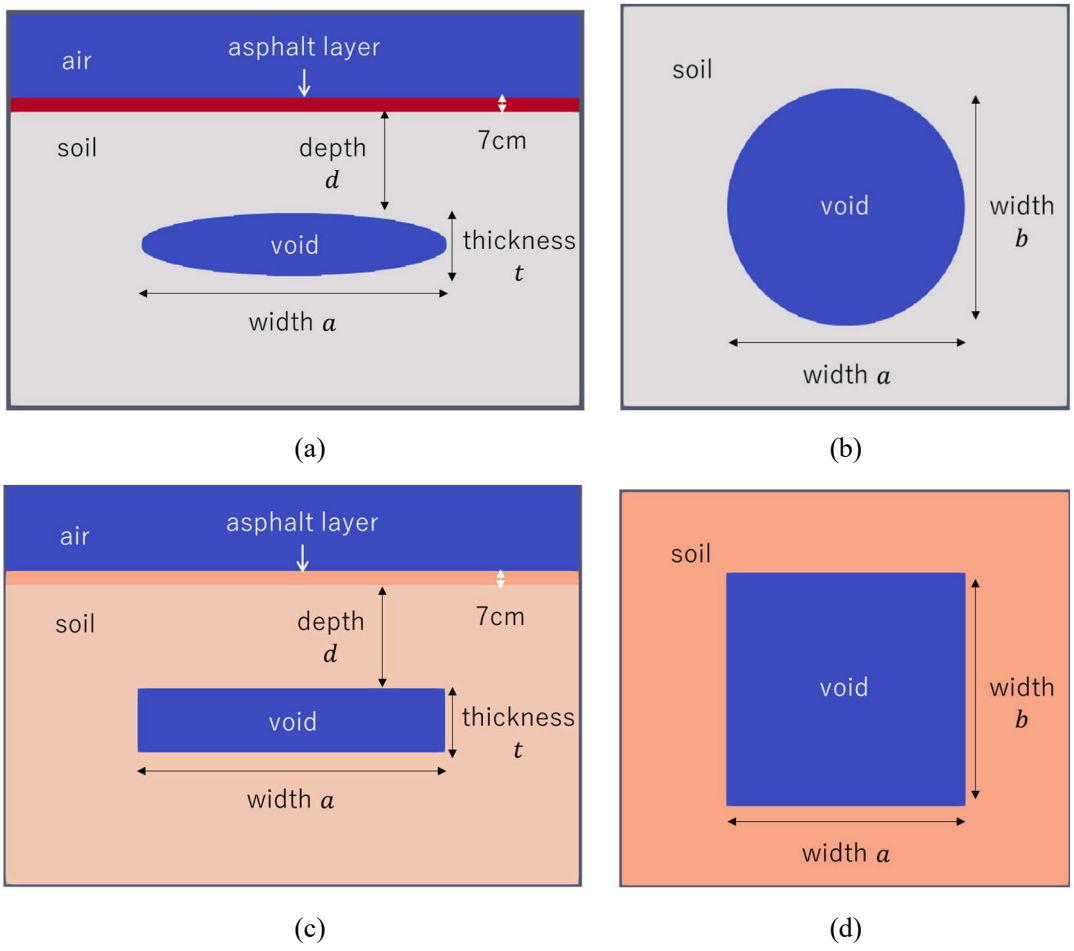


Fig. 5.39 Reproduction of measurement data (a) (b) measured and simulated reflection patterns (c) (d) 1/2 and 1/4 point signals



(a) (b)

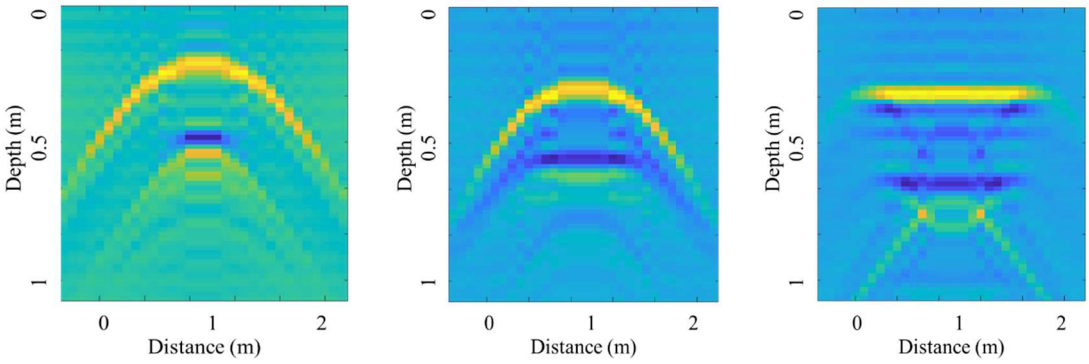
Fig. 5.40 2D void models (a) ellipsoid (b) cuboid



(a) (b)

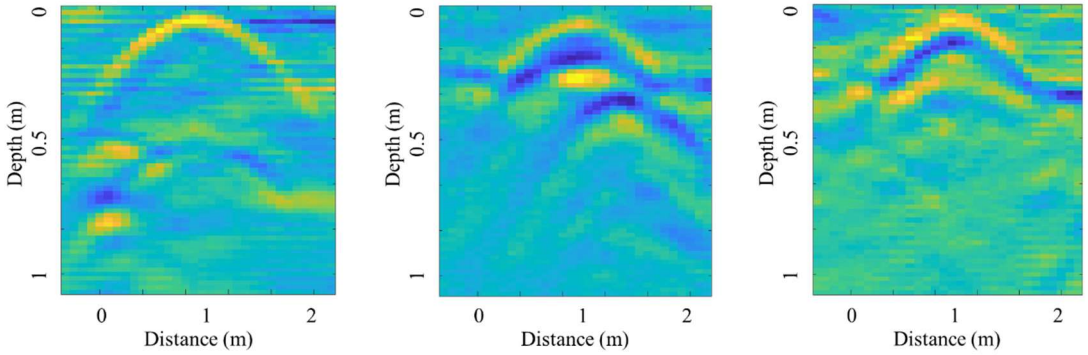
(c) (d)

Fig. 5.41 3D void models (a) (b) ellipsoid, vertical and horizontal sections respectively (c) (d) cuboid, vertical and horizontal sections respectively



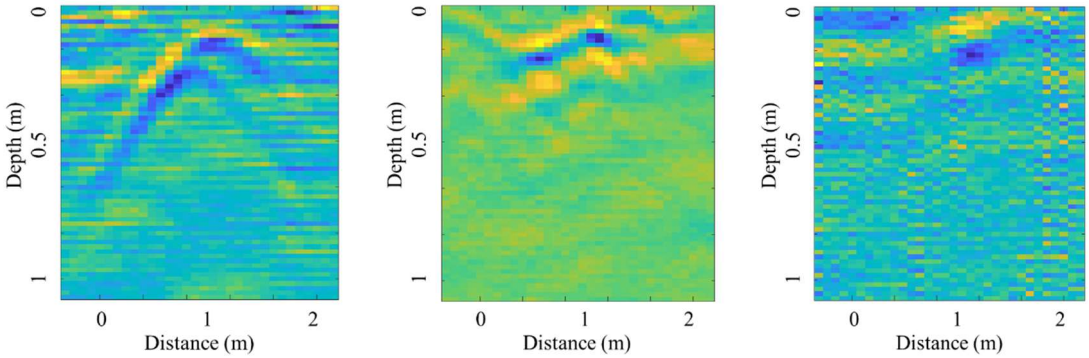
(a) (b) (c)

Fig. 5.42 Reflection patterns of simulation data



(a) (b) (c)

Fig. 5.43 Reflection patterns of experimental field data



(a) (b) (c)

Fig. 5.44 Reflection patterns of measurement data

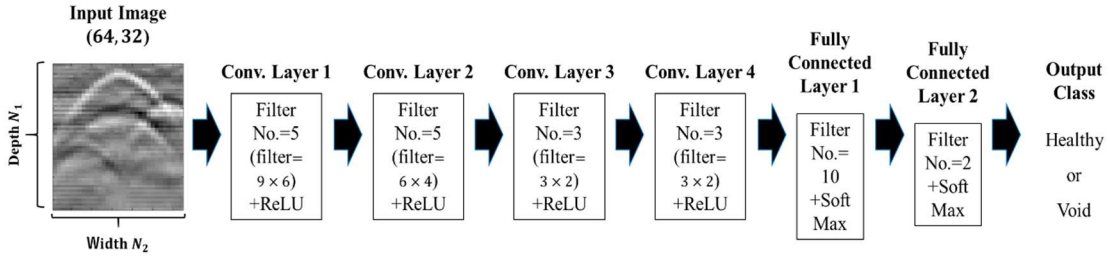
length in a scan direction. The parameters were generated from uniform random numbers. In terms of the 2D void model, from the discussions of Section 6.4, 2500 void training data was produced. 2,500 healthy section training data was randomly extracted from the measurement data. The total number of the training data was 5,000. In terms of the 3D void model, the 250 void and 250 healthy section, total 500 training data was produced.

Fig. 5.42 shows the examples of simulated reflection patterns. Fig. 5.43 and Fig. 5.44 show the examples of the experimental field and measurement data. Fig. 5.42 (a), (b) are the ellipsoid models with different thickness and lengths in a horizontal direction. Fig. 5.42 (c) is the cuboid model. Depending on the geometry of void, various reflection patterns were produced. However, most of the reflection patterns showed a hyperbolic shape. From Fig. 5.43, in the case of the experimental field data, clear hyperbolic shapes were observed. On the other hand, from Fig. 5.44, in the case of measurement data, the reflection patterns were complicated.

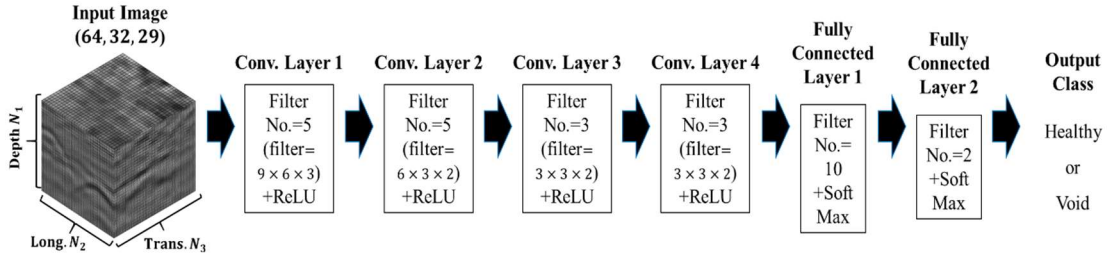
5.5.2 Training by transfer learning

Fig. 5.45 (a) shows the optimized 2D-CNN model for void classification. Fig. 5.45 (b) shows the optimized 3D-CNN model. From the discussions of previous research, a 4 convolution and 2 fully connected layers, total 6 layer model was adopted [83), 84)]. Maxpooling was adopted between each layer to reduce calculation cost. The size of input data is represented by pixels in depth, scan and channel directions (N_1, N_2, N_3). In terms of 2D-CNN, $(N_1, N_2) = (64,32)$. In terms of 3D-CNN, $(N_1, N_2, N_3) = (64,32,29)$. The effects of training method, sizes of input data and convolution filters were evaluated.

The number of the experimental field data is 20. The measurement data is 88. Therefore, deep learning models cannot be trained only by the experimental field or measurement data. Fig. 5.46 shows the results of



(a)



(b)

Fig. 5.45 Optimized CNN models (a) 2D-CNN (b) 3D-CNN

2D-CNN trained by the simulation data and validated by the measurement data. From Fig. 5.46 (a), the simulation data was classified with almost 100 % classification accuracy. However, from Fig. 5.46 (b), the training failed when the trained CNN model was applied to the measurement data. This is because the simulation data is different from the measurement data as discussed in Section 5.4.3. The CNN models were trained by transfer learning as stated in previous research [83), 84)]. The CNN models were trained by the simulation data. Then, the parameters of convolution filters shown in Fig. 5.45 were fixed. The two thirds of the measurement data were used for training and one third was used for test. Only the fully connected layers were updated to utilize the learned spatial features of convolution filters. As shown in Fig. 5.47 (a), the learning curve monotonically increased. From Fig. 5.47 (b), the validation curve converged to a certain value after 100 epochs, indicating the training succeeded. Therefore, the optimal parameters of convolution filters are considered to be in common between the simulation and measurement data.

Fig. 5.48 summarizes the effect of the training method on the classification accuracy of the experimental field and measurement data by 2D-CNN. In terms of the experimental field data, even without transfer

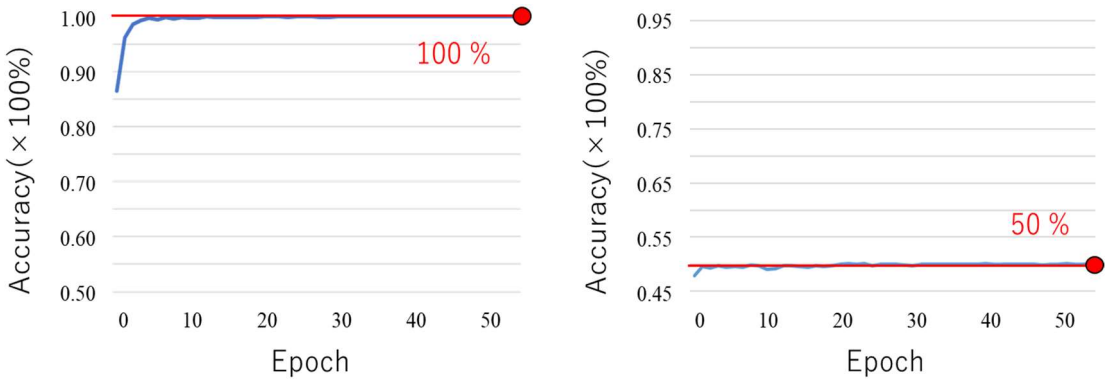


Fig. 5.46 Training by simulation data (2D-CNN) (a) learning curve (b) validation curve

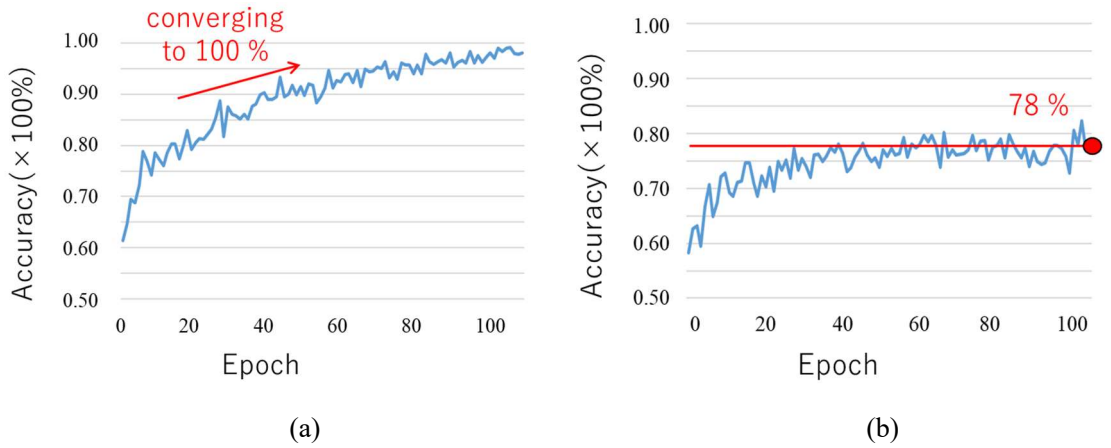
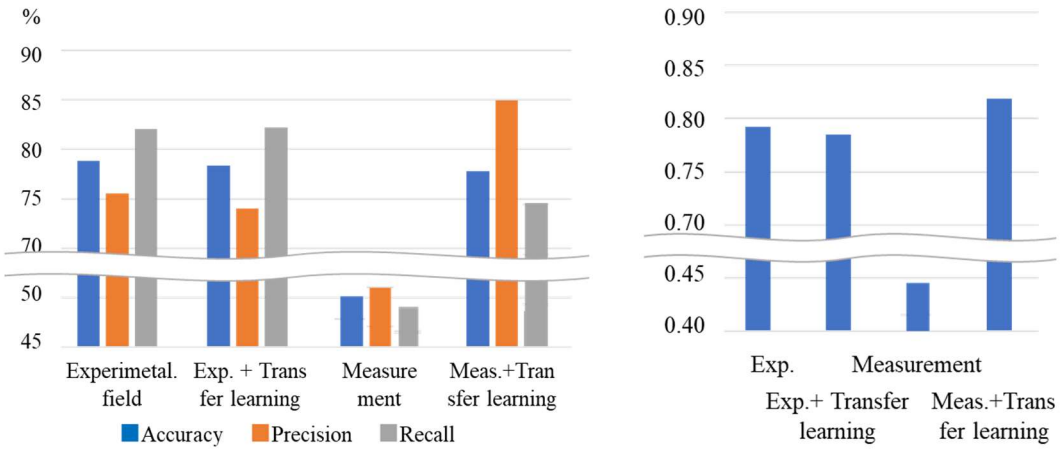
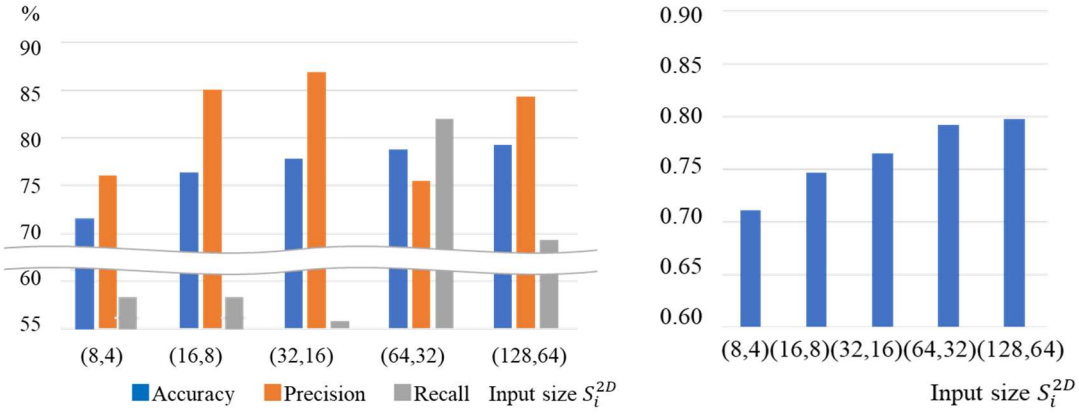


Fig. 5.47 Training by transfer learning (2D-CNN) (a) learning curve (b) validation curve



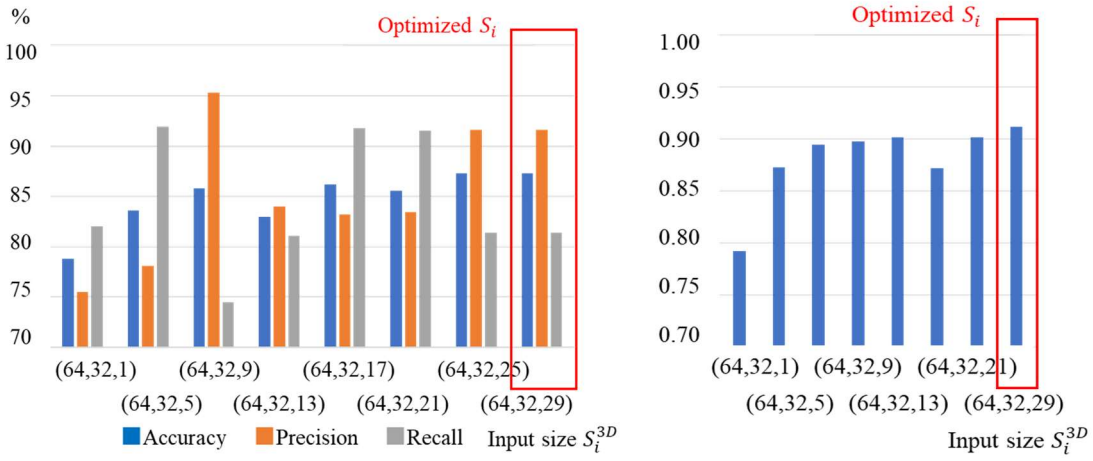
(a) (b)

Fig. 5.48 Comparison of training method (a) Accuracy, Precision, Recall (b) AUC



(a) (b)

Fig. 5.49 Effect of size of input data on 2D-CNN (measurement data) (a) Accuracy, Precision, Recall (b) AUC



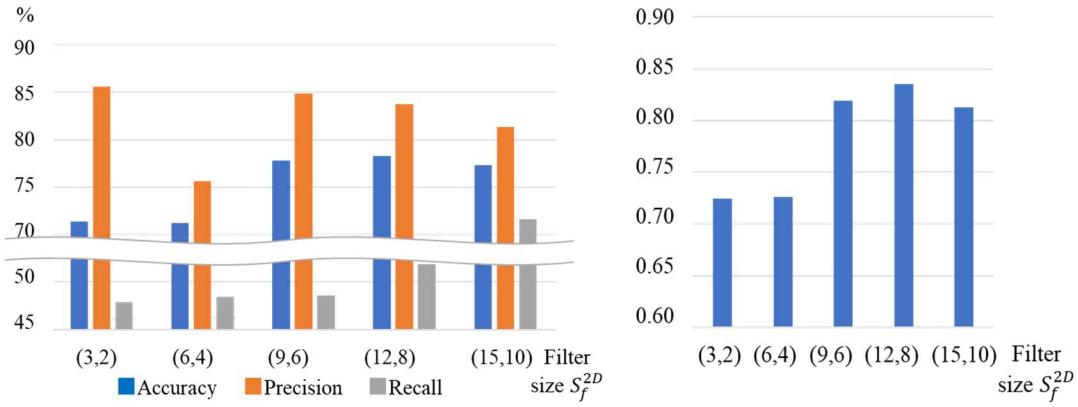
(a) (b)

Fig. 5.50 Effect of size of input data on 3D-CNN (measurement data) (a) Accuracy, Precision, Recall (b) AUC

learning classification accuracy was about 78 %. After transfer learning, classification accuracy did not change. Certain classification accuracy was obtained because the reflection patterns of the experimental field data were similar to the simulation data. This is consistent with previous research [84), 85)]. On the other hand, in terms of the measurement data, the same classification accuracy about 78 % was obtained only by transfer learning. The following discussions were based on the measurement data case after transfer learning.

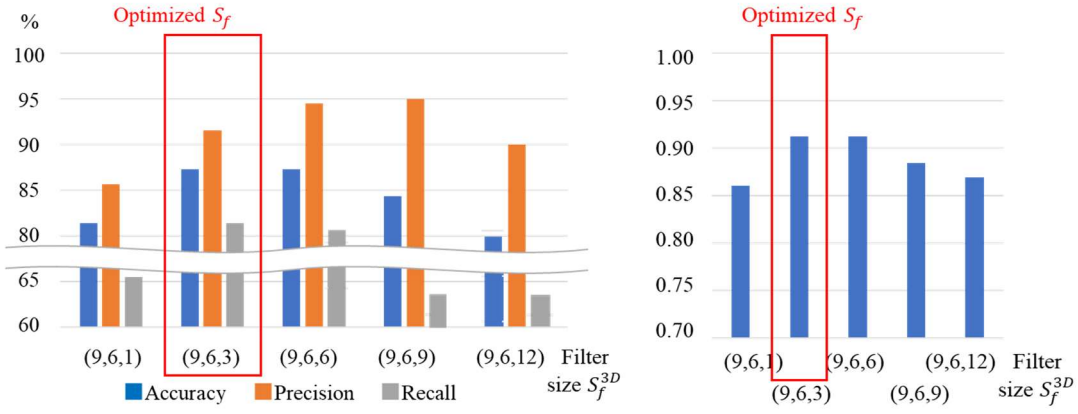
5.5.3 Optimization of the models

Fig. 5.49 shows the effect of the size of input data S_i on the classification accuracy of 2D-CNN. In all the cases, the sizes of convolution and max pooling filters were optimized. From Fig. 5.49, the classification accuracy at $S_i^{2D} = (64,32)$ was about 78 %, by about 7 % increase compared to the case $S_i^{2D} = (8,4)$. Classification accuracy converged over $S_i^{2D} = (64,32)$. $S_i^{2D} = (64,32)$ was considered to be enough to extract the reflection patterns of void. This tendency was the same as the subsurface pipes case. $S_i^{2D} =$



(a) (b)

Fig. 5.51 Effect of size of convolution filter on 2D-CNN (measurement data) (a) Accuracy, Precision, Recall (b) AUC



(a) (b)

Fig. 5.52 Effect of size of convolution filter on 3D-CNN (measurement data) (a) Accuracy, Precision, Recall (b) AUC

$(64,32)$ was adopted in the research. Fig. 5.50 shows the effect of the size of 3D input data S_i^{3D} on the classification accuracy of 3D-CNN. As shown in Fig. 5.50 (a), classification accuracy monotonically increased when the pixel number in a channel direction increased. This tendency was also the same as the pipes case. As stated above, classification accuracy was 78 % in the case $S_i^{3D} = (64,32,1)$ corresponds to 2D-CNN. The classification accuracy at $S_i^{3D} = (64,32,29)$ was about 87 %, by about 9 % increase compared to the 2D-CNN case. High accuracy classification was possible by considering the 3D hyperboloid reflection patterns of void by 3D input data. Therefore, $S_i^{3D} = (64,32,29)$ was considered to be the optimized size of input data.

Fig. 5.51 shows the effect of the size of convolution filters S_f^{2D} of the first layer on the classification accuracy of 2D-CNN at $S_i = (64,32)$. The sizes of the convolution filters of the other layers were also optimized. Classification accuracy was about 71 % at $S_f^{2D} = (3,2)$ and $S_f^{2D} = (6,4)$ while about 78 % at $S_f^{2D} = (9,6)$, by about 7 % increase. Classification accuracy converged over $S_f^{2D} = (9,6)$. This tendency was the same as the pipes case. The reflection patterns of void can be extracted over a certain S_f^{2D}. In the research, $S_f^{2D} = (9,6)$ was adopted. Fig. 5.52 shows the effect of the size of 3D convolution filters S_f^{3D} on the classification accuracy of 3D-CNN. Classification accuracy was about 81 % in the 2D filter case $S_f = (9,6,1)$. At $S_f = (9,6,3)$, classification accuracy was about 87 %, by about 6 % increase compared to the

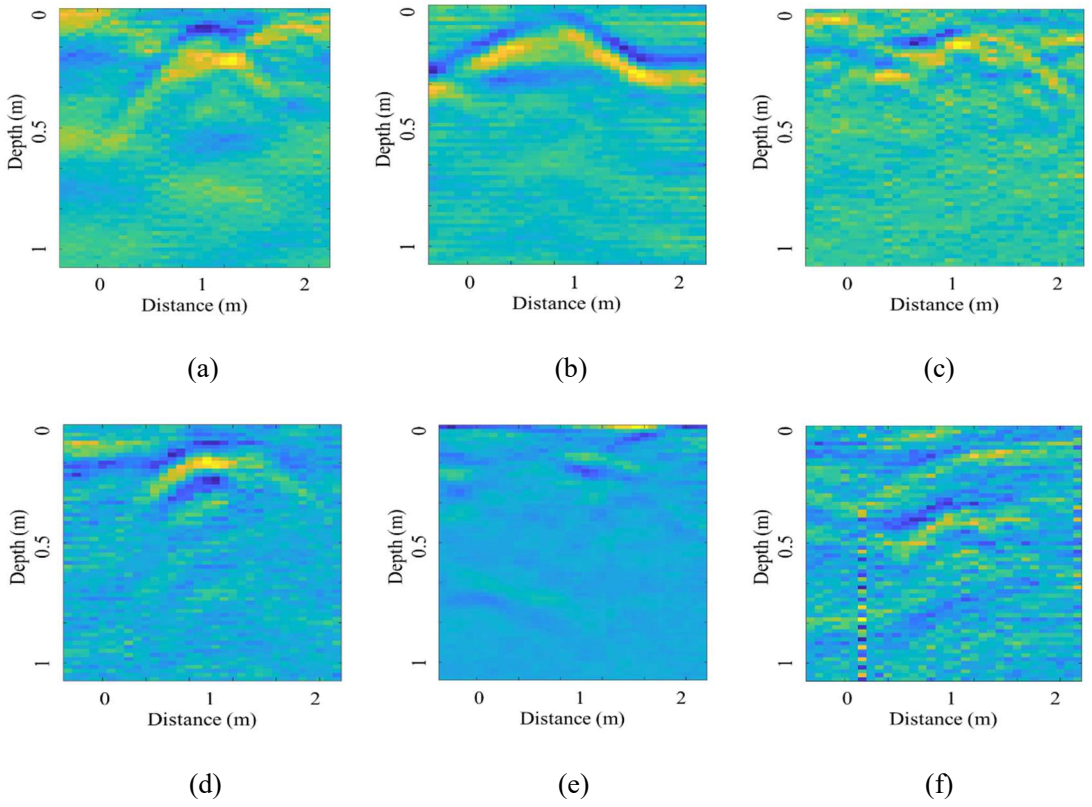


Fig. 5.53 Examples of detection (a) (b) detected by 2D-CNN (c) (d) only detected by 3D-CNN (e) (f) not detected by 2D and 3D-CNN

2D filter case. Different form the pipes case, because void has geometry in a channel direction, 3D convolution filters improved classification accuracy. In the research, $S_f = (9,6,3)$ was adopted. Classification accuracy decreased over $S_f = (9,6,9)$, possibly because overfitting occurred using too large filters for extracting target features.

Fig. 5.53 shows the examples of detected void. Clear reflection patterns were detected by 2D-CNN even if they were distorted from hyperbolic shapes as shown in Fig. 5.53 (a), (b). By 3D-CNN utilizing 3D features of the data, void can be detected even if the hyperbolic reflection patterns were further distorted and clutters were superposed as shown in Fig. 5.53 (c), (d). Void cannot be detected when the reflection patterns are weak and noise is dominant as shown in Fig. 5.53 (e), (f). The classification accuracy of the proposed 3D-CNN model was about 87 %, almost 90 %. The high accuracy 3D-CNN void model comparable with skilled inspectors' detection accuracy was developed.

5.6 Summary

The methodology to reproduce 3D void training data by 2D-FDTD method was proposed to develop a 3D-CNN model. The theory of FDTD method was introduced. The necessary time duration, cell size and time step width were theoretically derived. The directivity of antennas was reproduced by an array antenna model comparing to measured antenna patterns. The effect of the directivity on reflection patterns was small, at most 5 % difference of peak values. The effect of the variation of asphalt layer was also small. On the other hand, the electromagnetic characteristics of soil layer and geometry of void affected the results. Therefore, the parameters of soil and void were randomly assigned, in the former case referring to the previous research and latter case to the excavation survey results.

To reproduce 3D void data by 3D-FDTD method, calculation time was enormous, about 1 day for 1 data. On the other hand, because void exists in a deeper region and was flat, the 3D propagation phenomenon of electromagnetic waves can be ignored. A 3D void model was approximated by a 2D void model. A void model has symmetricity around a center axis. Therefore, calculation time can be reduced to 1/4. In cylindrical coordinates, a section image of each angle was simulated. Then, a region between the simulated sections were linearly interpolated. It was found that when angle pitch was 22.5°, high accuracy approximation was possible with less calculation cost. The calculation time for 1 data was about 10 minutes, reduced to 0.7 % by the proposed method compared to 3D-FDTD method. The method accurately reproduced the reflection patterns of the experimental field data. On the other hand, the measurement data was affected by undesired clutters, showing the complicated reflection patterns which were different from the simulations.

When the model was trained by the simulation data and validated by the experimental field data, classification accuracy was about 78 %. On the other hand, training by the simulation data and validation by the measurement data failed. Therefore, after training by the simulation data, the convolution filters were fixed and fully connected layers were updated, which is called transfer learning. The classification accuracy of the measurement data was the same, about 78 % after transfer learning. In terms of 2D-CNN, the sizes of input data and convolution filters affected the results. This tendency was the same as the pipes case. On the

other hand, in terms of 3D-CNN, the sizes of not only input data but also convolution filters in a channel direction affected classification accuracy. This is because void has geometry also in a channel direction. 3D reflection patterns can be extracted by certain sizes of input data and convolution filters. Compared to a 2D-CNN model, the classification accuracy of 3D-CNN was high. It was about 87 %, by about 9 % increase. In the following chapter, a 3D-CNN model for subsurface pipes was developed. Then the classification accuracy of 2D and 3D-CNN was compared by the same procedure.

Chapter 6: Pipe detection by 3D-CNN

In this chapter, a pipe classification model by 3D-CNN was developed. The theory and training method of deep learning models were stated. Training parameters, preprocessing filters, model architectures, sizes of input data and convolution filters were optimized for 2D-CNN. For comparison, classification accuracy was compared to the optimized SVM model proposed in Chapter 4. The sizes of input data and convolution filters of a 3D-CNN model were optimized. The classification accuracy of 3D-CNN was compared to 2D-CNN. Considering the optimized parameters of transverse and longitudinal pipes' cases, a three-category, transverse, longitudinal pipes and no pipe section classification model was developed.

6.1 Production of training data

Training data of pipes was produced referring to the points of apexes of hyperbolic reflection patterns extracted by the inspectors as discussed in Section 2.3.1. In the case of 2D images, the cross-section images perpendicular to the directions of pipes were utilized. In a horizontal direction, the apexes of hyperbolic reflection patterns lie in the center of the images. In a depth direction, there was a 5 % margin (offset) just in case because the whole hyperbolic reflection pattern should be included.

In the case of 2D images, because the number of channels was 29, 29 training images were produced from one transverse pipe. On the other hand, in the case of 3D data, assuming a transverse pipe exists in all the channels, a 3D region of 29 channels was extracted (Fig. 6.1). Assuming the pixel number of input data in a channel direction M, the number of produced data from one transverse pipe will be an integer closest to $29/M$. Depending on M, there should be overlapping or unused channels. To remove the effect of channel characteristics, overlapping and unused channels were randomly determined. In the case of longitudinal pipes, start and end points were provided. Assuming a pipe exists linearly between the start and end points, training data with a certain size was extracted.

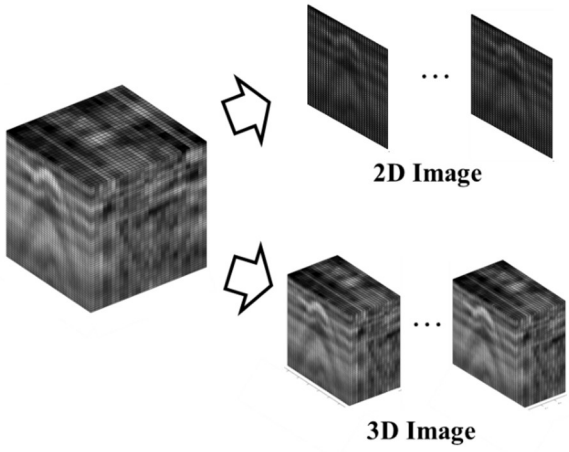
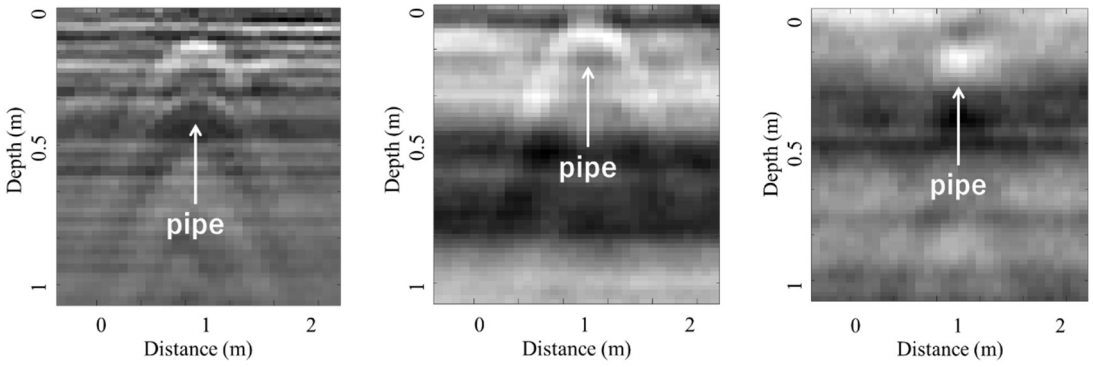


Fig. 6.1 Production of 2D and 3D training data

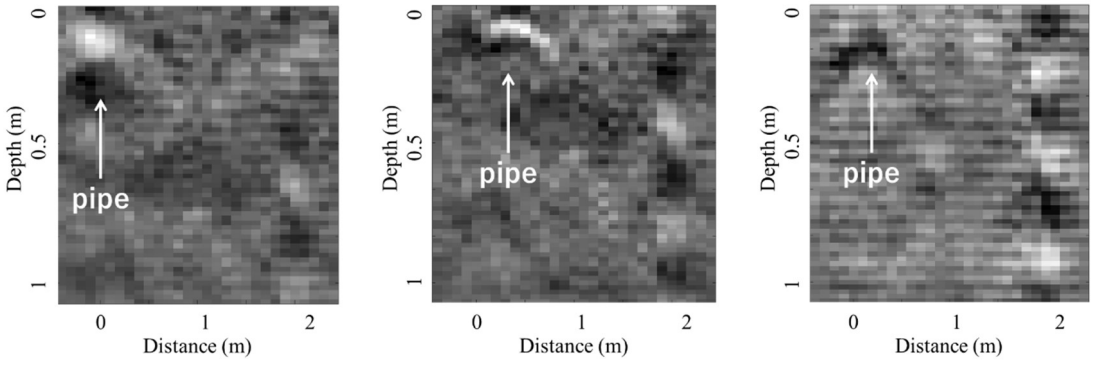
Fig. 6.2 and Fig. 6.3 show the examples of the cross-section images of transverse and longitudinal pipes. From Fig. 6.2, various reflection patterns appeared though most of the images show hyperbolic patterns. The phases and intensity of patterns depend on the depth, diameter, material of pipes and characteristics of surrounding soil. Diameter ranges from several cm to several tens cm. In some cases, diameter may be up to 1 m. Reflection patterns can be extracted by 64 pixels by 64 pixels (1 m by 2 m) from Fig. 6.2. From Fig. 6.3, hyperbolic reflection patterns were also observed in longitudinal pipes' cases. However, because of the difference of radar characteristics in longitudinal and transverse directions, the reflection patterns of each direction pipe were not the same. This fact affected the training results of each direction pipe.

Fig. 6.4 and Fig. 6.5 show the examples of 3D data of transverse and longitudinal pipes. Images are displayed every 5 channels (about 38 cm) in the case of a transverse pipe and 5 scans (about 35 cm) in the case of a longitudinal pipe. From Fig. 6.4, in the case of the transverse pipe, patterns were continuous in a channel direction though the positions of hyperbolic reflection patterns were gradually shifted. From Fig. 6.5, the patterns of the longitudinal pipe were also continuous in a scan direction. Utilizing the 3D spatial features of pipes is expected to improve the accuracy.



(a) (b) (c)

Fig. 6.2 Reflection patterns of transverse pipe

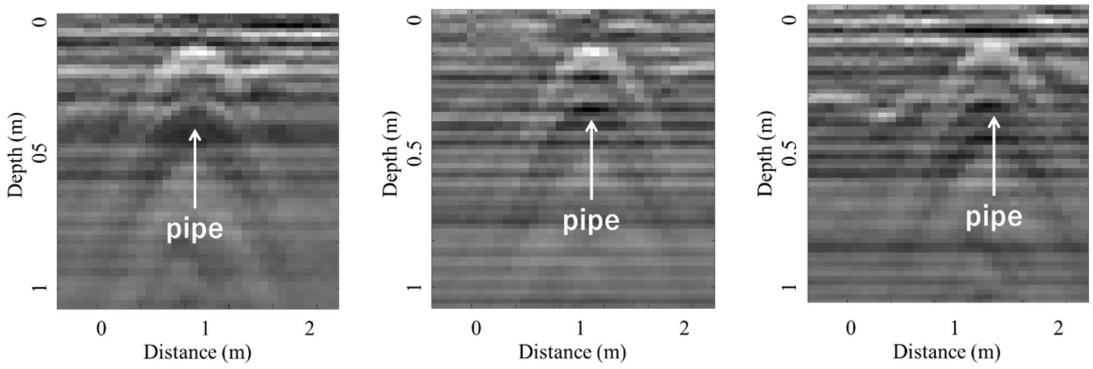


(a) (b) (c)

Fig. 6.3 Reflection patterns of longitudinal pipe

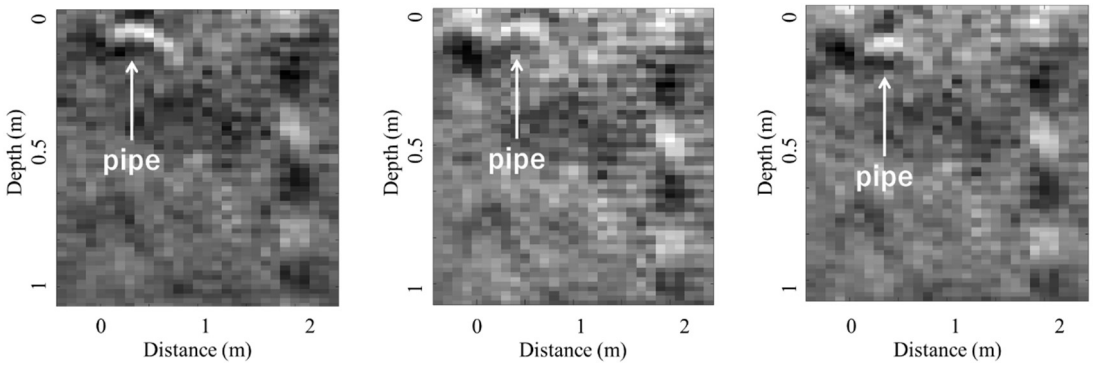
No pipe section training data was randomly extracted from regions without transverse and longitudinal pipes as shown in Fig. 6.6. No pipe section consists of not only noise but also clutters such as reflections from rocks and soil layer boundaries. The number of no pipe section data was the same as transverse and longitudinal pipes in the case of two-category classification and intermediate between each direction pipe in the case of three-category classification. The sizes of the data of all the categories were the same. The same preprocessing filters were applied. Training data was not normalized. Therefore, data type was double precision floating point with positive and negative values.

Fig. 6.7 and Fig. 6.8 show the examples of cross-sections of no pipe section 3D training data in scan and channel directions. From Fig. 6.7, the patterns of cross-section images were changed in a scan direction because of the effect of channel characteristics. On the other hand, from Fig. 6.8, the patterns of cross-section images were not changed in a channel direction. There is a continuity between two images in close distance in a longitudinal direction. Compared to the patterns shown in Fig. 6.7 and Fig. 6.8, most of the pipes' reflection patterns were obvious as shown in Fig. 6.2 - Fig. 6.5 so that a CNN model can easily detect pipes. A CNN model may also learn the channel characteristics to detect small changes in patterns.



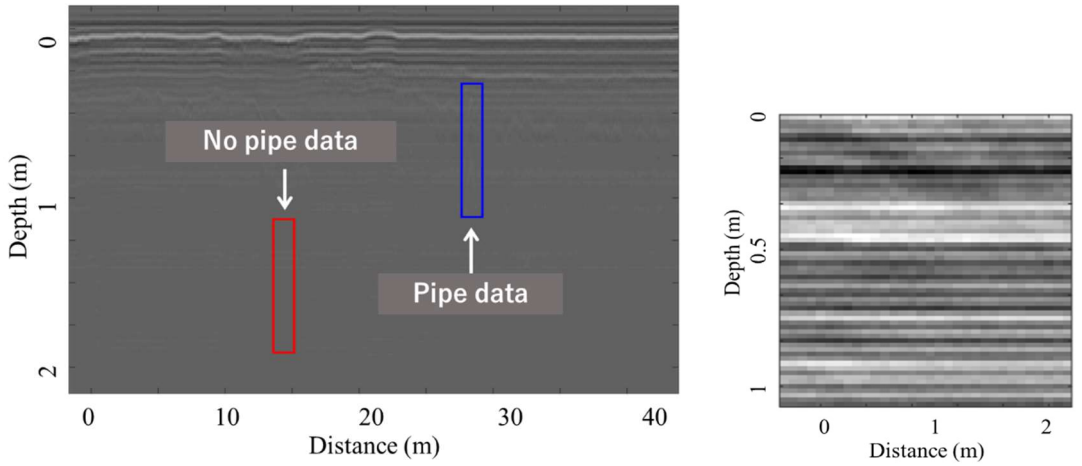
(a) (b) (c)

Fig. 6.4 Transverse pipe 3D data (scan direction cross-sections, every 5 channels)



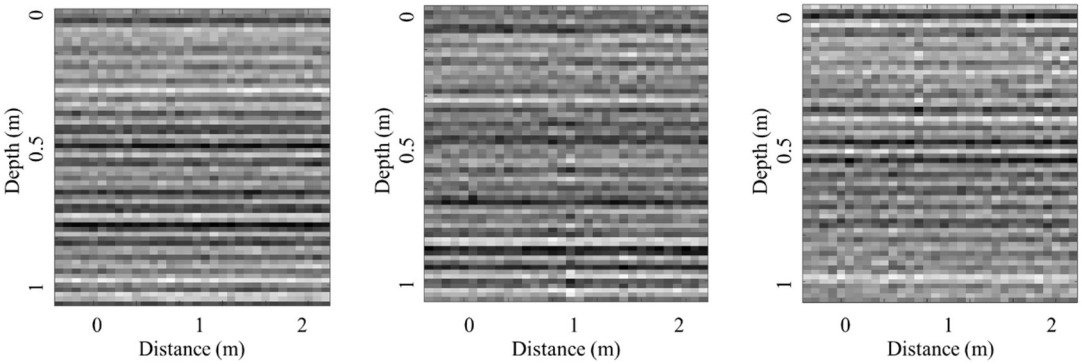
(a) (b) (c)

Fig. 6.5 Longitudinal pipe 3D data (channel direction cross-sections, every 5 scans)



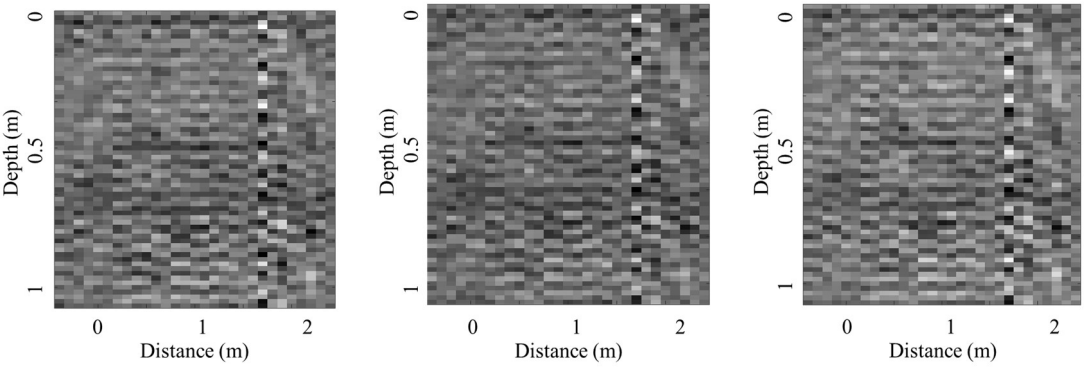
(a) (b)

Fig. 6.6 No pipe section (a) extraction of data (b) extracted image



(a) (b) (c)

Fig. 6.7 No pipe section data (scan direction cross-sections, every 5 channels)



(a) (b) (c)

Fig. 6.8 No pipe section data (channel direction cross-sections, every 5 scans)

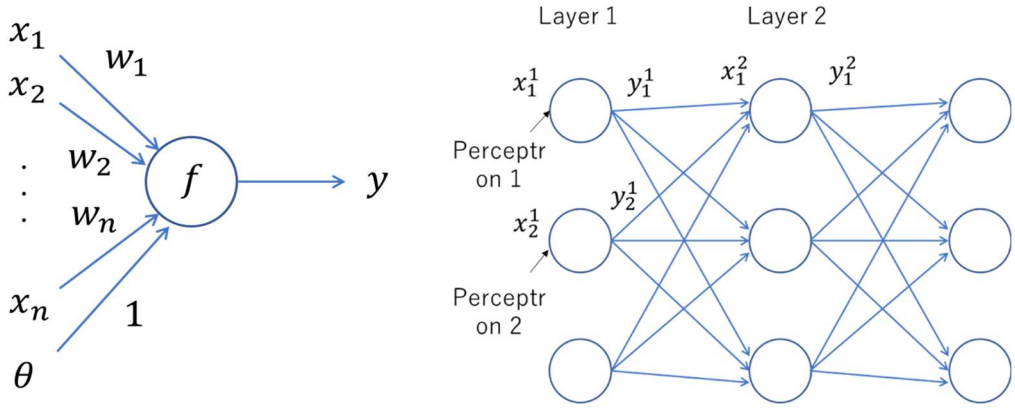
6.2 Theory of CNN

6.2.1 CNN model

Neural Network (NN) is one of the machine learning algorithms [139) - 143)]. NN is categorized into deterministic and probabilistic models, or deep layer and other models. Training method is categorized into supervised and unsupervised learning. In the research, a deterministic NN model was trained by supervised learning. Deterministic and supervised NN is further categorized into hierarchical NN and Recurrent Neural Network (RNN) for the classification of time series data. One example of RNN is Long Short Term Memory (LSTM). Models except for the simplest one-layer hierarchical NN are considered to be deep layer models. In the research, one of the hierarchical NN models which is designed for the classification of images, Convolutional Neural Network (CNN) was adopted.

In CNN, the parameters of convolution filters and weight and bias of fully connected layer perceptrons are to be tuned. The number of parameters of a NN model drastically increases depending on the number of layers. Early research focused on one-layer NN models. Hinton *et al.* (2006) summarized the idea of deep learning to train a Boltzmann machine model, one of the probabilistic NN models, for the classification of images [144)]. Recent years, because of the advent of GPU, several tens layer models can be trained. In deep learning, more abstract features can be obtained by deep layer architectures. In classical machine learning approaches, target features were extracted manually. A machine learning model is trained to classify extracted feature values. On the other hand, a deep learning model learns target features automatically from input data. A deep learning model extracts features and classifies input data at the same time. Deep learning is suited for the manual inspections by skilled inspectors such as radar images. Algorithms called AI in recent years are most likely deep layer hierarchical NN models.

The idea of NN is stemmed from the mathematical model of a nerve cell (neuron) of an animal, which is called a perceptron. A perceptron outputs y from input x utilizing a weight vector of the perceptron w, bias θ and activation function f as below. Fig. 6.9 (a) shows the concept of a perceptron.



(a) (b)

Fig. 6.9 NN model (a) perceptron (b) hierarchical NN

$$y = f(w \cdot x + \theta) \tag{6-1}$$

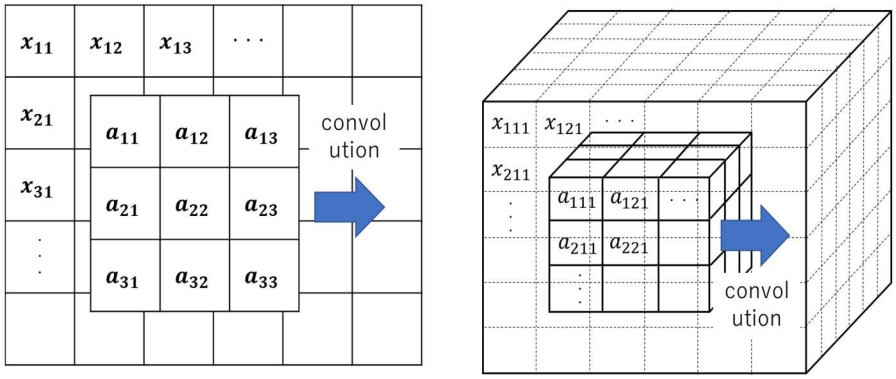
By adding a weight 1 to the vector x, θ can be included in w. The simplest f is a threshold function, though conventionally more complex nonlinear functions are utilized. The details of activation functions are explained in Appendix D.

A hierarchical NN model is composed of aligned perceptrons in each layer. An input of each perceptron is the vector of outputs of perceptrons in the previous layer. The number of perceptrons of the final layer equals to the number of output categories. The outputs of the final layer are normalized to relate the values to probability Fig. 6.9 (b) shows the example of hierarchical NN. In this section, n is the number of layers. m is the number of perceptrons. x_m^n corresponds to the input of the mth perceptrons of the nth layer. The perceptrons which affect the output of a certain perceptron are called a receptive field. For example, in Fig. 6.9 (b), all the perceptrons of the layers before a certain perceptron are a receptive field of the perceptron. A NN model which is composed of only one layer is not considered to be deep. A model whose number of layers is two or larger than two can be considered as a deep model.

CNN is a model to simulate optic nerve cells of humans, which was proposed by LeCun *et al.* (1998) [145]. CNN consists of fully connected and convolution layers which are added before the fully connected layers. CNN learns spatial features of images by convolution filters. The concept of 2D convolution is shown in Fig. 6.10 (a). Spatial filters are moved in each position of images. Assuming an image before convolution is x and convolution filter parameters are w, an image after convolution y is defined below. The filter size of w is denoted as (M, N).

$$y_{i,j} = \sum_{m=1}^{M} \sum_{n=1}^{N} w_{m,n} \cdot x_{i+\frac{2m-M-1}{2}, j+\frac{2n-N-1}{2}} \tag{6-2}$$

Eq. (6-2) is different from the definition of convolution operation in mathematics. Eq. (6-2) is common in image processing. After convolution operation, the size of a image is reduced by $(M-1, N-1)$. To



(a) (b)

Fig. 6.10 Convolution operation (a) 2D filter (b) 3D filter

maintain the size of the image, padding such as zero padding is applied to the convoluted image. Fig. 6.10 (b) shows the concept of 3D convolution. The idea is the same as 2D convolution. In 3D-CNN, 3D convolution filters are moved in 3D space. Assuming an image before convolution is x and convolution filter w, an image after convolution y is defined as below. The dimension of x, y, w is three. The size of a convolution filter w is denoted as (L, M, N).

$$y_{i,j,k} = \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{l=1}^{L} w_{l,m,n} \cdot x_{i+\frac{2l-L-1}{2}, j+\frac{2m-M-1}{2}, k+\frac{2n-N-1}{2}} \qquad (6\text{-}3)$$

The size of an image after convolution is reduced by $(L-1, M-1, N-1)$. Padding may be applied to maintain the size of the image.

As shown in Fig. 6.11, convolution layers are composed of multiple convolution filters. Images After applying filters to input images are called feature maps. Convolution filters are further applied to feature

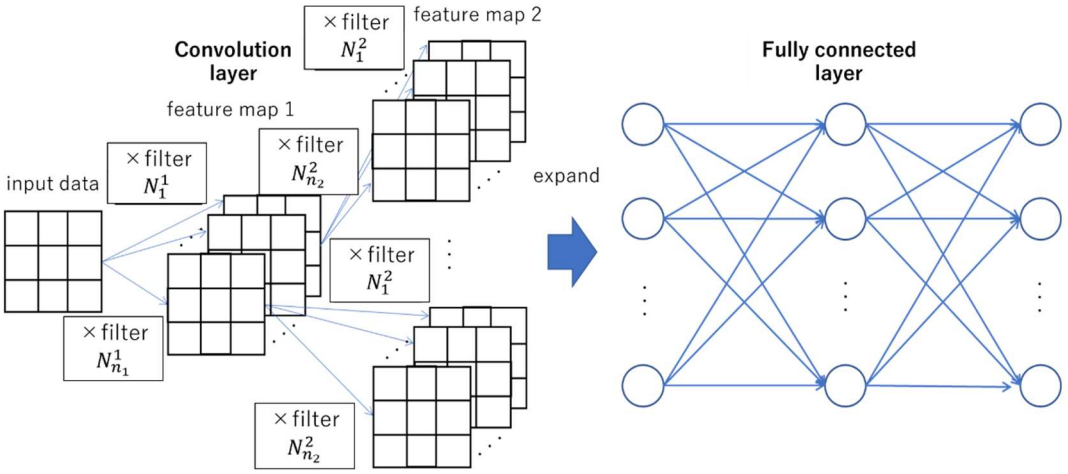


Fig. 6.11 2D-CNN model

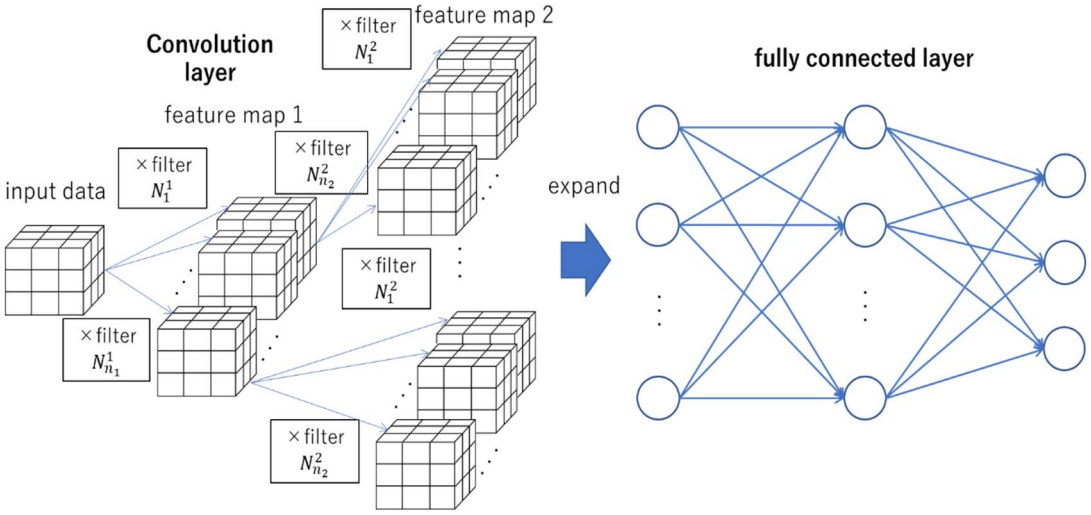


Fig. 6.12 3D-CNN model

maps of each layer. Feature maps are flattened to a vector to input to a fully connected layer. All the coefficients of feature maps before a certain coefficient are called a receptive field. By adding layers, more abstract spatial features are learned combining multiple features learned by previous layers. More the number of layers is, better the performance of the model becomes. However, the number of training parameters rapidly increases according to the number of layers. Furthermore, overfitting (overtraining) and vanishing/exploding gradient problem may occur. Therefore, to reduce calculation cost, pooling filters are added after convolution filters. Pooling filters are convoluted to feature maps to downsample the maps. As a pooling filter, conventionally Max Pooling, which extracts the maximum value of a target area or Average Pooling, which calculates the average of the target area is adopted. In the research, Max Pooling was adopted. By utilizing pooling filters, a model can also be robust to geometric transformations such as translation and scale change.

In the research, 3D-CNN was applied to the detection of pipes and void. Generally, CNN is applied to 2D images by utilizing 2D convolution filters. From Fig. 6.12, 3D-CNN utilizes 3D input data and 3D convolution and pooling filters to deal with 3D spatial features of the data. Feature maps are also 3D. The dimension of fully connected layers does not change. However, the number of parameters drastically increases according to the increase of the dimension of feature maps. Calculation cost is much higher. Training by CPU is not feasible for 3D-CNN. Another disadvantage is training tends to fail unless layer architectures and training method are deliberately chosen. Taining method was discussed in Section 6.3. In previous research, 3D-CNN was applied to the detection of tumors from MRI data (Kamnitsas *et al.*, 2017) and detection of human motion from video data (Ji *et al.*, 2013) [146) - 148)]. However, there is no research about the application of 3D-CNN to radar data. 3D-CNN can learn video data. On the other hand, RNN is targeted for time series data. Therefore, 3D-CNN is considered to connote the models for time series data.

6.2.2 Backpropagation

Rumelhart *et al.* (1986) proposed a training algorithm of NN, backpropagation [149)]. The number of categories is N. When input data is categorized as nth class, a one-hot vector t, whose nth value is 1 and other elements are 0, is output.

$$t = [0,0,0,\dots,1,\dots,0] \tag{6-4}$$

An optimization problem of NN parameters can be converted to a minimization problem of the difference of output and true labels. Normalized output labels are y. True labels are t. One of the typical objective functions for minimization, which is also called a loss function, is cross entropy L defined below.

$$L = -\sum_{i=1}^{N} t_i \log y_i \tag{6-5}$$

In the research, cross entropy was adopted.

In hierarchical NN, w_{ij}^n stands for the weight relating the ith node of the $n-1$th layer and jth node of the nth layer. In backpropagation, the derivative of L by w_{ij}^n is calculated. Then w_{ij}^n is updated by the

equation below, which is called gradient descent method. η is learning rate.

$$w_{ij}^{n} \leftarrow w_{ij}^{n} - \eta \frac{\partial L}{\partial w_{ij}^{n}} \tag{6-6}$$

The derivative $\frac{\partial L}{\partial w_{ij}^{N}}$ of the final Nth layer is written by the output of each node $y_j = f(v_j)$ and intermediate variable $v_j = \sum_i w_{ij}^{N} x_i$. The derivative is replaced by the equation below derived from the chain rule of derivative operation.

$$\frac{\partial L}{\partial w_{ij}^{N}} = \frac{\partial L}{\partial y_j} \cdot \frac{\partial y_j}{\partial v_j} \cdot \frac{\partial v_j}{\partial w_{ij}^{N}} \tag{6-7}$$

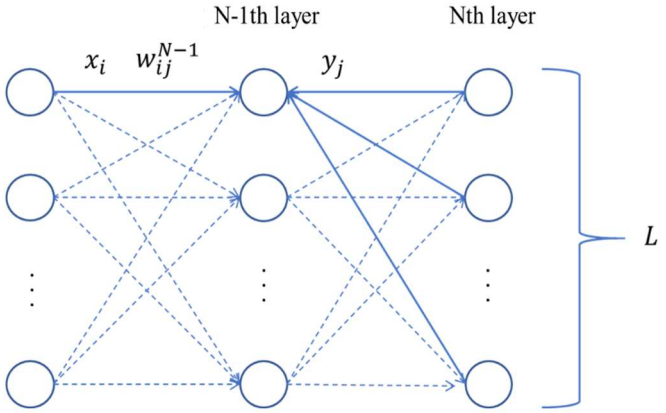
$$= \frac{\partial L}{\partial y_j} \cdot \frac{\partial y_j}{\partial v_j} \cdot x_i \tag{6-8}$$

$\frac{\partial L}{\partial y_j}$ is derived from the definition of L. $\frac{\partial y_j}{\partial v_j}$ is derived from the definition of f. From Eq. (6-8), $\frac{\partial L}{\partial w_{ij}^{N}}$ depends on the output of the previous layer x_i. The derivative of the $N-1$th layer $\frac{\partial L}{\partial w_{ij}^{N-1}}$ is derived by the same procedure. Because y_j affects the outputs of all the nodes of the Nth layer, $\frac{\partial L}{\partial y_j}$ of all the nodes should be summed up.

$$\frac{\partial L}{\partial w_{ij}^{N-1}} = \sum_k \frac{\partial L}{\partial y_k} \cdot \frac{\partial y_k}{\partial v_k} \cdot \frac{\partial v_k}{\partial y_j} \cdot \frac{\partial y_j}{\partial v_j} \cdot x_i \tag{6-9}$$

$$= \sum_k \frac{\partial L}{\partial y_k} \cdot \frac{\partial y_k}{\partial v_k} \cdot w_{jk}^{N} \cdot \frac{\partial y_j}{\partial v_j} \cdot x_i \tag{6-10}$$

The derivatives of the layers close to the input layer are related to more nodes. Fig. 6.13 shows the concept



F.13 Backpropagation algorithm

of backpropagation. Backpropagation starts from the output layer and sequentially calculates $\frac{\partial L}{\partial w_{ij}^n}$ of each layer saving calculated variables. Backpropagation of convolution layers is conducted by the same procedure. The calculation cost of backpropagation of convolution layers is large because convolution layers are usually close to the input layer. However, calculation cost is much reduced because the parameters of convolution filters are in common for all the feature maps of the same layer.

The optimization algorithm of 3D-CNN is the same as 2D-CNN. The derivative $\frac{\partial L}{\partial w_{ijk}^N}$ of the final Nth layer is written by the output of each node $y_{ijk} = f(v_{ijk})$, and intermediate variable $v_{ijk} = \sum w_{ijk}^N x_{ijk}$.

$$\frac{\partial L}{\partial w_{ijk}^N} = \frac{\partial L}{\partial y_{ijk}} \cdot \frac{\partial y_{ijk}}{\partial v_{ijk}} \cdot x_{ijk} \tag{6-11}$$

The derivative $\frac{\partial L}{\partial w_{ijk}^{N-1}}$ of $N-1$th layer is written below.

$$\frac{\partial L}{\partial w_{ijk}^{N-1}} = \sum_l \sum_m \sum_n \frac{\partial L}{\partial y_{l,m,n}} \cdot \frac{\partial y_{l,m,n}}{\partial v_{l,m,n}} \cdot w_{ijk}^N \cdot \frac{\partial y_{ijk}}{\partial v_{ijk}} \cdot x_{ijk} \tag{6-12}$$

From Eq. (6-12), the receptive field of 3D-CNN is 3D. Therefore, the number of the nodes related to the node of the input layer rapidly increases. However, also in this case, because the parameters of convolution filters are in common, calculation cost is much reduced.

Fig. 6.14 shows the concept of gradient descent method. Initial values of w_{ij}^n in Eq. (6-6) are randomly provided. In steepest descent method, the average of the derivatives for all the input data is calculated. In stochastic descent method (online training), one randomly extracted data is utilized. As an intermediate method of the two methods, the parameters are updated by data sets (batches) randomly extracted from the whole data, which is called minibatch training. In minibatch training, training was finished when all the batches are used for training. Then, training restarts from the first minibatch newly extracted from the

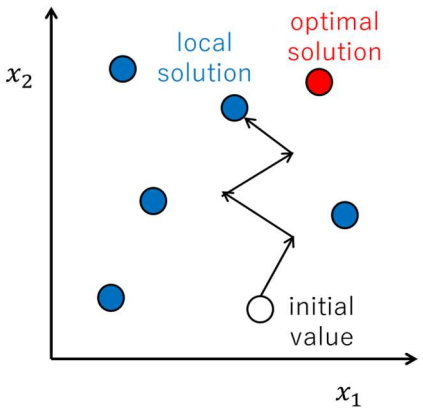


Fig. 6.14 Gradient descent method

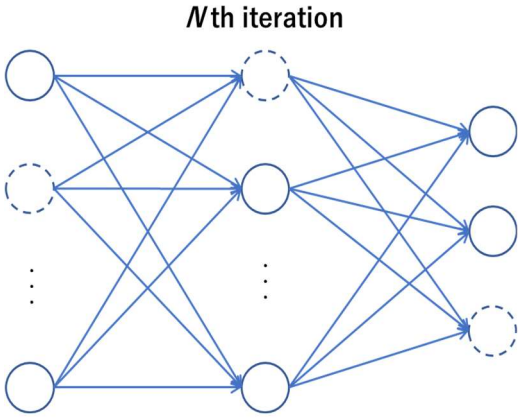
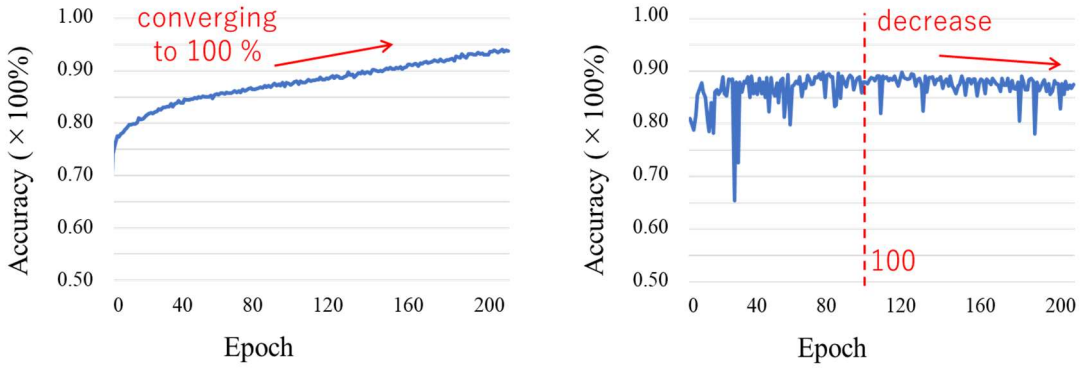


Fig. 6.15 Dropout

training data. The number of data in one batch is called a batch size. The number of completed trainings (completed batch) is called an iteration number. The number of completed trainings for all the batches is called an epoch number. In steepest descent method, because the number of parameters w_{ij}^{n} is enormous, results tend to fall in local solutions. Furthermore, the number of data used for one update is large, resulting in large calculation cost. In stochastic descent method, randomly extracting training data prevents results from falling in local solutions. Calculation cost is drastically reduced. The disadvantage is, it converges slowly. Results may oscillate around the optimal solutions. Therefore, training can be accelerated by adjusting the batch size of minibatch training. Results may be improved. In the research, minibatch training was adopted. Learning rate also affects training. The effects of batch size and learning rate on training time and classification accuracy are evaluated in the following section.

Other important techniques are, pretraining, dropout, transfer learning and visualization of learned features. Pretraining is obtaining appropriate initial values of parameters by unsupervised learning such as autoencoder. Because initial values are already close to the optimal one, it converges rapidly. Furthermore, pretraining prevents results from falling in local solutions. Failure of training such as the vanishing/exploding gradient problem less likely occurs. On the other hand, the training time for pretraining is needed. Sometimes obtained initial values are not appropriate. The problems of local solutions and failure of training are also avoided by minibatch training and dropout. Therefore, pretraining was not adopted in the research. Dropout deactivates some nodes randomly at each iteration to train only the other activated nodes (Fig. 6.15). Therefore, a model to be trained is different at each iteration. A parameter is the activation probability p for each node. Dropout prevents overfitting [150]. Dropout is considered to be similar to ensemble learning, where the combination of multiple models is utilized to increase performance. In the research, the effect of dropout was evaluated by adding dropout to convolution and fully connected layers in Appendix D. Transfer learning is conducted by introducing a conventional NN model architecture and learned parameters. The effect of transfer learning was discussed in Section 6.5.2 by 2D-CNN. Transfer learning of 3D-CNN is not possible because there is no commercial model. The visualization of learned features was discussed in Appendix E. Visualization may



(a) (b)

Fig. 6.16 Training curves (a) learning curve (b) validation curve

be helpful for discussing the characteristics of reflection patterns. However, visualization method is discussed in few research and not the target of this research [146].

6.3 Training method

In this section, the effect of training method on training time and training and test accuracy was discussed. All the results shown in this section are a transverse pipe case. Fig. 6.16 shows the examples of learning and validation curves. Training time depends on data volume, number of layers and sizes of filters. From Fig. 6.16, the learning curve monotonically increased in 200 epochs while the validation curve converged over 50 epochs. Generally, a learning curve converges to 100 % training accuracy. On the other hand, a validation curve reaches the maximum test accuracy at certain epochs. The validation curve slightly decreased after 50 epochs because of overfitting. Therefore, there is an optimal epoch number to stop training. The optimal epoch number and test accuracy at that epoch was discussed hereafter showing validation curves.

Fig. 6.17 shows the effect of a batch size S_b on convergence. Fig. 6.16 (b) corresponds to the case $S_b = 32$ and learning rate $lr = 0.01$. Larger S_b was, more rapidly a curve converged. At $S_b = 8$, the test

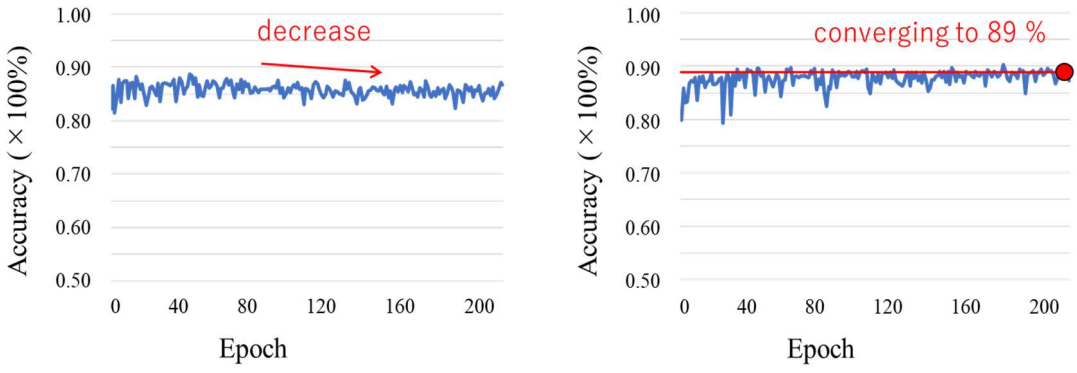


Fig. 6.17 Effect of batch size (a) $S_b = 8$ (b) $S_b = 128$

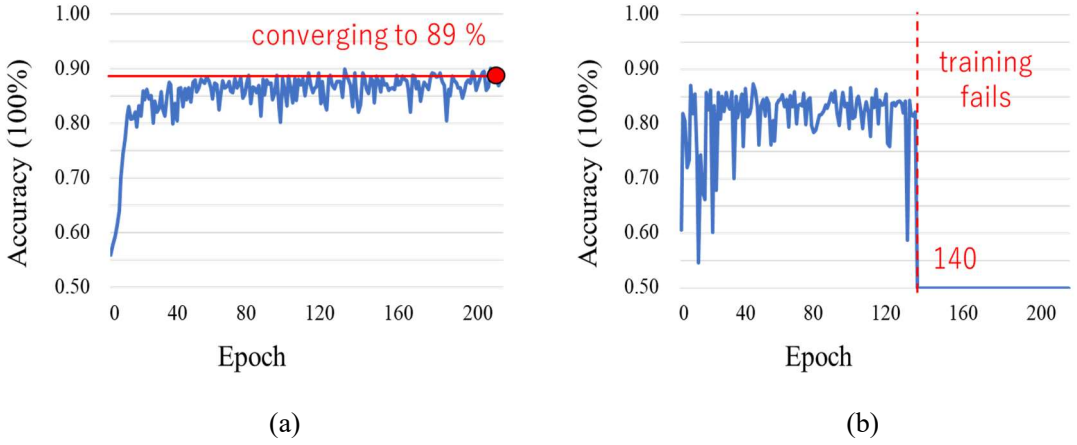


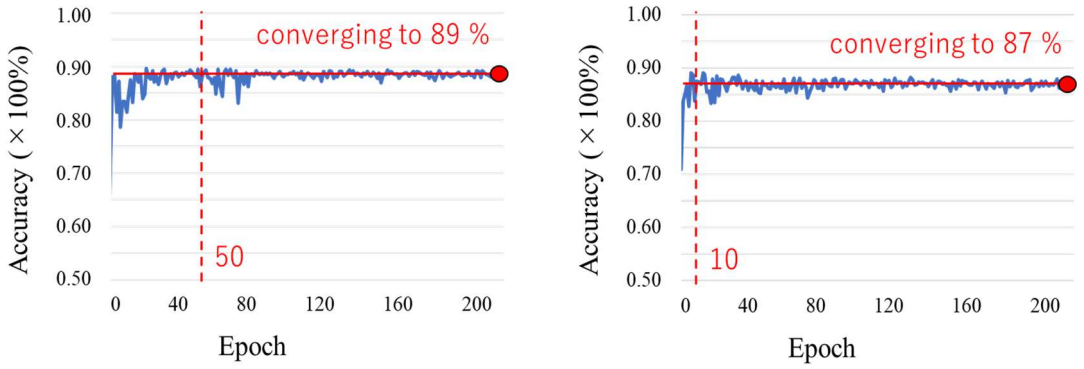
Fig. 6.18 Effect of learning rate (a) $lr = 0.001$ (b) $lr = 0.1$

accuracy reached the maximum at the first epoch and decreased from Fig. 6.17 (a). The test accuracy at the first epoch was lower than the maximum test accuracy 89 % at $S_b = 32$, suggesting the curve may converge to a local solution. At $S_b = 126$, the curve asymptotically reached the maximum test accuracy in 200 epochs from Fig. 6.17 (b). Therefore, in the research, $S_b = 32$ was adopted. Fig. 6.18 shows the effect of lr on convergence at $S_b = 32$. Compared to Fig. 6.16 (b) at $lr = 0.01$, the case $lr = 0.001$ required 200 epochs to converge from Fig. 6.18 (a). On the other hand, at $lr = 0.1$, the oscillation of the curve was large from Fig. 6.18 (b). Furthermore, the training failed after 140 epochs. Because of too large learning rate, the vanishing/exploding gradient problem may occur.

To converge to high test accuracy with less epoch number, the weight decay of learning rate may be adopted. A decay factor is d. An iteration number is N_d. lr at each epoch is updated below.

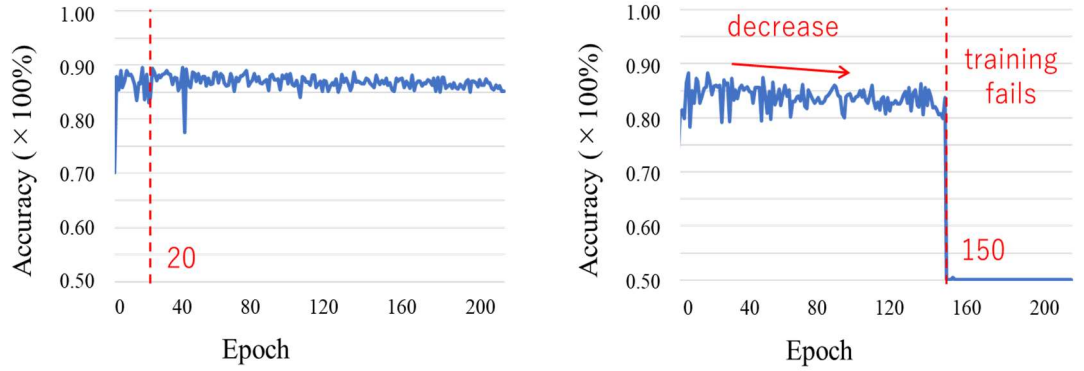
$$lr \leftarrow lr \cdot \frac{1}{1 + d \cdot N_d} \tag{6-13}$$

At $S_b = 32$, N_d is about 5000. lr decreases to 3/4 - 2/3 per epoch at $d = 0.001$ and 1/4 - 1/5 per epoch at $d = 0.01$. Fig. 6.19 shows the effect of d at initial learning rate $lri = 0.1$. At $d = 0.001$ the curve



(a) (b)

Fig. 6.19 Effect of weight decay (a) $d = 0.001$ (b) $d = 0.01$



(a) (b)

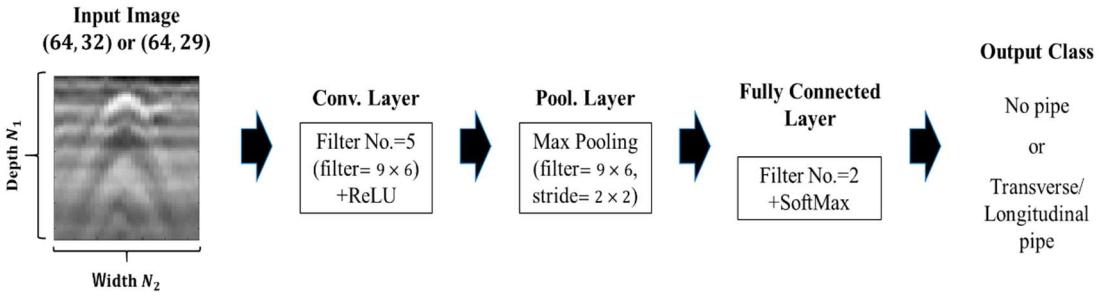
Fig. 6.20 Effect of momentum (a) $\alpha = 0.5$ (b) $\alpha = 0.9$

converged in 50 epochs from Fig. 6.19 (a). Because of the decrease of learning rate, the oscillation of the curve was small, smoothly converging to the maximum test accuracy. On the other hand, at $d = 0.01$, because of the rapid decrease of learning rate, the curve converged in 10 epochs from Fig. 6.19 (b). Furthermore, the converged accuracy was about 87 %, 2 % smaller than the case of Fig. 6.19 (a) possibly because it fell in a local solution. Therefore, in the research, weight decay with $lri = 0.1$ and $d = 0.001$ was adopted.

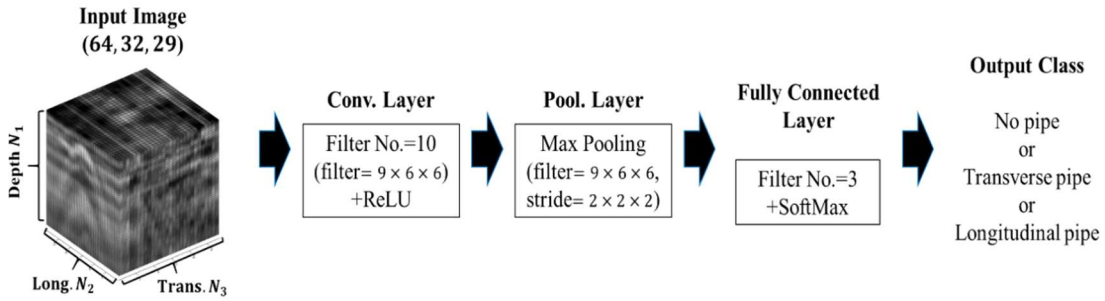
To accelerate training and suppress the oscillation of a validation curve, momentum may be introduced to the update of parameters. A factor of momentum is α. The updates of the parameters are Δw_{ij}^n. Eq. (6-6) is rewritten as below.

$$w_{ij}^n \leftarrow w_{ij}^n - \eta \frac{\partial L}{\partial w_{ij}^n} + \alpha \Delta w_{ij}^n \tag{6-14}$$

From Eq. (6-14), depending on α, the update of w_{ij}^n is suppressed by the effect of a momentum term $\alpha \Delta w_{ij}^n$. Fig. 6.20 shows the effect of α at $lr = 0.01$. At $\alpha = 0.5$, the oscillation of the validation curve was reduced from Fig. 6.20 (a) compared to Fig. 6.16 (b). Furthermore, the curve converged more rapidly in 20 epochs. At $\alpha = 0.9$, the maximum test accuracy was obtained at several epochs and lower than the case shown in Fig. 6.16 (b). The failure of training occurred after 150 epochs. In the research, the combination of weight decay and momentum $\alpha = 0.5$ was adopted. Training was automatically stopped when the decrease of a loss function was below a threshold, which is called early stopping. The increase of test accuracy was 0.1 % order. The training of 2D-CNN finished after 20 - 50 epochs in several to several tens minutes.



(a)



(b)

Fig. 6.21 Optimized models (a) 2D-CNN (b) 3D-CNN

In 3D-CNN, the same parameters, a batch size $S_b = 32$, initial learning rate $lri = 0.1$, weight decay $d = 0.001$ and momentum $\alpha = 0.5$ were adopted. Early Stopping was also adopted. In all the cases, the training finished after 10 - 20 epochs. The volume of input data did not change. The training time of 3D-CNN for one epoch was expected to be 5 - 10 times larger than 2D-CNN because the number of parameters of convolution filters and fully connected layers was 5 - 10 times larger owing to the increase of the dimension. However, the training time of one epoch was several times larger than 2D-CNN. The total training time was not much different from 2D-CNN, ten to at most 30 minutes by GPU. The backpropagation algorithm of TensorFlow may be designed to calculate the updates efficiently. Training time was more sensitive to the number of layers, convolution filters and sizes of input data and convolution filters.

6.4 Repeatability of training and number of data

In the following section, the effects of preprocessing filters and model architectures on classification accuracy were discussed. On the other hand, because of the variation of validation curves and no pipe section data sets, classification accuracy varied. Therefore, these effects were evaluated in this section. The convergence of classification accuracy in terms of the number of training data was also discussed. In this section, the effects were evaluated based on the optimized 2D and 3D-CNN models shown in Fig. 6.21. The optimized 2D-CNN model is a 2 layer deep learning model consisted of a 1 convolution filter, 1 pooling and 1 fully connected layers. The output is two categories. The models shown in Fig. 6.21 are relatively simple compared to conventional 2D-CNN models for image classification. The 3D-CNN model shown in Fig. 6.21 (b) was a 3D version of Fig. 6.21 (a). The dimension of input data, convolution and pooling filters of a 2D-CNN model were augmented. The output is each direction pipe and no pipe section for two-category classification and transverse, longitudinal pipes and no pipe section for three-category classification.

Table 6.1 Repeatability of validation accuracy

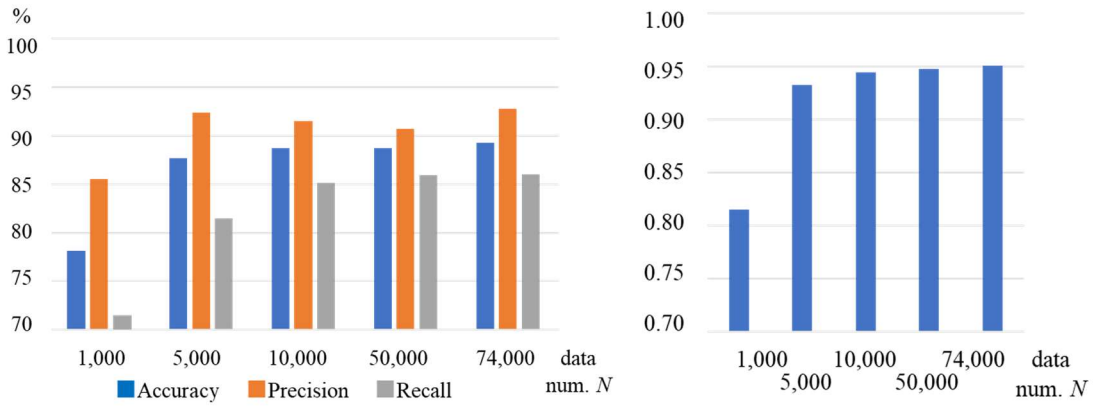
Trial	1	2	3	4	5
Accuracy	89.4	89.5	88.3	89	89.4
Precision	92.9	91.5	93.1	91.2	92.1
Recall	84.5	86.8	82.2	86	86.1
AUC	0.95	0.954	0.942	0.949	0.951

Table 6.2 Variation of validation accuracy of different no pipe section data set

Trial	1	2	3	4	5
Accuracy	89.2	88	88.1	89.7	89.7
Precision	92.9	88.3	92	84	92.5
Recall	84.5	87.4	85.1	84.2	86.1
AUC	0.947	0.95	0.957	0.948	0.955

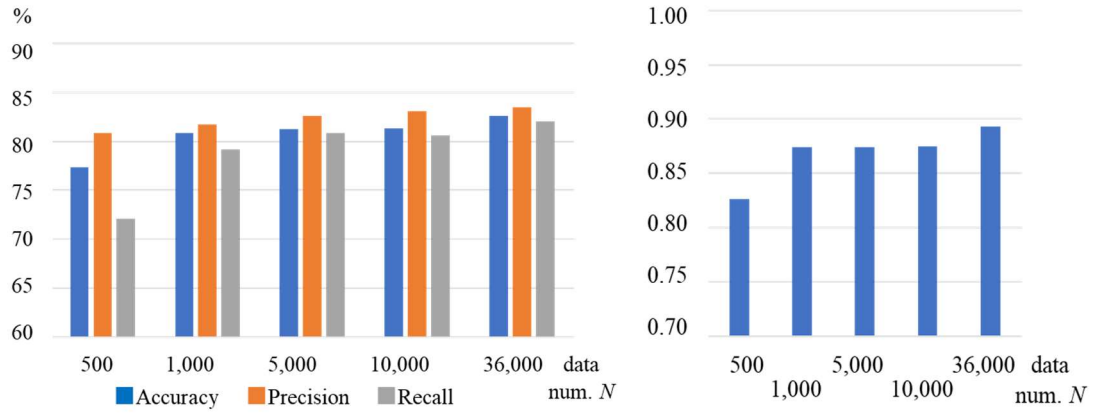
Table 6.1 shows the variation of test accuracy trained by the same data set. Table 6.2 shows the test accuracy of 5 different data sets by newly extracting no pipe section data for each data set. From Table 6.1, the variation of test accuracy was about 1 %. From Table 6.2, the variation of test accuracy was about 2 %. The variation of Table 6.2 was affected by the variation in training process and different data sets. The variation in training process is considered to be 1 % and different data sets is 1 %. The variation was at most 2 % in terms of classification accuracy and 0.01 in terms of AUC from Table 6.2. Therefore, if the difference of classification accuracy is over 2 % and AUC is over 0.1, the difference is significant.

In terms of 2D-CNN, Fig. 6.22 and Fig. 6.23 show the effect of the number of training data N on the classification accuracy of transverse and longitudinal pipes. From Fig. 6.22, In terms of transverse pipes, classification accuracy was about 78 % at $N = 1{,}000$. At $N = 10{,}000$, classification accuracy converged to about 87 %, by about 9 % increase. From Fig. 6.23, in terms of longitudinal pipes, classification accuracy



(a) (b)

Fig. 6.22 Effect of number of data (2D-CNN, transverse pipes) (a) Accuracy, Precision, Recall (b) AUC

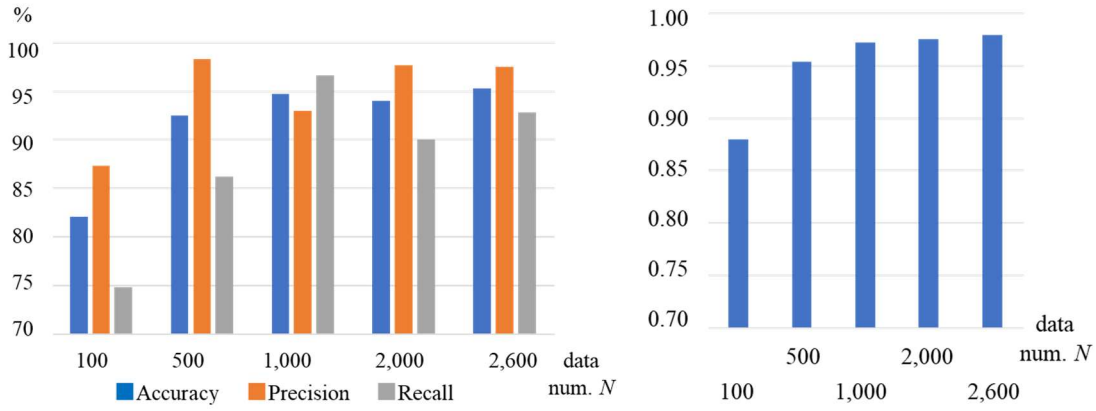


(a) (b)

Fig. 6.23 Effect of number of data (2D-CNN, longitudinal pipes) (a) Accuracy, Precision, Recall (b) AUC

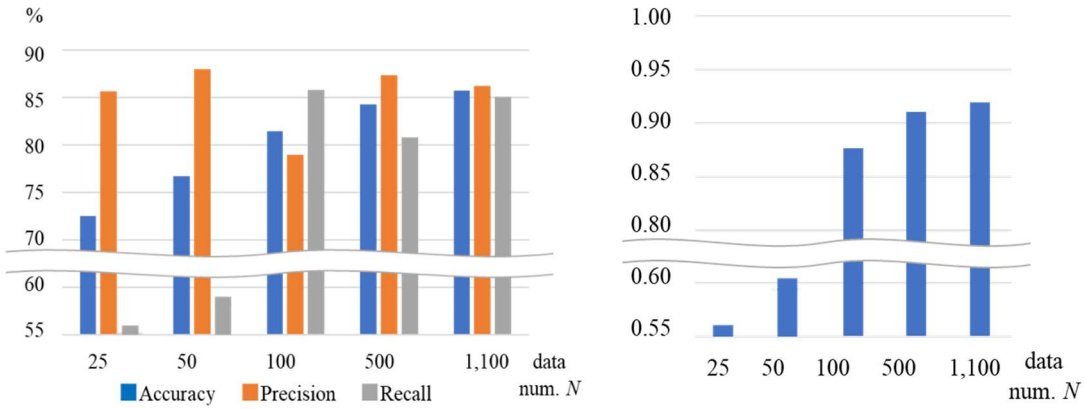
was about 77 % at $N = 500$. Classification accuracy was about 83 % at $N = 3{,}6000$, by about 6 % increase. In the research, the number of transverse pipe data was about $74{,}000$, which was considered to be enough. The number of longitudinal pipe data was about $36{,}000$. The classification accuracy of longitudinal pipes may further increase by several % increasing the number of data. In the research, the classification accuracy of the models was compared using the same training data. Therefore, the number of the data will not affect the discussions of this chapter. Conventional CNN models were over several tens layers. There is a possibility that the number of data may not be enough for conventional CNN models. Therefore, the effect of fine tuning and transfer learning was evaluated in Section 6.5.1.

In terms of 3D-CNN, Fig. 6.24 and Fig. 6.25 show the effect of N on the classification accuracy of transverse and longitudinal pipes. From Fig. 6.24 and Fig. 6.25, at $N = 100$ and $N = 25$, because the numbers of two direction pipes data are not enough, classification accuracy was low, about 82 % and 72 %.



(a) (b)

Fig. 6.24 Effect of number of data (3D-CNN, transverse pipes) (a) Accuracy, Precision, Recall (b) AUC



(a) (b)

Fig. 6.25 Effect of number of data (3D-CNN, longitudinal pipes) (a) Accuracy, Precision, Recall (b) AUC

Over $N = 1000$ the classification accuracy of transverse pipes converged to about 92 %, by about 10 % increase. The classification accuracy of longitudinal pipes converged to about 85 % over $N = 500$, by about 13 % increase. The number of transverse pipe data about $2{,}600$ was considered to be enough. The number of longitudinal pipe data was about $1{,}100$. The classification accuracy of longitudinal pipes may further increase by several % also in this case. In terms of 3D-CNN compared to 2D-CNN, the total volume of training data was not changed. The number of training data was reduced to 1/29 (the reciprocal of the number of channels) by increasing the dimension. The number of parameters of convolution filters was about 5 times larger than 2D-CNN. Therefore, the required number of training data may be about 1/6 smaller than 2D-CNN, which was consistent with the results of Fig. 6.24 and Fig. 6.25.

6.5 Optimization of CNN

6.5.1 Preprocessing filters

In Section 3.4, common preprocessing filters in GPR method practice such as a gain filter, BGR were introduced. In Section 4.4.2, the classification accuracy of SVM was improved by applying a Laplacian filter for edge enhancement method. Therefore, in this section, the effect of preprocessing filters on classification accuracy was evaluated. Fig. 6.26 shows the images after preprocessing filters. From 6.26 (c), BGR improved

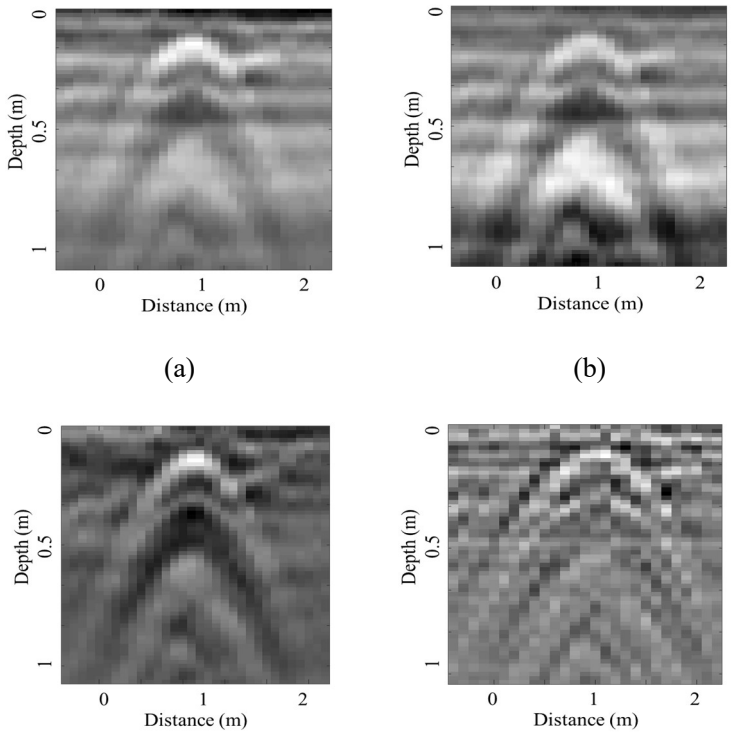
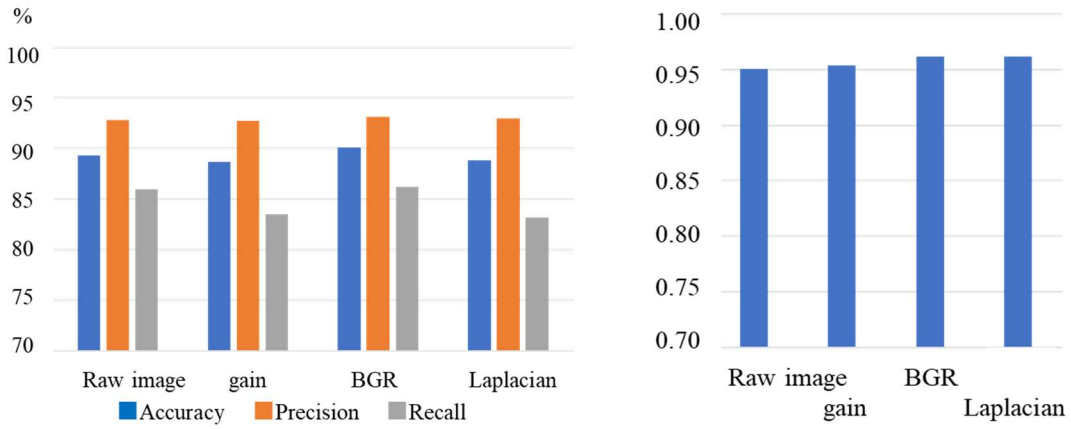


Fig. 6.26 Comparison of preprocessing filters (a) raw image (b) gain function (c) BGR (d) Laplacian filter

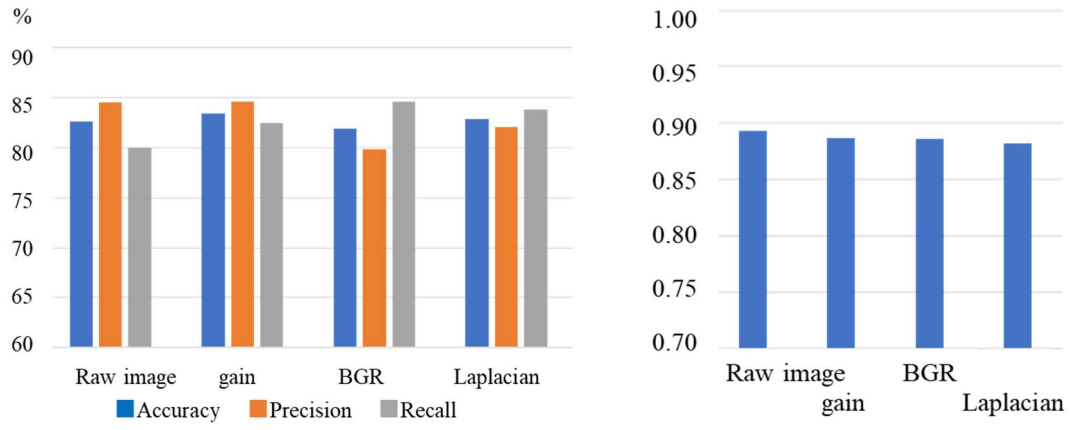
the quality of the image. However, from Fig. 6.26 (b) (d), the images were not improved by a gain filter and Laplacian filter. The effect of the gain filter was small in extracted images. Edge enhancement filters have the characteristics to amplify high frequency noise, which may not be favorable in this case in terms of reflection pattern detection from radar data.

From Fig. 6.27 and Fig. 6.28, all the preprocessing filters did not improve the classification accuracy of the two direction pipes. CNN automatically learns spatial features by convolution filters. Therefore, target features can be extracted without preprocessing filters such as BGR. Convolution filters serve as spatial filters to enhance the reflection patterns of pipes and suppress the multiple reflections and other clutters. In the research, preprocessing filters were not adopted. In manual inspection, the parameters of a gain filter and BGR are adjusted for each data and target features. It is favorable to detect features only by CNN models without preprocessing filters.



(a) (b)

Fig. 6.27 Effect of preprocessing filters (transverse pipes) (a) Accuracy, Precision, Recall (b) AUC



(a) (b)

Fig. 6.28 Effect of preprocessing filters (longitudinal pipes) (a) Accuracy, Precision, Recall (b) AUC

6.5.2 Model architecture

By utilizing the learned parameters of conventional CNN models, accuracy may be high with less data. Training can be accelerated. Three training methods are possible: (i) utilizing imported parameters as initial values (fine tuning) (ii) importing learned parameters and fixing layers (transfer learning) (iii) only adopting model architectures without utilizing the learned parameters (from scratch). All the fully connected layers of the latter part of imported models were replaced by a 1 fully connected layer to adjust the number of output categories to target problem. Fig. 6.29 and Fig. 6.30 show the comparison of training methods in terms of the classification accuracy of transverse and longitudinal pipes by a VGG16 model. From Fig. 6.29 and Fig. 6.30, in the both direction pipes, no significant difference was observed. Convolution filters trained by ordinary scenery photo images were also effective in pipe detection from radar images. However, because the number of training data was enough, classification accuracy was not improved by fine tuning or transfer learning. Training time was almost the same because the time to tune fully connected layers was needed.

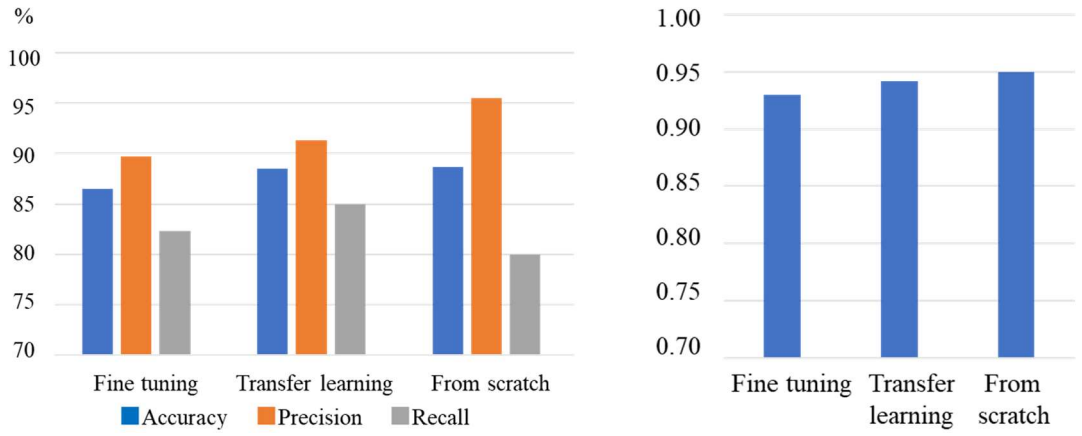


Fig. 6.29 Effect of training method (transverse pipes) (a) Accuracy, Precision, Recall (b) AUC

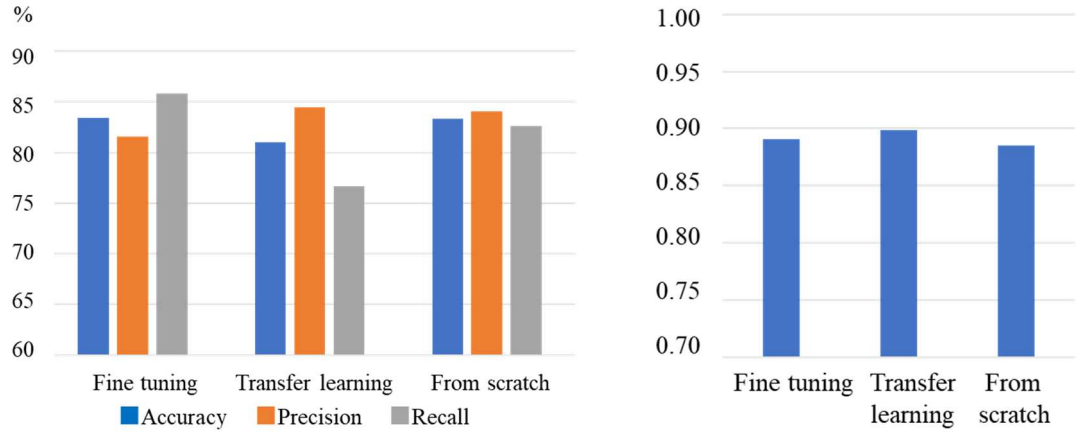
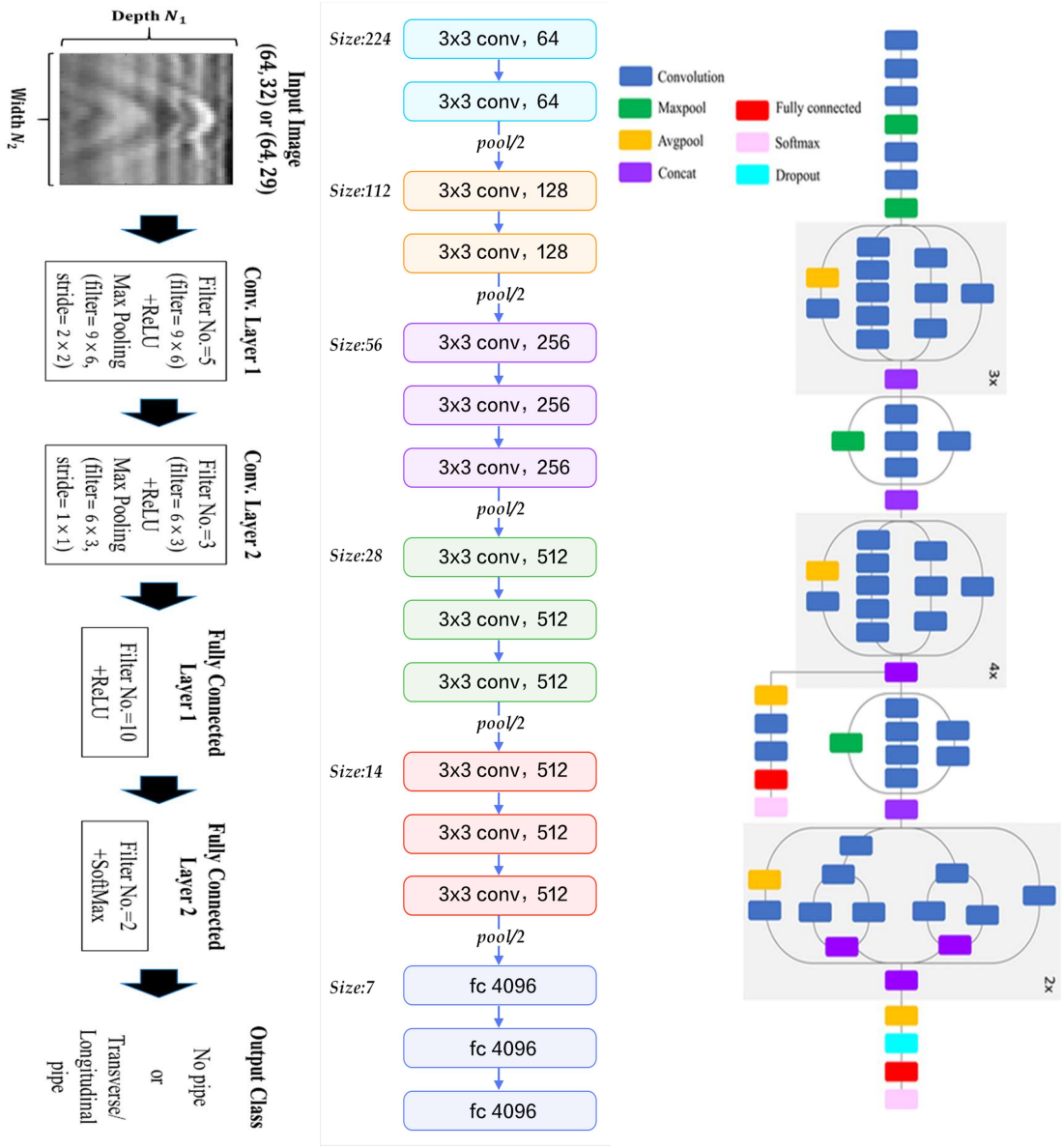


Fig. 6.30 Effect of training method (longitudinal pipes) (a) Accuracy, Precision, Recall (b) AUC

Hereafter, method (i) was adopted for the simplicity of the programs when the conventional CNN models were utilized.

To optimize model architectures, the numbers of model layers and filters are the important parameters. Therefore, classification accuracy was compared. A 2 layer custom net was proposed by the author as shown in Fig. 6.21. A 1 layer model was proposed in previous research. A 1 layer model is a simple NN model produced by removing the convolution layer of the 2 layer model. An SVM model proposed in the previous chapter was also utilized for comparison. After applying a Laplacian filter, HOG features were extracted to input the SVM model with an RBF kernel. Fig. 6.31 shows the details of other CNN models. Fig. 6.31 (a) shows a 4 layer model. The 4 layer (3 hidden layer) custom net was produced by adding a 1 convolution

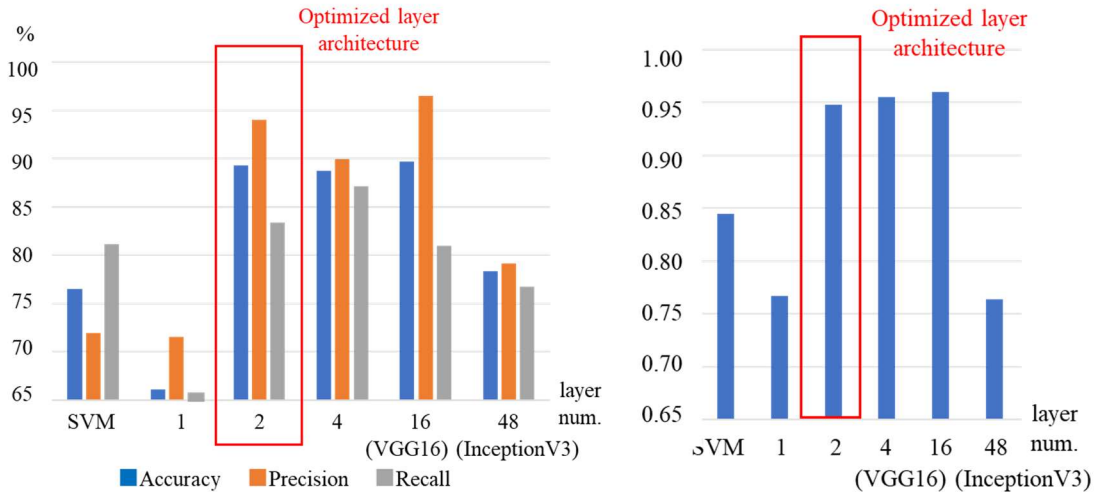


(a) (b) (c)

Fig. 6.31 Utilized models (a) 4 layer custom model (b) VGG16 (c) InceptionV3

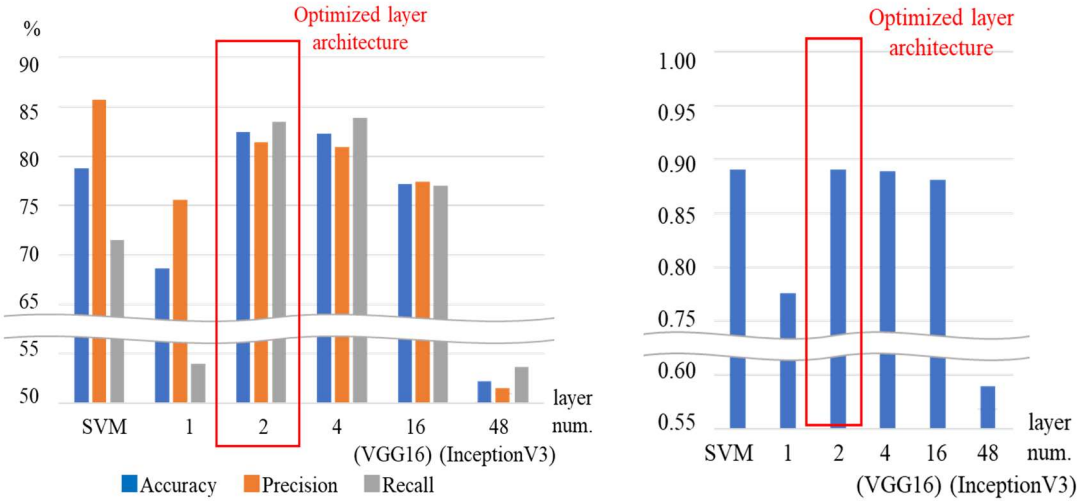
layer after the convolution layer of the 2 layer model and 1 fully connected layer before the fully connected layer. In the research, as a conventional model, VGG16 shown in Fig. 6.31 (b) and InceptionV3 (GoogLeNet) shown in Fig. 6.31 (c) were utilized [151) - 154)]. VGG16 was proposed by Simonyan *et al.* (2012). VGG16 is composed of 13 convolution layers and 3 fully connected layers. By adding many small size convolution filters, the performance of the model is improved by obtaining more abstract features with less training data. InceptionV3 was proposed by Szegedy *et al.* (2014). The model consists of total 48 layers. By utilizing the architecture with different size convolution filters in parallel, the performance of the model was enhanced reducing calculation cost.

Fig. 6.32 and Fig. 6.33 show the effect of the number of layers on the classification accuracy of transverse



(a) (b)

Fig. 6.32 Effect of number of layers (transverse pipes) (a) Accuracy, Precision, Recall (b) AUC



(a) (b)

Fig. 6.33 Effect of number of layers (longitudinal pipes) (a) Accuracy, Precision, Recall (b) AUC

and longitudinal pipes. The highest classification accuracy about 89 % and 82 % was obtained by the 2 layer model in the both cases of transverse and longitudinal pipes. The classification accuracy of the 1 layer model was about 66 % and 68 %, by about 23 % decrease compared to the 2 layer model in the case of transverse pipes and 14 % decrease in the case of longitudinal pipes. The performance of the model was improved by adding a convolutional layer utilizing spatial features of the data. Classification accuracy did not change over 4 layers. Highly abstract features may not be utilized in the detection of pipe reflection patterns from radar images. The classification accuracy of InceptionV3 was decreased by about 20 % - 30 %, possibly because too complicated models cause overfitting.

The performance of the SVM model was higher than the 1 layer model because spatial features were utilized by feature values. However, the classification accuracy of the SVM model was about 76 % and 78 %, by about 13 % decrease compared to the 2 layer model in the case of transverse pipes and 4 % decrease in the case of longitudinal pipes. The performance of the model was improved by deep layer architectures

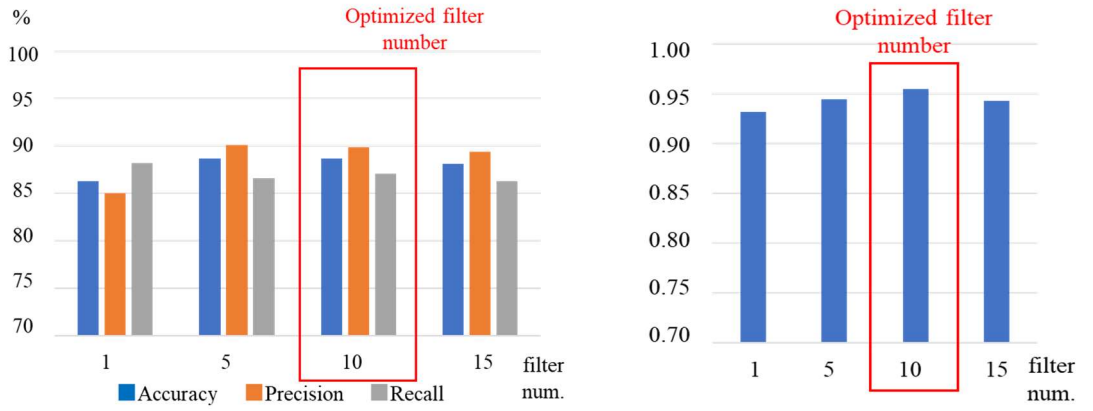


Fig. 6.34 Effect of number of filters (transverse pipes) (a) Accuracy, Precision, Recall (b) AUC

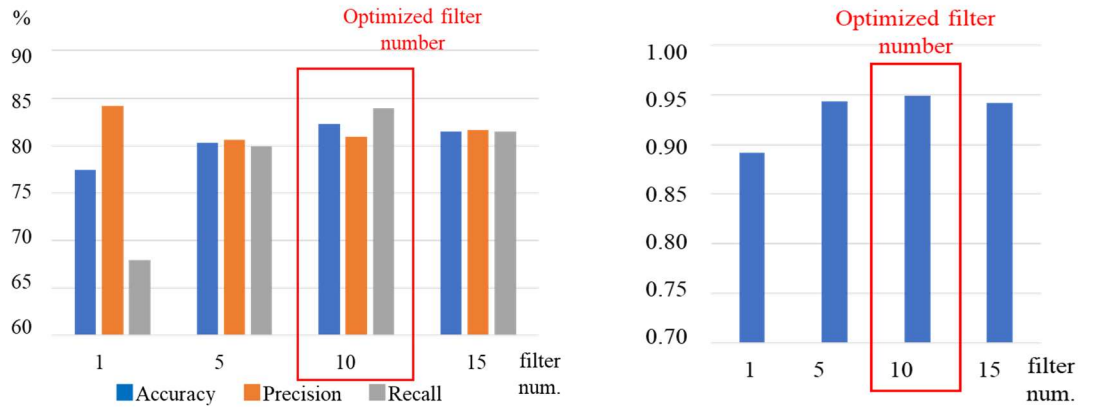


Fig. 6.35 Effect of number of filters (longitudinal pipes) (a) Accuracy, Precision, Recall (b) AUC

utilizing more abstract features. There is previous research about the comparison between SVM and CNN models [155) - 156)]. The abstract features may correspond to the knowledge of skilled inspectors. Generally, the optimal layer architectures depend on the problem. In the research, the 2 layer model was adopted.

Fig. 6.34 and Fig. 6.35 show the effect of the number of convolution filters N_f of the 2 layer model on the classification accuracy of transverse and longitudinal pipes. From Fig. 6.34 and Fig. 6.35, even at $N_f = 1$, classification accuracy was high. It was about 86 % and 77 % possibly because most of the reflection patterns of the pipes of the data may be hyperbolics and similar. The number of utilized features may not be large. Classification accuracy converged to about 89 % and 82 % over $N_f = 5$ and 10. In the research, $N_f = 10$ was adopted.

6.5.3 Sizes of input data and convolution filters

The sizes of input data and convolution filters are expected to affect the results. Larger the sizes of input data and convolution filters are, larger calculation cost becomes. However, smaller the sizes are, lower classification accuracy may be. There may be required sizes to extract target features. The optimal sizes of

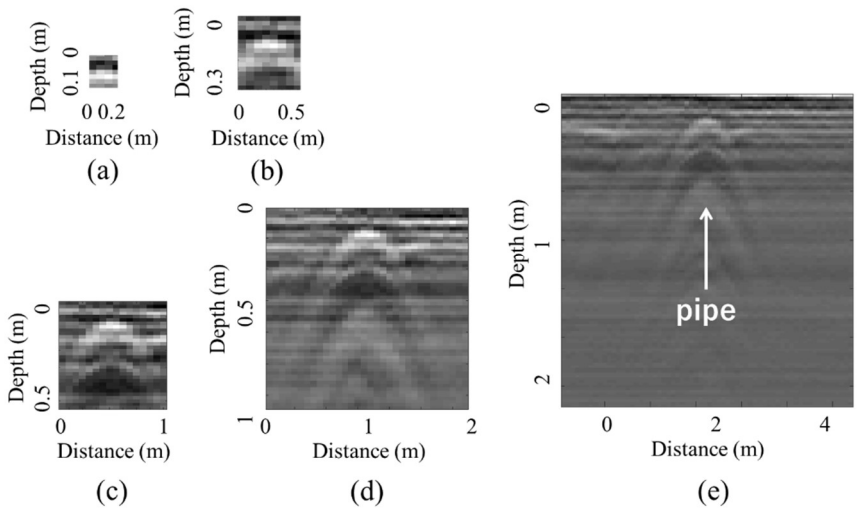


Fig. 6.36 Transverse pipe (a) (8,4) (b) (16,8) (c) (32,16) (d) (64,32) (e) (128,64)

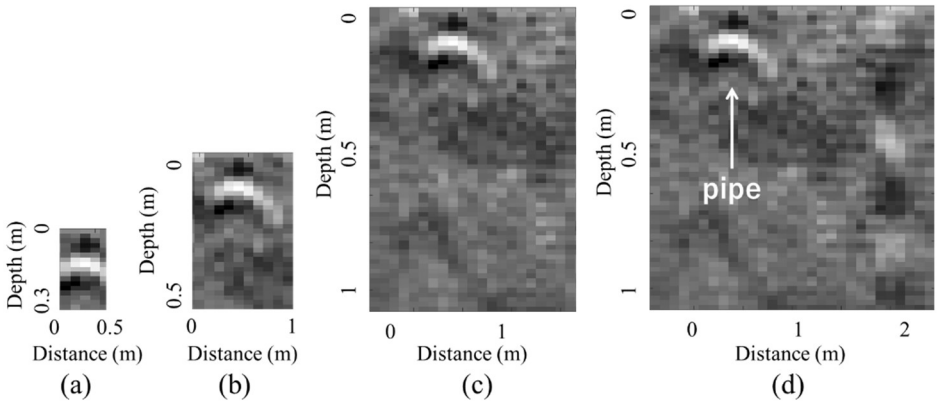
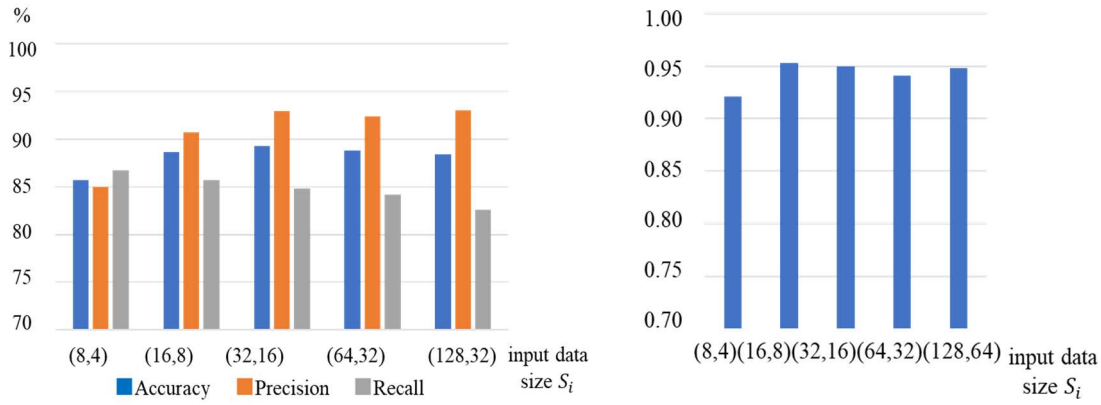


Fig. 6.37 Longitudinal pipe (a) (16,5) (b) (32,13) (c) (64,21) (d) (64,29)

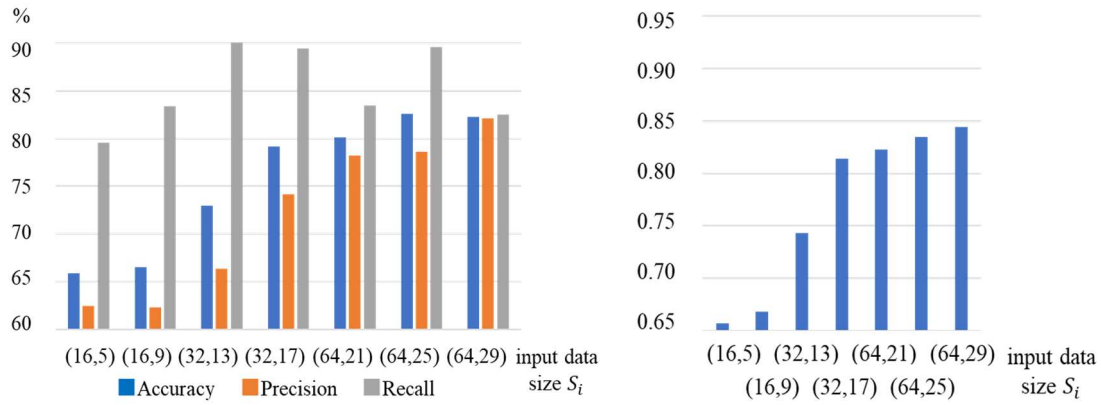
input data and convolution filters depend on the scale of target features. Fig. 6.36 and Fig. 6.37 show the examples of the images of transverse and longitudinal pipes with a certain size of input data $S_i = (x, y)$. x and y represent the pixel numbers in depth and longitudinal or transverse directions. From Fig. 6.36 (a) and Fig. 6.37 (a), certain features may be extracted even by small S_i. However, S_i should be large enough to extract the whole reflection patterns to obtain high classification accuracy. From Fig. 6.36 (d) and Fig. 6.37 (d), too large S_i will not improve the results because the scale of target features was constant. The optimal size of a convolution filter S_f may depend on S_i. Therefore, in the research, S_f was optimized for each S_i. The effect and optimal parameters of S_i and S_f were summarized below.

Fig. 6.38 and Fig. 6.39 show the effect of S_i on the classification accuracy of transverse and longitudinal pipes. In terms of transverse pipes $S_f = (9,6)$ was assumed. In terms of longitudinal pipes, $S_f = (6,4)$ at $S_i = (16,5)$ and in other cases $S_f = (9,6)$ were assumed. The classification accuracy of transverse pipes



(a) (b)

Fig. 6.38 Effect of size of input data (2D-CNN, transverse pipes) (a) Accuracy, Precision, Recall (b) AUC

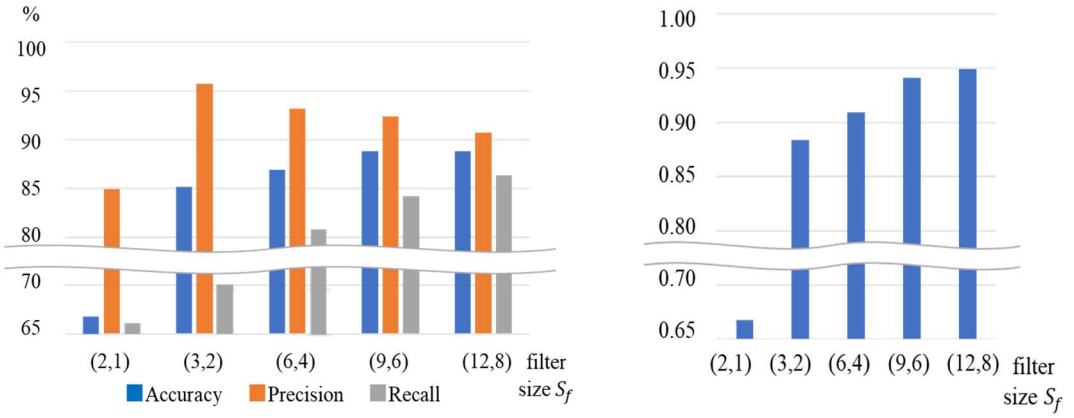


(a) (b)

Fig. 6.39 Effect of size of input data (2D-CNN, longitudinal pipes) (a) Accuracy, Precision, Recall (b) AUC

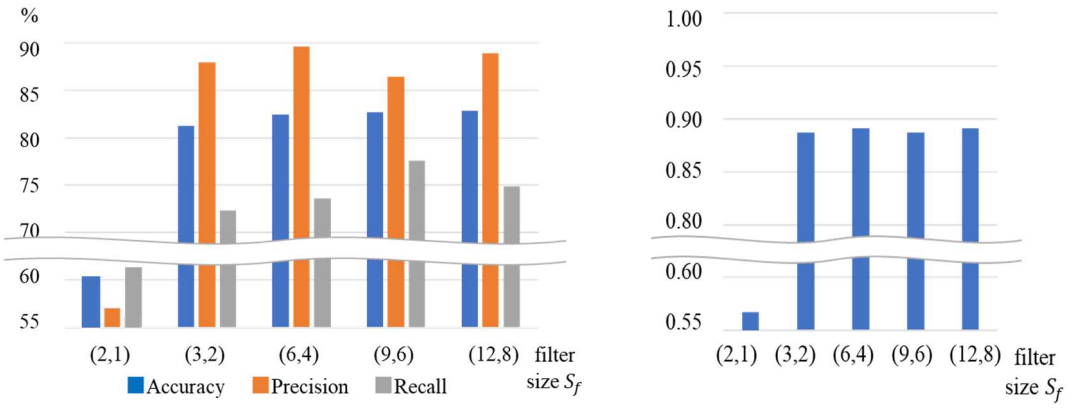
converged to 89 % over $S_i = (16,8)$, by about 3 % increase compared to the case $S_i = (8,4)$. The classification accuracy of longitudinal pipes converged to 82 % over $S_i = (64,25)$, by about 16 % increase compared to the case $S_i = (16,5)$. The optimal sizes depend on the characteristics of the system. The parameters may be different for two direction pipes. Features can be extracted when S_i is over $(16,8)$ in the case of transverse pipes and $(64,25)$ in the case of longitudinal pipes.

Fig. 6.40 and Fig. 6.41 show the effect of S_f on the classification accuracy of transverse and longitudinal pipes. $S_i = (64,32)$ in the case of transverse pipes and $S_i = (64,29)$ in the case of longitudinal pipes were assumed. From Fig. 6.40, in terms of transverse pipes classification accuracy converged to 89 % over $S_f = (9,6)$, by about 4 % increase compared to the case $S_f = (3,2)$. The classification accuracy of longitudinal pipes converged to 82 % over $S_f = (6,4)$ by about 22 % increase compared to the case $S_f = (3,2)$. S_f should be larger than the scale of target features. Features can be extracted when S_f is over $(9,6)$ in the



(a) (b)

Fig. 6.40 Effect of size of filter (2D-CNN, transverse pipes) (a) Accuracy, Precision, Recall (b) AUC



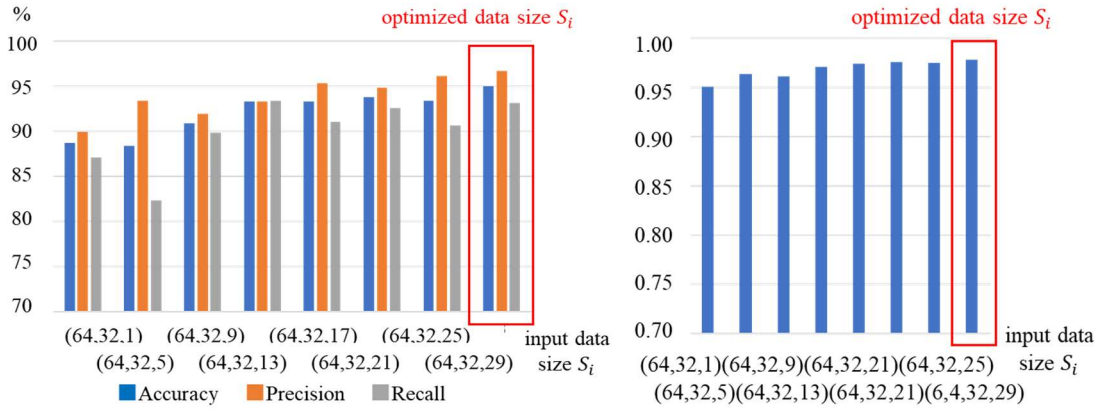
(a) (b)

Fig. 6.41 Effect of size of filter (2D-CNN, longitudinal pipes) (a) Accuracy, Precision, Recall (b) AUC

case of transverse pipes and $(6,4)$ in the case of longitudinal pipes.

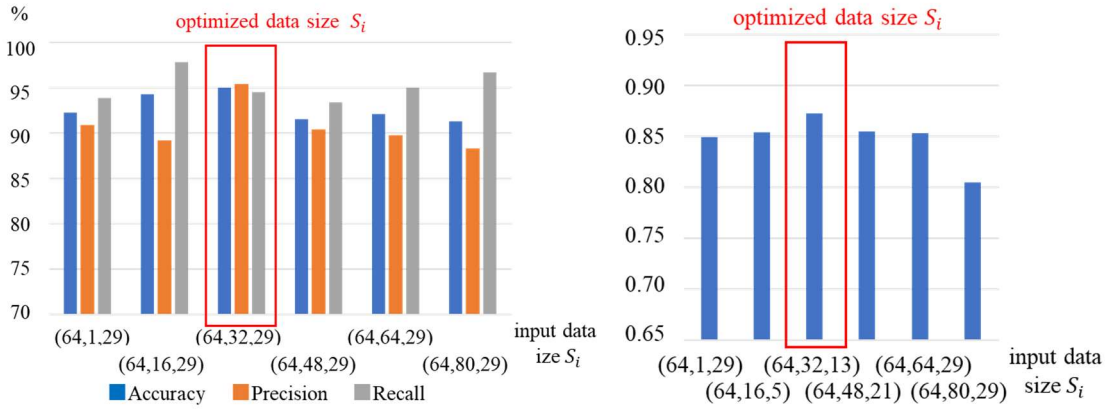
To summarize the results, the smallest S_i and S_f for each direction pipe to obtain the highest classification accuracy was derived. In the research, to conduct three-category classification by 3D-CNN, in the case of transverse pipes $S_i = (64,32)$, $S_f = (9,6)$ and in the case of longitudinal pipes $S_i = (64,29)$, $S_f = (9,6)$ were assumed. Assuming the relative permittivity of soil is 3, S_i corresponds to a 1 m by 2 m region. The classification accuracy of transverse pipes was about 89 % and longitudinal pipes was about 82 % with the optimal S_i and S_f by 2D-CNN.

Also, in the case of 3D-CNN, considering the scale of target features, there may be the optimal sizes of input data and convolution filters. Fig. 6.42 and Fig. 6.43 show the effect of the size of 3D input data $S_i = (x, y, z)$ on the classification accuracy of transverse and longitudinal pipes. x, y and z represents the pixel numbers in depth, scan and channel directions. From Fig. 6.42, in terms of transverse pipes, the



(a) (b)

Fig. 6.42 Effect of size of input data (3D-CNN, transverse pipes) (a) Accuracy, Precision, Recall (b) AUC

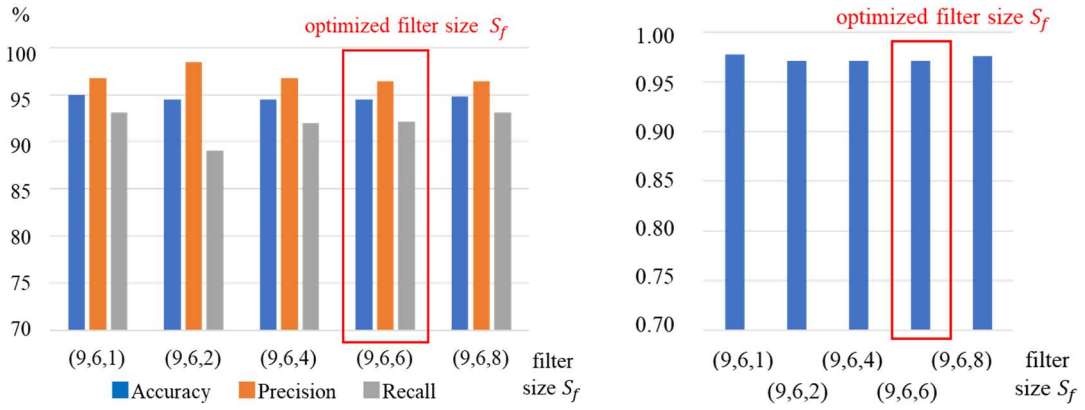


(a) (b)

Fig. 6.43 Effect of size of input data (3D-CNN, longitudinal pipes) (a) Accuracy, Precision, Recall (b) AUC

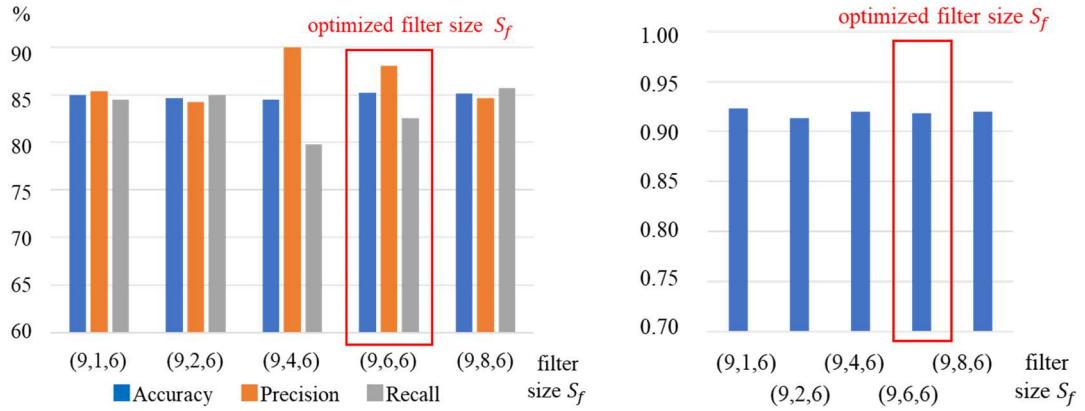
classification accuracy of $S_i = (64,32,29)$ was highest. It was about 95 %, by about 6 % increase by utilizing 3D features compared to the case $S_i = (64,32,1)$ of 2D-CNN. From Fig. 6.43, in terms of longitudinal pipes, classification accuracy was about 85 % at $S_i = (64,32,29)$, by about 3 % increase compared to the case $S_i = (64,1,29)$. However, classification accuracy decreased by about 4 % at $S_i = (64,80,29)$ compared to the case $S_i = (64,32,29)$. From the discussions of Section 2.3.2, the lengths of most of longitudinal pipes are under 0.5 m. Furthermore, the maximum inclination angle is over 20°, which corresponds to 30 pixel shift in 80 pixel scan section. Therefore, the pipes are out of 3D regions, resulting in the decrease of classification accuracy. Considering the optimal size in a channel direction derived by transverse pipes and scan direction by longitudinal pipes, $S_i = (64,32,29)$ is optimal.

Fig. 6.44 and Fig. 6.45 show the effect of the size of a filter S_f on the classification accuracy of transverse



(a) (b)

Fig. 6.44 Effect of size of filter (3D-CNN, transverse pipes) (a) Accuracy, Precision, Recall (b) AUC



(a) (b)

Fig. 6.45 Effect of size of filter (3D-CNN, longitudinal pipes) (a) Accuracy, Precision, Recall (b) AUC

and longitudinal pipes. From Fig. 6.44 and Fig. 6.45, the classification accuracy of the both directions pipes did not change irrespective of S_f. $S_f = (9,6,1)$ and $(9,1,6)$ correspond to the 2D convolution filters. In fully connected layers, the continuity of reflection patterns may be considred. To detect each direction pipe, 3D convolution filters may not be needed. However, to estimate the direction of pipes in three-category classification and detect void, 3D filters are necessary.

Summarizing the results, the optimal size in a channel direction was derived from the case of transverse pipes and scan direction from the case of longitudinal pipes. In the research, to conduct three-category classification, $S_i = (64,32,29)$ and $S_f = (9,6,6)$ were adopted for the both direction pipes. The classification accuracy of transverse pipes by the optimized 3D-CNN model was about 95 % and longitudinal pipes was about 85 %, by 3 % - 6 % increase compared to the optimized 2D-CNN model. $S_i = (64,32,29)$ corresponds to a 1 m by 2 m by 2 m 3D region assuming the relative permittivity of soil is 3.

Fig. 6.46 shows the examples of detected pipes. Clear hyperbolic reflection patterns were detectable even by 2D-CNN as shown in Fig. 6.46 (a), (b). By 3D-CNN utilizing the continuity of reflection patterns, weak and distorted reflection patterns can be detected as shown in Fig. 6.46 (c), (d) which were not detected by 2D-CNN. Pipes cannot be detected even by 3D-CNN when the reflection patterns are too weak and covered by noise as shown in Fig. 6.46 (e), (f).

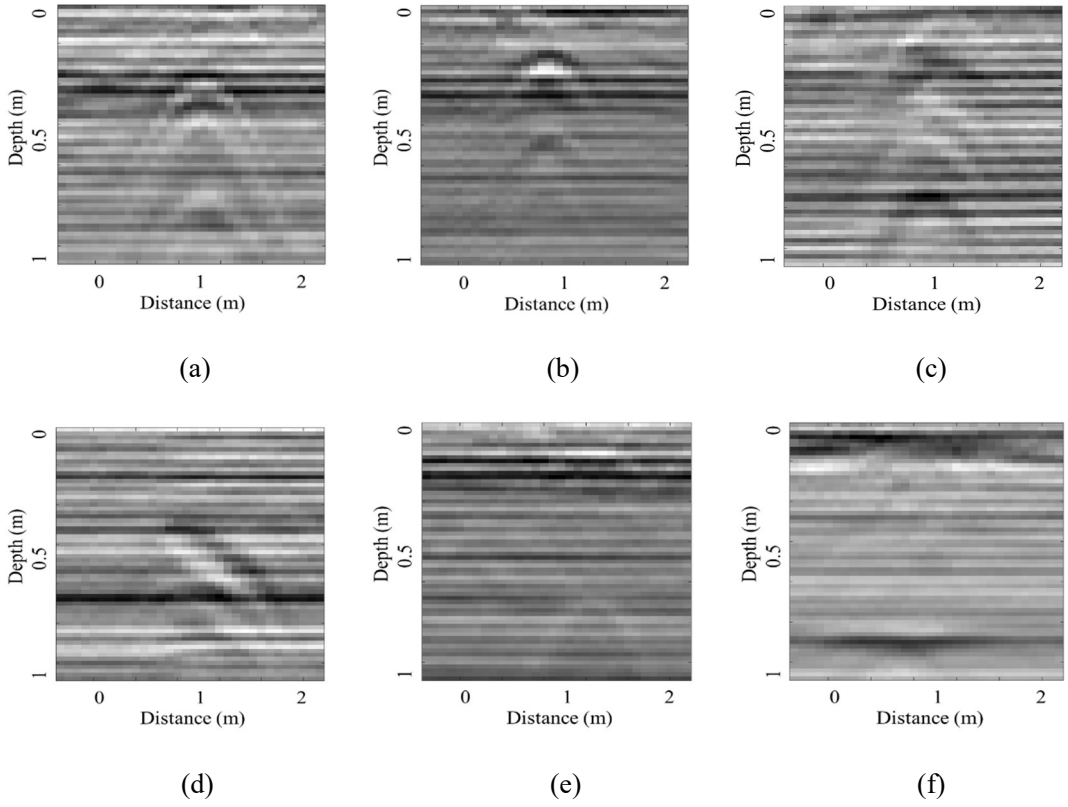
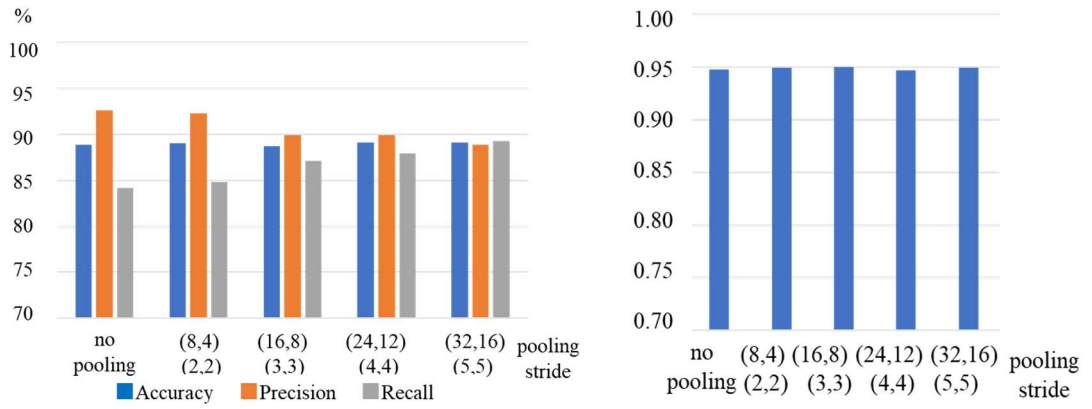


Fig. 6.46 Examples of detection (a) (b) detected by 2D-CNN (c) (d) only detected by 3D-CNN (e) (f) not detected by 2D and 3D-CNN

6.5.4 Pooling layer

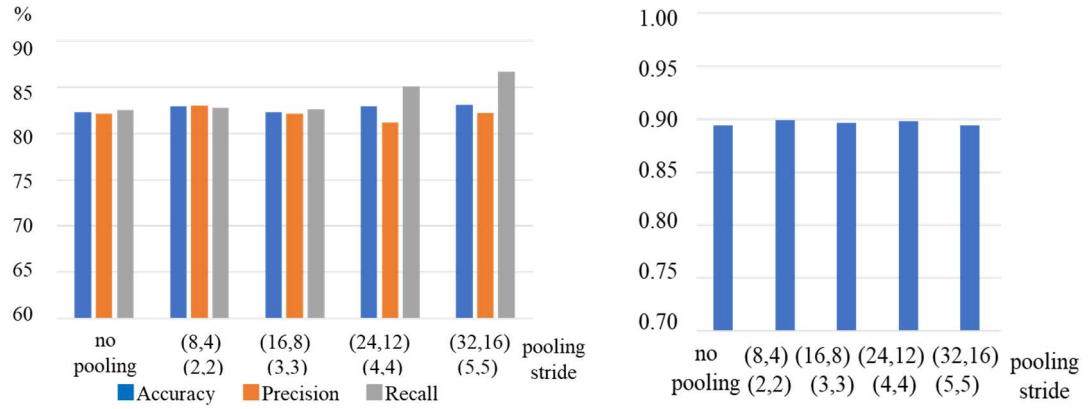
The number of feature maps of a CNN model rapidly increases by adding convolution layers. Therefore, by utilizing pooling layers, calculation cost is drastically reduced. In the research, Max Pooling was adopted. Smaller the size of pooling filters and strides are, larger calculation cost is. However, important features of images may be lost by too large filters and strides. In the research, a pooling layer was added to the convolution layer of Fig. 6.21. The effect of the size of a pooling filter and stride on classification accuracy was evaluated.

Fig. 6.47 and Fig. 6.48 show the effects of the size of a pooling filter S_f and stride S_s on the classification accuracy of transverse and longitudinal pipes. The classification accuracy of the both directions pipes did not change irrespective of S_f and S_s. In the case $S_f = (32,16)$ and $S_s = (5,5)$, the number of parameters of a fully connected layer was reduced to about 1/100. The training time of 1 epoch was reduced to about



(a) (b)

Fig. 6.47 Effect of pooling layer (2D-CNN, transverse pipes) (a) Accuracy, Precision, Recall (b) AUC

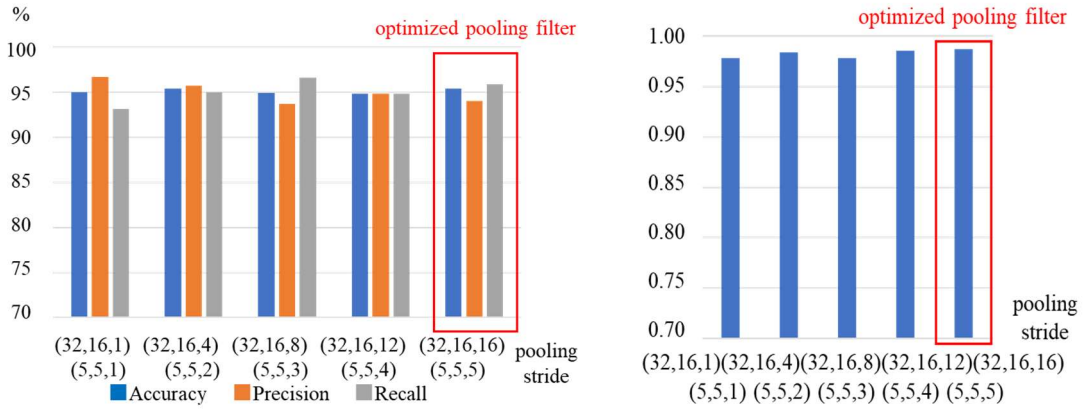


(a) (b)

Fig. 6.48 Effect of pooling layer (2D-CNN, longitudinal pipes) (a) Accuracy, Precision, Recall (b) AUC

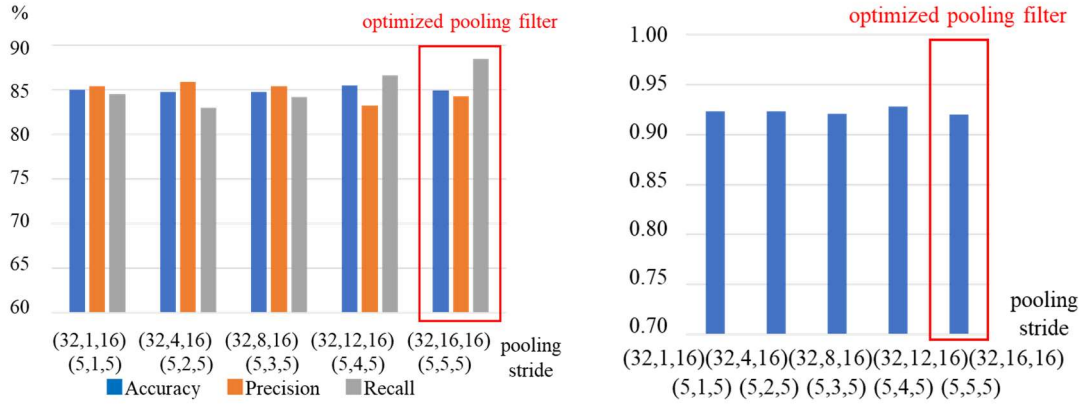
1/5 - 1/10. $S_f = (32,16)$ and $S_s = (5,5)$ were considered to be the possible largest parameters. Even after applying a pooling layer with a large size of filter size and stride, target features can be extracted. Therefore, in 2D-CNN $S_f = (32,16)$ and $S_s = (5,5)$ were adopted.

Fig. 6.49 and Fig. 6.50 show the effects of S_f and S_s on the classification accuracy of 3D-CNN in terms of transverse and longitudinal pipes. The classification accuracy of the both direction pipes did not change irrespective of S_f and S_s also in this case. In the case $S_f = (32,16,16)$ and $S_s = (5,5,5)$, the number of parameters of a fully connected layer was reduced to about 1/1000. The training time of 1 epoch was reduced to 1/10 - 1/20. The reduction rate of 3D-CNN was higher than 2D-CNN. The pooling layer works efficiently in the research especially in a 3D-CNN case. The parameters $S_f = (32,16,16)$ and $S_s = (5,5,5)$ were adopted.



(a) (b)

Fig. 6.49 Effect of pooling layer (3D-CNN, transverse pipes) (a) Accuracy, Precision, Recall (b) AUC

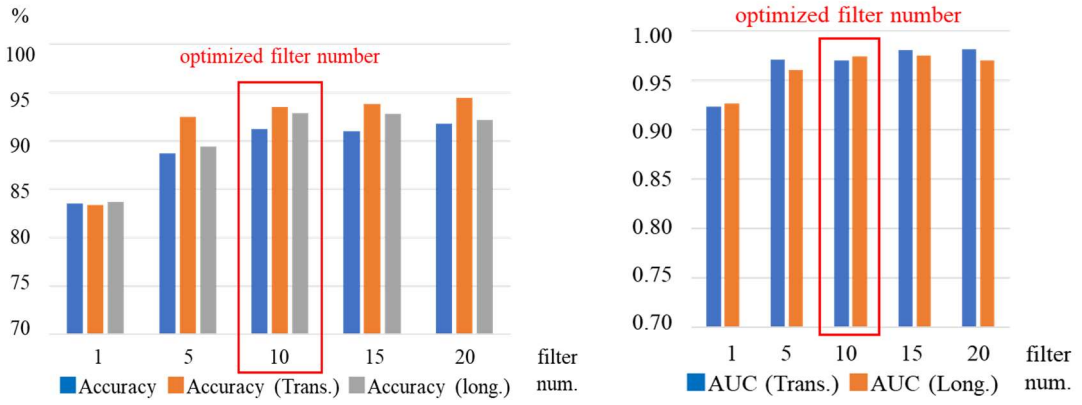


(a) (b)

Fig. 6.50 Effect of pooling layer (3D-CNN, longitudinal pipes) (a) Accuracy, Precision, Recall (b) AUC

6.6 Three-category classification by 3D-CNN

To classify transverse, longitudinal pipes and no pipe section, a three-category classification model by 3D-CNN was developed. The 3D-CNN model for each direction pipe was developed in the previous section. The optimal sizes of input data, convolution filters and pooling layers of two direction pipes were in common.



(a) (b)

Fig. 6.51 Effect of number of convolution filters on accuracy and AUC (three-category classification) (a) Accuracy (b) AUC

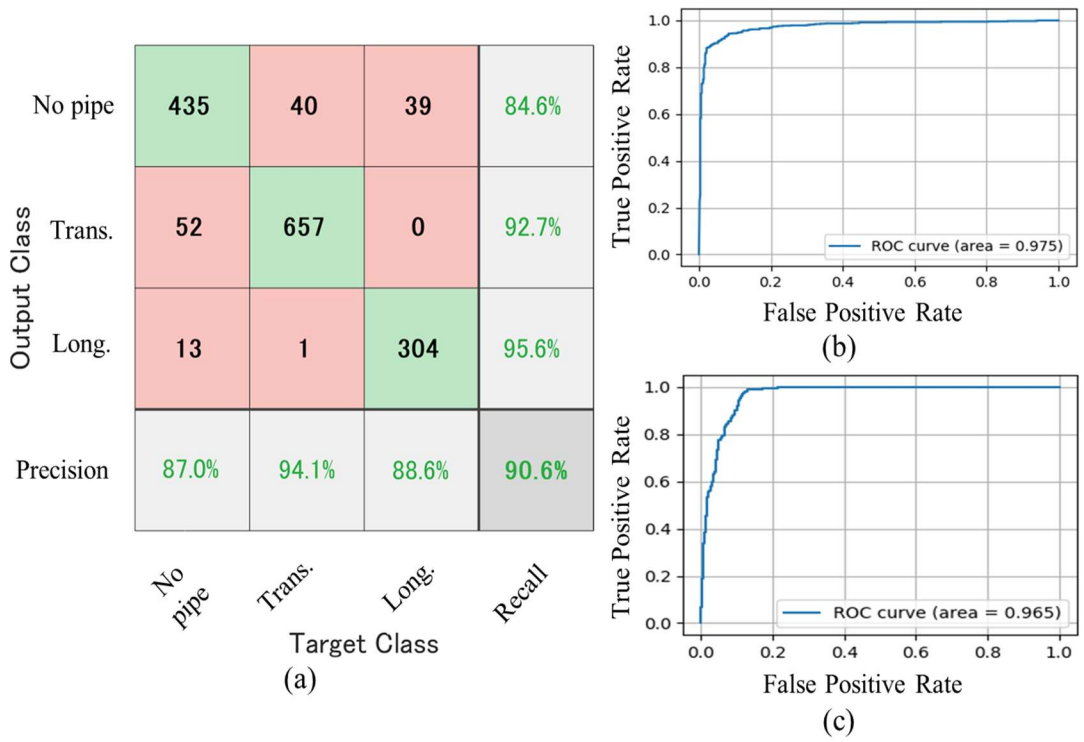


Fig. 6.52 Three-category classification of transverse and longitudinal pipes (a) confusion matrix (b) transverse pipe ROC (c) longitudinal pipe ROC

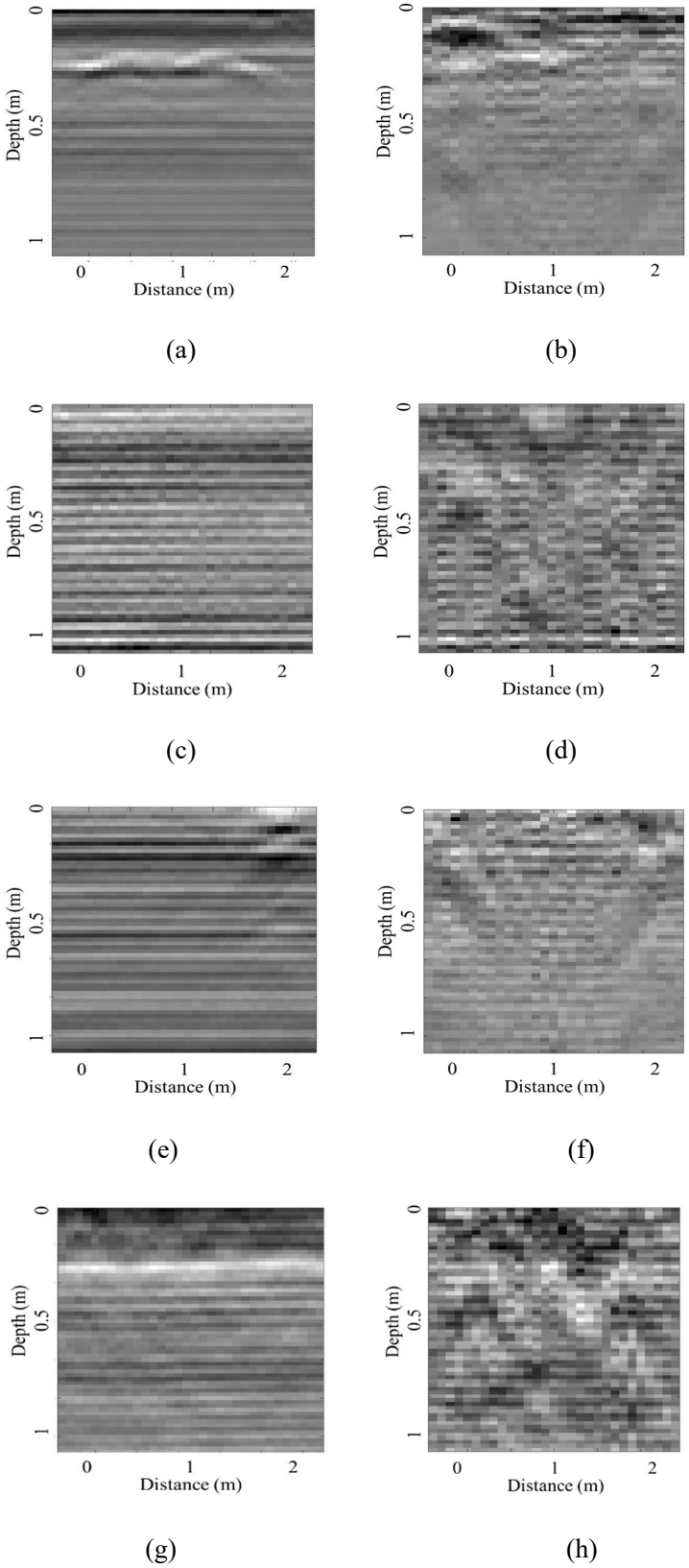


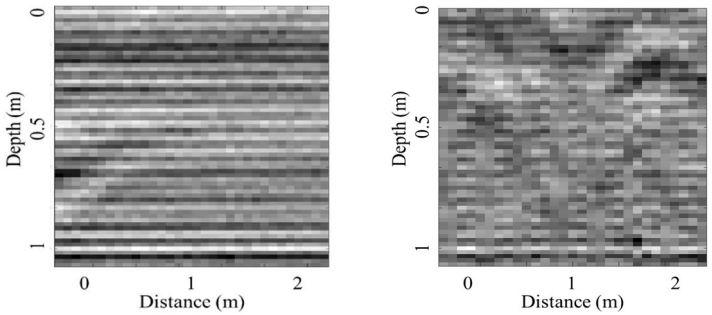
Fig. 6.53 False detection examples, scan and channel sections (true, output) (a) (b) (no pipe, transverse) (c) (d) (no pipe, longitudinal) (e) (f) (transverse, no pipe) (g) (h) (longitudinal, no pipe)

The adopted parameters are shown in Fig. 6.21 (b). Then, the number of the output categories of the fully connected layer was increased to 3. The output categories were transverse, longitudinal pipes and no pipe section. By augmenting the number of categories, the required number of convolution filters may be increased. The effect of the number of convolution filters was evaluated.

Fig. 6.51 shows the effect of the number of convolution filters N_f on classification accuracy. From Fig. 6.51, classification accuracy converged to 91 % over $N_f = 10$, by about 8 % increase compared to $N_f = 1$ and 3 % increase compared to $N_f = 5$. The required N_f for transverse and longitudinal pipes was around 5 and 10 respectively from the discussions of Section 6.5.2. Therefore, the results were consistent with the results of each direction pipe. Some features may be in common for the both direction pipes. Target features can be extracted at $N_f = 10$, which was adopted in the research.

Fig. 6.52 shows the confusion matrix and ROC curves of three-category classification. A ROC curve of each direction pipe was derived by considering the other direction pipe as negative data. From Fig. 6.52 (a), the numbers of false detection of transverse pipes as longitudinal pipes and vice versa were much less. The proposed 3D-CNN model accurately estimated the direction of a pipe when there was a pipe. Considering the numbers of the other elements of the confusion matrix, the detection rate of each direction pipe was consistent with each direction pipe 3D-CNN model. The total classification accuracy of three categories was about 91 %, intermediate between each 3D-CNN model. The conclusion is, a 3D-CNN model which accurately classifies the input data into three categories was developed.

Fig. 6.53 shows the examples of falsely classified data. Fig. 6.53 (a), (c), (e), (g) and (b), (d), (f), (h) show the scan and channel direction sections of the data respectively. Each caption stands for (true class, output class). From Fig. 6.53 (a) (b), even at no pipe section, because of clutters possibly reflected from soil layer boundaries and buried objects, the reflection patterns were falsely detected as pipes. From Fig. 6.53 (c) (d), noise may be detected as pipes. On the other hand, from Fig. 6.53 (e) - (h), when the reflections from the pipes were weak, the two direction pipes were detected as no pipe section. Fig. 6.54 shows the example of a longitudinal pipe falsely detected as a transverse pipe. Because the reflection from the pipe was weak, the reflection pattern may be misclassified.



(a) (b)

Fig. 6.54 Longitudinal pipe falsely detected as transverse pipe, scan and channel sections

In the research, by adopting a 3D-CNN model, 3D features were utilized to accurately estimate the existence and directions of pipes compared to 2D-CNN. However, detected regions were 1 m by 2 m by 2 m 3D boxes. The positions of pipes in detected regions should be estimated for practical mapping of pipes. In the following chapter, the 3D-CNN models were applied to the measurement and experimental field data of pipes and void. Furthermore, the positions of pipes were accurately estimated by inverse analysis (migration). The accuracy of the estimation algorithm was quantitatively evaluated by the experimental field data.

6.7 Summary

The three-category classification, transverse and longitudinal pipes and no pipe section was conducted by 3D-CNN. The methodology to process training data was explained. The theory of CNN and backpropagation were also introduced. In the research, minibatch training was adopted. Weight decay of learning rate and momentum were introduced. Early stopping was adopted to reduce training time checking the convergence of a validation curve. The variation of test accuracy between different no pipe section data sets was about 2 %. Therefore, the 2 % difference of classification accuracy was considered to be significant.

Preprocessing filters did not affect the results. The parameters of preprocessing filters are manually adjusted by inspectors. In a CNN model, target features were automatically extracted by convolution filters. Therefore, preprocessing filters were not necessary. In terms of CNN model optimization, the number of model layers is considered to be an important parameter. In the research, the classification accuracy of custom net defined by the authors was compared to existing CNN models such as VGG16 and InceptionV3. The 2 layer simple CNN model showed the highest classification accuracy. Classification accuracy increased compared to 1 layer NN and SVM models. The performance of the model was improved by utilizing convolution filters. More abstract features were utilized by a deep layer architecture. However, too deep layer models do not have an advantage.

3D-CNN utilizes 3D input data and convolution filters. A training algorithm is the same as 2D-CNN. Training time is at most several times larger than 2D-CNN. In the case of transverse pipes, when utilizing maximum 29 channels, 3D-CNN improved classification accuracy by about 6 % compared to 2D-CNN. In the case of longitudinal pipes, considering the lengths and inclination of pipes, there was an optimal section number. The classification accuracy of longitudinal pipes by 3D-CNN was about 3 % higher than 2D-CNN. In either case classification accuracy improved by utilizing the 3D features of input data by 3D-CNN. There are required sizes for convolution filters of 2D-CNN. Classification accuracy did not improve by 3D convolution filters. However, to conduct three-category classification, 3D convolution filters were necessary to estimate the directions of pipes.

The optimal parameters for each direction pipe were integrated to conduct three-category classification, transverse, longitudinal pipes and no pipe section by 3D-CNN. The advantages of three-category classification compared to each two-category classification are to reduce calculation time and conduct the integrated evaluation of two direction pipes. The optimal number of sections in a channel direction was derived by transvers pipes and scan direction by longitudinal pipes. In the 3D-CNN model, a pooling layer

was adopted to reduce calculation cost. Total calculation accuracy was high. It was about 91 %. Furthermore, the directions of pipes were accurately estimated when there were pipes. However, detected regions were 1 m by 2 m by 2 m 3D boxes in which pipes may exist. In the following chapter, the 3D positions of pipes were estimated by inverse analysis (migration). The detection of void by 3D-CNN were also discussed in the following chapter.

Chapter 7: 3D mapping of pipes and void

In this chapter, the 3D positions of pipes and void were estimated by migration, one of the inverse analysis methods. The regions of manhole, joint, pipes and void were localized by the SVM and 3D-CNN models. The SVM and CNN models developed in Chapter 4 - 6 were for classifying input data. In terms of the application of SVM and CNN models, the whole region was searched by box-by-box search. The overlap and probability threshold of categories were optimized. Then, the circles and lines of manhole and joint were detected by Hough transform. The positions of pipes and void were estimated by Kirchhoff migration. The parameters of Hough transform and Kirchhoff migration were optimized. The 3D positions were mapped to visualize the arrangement of target objects. The proposed algorithm was applied to the measurement and experimental field data. The results of other measurement data were shown in Appendix F.

7.1 Application of SVM and 3D-CNN

7.1.1 Comparison of region proposal methods

Among the region proposal methods, box-by-box search (exhaustive search) is the most fundamental method. Fig. 7.1 shows the concept of 2D and 3D box-by-box search. A target region is a 2D or 3D box with a certain size. The target region is shifted in a certain direction with a certain overlap to scan the whole region. An output is the probability of each category at each position. The parameters of box-by-box search are the overlap in each direction and probability threshold of each category. Many up-to-date region proposal methods other than box-by-box search were proposed in previous research. In terms of the object detection from photo images, Fast R-CNN, Faster R-CNN, You Look Only Once (YOLO) and Single Shot MultiBox

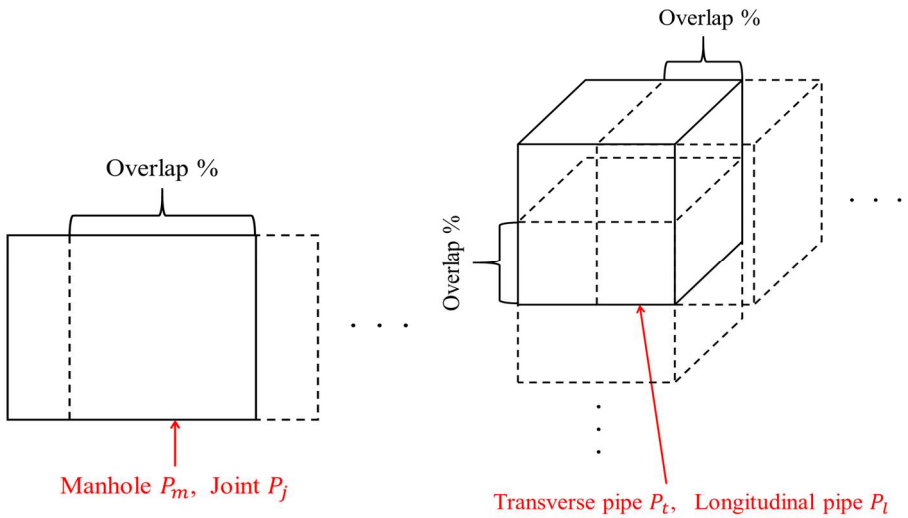


Fig. 7.1 Detection by box-by-box search (a) 2D (b) 3D

Detector (SSD) were notable [155) - 161)]. Calculation cost and robustness are the keywords. In selective search, neighboring regions are integrated comparing the similarity of color and intensity. The criteria and extent of similarity to integrate the regions are learned by comparing results and ground truth data. In fast and faster R-CNN, the positions and sizes of targets are directly output by the networks as the regression problem. In SSD, features are effectively extracted by multiscale feature maps. Later algorithms need lower calculation cost by special layer architectures. Furthermore, later algorithms are considered to be more robust to geometric transformation.

In terms of radar data, images are grayscale and relatively large including many reflection patterns and clutters. The scale of target patterns is almost constant. The situation is completely different from colored scenery photos. Selective search and faster R-CNN were applied to radar images in the research. The algorithms failed because too many regions were proposed to require high calculation cost. Furthermore, target data was 3D in the research. Therefore, the simplest box-by-box search was adopted. Segmentation is to assign categories by pixel level. The application of other up-to-date region proposal and segmentation methods are left as future work. In the research, the positions of manhole, joint, pipes and void were estimated by Hough transform and migration method. Therefore, the evaluation criterion is not the rate of overlapping areas but estimation error of target positions. Estimation error was discussed in the application to the experimental field data.

7.1.2 Localization by box-by-box search

In terms of manhole and joint detection, the SVM model developed in Chapter 4 was applied to the surface reflection of asphalt pavement by box-by-box search. The size of input data was optimized in Section 4.4.1. A target region was 29 pixels by 29 pixels (about 2 m by 2 m). 5 images in a depth direction were utilized to increase classification accuracy. 3D $(29, 29, 5)$ pixel data was converted to a one-dimensional feature vector to input the SVM model. Overlap in a scan (longitudinal) direction and probability threshold of each category were the parameters. In a channel (transverse) direction all the channels were extracted. The overlap in a scan direction was 75 % considering the following discussions of pipes case. The overlap 75 % corresponds to the 0.5 m shift of images. Because the sizes of manhole and joint were at most 1.5 m, the overlap was appropriate considering the reflection patterns. The probability threshold of manhole P_m and joint P_j were both 0.95. The probability P tended to range from 0.98 to 0.99 when there were manhole and joint. In pavement section, P was under 0.5. Therefore, $P_m = P_j = 0.95$ were appropriate threshold. Several boxes including target objects were successively detected in a scan direction. To decrease the calculation cost of Hough transform, only the region with the smallest distance was extracted.

In terms of pipes and void detection, the 3D-CNN models developed in Chapter 5 and 6 were utilized. 3D region was searched by box-by-box search. The size of input data was 64 pixels by 32 pixels by 29 pixels (about 1 m by 2 m by 2 m) in depth, scan and transverse directions respectively. Overlap and probability threshold of each category were the parameters. In a transverse direction, all the channels were extracted. Overlap in a depth direction Ol_d and scan direction Ol_s were the parameters. The probability threshold of transverse pipes P_t, longitudinal pipes P_l and void P_v were also the parameters. The effects of these

parameters were discussed below. The box with the smallest distance was adopted also in this case.

Fig. 7.2 shows the 3D maps of three transverse pipes case of the measurement data used as test data changing overlap Ol_t, Ol_l with constant probability threshold $P_t = P_l = 0.995$. The red and blue boxes correspond to transverse and longitudinal pipes respectively. In Fig. 7.2 (d), the pipes were successively detected in a depth direction because of the multiple reflections between the antennas and pipes. The actual locations of pipes are in the top of the aligned 3D boxes. The boxes at the same distance in a deeper region were automatically removed in the following migration process. False signals behind anomalies need to be removed beforehand, which is out of scope of the research.

From Fig. 7.2 (a) (b), under $Ol_t = Ol_l = 50\ \%$, some pipes were missing. At $Ol_t = Ol_l = 50\ \%$, the shift of the images was 50 %, which was not allowable for the classification by CNN. At $Ol_t = Ol_l = 25\ \%$, all the pipes were not detected because most of the hyperbolic reflection patterns were outside the target areas. From Fig. 7.2 (c) (d), the three pipes were successfully detected over $Ol_t = Ol_l = 75\ \%$. The shift of the images was 25 %, which was small enough for the detection of the pipes. Considering calculation cost, $Ol_t = Ol_l = 90\ \%$ or larger overlap were not feasible because the total amount of data becomes

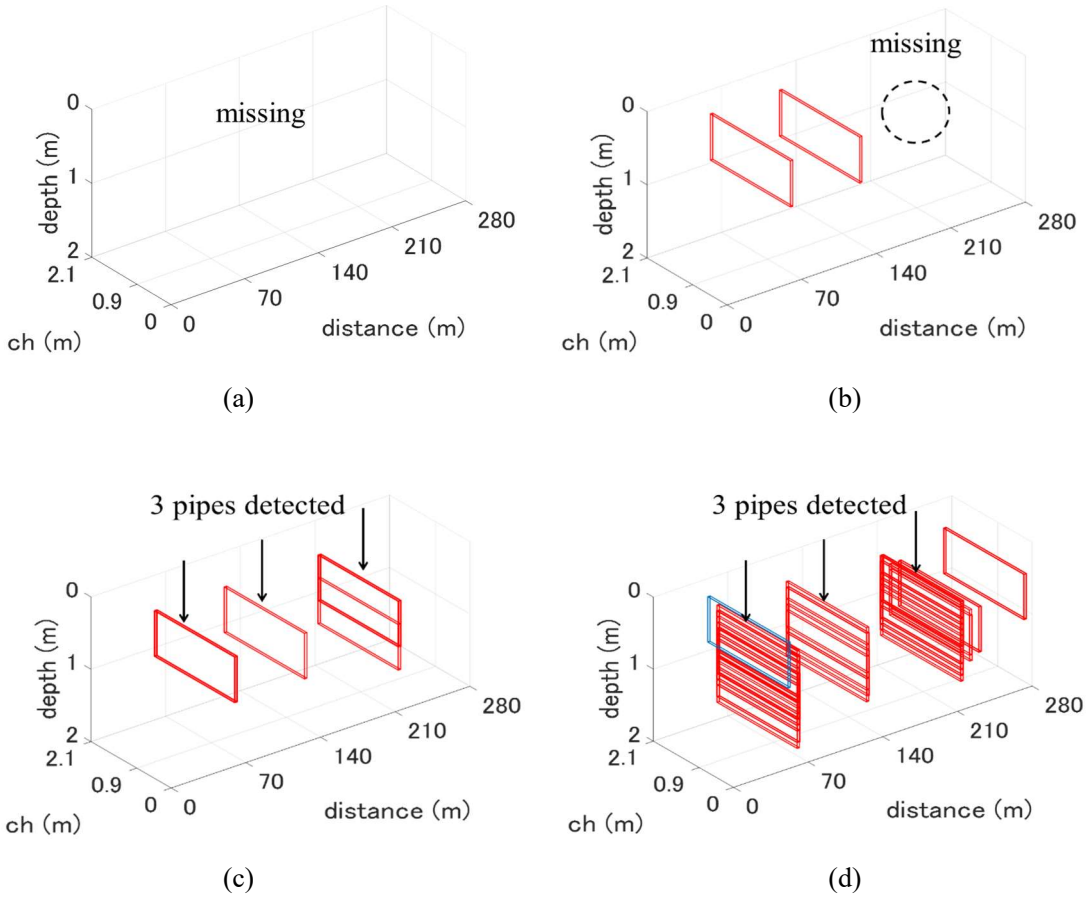


Fig. 7.2 Effect of overlap Ol at $P = 0.999$ on detection of pipes (red boxes correspond to transverse pipes and blue boxes to longitudinal pipes) (a) $Ol = 25\ \%$ (b) $Ol = 50\ \%$ (c) $Ol = 75\ \%$ (d) $Ol = 90\ \%$

hundred or more times larger than the original data. Calculation time also increased proportional to the amount of data. $Ol_t = Ol_l = 75\ \%$ took about 30 minutes for 1 km section by ordinary laptop PC CPU cores, small enough comparing to the manual inspection. Calculation time decreased to about 5 minutes with GPU as discussed in Section 8.1.3.

Fig. 7.3 exhibits the same pipes 3D maps by changing P_t and P_l from 0.99 to 0.99999 with constant $Ol_t = Ol_l = 75\ \%$. Compared to Fig. 7.3 (b) at $P_t = P_l = 0.999$, false detection increased in the case $P_t = P_l = 0.99$ from Fig. 7.3 (a). From Fig. 7.3 (c) (d), the one pipe was missing in the case $P_t = P_l = 0.9999$ while all the pipes were missing in the case $P_t = P_l = 0.99999$. The results were reasonable because larger probability threshold is, lesser false detection is and smaller a detection rate becomes. Because no pipe areas in the whole data was much larger than the areas of pipe reflections, 3D maps improved with relatively high P_t and P_l. The test data included pipes of various material and diameter buried in wide areas. Therefore, appropriate P_t and P_l for target objects only depend on the measurement system. Practically, P_t and P_l may be determined by the acceptable false negative rate required for application purposes. P_t and P_l may be independently optimized. For simplicity, $P_t = P_l = 0.999$ were adopted in the research.

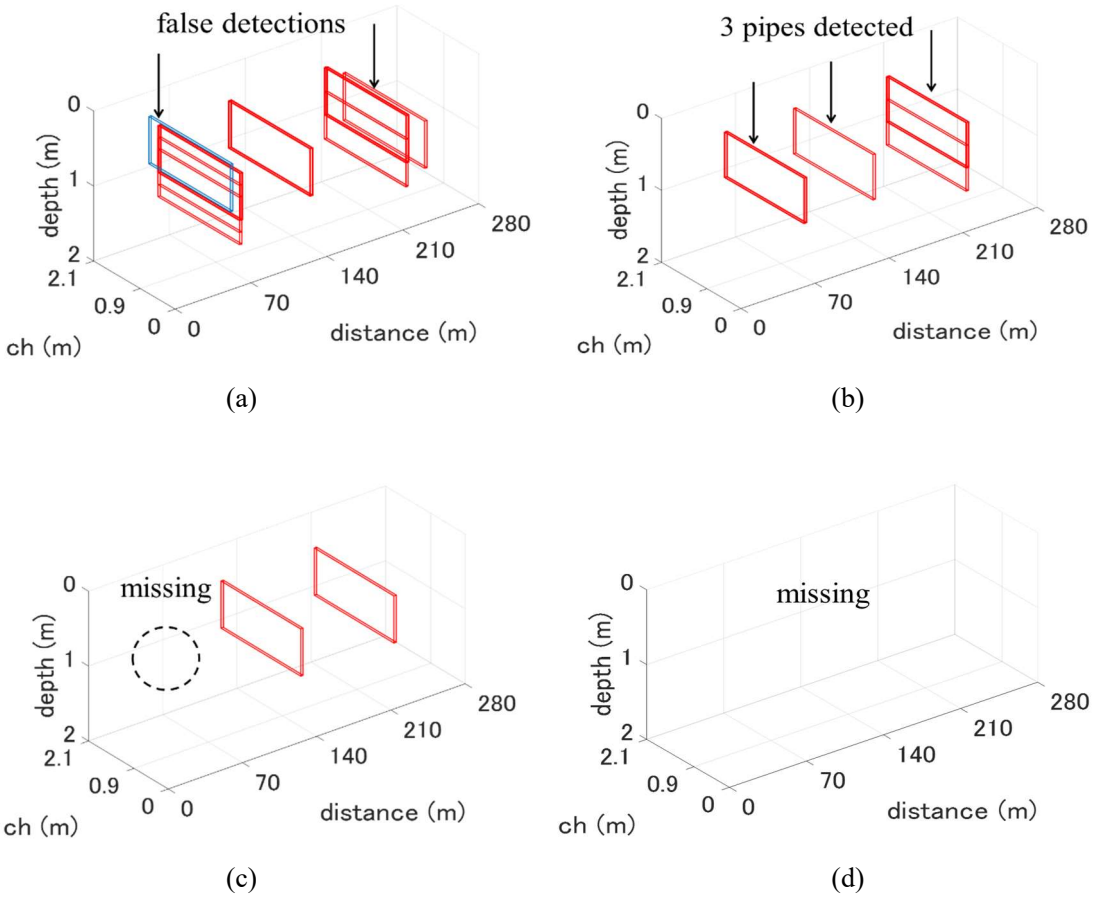


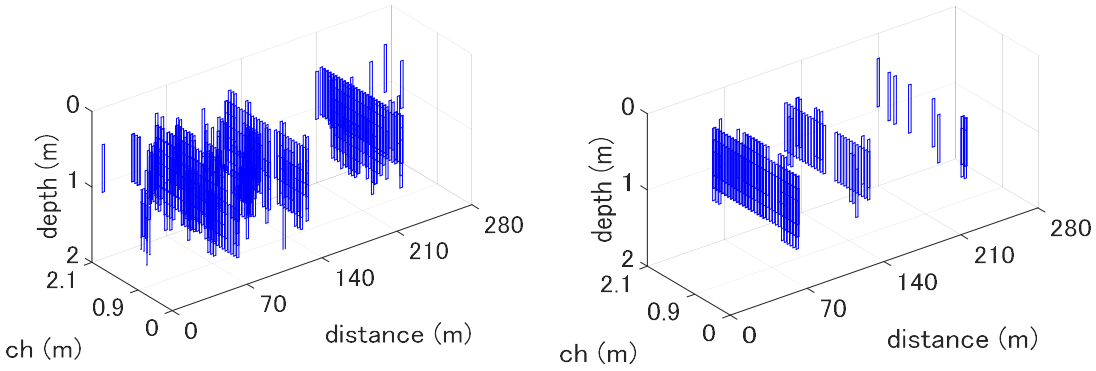
Fig. 7.3 Effect of probability threshold P at $Ol = 75\ \%$ on detection of pipes (red boxes correspond to transverse pipes and blue boxes to longitudinal pipes) (a) $P = 0.99$ (b) $P = 0.999$ (c) $P = 0.9999$ (d) $P = 0.99999$

Fig. 7.4 shows the results of a 3 transverse pipes case shown in Fig. 7.2 and Fig. 7.3 by 2D-CNN with the same overlap. In 2D-CNN, each direction pipe was separately detected by each direction pipe model. From Fig. 7.4, many false detection was observed at $P = 0.99$ while one pipe was missing at $P = 0.999$. By utilizing the 3D features of the data, 3D-CNN accurately detected the pipes. The continuity of hyperbolic reflection patterns may help the detection even when the reflection patterns were weak or covered by noise or clutters. Fig. 7.5 shows the results of a one longitudinal pipe case detected by 3D-CNN with the same overlap and probability threshold as the transverse pipe case. The longitudinal pipe was successively detected in a scan direction. In the case of a long longitudinal pipe, pipes may be detected even by smaller overlap in a scan direction. The effect of probability threshold on the results was the same as the transverse pipe case. Fig. 7.6 shows a void case. 3D-CNN accurately detected the transverse, longitudinal pipes and void without false detection.

7.2 Manhole and joint detection

7.2.1 Comparison of detection methods

Edges of typical manhole and joint are a circle and set of lines respectively. Hough transform is utilized



(a) (b)

Fig. 7.4 Detection by 2D-CNN (a) $P = 0.99$ (b) $P = 0.999$

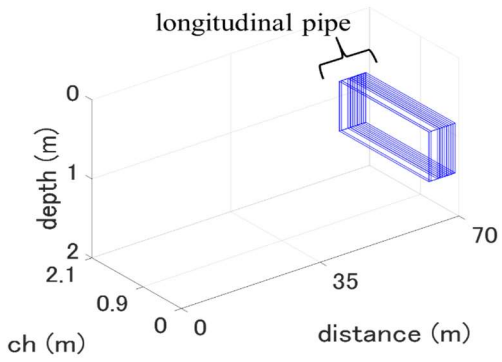


Fig. 7.5 Detection of longitudinal pipe

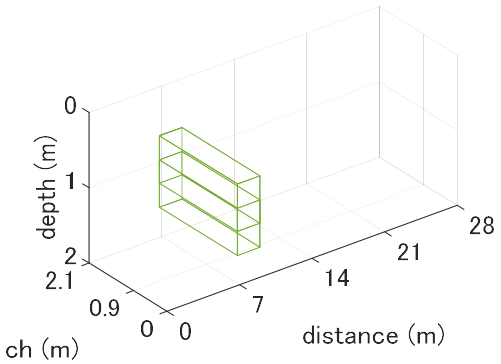
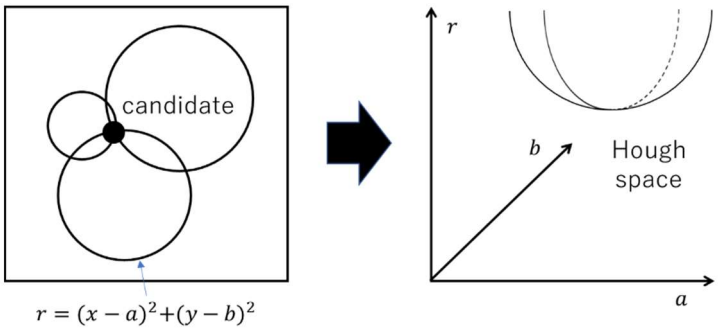


Fig. 7.6 Detection of void

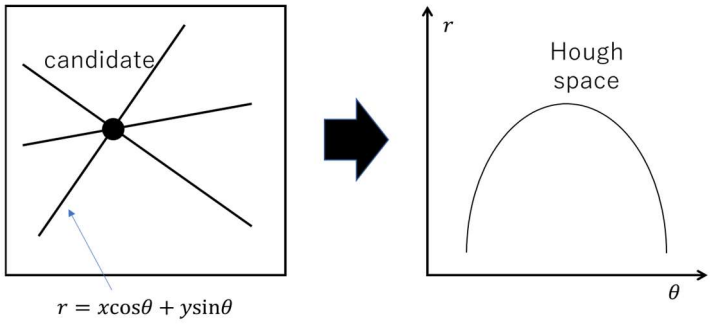
to detect circles and lines from images. Fig. 7.7 shows the concept of Hough transform. An image was converted to binary data by edge detection method. Edge detection method was discussed in Section 4.4.2. The points on the detected edges are the candidates for Hough transform. To detect circles, considering all the circles which pass through the candidate point, the center of circles (a, b) and radius r are plotted on Hough space (a, b, r). One candidate point is converted to one quadric surface in Hough space. To detect lines, considering all the lines which pass through the candidate point, inclination θ and distance from the origin r are plotted on Hough space (θ, r). One line is converted to one hyperbolic. The points in Hough space where most quadric surfaces or hyperbolics pass through are detected as circles or lines. The algorithm is efficient when the number of candidate points is small. Another possible circle and line detection algorithm is pattern matching. However, to apply pattern matching, many possible circles and lines should be searched by trial and error. Calculation cost is generally high. Therefore, Hough transform was adopted.

7.2.2 Detection by Hough transform

The surface reflection of localized areas by SVM were extracted. Hough transform was applied to the areas to detect circles of manhole and lines of joint. In Hough transform, the range of radius of circles and minimum length of lines were the parameters. In the research, considering the geometry of manhole and joint, the range of radius r was 5 to 15 pixels (about 0.4 m to 1.1 m). The minimum length l was 10 pixels



(a)



(b)

Fig. 7.7 Hough transform (a) detection of circles (b) detection of lines

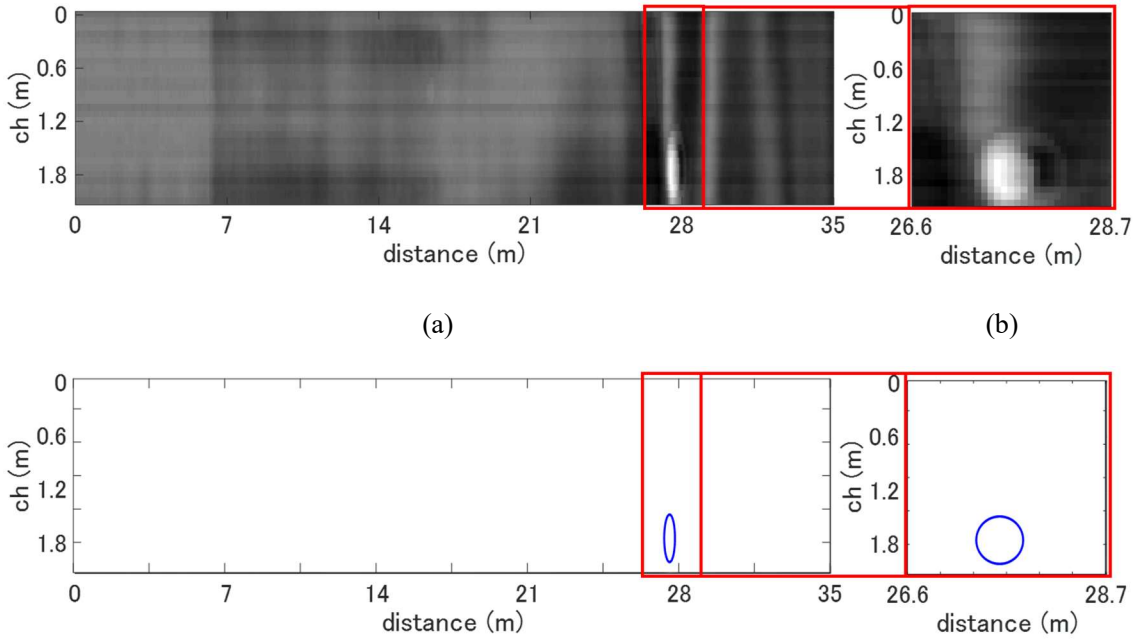


Fig. 7.8 Manhole detection result I (a) (b) raw images (c) (d) detection of manhole

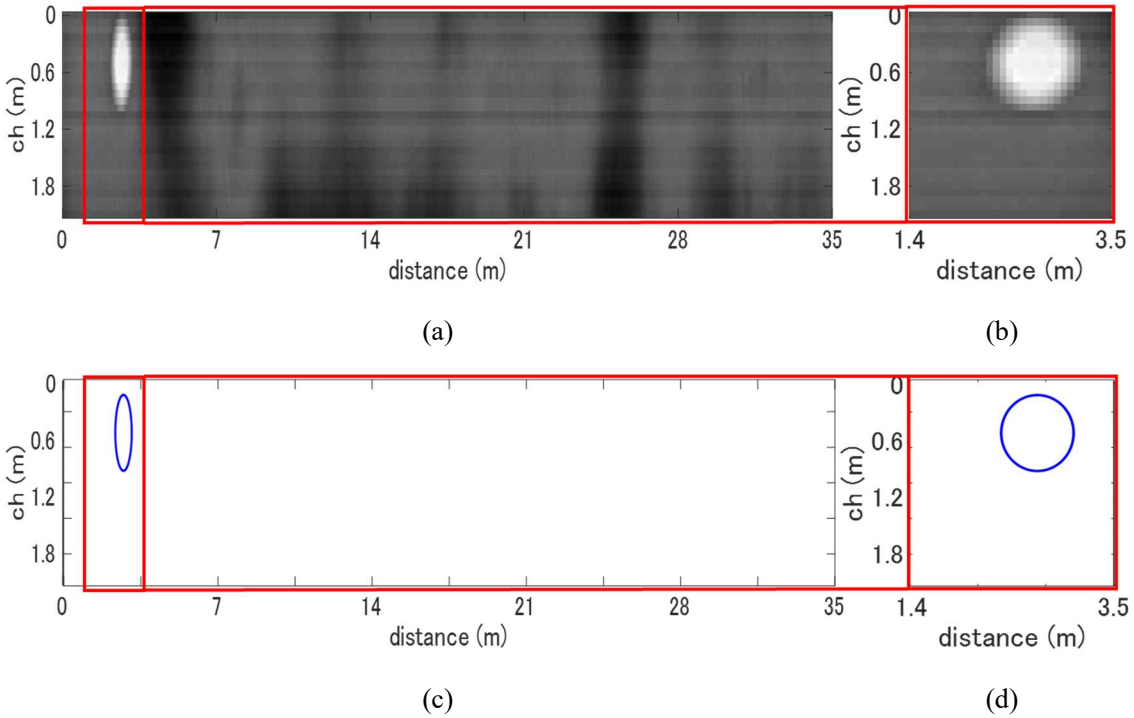


Fig. 7.9 Manhole detection result II (a) (b) raw images (c) (d) detection of manhole

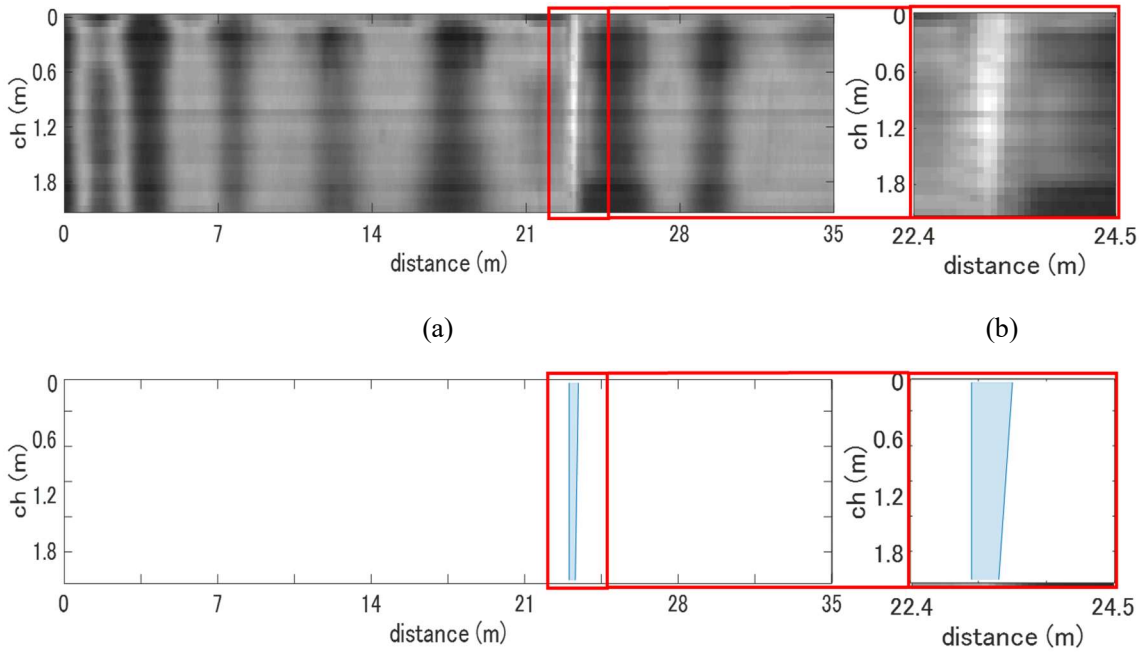


Fig. 7.10 Joint detection result I (a) (b) raw images (c) (d) detection of manhole

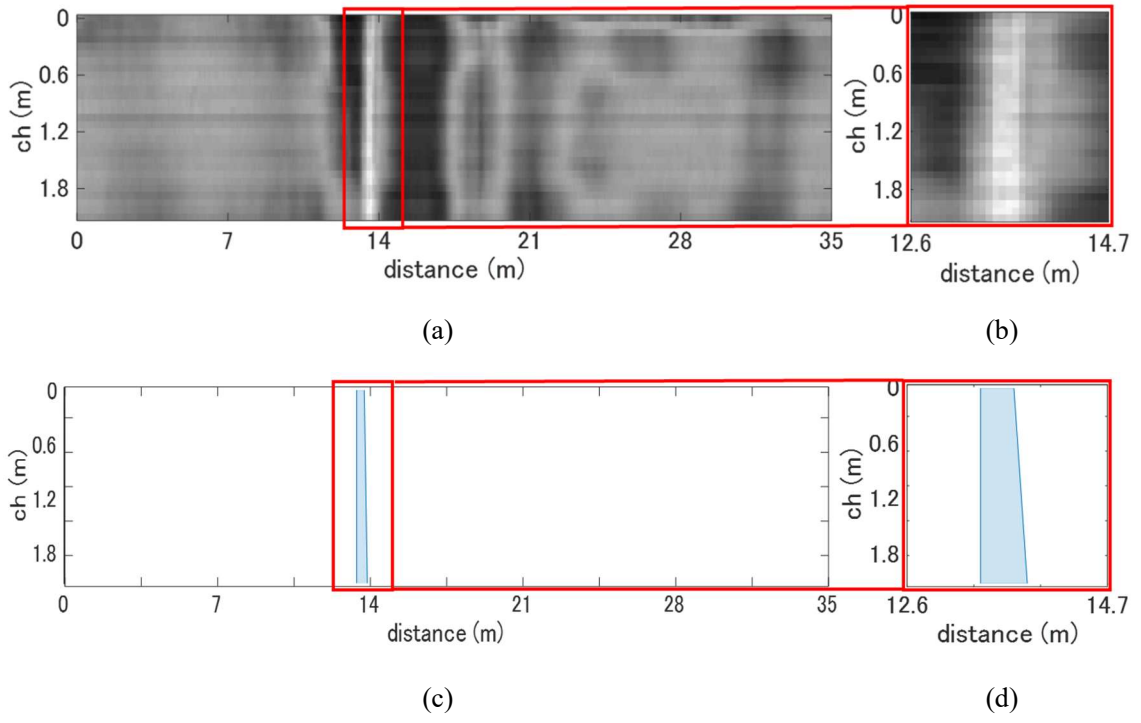


Fig. 7.11 Joint detection result II (a) (b) raw images (c) (d) detection of manhole

(about 0.7 m). Edge detection was conducted by Canny method introduced in Section 4.4.2. Manhole was represented by the center point and radius of the detected circle. Joint was represented by a set of two lines with positions in a scan direction and inclination.

Fig. 7.8 and Fig. 7.9 show the results of manhole detection. From Fig. 7.8 and Fig. 7.9, two manholes were appropriately detected without false detection. Then, all the manholes were accurately visualized by circles. Fig. 7.10 and Fig. 7.11 show the results of joint detection. The area between two lines were colored blue. As shown in Fig. 7.10 and Fig. 7.11, two joints were appropriately detected without false detection. Then, all the joints were accurately visualized by the sets of lines. In the detection of manhole and joint, information about the sizes of targets was automatically obtained, which may be useful from a practical point of view. From Fig. 7.8 - Fig. 7.11, by the combination of SVM and Hough transform, manhole and joint were successfully localized. Then, their positions were estimated. The estimation accuracy was pixel order (several to several tens cm) from Fig. 7.8 - Fig. 7.11, which was accurate enough considering the sizes of manhole and joint are several tens cm to 1 m.

7.3 Pipes and void detection

7.3.1 Comparison of migration methods

Migration (aperture synthesis method) is one of the inverse analysis methods [162) - 164)]. Fig. 7.12 shows the concept of migration. The shape of a scatterer is reconstructed from a reflection pattern. Reflected waves are observed not only just above the scatterer but also other positions, forming unique reflection patterns. Assuming an infinitely small scatterer, a reflection pattern is a hyperbolic shape whose apex corresponds to the position of the scatterer. In the case of a small pipe, an apex corresponds to the upper surface of the pipe. Migration depends on the velocity of propagating waves. In time migration, migration is conducted in a time direction assuming a uniform medium. In depth migration, considering propagating velocity in each medium, time is converted to depth. In the research, time migration was adopted. In an actual situation, if the permittivity of a medium is known, depth migration may further improve the results.

Kirchhoff migration, f-k migration and reverse time migration are the widely used migration algorithms

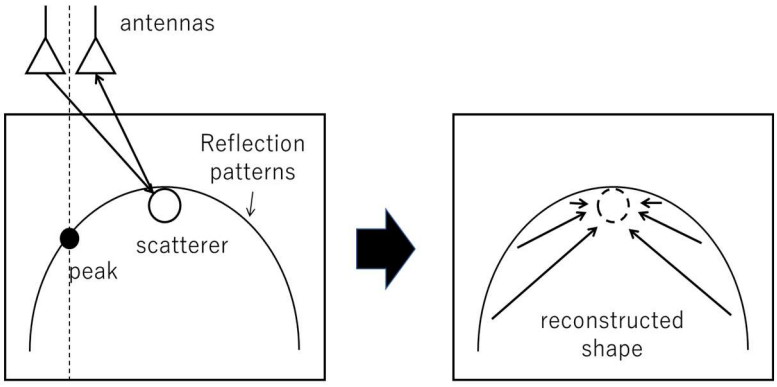


Fig. 7.12 Migration

in NDT practice. Kirchhoff migration is the most accurate and simplest migration algorithm. In homogeneous medium, electromagnetic waves diffuse from antennas and scatterers concentrically. Therefore, the received electric field in certain time is affected by the scatterers in the same distance from the antennas.

Fig. 7.13 shows the section of a target 3D region in a scan direction. The top left of the target region is the origin of xz space. The midpoint of the transmitting and receiving antennas positions at x'. The depth of the top left of the detected region (depth offset) is dz. The distance between two the antennas is $2h$. The distance from the antennas to a scatterer $L_{\pm}$ is written below.

$$L_{\pm} = \sqrt{(z + dz)^2 + (x - x' \pm h)^2} \tag{7-1}$$

The difference from ordinary Kirchhoff migration is the additional term dz and limited target region. dz was obtained when applying 3D-CNN. An image after Kirchhoff migration I_{m_k} is derived from the integration of a raw image $I_o(L)$ by an angle θ at constant distance $L = L_- + L_+$. Considering the angle range of migration $\theta = -\varphi \sim +\varphi$, $I_{m\_k}(L)$ is written below.

$$I_{m\_k} = \int_{-\varphi}^{+\varphi} I_o(L)\, d\theta \tag{7-2}$$

Electromagnetic waves dissipate in free space. Therefore, loss was compensated before migration. A compensation function depends on L. When the size of I_o is large, the calculation cost of Eq. (7-2) is extremely large. However, in the research, the regions were localized by 3D-CNN. A target region was limited to 64 pixels by 32 pixels images. Therefore, calculation cost was not high even by Kirchhoff migration. Calculation time was discussed in Section 8.1.3.

f-k migration is based on 2D Fourier transform. The Fourier transform of a 2D image $I(x, z)$ is written by angular frequency ω_x, ω_z below.

$$I(\omega_x, \omega_z) = \frac{1}{2\pi} \iint I(x, z) e^{-j(\omega_x x + \omega_z z)}\, dx dz \tag{7-3}$$

Time histories of waves at $z = 0$, $I(x, 0)$ is converted to a spectrum $J(\omega_x, \omega_t)$. Assuming no dispersion of

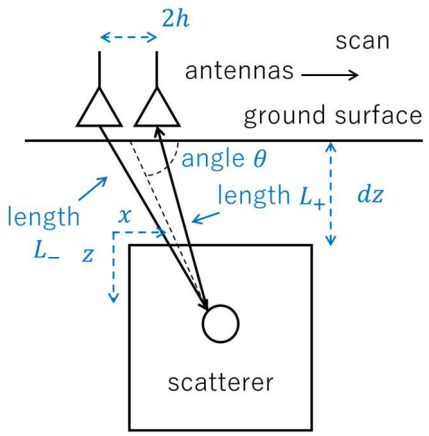


Fig. 7.13 Kirchhoff migration

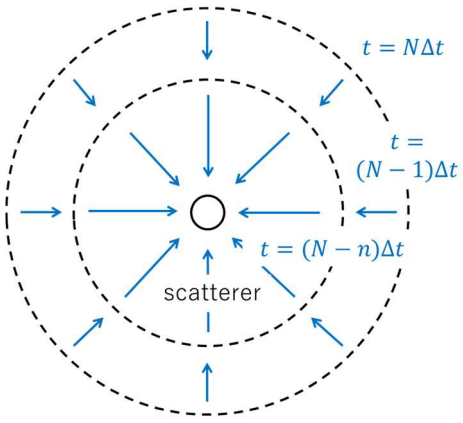


Fig. 7.14 Reverse time migration

propagating velocity v, wave numbers k_x, k_z are related to ω_t by the equation below.

$$\omega_t = v\sqrt{k_x^2 + k_z^2} \tag{7-4}$$

From Eq. (7-4) and nature of plane waves, $I(\omega_x, \omega_z)$ is written by $J(\omega_x, \omega_t)$ below.

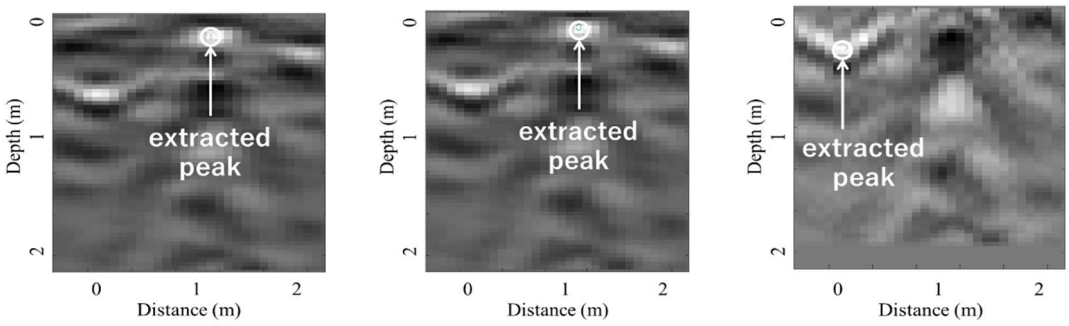
$$I(\omega_x, \omega_z) = J(\omega_x, \omega_t) \cdot \frac{\omega_z}{\sqrt{\omega_x^2 + \omega_z^2}} \tag{7-5}$$

Because fast Fourier transform of Eq. (7-3) is utilized, f-k migration is faster than Kirchhoff migration. To obtain $I(\omega_x, \omega_z)$ from $J(\omega_x, \omega_t)$, data should be interpolated in time and space directions by Stolt transform, which decreases the accuracy.

Reverse time migration utilizes the difference equations of FDTD method. Time is literally reversed as shown in Fig. 7.14 by FDTD method. Reverse time migration can consider the inhomogeneity of a medium. However, calculation cost is high to conduct high accuracy simulations. Therefore, in the research, considering accuracy and calculation cost, Kirchhoff migration was adopted. The estimation accuracy of positions of pipes was evaluated in the following section.

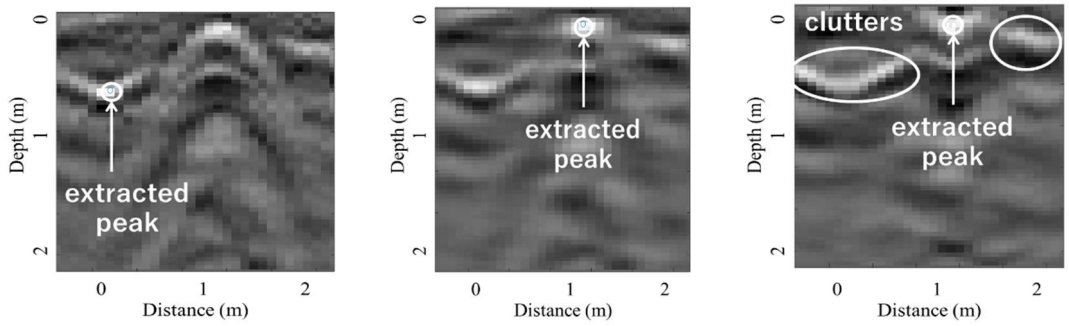
7.3.2 Pipe detection by Kirchhoff migration

3D regions were localized by 3D-CNN. The direction of a pipe was estimated by 3D-CNN. Then, the



(a) (b) (c)

Fig. 7.15 Migrated images (a) relative permittivity of soil $\varepsilon_s = 1$ (b) $\varepsilon_s = 3$ (c) $\varepsilon_s = 20$



(a) (b) (c)

Fig. 7.16 Migrated images (a) aperture length $l = 35$ cm (b) $l = 70$ cm (c) $l = 210$ cm

cross-section images perpendicular to the pipe was extracted. The maximum peak of the cross-section image was detected as the upper surface of the pipe. In the application of Kirchhoff migration, relative permittivity of soil and synthesis aperture length are the important parameters. Fig. 7.15 shows the migrated images with the different relative permittivity of soil ε_s. The images include the hyperbolic reflection patterns of pipes. With too small ε_s, hyperbolic patterns remain, which is called under-migration [62), 63)]. On the other hand, with too large ε_s, downward convex patterns appear and is enhanced, which is called over-migration. From the results of Fig. 7.15, when $\varepsilon_s = 1$ and 3, migration was successfully conducted. However, when $\varepsilon_s = 20$, over migration occurred. From the results, ε_s was considered to be small from 1 to 3. In the research, considering the discussions of the experimental field case, $\varepsilon_s = 3$ was assumed. Synthesis aperture length l was the length in a horizontal direction to migrate the images, which corresponds to an interval of integration in Eq. (7-2). From the result of Fig. 7.16 (a), with too small l, the effect of migration was small. The migrated image was similar to the raw image. On the other hand, from Fig. 7.16 (c), with too large l, the reflection patterns were covered by the clutters because the target patterns were filtered out after integration. l depends on the characteristics of antennas. In the research, $l = 0.70$ cm was adopted for all the cases, which was optimized by the several measurement data shown in Appendix F.

Fig. 7.17 shows the examples of images before and after migration. From the results of Fig. 7.17 (a) (b), before migration, a clutter was falsely detected as the peak of a pipe. After migration, because the energy of the reflection pattern of the pipe was concentrated on the peak, the pipe's position was successfully extracted. From the results of Fig. 7.17 (c) (d), before migration, a sidelobe was detected. After migration, the apex of

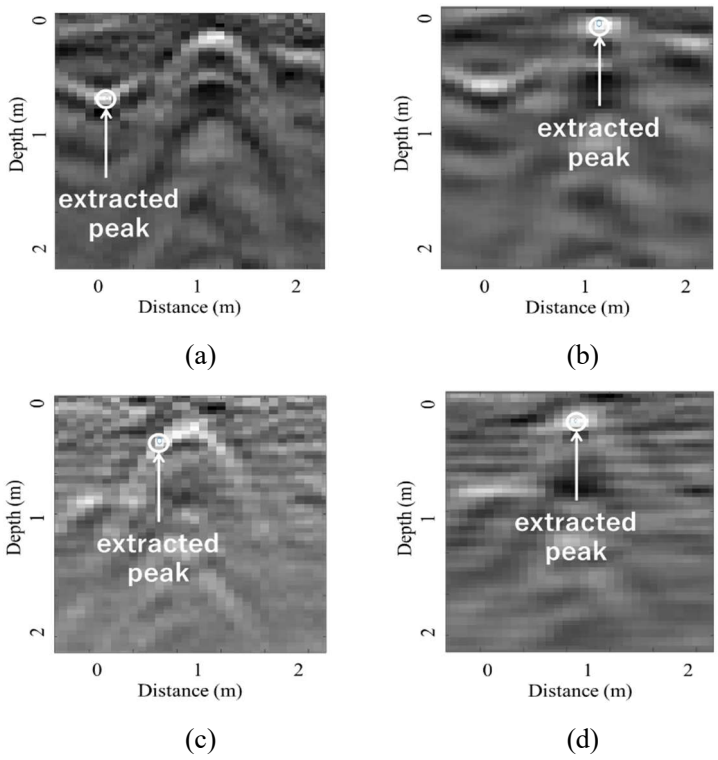


Fig. 7.17 Images before and after migration (a) (b) image I, before and after (c) (d) image II

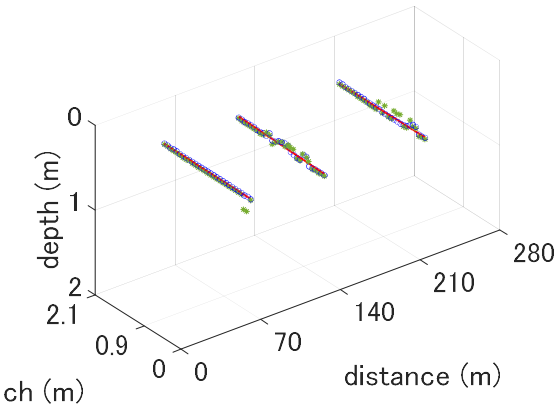
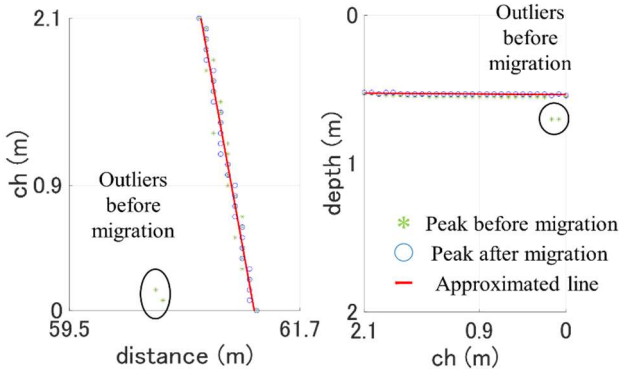


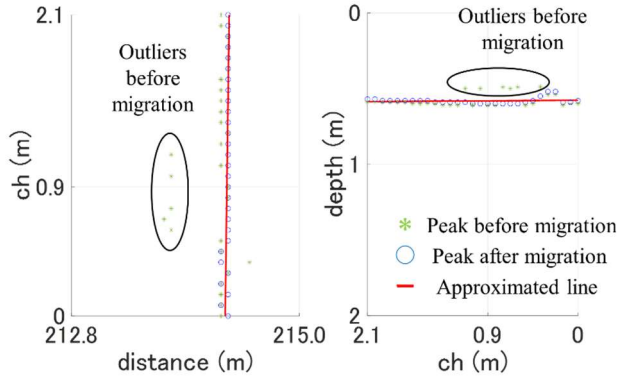
Fig. 7.18 3D map of 3 transverse pipes before and after migration

Table 7.1 Standard deviations of 3 pipes from approximated lines

	Before migration	After migration
Pipe 1 (60 m)	4.52 pixel	0.46 pixel
Pipe 2 (130 m)	5.03 pixel	3.04 pixel
Pipe 3 (210 m)	5.18 pixel	2.02 pixel



(a) (b)



(c) (d)

Fig. 7.19 Visualized transverse pipes (a) (b) top and side views at 60 m (c) (d) 214 m

the hyperbolic reflection pattern was precisely detected. After migration, the estimation accuracy of the pipes increased without false detection.

Fig. 7.18 shows the 3D map of the three transverse pipes case shown in Fig. 7.2 and Fig. 7.3. The approximated lines were derived by the detected points. From 7.18, the transverse pipes were successfully visualized by the aligned points in a transverse direction and approximated lines. Table 7.1 shows the standard deviations of the pipes' points from the approximated lines before and after migration. Fig. 7.19 shows the zoomed figures of the pipes at 60 m and 214 m of Fig. 7.18 with the points before and after migration. From Table 7.1, before migration, the maximum standard deviation was 5.2 pixel, corresponding to about 40 cm error. After migration, standard deviations decreased to several cm to at most 20 cm. Considering the diameter of pipes ranges from several cm to several tens cm, the estimation accuracy was enough from a practical point of view. From Fig. 7.19, outliers were observed before migration. After migration, the points were precisely aligned on the approximated lines. From Fig. 7.19 (a), not only the position but also the inclination of a pipe was reproduced because the estimation error was under the diameter of pipes.

Fig. 7.20 shows the 3D map of a one longitudinal pipe case shown in Fig. 7.5. Fig. 7.21 shows the zoomed

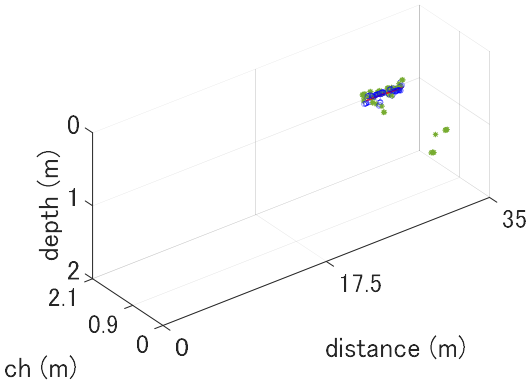
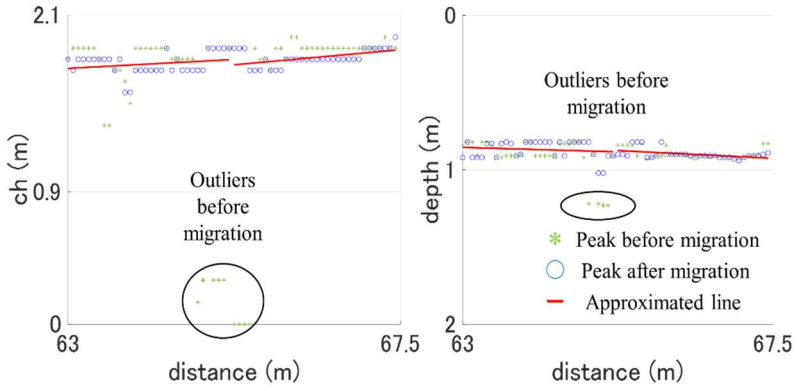


Fig. 7.20 3D map of one longitudinal pipe before and after migration



(a) (b)

Fig. 7.21 Visualized longitudinal pipe (a) (b) top and side views at 65 m

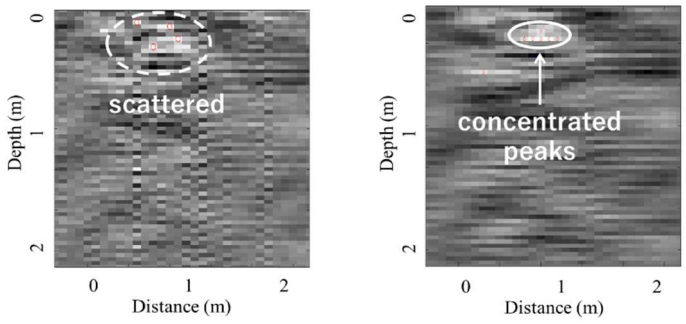
figures of Fig. 7.20. From Fig. 7.21, after migration, the points were precisely aligned on the approximated lines also in the case of the longitudinal pipe. Summarizing the results of Fig. 7.18 - Fig. 7.21, the positions, directions and inclination of pipes were successfully visualized in a 3D map after migration. The arrangement of pipes was realistic and useful from a practical point of view. The results of other pipe measurement data were introduced in Appendix F.

7.3.3 Void detection by Kirchhoff migration

3D regions of void were also localized by 3D-CNN. In the case of void, a void area spreads in horizontal and depth directions. Therefore, after migration in cross-section images in a scan direction, the maximum peak in the detected 3D area was extracted. The absolute intensity of the maximum peak was E. Then, all the points whose absolute intensity was over $T \times E$ were extracted as a void area with threshold T. The absolute intensity of the edge of void decreases compared to the center because of the difference of the thickness of void assuming an ellipsoid body. By adjusting T, a void area can be appropriately extracted. T depends on the characteristics of the antennas, which was optimized by the several measurement data shown in Appendix F. Other Kirchhoff migration parameters were the same as the pipe case.

Fig. 7.22 shows the images before and after migration. Before migration, the reflections from the upper surface of void was not clear. After migration, the energy was concentrated on a certain area, showing clear positive peaks. The reflections from the lower surface of the void was not clear even after migration possibly because most of the electromagnetic waves reflected back at the upper surface. In some cases, lower surface can be observed as negative peaks as shown in Appendix F. In that case, the distance between the upper and lower surface corresponds to the thickness of void.

Fig. 7.23 shows the detected area changing T. Red and blue points represent positive and negative peaks respectively. With smaller T from Fig. 7.23 (a), other areas were falsely detected as void. With larger T from Fig. 7.23 (c), a void area was underestimated. From the results of Fig. 7.23, $T = 0.5$ was considered to be appropriate threshold. Fig. 7.24 shows the 3D map and side view at $T = 0.5$. From Fig. 7.24 (a), the area was successfully detected. From Fig. 7.23 and Fig. 7.24, the void was successfully visualized as the cluster of the extracted points after migration. The results of other measurement data with the same parameters were introduced in Appendix F.



(a) (b)

Fig. 7.22 Void images (a) before migration (b) after migration

7.4 Experimental field results

7.4.1 Pipe detection results

The configuration of the experimental field pipe data was shown in Section 2.4.1 and Appendix A. Fig. 7.25 - Fig. 7.28 show the results of the parts of the measurement course 2 in Appendix A including the two transverse pipes. Fig. 7.25 and Fig. 7.27 are the geometry of the buried pipes. Fig. 7.26 (a) - (c) and Fig. 7.28 (a) - (c) show the 3D maps, top and side views respectively. From Fig. 7.26 and Fig. 7.28, both the transverse and longitudinal pipes were recognizable by the aligned points. The two pipes were plotted in appropriate depth, distance and channel. Furthermore, the 3D arrangement of the two pipes were accurately visualized. The transverse pipes located in shallower depth than the longitudinal pipe. The transverse and longitudinal pipes were in perpendicular directions. The pipes of various material and diameter were included in the measurement and experimental field data. Therefore, the proposed algorithm is applicable to any data with the same measurement system. Some missing points of the longitudinal pipes may be the effects of the above

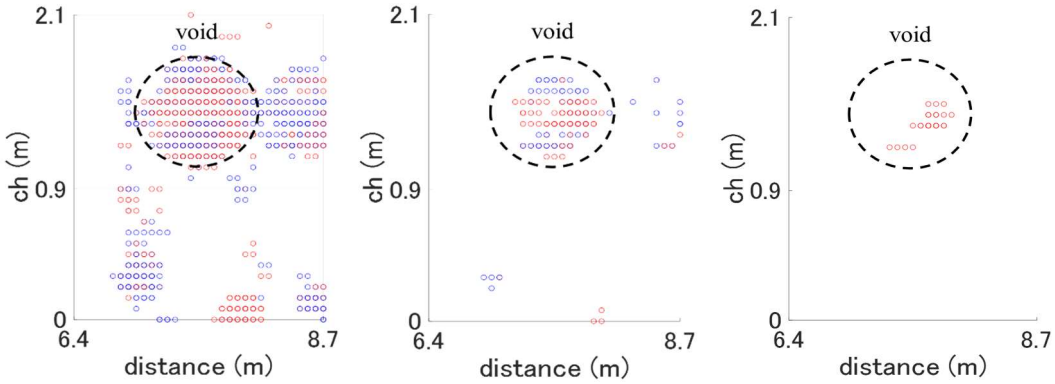


Fig. 7.23 Extracted void area (a) $T = 0.3$ (b) $T = 0.5$ (c) $T = 0.7$

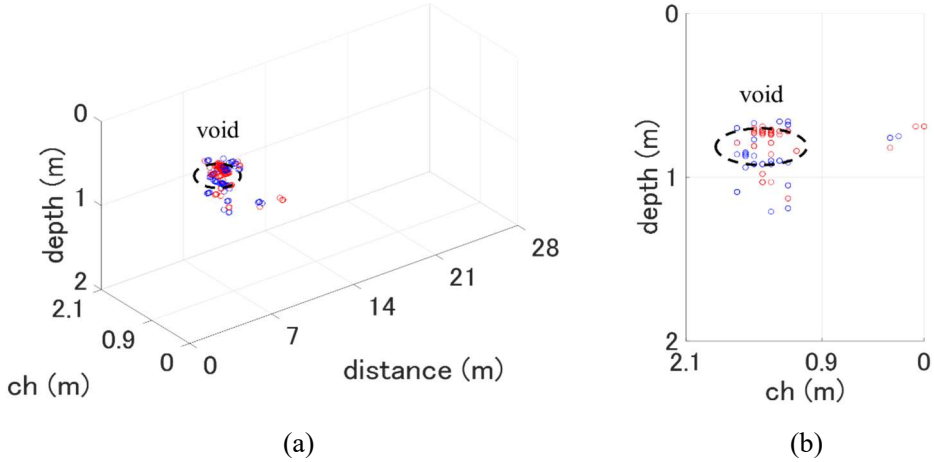
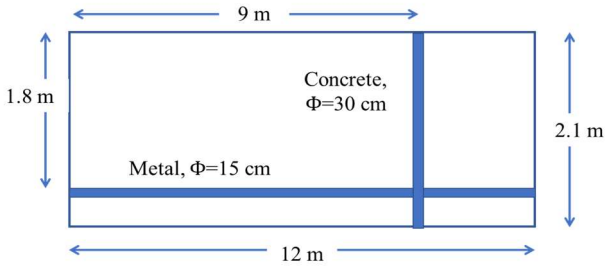
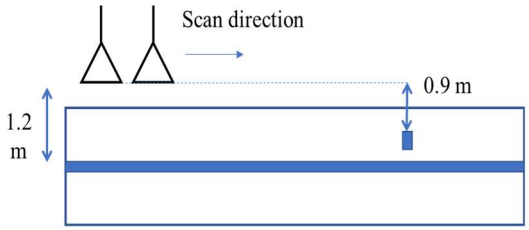


Fig. 7.24 3D map of void after migration (a) 3D map (b) side view

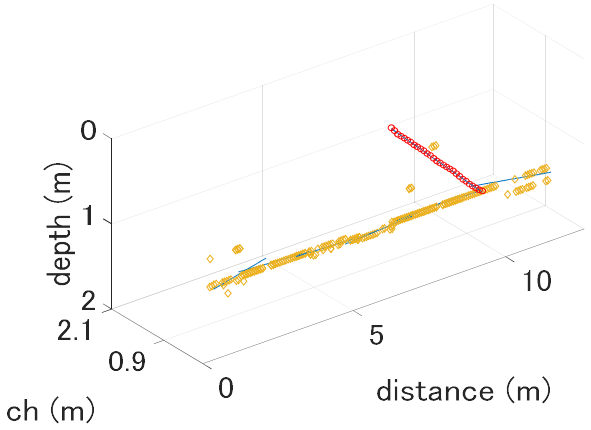


(a)

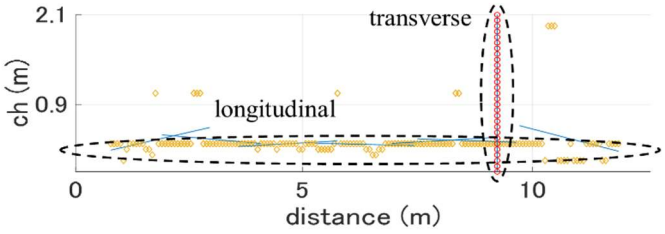


(b)

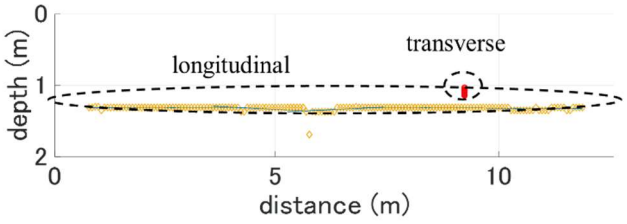
Fig. 7.25 Geometry of buried pipes I (a) (b) top and side views



(a)

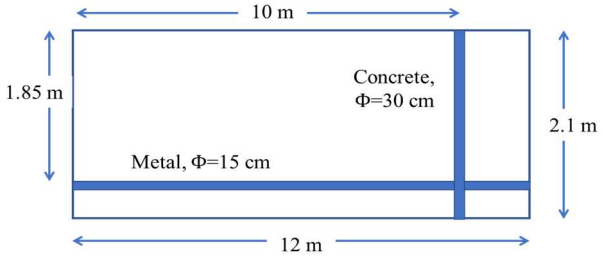


(b)

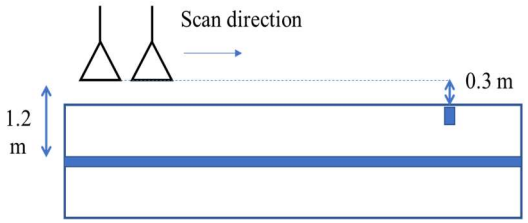


(c)

Fig. 7.26 Detection results of pipes I (a) (b) (c) 3D map, top and side views

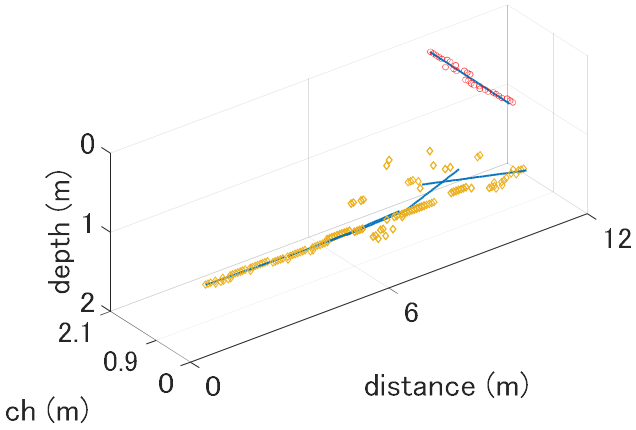


(a)

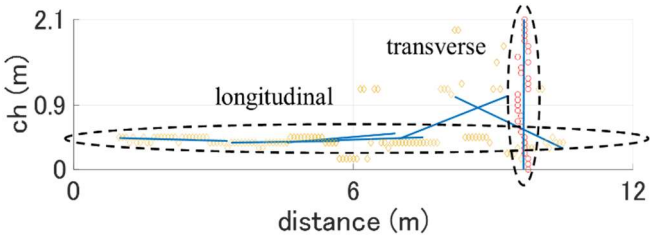


(b)

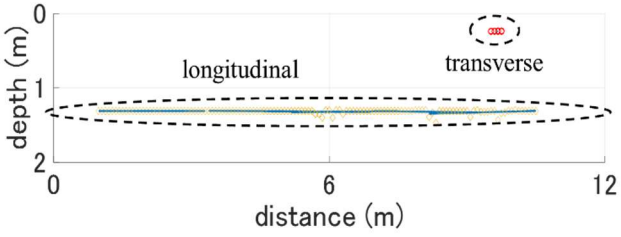
Fig. 7.27 Geometry of buried pipes II (a) (b) top and side views



(a)

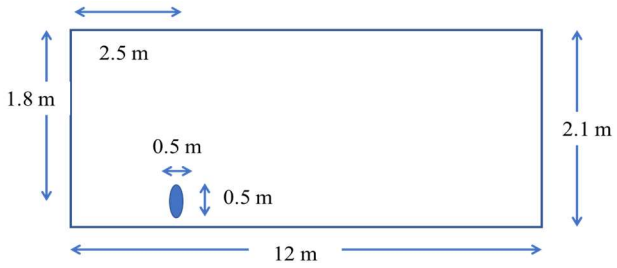


(b)

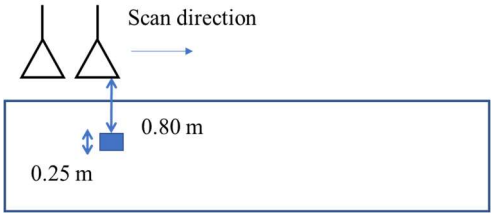


(c)

Fig. 7.28 Detection results of pipes II (a) (b) (c) 3D map, top and side views

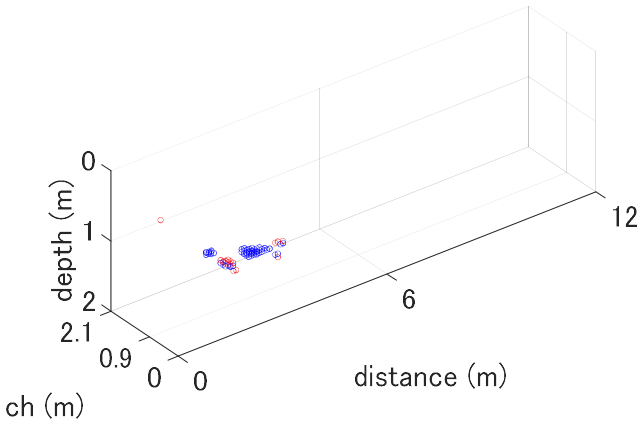


(a)

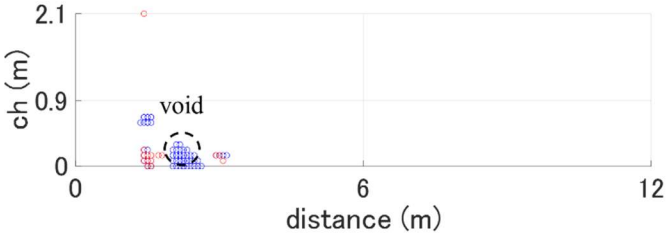


(b)

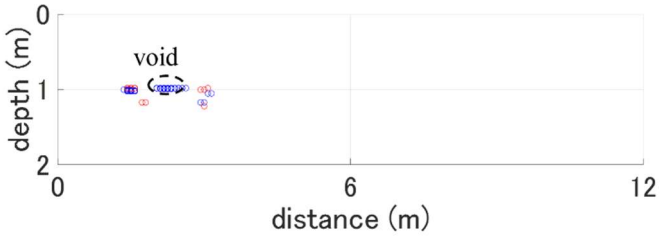
Fig. 7.29 Geometry of void I (a) (b) top and side views



(a)

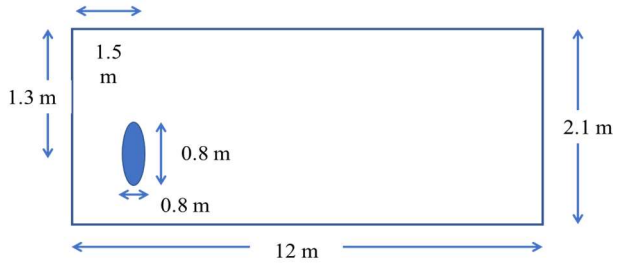


(b)

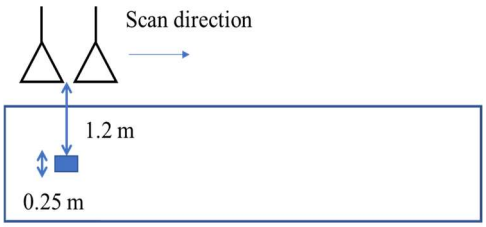


(c)

Fig. 7.30 Detection results of void I (a) (b) (c) 3D map, top and side views

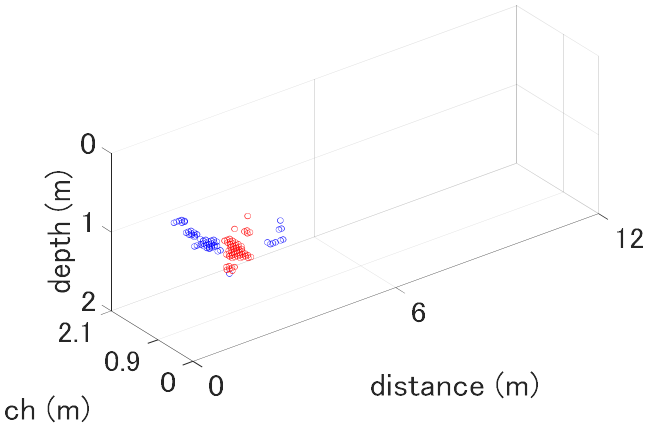


(a)

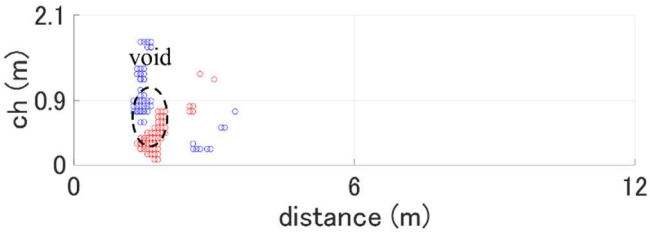


(b)

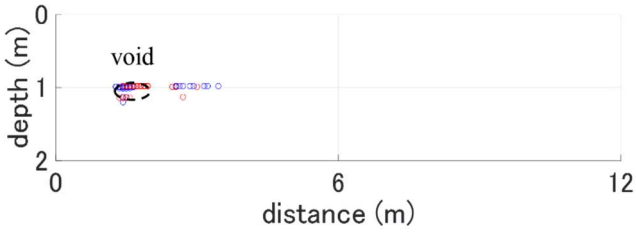
Fig. 7.31 Geometry of void II (a) (b) top and side views



(a)



(b)



(c)

Fig. 7.32 Detection results of void II (a) (b) (c) 3D map, top and side views

transverse pipes. Outliers may be due to the effects of embedded Styrofoam outside the region.

In the experimental field case, the estimation accuracy can quantitatively be evaluated. The scan and channel direction errors were within 20 cm. The depth error was also within 20 cm in the case of Fig. 16 assuming the relative permittivity of soil is 3. However, to obtain accurate depth, permittivity should be estimated from radar images, which is left as future work. The estimation accuracy was enough from a practical point of view because actual pipes' diameter is several cm to several tens cm. Total calculation time was less than five minutes for 1 km section because of the combination of 3D-CNN (AI) and Kirchhoff migration (inverse analysis). Ordinary migration methods need several days to analyze 1 km section. Calculation speed is extremely high compared to conventional migration research framework.

7.4.2 Void detection results

The configuration of the experimental field void data was shown in Section 2.4.2. Fig. 7.29 - Fig. 7.32 show the results of two cases. Fig. 7.29 and Fig. 7.31 are the geometry of the void. Fig. 7.30 and Fig. 7.32 show the 3D maps. From Fig. 7.30 and Fig. 7.32, the void was visualized by the clusters of the detected points. The void was plotted in appropriate depth, distance and channel. The proposed algorithm is applicable to any data with the same measurement system irrespective of the size and depth of void. The experimental field of void consists of many pipes and void in small areas. Some missing points and false detection may be due to the effects of other buried objects.

7.5 Summary

The methodology to estimate the positions of manhole, joint, pipes and void by Hough transform and Kirchhoff migration was proposed to visualize the objects in a 3D map. The developed SVM and 3D-CNN models were applied to the whole data by box-by-box search. Box-by-box search is the simplest localization method. Overlap and probability threshold of each category were the parameters. In the research, parameters were optimized for several cases. Objects were missing when overlap was too small because the shift of images was not allowable for the detection by CNN. The overlap 75 % was enough for detecting pipes considering calculation cost. Too small probability threshold causes false detection. On the other hand, too large probability threshold leads to missing pipes. These parameters depended only on the measurement system. The parameters were also optimized for manhole, joint, longitudinal pipes and void.

Manhole and joint were detected by Hough transform. Hough transform detects the circles of manhole and lines of edges of joint. The positions of the targets and diameter of manhole were accurately estimated by several to several tens cm error. The positions of pipes were estimated by Kirchhoff migration, which is the most fundamental and accurate migration algorithm. The direction of pipes was estimated by 3D-CNN. The cross-section images perpendicular to a pipe were migrated to estimate the pipes' positions and suppress other clutters and noise. The relative permittivity of soil and synthesis aperture length were the parameters. In the research, the relative permittivity of soil was assumed 3. Synthesis aperture length depends on the antennas, which was tuned by several measurement data cases. The peak of maximum absolute intensity was extracted as the upper surface of the pipe to plot in a 3D map. The positions, directions and inclination of the

pipes were successfully visualized as the aligned points and approximated lines. The variation of the points from the approximated lines was within 20 cm, which was below the diameter of pipes. The void was visualized by extracting areas over certain absolute intensity after migration. Parameters depend only on the measurement system, which were optimized for several measurement data cases. The void was represented by the clusters of detected points in certain depth, distance and channel.

The proposed algorithm was validated by the experimental field data. From the results, the 3D arrangement of the pipes was accurately visualized with several to several tens cm error. The void was also visualized by the clusters of detected points. The measurement and experimental field data consist of pipes of various material and diameter and void of various sizes in various depth. Therefore, the algorithm is applicable to any data with the same measurement system, which is useful from a practical point of view. The following chapter summarizes the proposed algorithm with a flow chart. The discussions were provided in terms of calculation time. Expected impacts and future work were also pointed out.

Chapter 8: Conclusion

8.1 Conclusions

8.1.1 Objective

Aging of subsurface infrastructures is becoming an important social issue. Unknown utility pipes cause construction delays. Sometimes damaged pipes bring economic loss to surrounding areas. For road administrators, proper understanding of the existence and positions of subsurface pipes is important. Deteriorated pipes cause void, leading to a road collapse. Early detection of void enables road administrators to take countermeasures.

This research focusses on GPR method. The system is highspeed, up to 80 km/h. Therefore, GPR method does not need traffic regulation. The system scans under the road with dense measurement points. Target objects include the both subsurface pipes and void. The problem is, radar images are usually difficult to interpret. Obtained data is 3D. Total volume of all the channels is up to several TB for one measurement of several tens km section. Skilled inspectors are trying to detect buried objects and anomalies manually from the whole amount of data. Usually, one inspection takes several months and million yen (tens thousand dollar). Therefore, an automatic and highspeed detection algorithm may reduce significant inspection time and cost. In terms of void detection, the number of measurement data by GPR method is few. Therefore, training data should be produced by electromagnetic simulations.

In terms of previous research of subsurface pipes, migration and inverse analysis need high calculation cost. Therefore, it is not feasible to simply apply these methods on the whole measurement data. The detection rate of hyperbolic reflection patterns by Hough transform, simple 1 layer NN and other pattern recognition approaches was about 70 - 80 %. The accuracy of previous research is not considered to be enough. Utilizing the 3D features of reflection patterns and deep layer architectures may improve the results. In terms of void, the classification accuracy of an experimental field model was about 80 %. The accuracy may increase also in this case by utilizing the 3D reflection patterns of void. However, even by GPU, producing several hundred training data by 3D-FDTD is not feasible.

The research objective is the development of an automatic 3D subsurface pipes and void detection algorithm by deep learning from radar images. The methodology to reproduce 3D void data by 2D-FDTD was proposed. The automatic detection algorithm of manhole and joint from surface reflection by SVM was developed to remove from the target regions. Then, the reflection patterns of pipes and void was accurately detected by a deep 3D-CNN model. The developed model can estimate the existence and directions of transverse and longitudinal pipes by three-category classification. After the detection by 3D-CNN, migration method was applied to the detected regions to estimate the pipes' positions highspeed and accurately.

8.1.2 Data and system

In the research, a vehicle-installed type multi-channel system was utilized. The system records a position

by GPS and odometer. 3D-Radar adopts SF-CW method. An operating frequency range and step width are optimized considering the desired resolution, maximum depth and vehicle speed. In the research, GPU computing environment was developed to train a deep layer model and conduct electromagnetic simulations. MATLAB was used to produce training data and plot detection results. gprMax was used for electromagnetic simulations by FDTD method. For a training program, TensorFlow was adopted.

In the research, about 3,000 pipe data measured in Nagano prefecture, 88 void data in Aichi and Nagano prefecture and 8,000 manhole and joint data in Nagano prefecture were used. The accuracy of the developed classification algorithms was evaluated by the inspection results because the objective of the research is to develop a reliable algorithm which can replace the decision of skilled inspectors. Training data of void was produced by electromagnetic simulations. From the results of the excavation survey, ellipsoid and cuboid bodies were assumed with the randomness of depth, channel number, horizontal length and thickness in a depth direction. In the research, the experimental field site of pipes was newly constructed by the authors in Gunma prefecture. The site consists of transverse and longitudinal pipes with various types of material, diameter and depth to reproduce the real 3D arrangement of pipes. The experimental field site of void consists of ellipsoid and cuboid Styrofoam with various sizes buried in soil at various depth. The objective of the experimental fields is to quantitatively evaluate the estimation accuracy of the positions of pipes and void by the proposed algorithm. The void data was also utilized to validate simulation models.

In the research, because the number of each category was comparable, accuracy, precision and recall show similar values. AUC was also evaluated in all the cases. Classification accuracy converged in the cases of manhole, joint and transverse pipes in terms of the number of training data. In the case of longitudinal pipes, there is a possibility that classification accuracy may increase by several % though the number of the training data will not affect the conclusions. Cross validation was conducted in the cases of manhole, joint and void. Cross validation was necessary especially in the case of void because the number of test data was limited.

8.1.3 Proposed algorithm

The proposed algorithm consists of radar data collection and preprocessing; manhole and joint detection by SVM; pipes and void detection by 3D-CNN; 3D mapping by Kirchhoff migration (Fig. 8.1). After data collection by the measurement system, preprocessing was conducted. Manhole and joint were detected from surface reflection by SVM. Then, pipes and void were searched by 3D-CNN in the region without manhole and joint. In the case of pipes, the positions were further estimated by Kirchhoff migration. Each object was visualized in a 3D map. Total calculation time was less than 5 minutes for 1 km section with GPU by the combination of 3D-CNN (AI) and Kirchhoff migration (inverse analysis) as shown in Table 8.1. If migration method is applied to the whole data, calculation time will increase by the square of area of images. It is expected to be several tens hours (several days) for 1 km section. The proposed algorithm was extremely high-speed. Below are the characteristics of each algorithm component.

Considering the fundamental electromagnetic characteristics of subsurface sensing environment, the basic preprocessing algorithms were studied. In the research, direct coupling and transmitted impulse waves were measured to remove the effects from measured signals. Raw data consists of real and imaginary parts of

Fourier coefficients at each scan and frequency. After IFFT and other fundamental preprocessing algorithms, data was converted to 3D time domain data with depth, channel and scan dimensions.

The proposed SVM model consists of three steps: preprocessing filter, feature extraction, application of the SVM model. A Laplacian filter was applied to input images to enhance the edges of manhole and joint. Then, HOG features of input images were calculated. A nonlinear dividing plane of an SVM model was introduced by an RBF kernel. By utilizing HOG features and an RBF kernel, the edges of manhole and joint were efficiently extracted. Multi-class classification was conducted by ECOC method. The output of the SVM model was the probability of the three categories: manhole, joint and pavement section. The total classification accuracy of the three categories was about 98 %. The developed SVM model was accurate. The optimized SVM and 3D-CNN models showed the same level performance, though the calculation cost of the SVM model was much lower than the CNN model owing to its simple model structure.

In terms of void detection, the methodology to reproduce 3D void data by 2D-FDTD method was proposed. Among the electromagnetic simulation methods, FDTD method was adopted in the research. Simulation conditions such as the size of cells, time step width and time duration were appropriately determined. A PML boundary was utilized for absorbing boundary condition. Measurement conditions such as antenna distance and height were according to the actual conditions. Point source was utilized. The void model consists of air, asphalt and soil layers and void. In the research, referring to the previous research, the soil layer with 2 - 20 relative permittivity and 0.001 s/m - 0.01s/m conductivity randomness was assumed. The void was ellipsoid and cuboid bodies, whose range of depth and horizontal direction lengths and thickness are 0.2 m - 0.5 m,

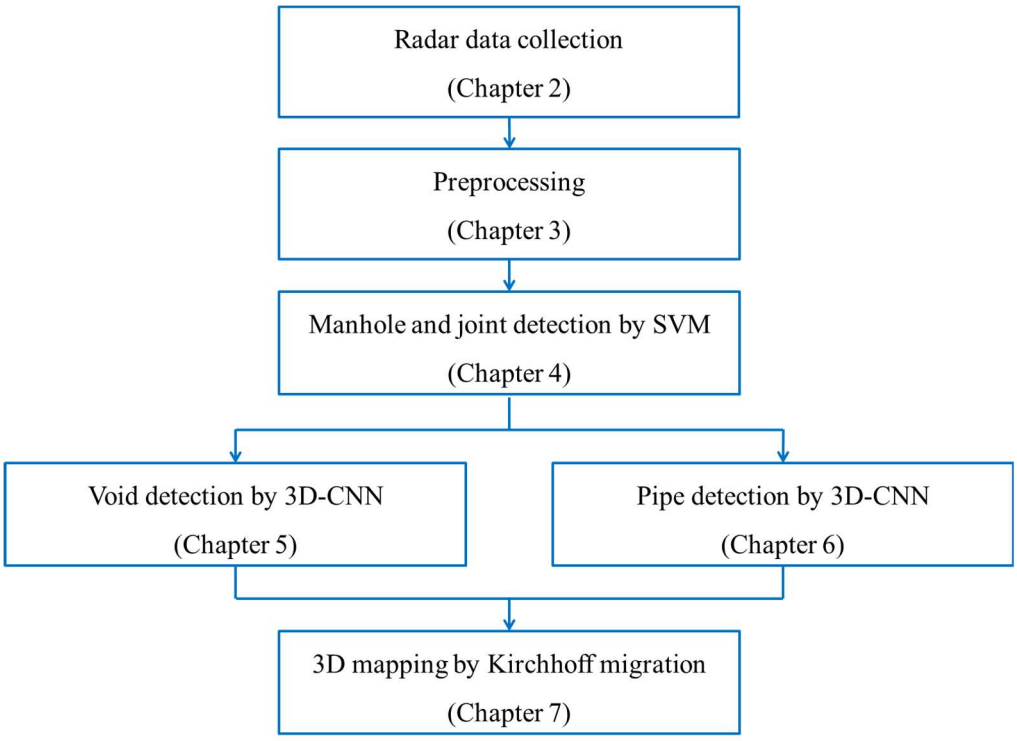


Fig. 8.1 Flow chart of proposed algorithm

0.5 m - 1.5 m, 0.1 m - 0.5 m respectively. These parameters were derived from the excavation survey results. The sections of the reflection patterns of a 3D model were approximated by 2D models. Then, the 3D reflection pattern was reconstructed by linearly interpolating the sections in cylindrical coordinates. Even with 22.5° angle pitch the reflection pattern was accurately reproduced. The models were validated by the measurement and experimental field data. From the discussions of previous research, a 6 layer CNN model was adopted. Transfer learning was utilized. The classification accuracy of the 3D-CNN model was high. It was about 87 %, by about 9 % increase compared to a 2D-CNN model. By 3D convolution filters, classification accuracy increased by about 6 %, indicating the 3D-CN model utilizes the 3D features of the reflection patterns.

In terms of pipe detection, a classification model by 3D-CNN was developed. Parameters were updated by gradient descent method. Minibatch training and early stopping were adopted. The variation of test accuracy at each training was at most 2 %. The preprocessing filters such as a gain filter and BGR did not improve the classification accuracy of the pipes. CNN automatically learns the spatial features by convolution filters. The highest classification accuracy, about 83 % - 89 %, was obtained by a 2 layer simple custom model (1 convolution and 1 fully connected layers) in the both cases of transverse and longitudinal pipes. Classification accuracy increased compared to a 1 layer NN model because of the utilization of convolution filters. Classification accuracy also increased compared to an SVM model because the performance of the model improved by deep layer architectures. In terms of 2D data, Because of the difference of radar and channel characteristics, the reflection patterns of transverse and longitudinal pipes were not the same. There were required sizes of input data and convolution filters to extract features for each direction pipe. In terms of 3D data, 3D-CNN utilizes 3D input data and convolution filters to handle 3D spatial features of the data. Considering the lengths and inclination of pipes, there were optimal 3D sizes for each direction pipe. Classification accuracy increased to 85 % - 95 %, by about 3 % - 6 % increase utilizing the 3D features of

Table 8.1 Calculation time of proposed algorithm

Calculation step	CPU (s/km)	GPU (=CPU/5, s/km)
Data collection and preprocessing	300	60
Manhole and joint detection by SVM	100	20
Pipes and void detection by 3D-CNN	1000	200
3D mapping by Kirchhoff migration	10 - 100	2 - 20
Total	1500	300

the pipes' reflection patterns. The optimal parameters were integrated to consider 3D regions with 1 m depth by 2 m scan distance by 2 m channel distance. The three-category classification, transverse, longitudinal pipes and no pipe section model by 3D-CNN was developed. To estimate the directions of pipes, 3D convolution filters were adopted. The total classification accuracy of the three categories was about 91 %. The developed 3D-CNN model accurately estimates the existence and directions of pipes.

The regions of manhole, joint, pipes and void were localized by the developed SVM and 3D-CNN models by box-by-box search. The parameters were overlap in each direction and probability threshold of each category. Overlap was determined not to miss target objects and reduce calculation cost. Probability threshold was determined not to miss target objects without false detection. These parameters were optimized for the several cases of the measurement data and validated by the experimental field data. Therefore, the parameters are in common when the target objects and measurement system are the same. The problem is, especially in terms of pipe case, the positions of pipes in detected regions should further be estimated from a practical point of view. In the research, Kirchhoff migration was adopted as one of the most accurate and simplest migration algorithms. The direction of a pipe was already estimated by 3D-CNN. The cross-section images perpendicular to the pipe was extracted. Maximum peaks were detected from the images after migration as the upper surface of the pipe. The pipes were visualized by the aligned points in each direction and approximated lines. After migration, the variation around the lines decreased from 40 cm to at most 20 cm. Even the inclination of the pipes was successfully estimated. From the experimental field results, the pipes of various material and diameter were detected by the proposed method. The 3D arrangement of the pipes was accurately visualized. The estimation error of the positions was 10 cm - 20 cm assuming the appropriate relative permittivity of soil. The estimation accuracy was high enough from a practical point of view because actual pipes' diameter is several cm to several tens cm.

8.1.4 Parameters

Table 8.2 shows the important parameters related to the proposed algorithm. In terms of data collection and preprocessing, a frequency range, frequency pitch and scan pitch were determined considering the tradeoffs between the required maximum depth, resolutions and vehicle speed. Depth limitation may be applied to reduce calculation cost only when target objects are in shallow depth. A channel number and length may be troublesome parameters. In the research, it was pointed out that increasing the number of channels may increase the classification accuracy of the 3D-CNN model. However, increasing the number decreases the vehicle speed because the required time to sweep electromagnetic waves is proportional to the number of channels. The optimized parameters were implicitly limited by the utilized hardware, which was not discussed in the thesis. The hardware may further be customized for the proposed algorithm.

In terms of the detection of manhole, joint, pipes and void, probability threshold was optimized for each target object. The required size of input data was derived maximizing classification accuracy. The size of input data can be increased if calculation time is not a bottleneck. For example, if the number of channels becomes 1.5 times larger, calculation time also becomes 1.5 times larger, which may be possible because the proposed algorithm was highspeed. The overlap in scan and depth directions strongly affect calculation time.

However, overlap may not be changed unless target objects or characteristics of the measurement system are totally changed. Migration parameters such as synthesis aperture length and soil permittivity depend on the actual conditions of soil. Therefore, these parameters should be close to measurement environment.

8.1.5 Expected Impacts

The output of the proposed algorithm is a 3D map of various objects including manhole, joint, pipes and void. As stated in Section 8.1.1, radar images are difficult to interpret. Skilled inspectors manually detect anomalies from images one by one. In the proposed algorithm, the AI models already trained by inspection results are stored in PC. After road administrators or consulting company persons collect measurement data, the positions of the objects are automatically plotted in a 3D map by the proposed algorithm. Especially, in terms of pipes, positions and inclination are visualized by approximation lines. Obtained 3D map is realistic and has much useful information for road administrators who are planning a road construction. This information cannot be obtained by visual inspection or other existing algorithms, which may have a great impact on the practice of subsurface sensing of infrastructure utilities and anomalies.

Calculation time is about 5 minutes for 1 km section, which is extremely highspeed compared to the migration methods of previous research. The time corresponds to about 10 km/h, which enables on-site analysis for further measurement and decision making. If the measurement conditions are further customized as discussed in Section 8.1.4, calculation time may further be reduced, for example by half or one third. The

Table 8.2 Parameters of proposed algorithm

Calculation step	Parameters
Data collection and preprocessing	- frequency range - frequency pitch - scan pitch - channel number / length - vehicle speed / time duration - depth range
Manhole and joint detection by SVM	- input data size - scan overlap - probability thresholds
Pipes and void detection by 3D-CNN	- input data size - scan / depth overlap - probability thresholds
3D mapping by Kirchhoff migration	- range of manhole diameter / joint length - migration area (length of synthesis aperture) - soil permittivity - threshold for void visualization

final goal is real-time analysis by the proposed algorithm.

8.2 Future work

About future work of the research, three topics were pointed out in this section. Some topics are related to the accuracy improvement of the algorithm. Other topics are related to the additional analysis which may be useful from a practical point of view.

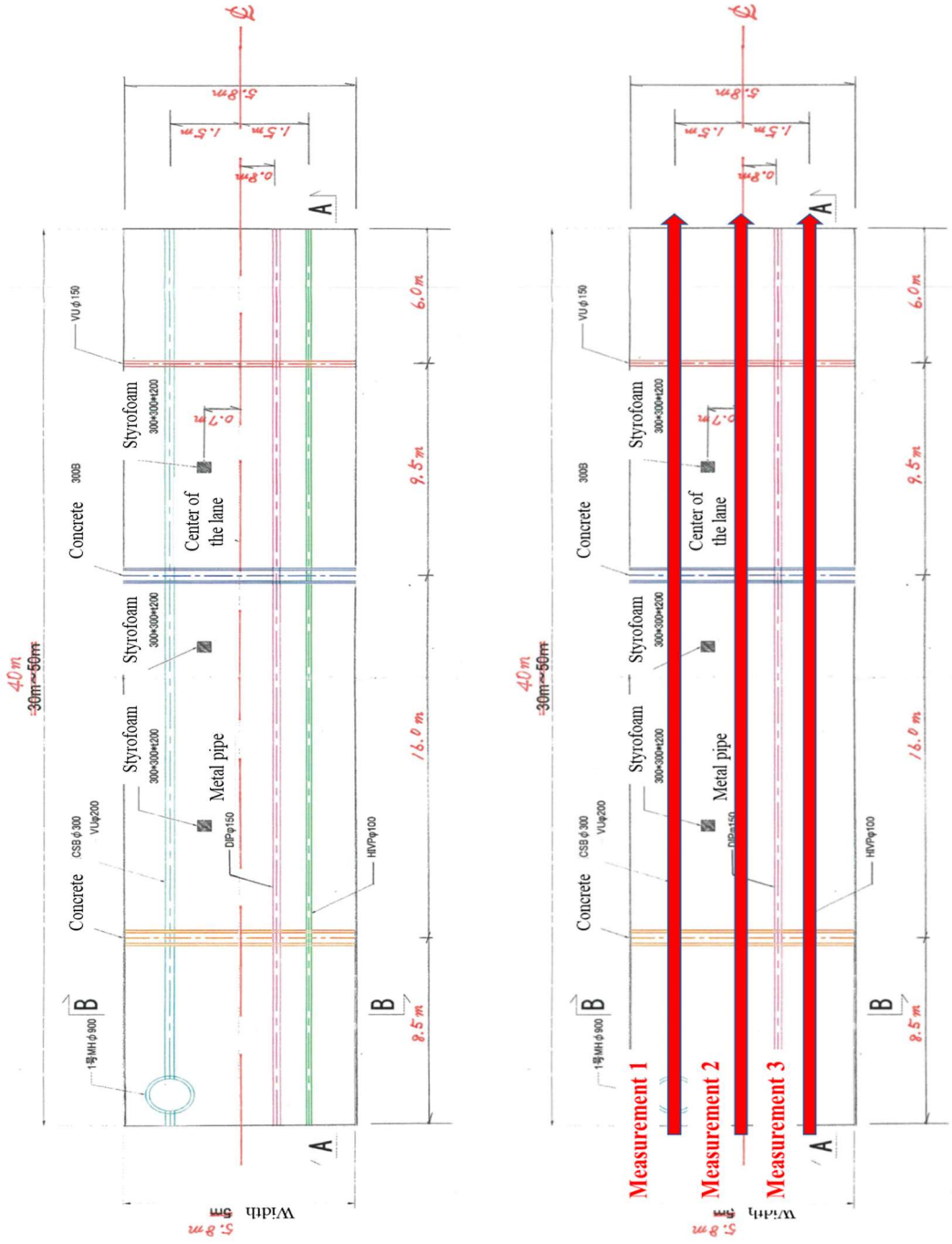
The first topic is the estimation of the permittivity of soil, material and diameter of subsurface pipes from radar images. In the research, the reflection patterns of various types of soil and pipes were included in the training data. The AI models learn these patterns one by one to deal with unknown input data. However, the positions, shapes and amplitude of hyperbolic reflection patterns were theoretically related to the positions and diameter of pipes and permittivity of soil and pipes. These parameters may be estimated in migration process or by the 3D-CNN models to help the migration process. The parameters provide the useful information of target objects. The estimation algorithms were discussed in previous research. The existing algorithms may directly be appliable to the detected regions by 3D-CNN.

The second topic is about the advancement of the proposed algorithm. In recent years, many up-to-date deep learning models and localization algorithms were proposed. In this research, VGG16 and GoogLeNet (InceptionV3) were considered. Other novel architectures such as ResNet, VGG and Inception series may be considered. State-of-the-art localization algorithms such as faster R-CNN and SSD may improve the localization accuracy. However, as discussed in Section 7.1.1, these algorithms are targeted for colored scenery photos and do not correspond to 3D data. Some improvement may be needed to apply these algorithms. Other state-of-the-art migration algorithms such as inverse scattering analysis and tomographic reconstruction may improve the results.

The third topic is the application of the proposed algorithm to large measurement data. The training data was collected in wide areas of Nagano and Aichi prefectures. For a future plan, large measurement data may be collected to demonstrate the performance of the algorithm. The proposed algorithm is expected to be applicable to unknown data of other areas. The accuracy of an AI model depends on the amount and quality of training data. Especially in terms of void, utilizing larger amount of training data may further improve the accuracy even though the amount of data is still not enough without simulations.

Appendix A: Experimental field site

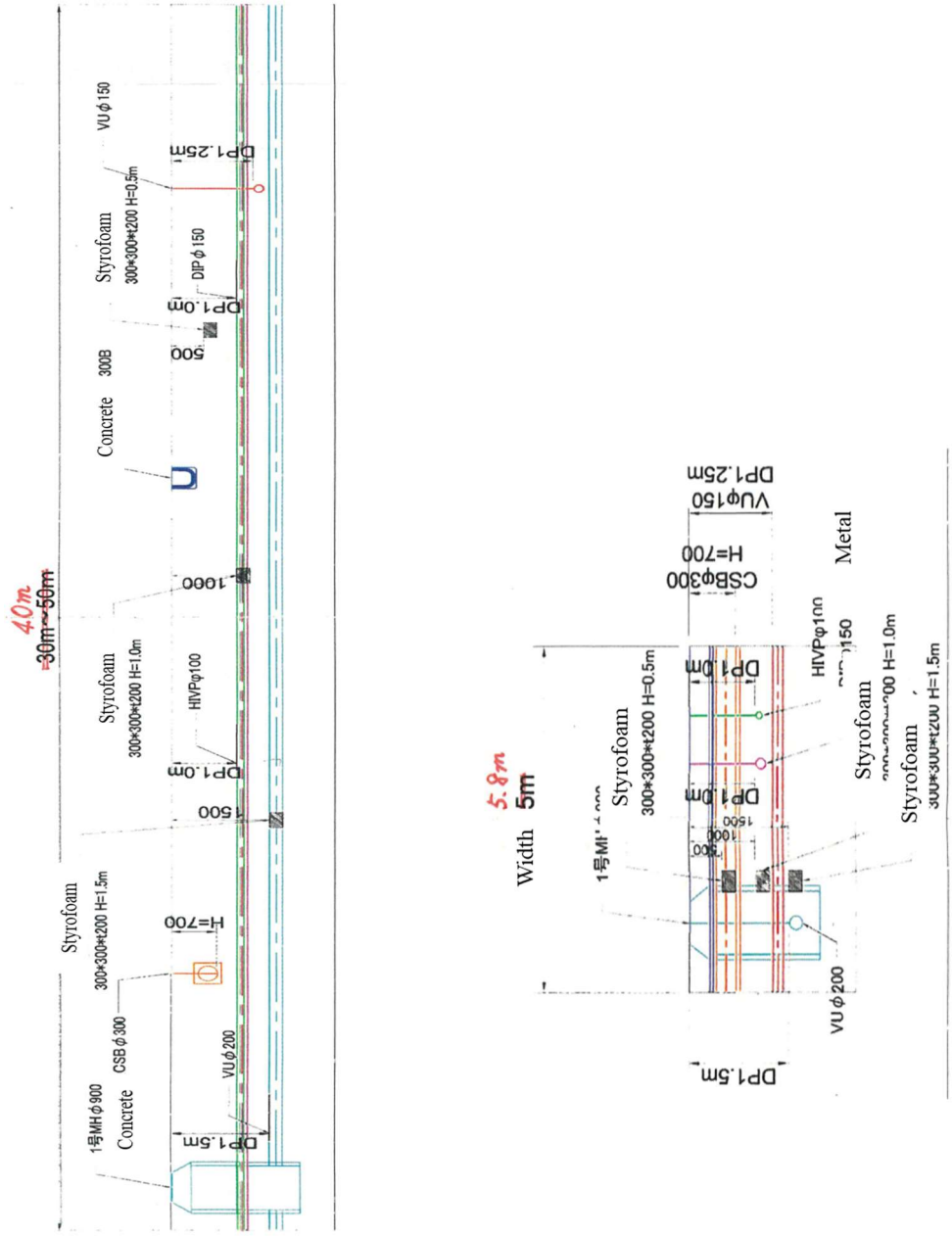
Fig. 1 and Fig. 2 show the design drawings of the experimental field site of pipes. Fig. 1 also shows the three measurement courses. Fig. 1 shows the top view of the site. Fig. 2 shows the elevation view of



(a) (b)

Fig. 1 Design drawings (a) top view (b) measurement course

transverse and longitudinal sections. Material and diameter of the pipes are also written, though the estimation of these parameters is out of scope of the research. There are three transverse pipes: two concrete pipes in 0.7 m depth and just below the road surface with 30 cm diameter; one plastic pipe in 1.25 m depth with 15 cm diameter. There are three longitudinal pipes: one plastic pipe in 1.5 m depth with 20 cm diameter, one metal pipe in 1.0 m depth with 15 cm diameter and one plastic pipe in 1.0 m depth with 10 cm diameter. The effects of these parameters on detection accuracy was discussed in Section 7.4. The configuration is according to the practical situations of subsurface pipes.

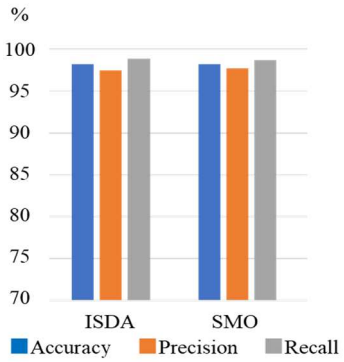


(a) (b)

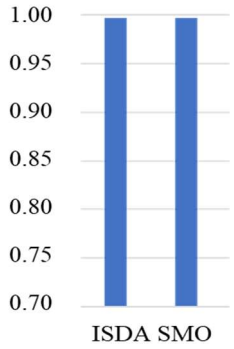
Fig. 2 Elevation views (a) transverse (b) longitudinal

Appendix B: SVM parameters

Optimization method of SVM did not affect classification accuracy and training time in the research. Fig. 1 and Fig. 2 show the effect of optimization method on manhole and joint classification. Classification accuracy was almost the same. The training time of SMO was slightly smaller than ISDA. SMO accelerates training by utilizing the analytical solution of a set of two data. However, total training time was at most 5 minutes for one case. The effect of SMO was limited because the total amount of the training data was small. Larger the volume of the data is, larger the effect of SMO becomes.



(a)

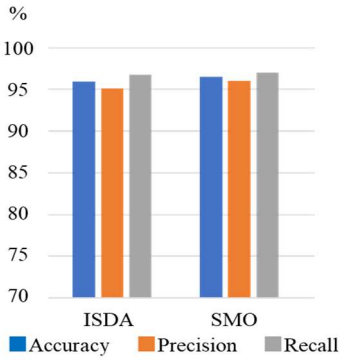


(b)

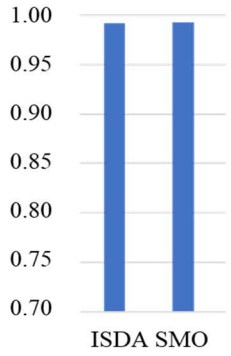
ISDA	271 s
SMO	264 s

(c)

Fig. 1 Comparison of optimization method (manhole) (a) Accuracy, Precision, Recall (b) AUC (c) training time for one case



(a)



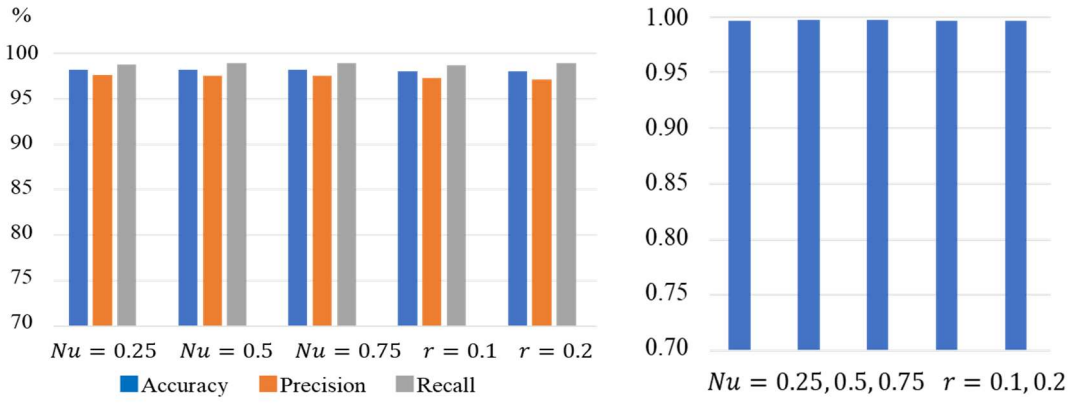
(b)

ISDA	55 s
SMO	54 s

(c)

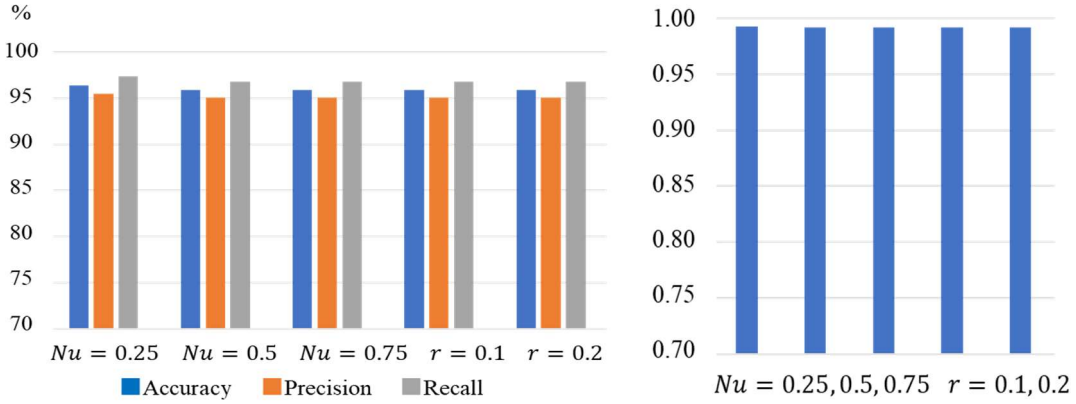
Fig. 2 Comparison of optimization method (joint) (a) Accuracy, Precision, Recall (b) AUC (c) training time for one case

The effects of regularization factor and outlier removal were also negligible. Fig. 3 and Fig. 4 show the effects of regularization factor N_u and outlier rate r on classification accuracy. Fig. 3 shows the cases $N_u = 0.25, 0.5, 0.75$ at $r = 0$. Fig.4 shows the cases $r = 0.1, 0.2$ at $N_u = 0.5$. From Fig. 3 and Fig. 4, classification accuracy did not depend on N_u and r. Overfitting may not occur in the research. The number of outlier data may be small. Generally, only support vectors near a diving hyperplane affect the results. Therefore, neither regularization nor outlier removal improved the results. In the research, regularization factor $u = 0.5$ was adopted. Outlier removal was not conducted.



(a) (b)

Fig. 3 Effects of regularization factor and outlier removal (manhole) (a) Accuracy, Precision, Recall (b) AUC



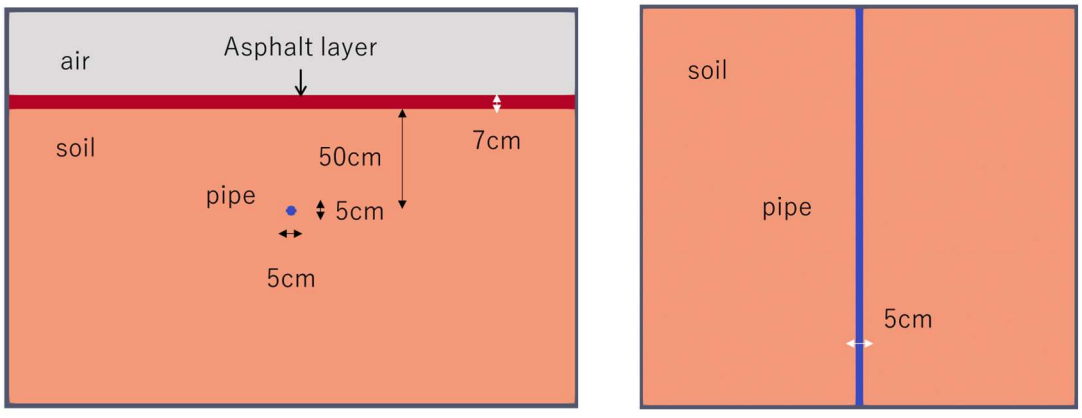
(a) (b)

Fig. 4 Effects of regularization factor and outlier removal (joint) (a) Accuracy, Precision, Recall (b) AUC

Appendix C: Pipe simulations by FDTD method

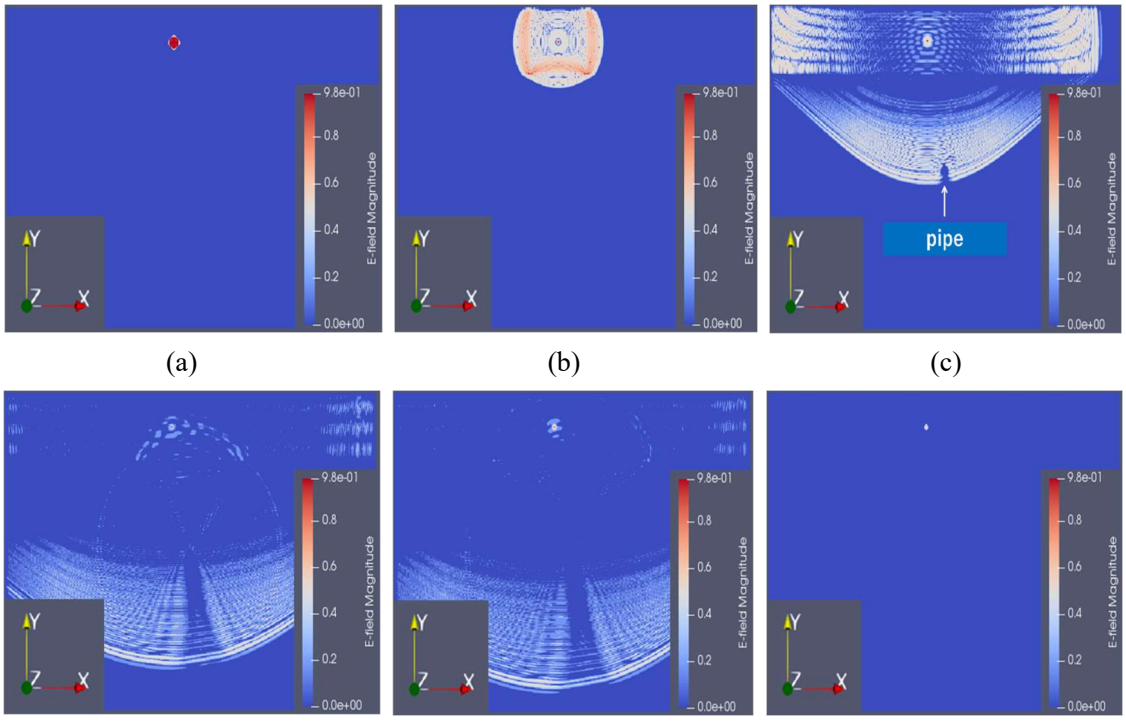
In this appendix, the effects of geometry of pipes and electromagnetic characteristics of soil on reflection patterns are discussed by FDTD method. In the research, training data of pipes was not produced by FDTD method because there were enough training data. This appendix is expected to help the discussions of Chapter 6 and Chapter 7. Fig. 1 shows a typical 3D-FDTD pipe model. The pipe model consists of a 40 cm air layer, 7 cm asphalt layer and 50 cm soil layer, the same as the void model. Pipes were made of metal, plastic or concrete. Diameter was assumed 5 cm for metal and plastic pipes and 50 cm for a concrete pipe considering typical cases of subsurface pipes. Inclination of pipes was assumed 0° considering the discussion of Section 2.3.2. The longitudinal section of the 3D model (Fig. 1 (a)) was the same as the 2D model. Therefore, the difference of the 2D and 3D models was the propagation of electromagnetic waves, plane or spherical waves. A PML layer, transmitting and receiving antennas were the same as the void model. Considering the consistency with the measurement data, phases of signals was reversed. Therefore, the upper surface of a metal pipe shows a positive peak and plastic pipe a negative peak.

Fig. 2 and Fig. 3 display the snapshots of 3D-FDTD results of the metal pipe. Each figure shows the electric field of longitudinal and horizontal sections. Longitudinal and horizontal sections pass through the middle of the pipe in a channel direction and center axis of the pipe respectively. Waves reached the surface of an asphalt layer in Fig. 2 (b) and pipe in Fig. 2 (c). The maximum peak reflected from the surface of the pipe reached the antennas in Fig. 2 (d). From Fig. 2 (d), waves diffused from the pipe in a circular shape. Small diameter metal pipes are considered to be a scatterer in a longitudinal section. After sufficient time, the intensity of reflected waves was much weaker in Fig. 2 (f). In a horizontal section, waves reached the surface of the pipe in Fig. 3 (c) and diffused in Fig. 3 (d), (e). Waves cannot penetrate through the metal pipe. The snapshots of the 2D model was similar to the 3D model.



(a) (b)

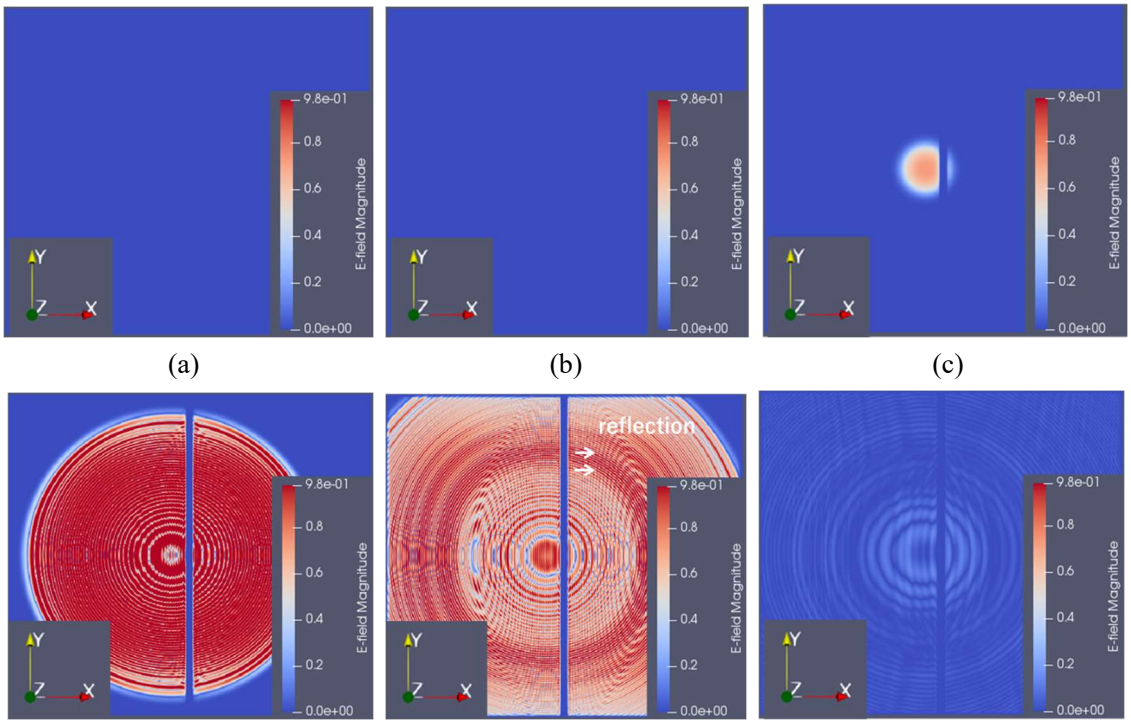
Fig. 1 3D pipe model (a) longitudinal section (b) horizontal section



(a) (b) (c)

(d) (e) (f)

Fig. 2 Snapshots of 3D reflection pattern of metal pipe (vertical section) (a) 0 s (b) 2.5×10^{-9} s (c) 5.0×10^{-9} s (d) 7.5×10^{-9} s (e) 10×10^{-9} s (f) 20×10^{-9} s



(a) (b) (c)

(d) (e) (f)

Fig. 3 Snapshots of 3D reflection pattern of metal pipe (horizontal section) (a) 0 s (b) 2.5×10^{-9} s (c) 5.0×10^{-9} s (d) 7.5×10^{-9} s (e) 10×10^{-9} s (f) 20×10^{-9} s

The difference of 2D and 3D reflection patterns was negligible because the pipe model was uniform in a transverse direction. As shown in Fig. 4, reflection patterns were almost the same. Furthermore, as shown in Table 1, the difference of peak values of 2D and 3D reflection patterns was at most 12 %. Difference of 2D and 3D models was smaller compared to the void case shown in Table 5.2 of Section 5.4.1 because the void model has geometry also in a transverse direction.

The relative permittivity ε_s and conductivity σ_s of soil have large effects on reflection patterns. Fig. 5 shows the results of the 2D metal pipe case. From Fig. 5 (a) and (b), compared to the case $\varepsilon_s = 2$, the reflection pattern delayed in the case $\varepsilon_s = 10$ because of the decrease of propagation velocity. The pattern showed the narrower hyperbolics because of the increase of relative permittivity. Peak values did not change

Table 1　Comparison of 2D and 3D models

	3D model Peak value $\times 10^{-4}$	2D model Difference (%)
1/2 point peak	4.71	-9.8
1/4 point peak	2.29	+11.8

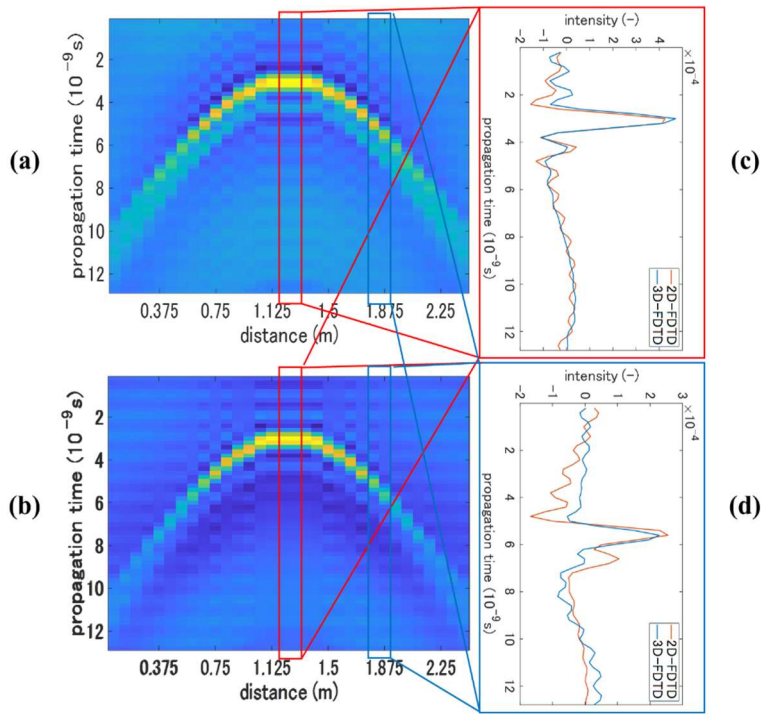


Fig. 4　Comparison of 2D and 3D reflection patterns (a) 2D model (b) 3D model (c)1/2 point signal (e) 1/4 point signal

irrespective of the relative permittivity of soil because the relative permittivity of metal pipes is infinitely large. Fig. 5 (c), (d) show the effect of σ_s. Because of the attenuation of input waves, a reflection pattern was weaker in the case $\sigma_s = 0.01$ s/m compared to $\sigma_s = 0.001$ s/m.

Depth d and diameter t of the pipe also have large effects on reflection patterns. Fig. 6 (a), (b) show the effect of d. Compared to the case $d = 0.3$ m, the reflection pattern delayed in the case $d = 0.7$ m. Fig. 6 (c), (d) show the effect of t. The reflection patterns were almost the same in the both cases. However, the peak value was almost doubled. The reflection pattern did not change because the diameter of the pipe was much smaller than the wavelength of input waves. However, assuming the pipe as a scatterer, peak values can change considering the scattering cross-section of the pipe.

In the case of plastic and concrete pipes, the thickness of the pipes (the difference between outer and inner diameter) and condition inside the pipes (dried or with water) were important because waves penetrate through the pipes. Fig. 7 shows the effect of thickness t_p and condition of the pipe on reflection patterns. From the comparison between Fig. 7 (a) and (b), (c) and (d), the effect of t_p was small because t_p was much smaller than the wavelength of input waves. On the other hand, from the comparison between Fig. 7 (a) and (c), (b) and (d), the condition of the pipes have a large effect on reflection patterns. Considering the

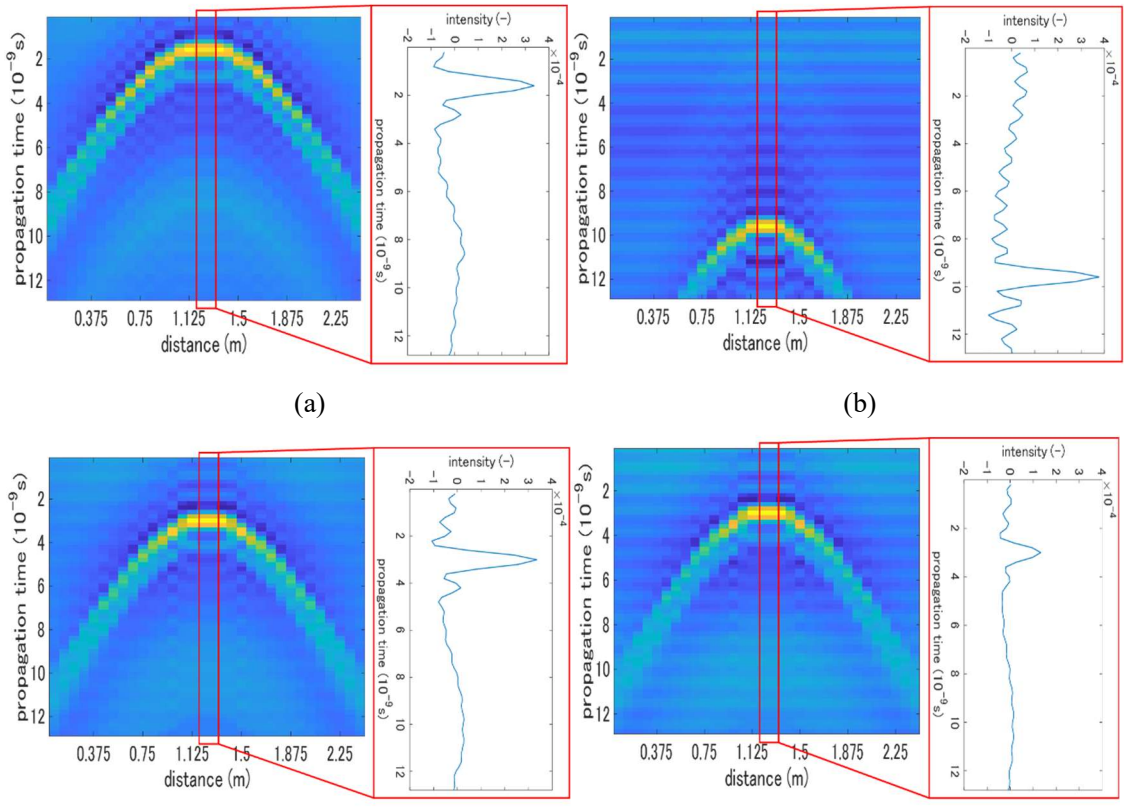


Fig. 5 Effect of electromagnetic characteristics of soil on reflection patterns (a) (b) relative permittivity 2, 10 (c) (d) conductivity 0.001 s/m, 0.01 s/m

relative permittivity of air and water, reflections at the inner surface of the pipe was dominant. From Fig. 7 (a), (b), a negative peak appeared at the upper surface and positive peak at the lower surface in dried condition. However, from Fig. 7 (c), (d) in the with-water cases, because the relative permittivity of water is much larger than the pipe, the peak was amplified and phases of the waves were reversed. Moreover, a complicated reflection pattern appeared at a deeper region because of multiple reflections inside the pipe. Fig. 8 shows the reflection pattern of the concrete pipe with diameter 50 cm and thickness 5 cm in dried condition and with water. From Fig. 8, a negative-positive peak appeared in dried condition and a strong positive peak appeared in the with-water case. Because the diameter of the concrete pipe was large, waves attenuated inside the pipe in the with-water case.

Summarizing the results, the 2D model can produce reflection patterns with enough validity because the effect of the 3D geometry of the model is negligible. Considering the reflection patterns of pipes, relative permittivity and conductivity of soil are the important parameters. Larger the difference of relative permittivity between soil and pipes is, and smaller conductivity of soil is, easier the detection of pipes reflection patterns is. The electromagnetic characteristics of actual soil strongly depends on water ratio. The estimation of these parameters from radar images is left as future work though it is difficult as stated in

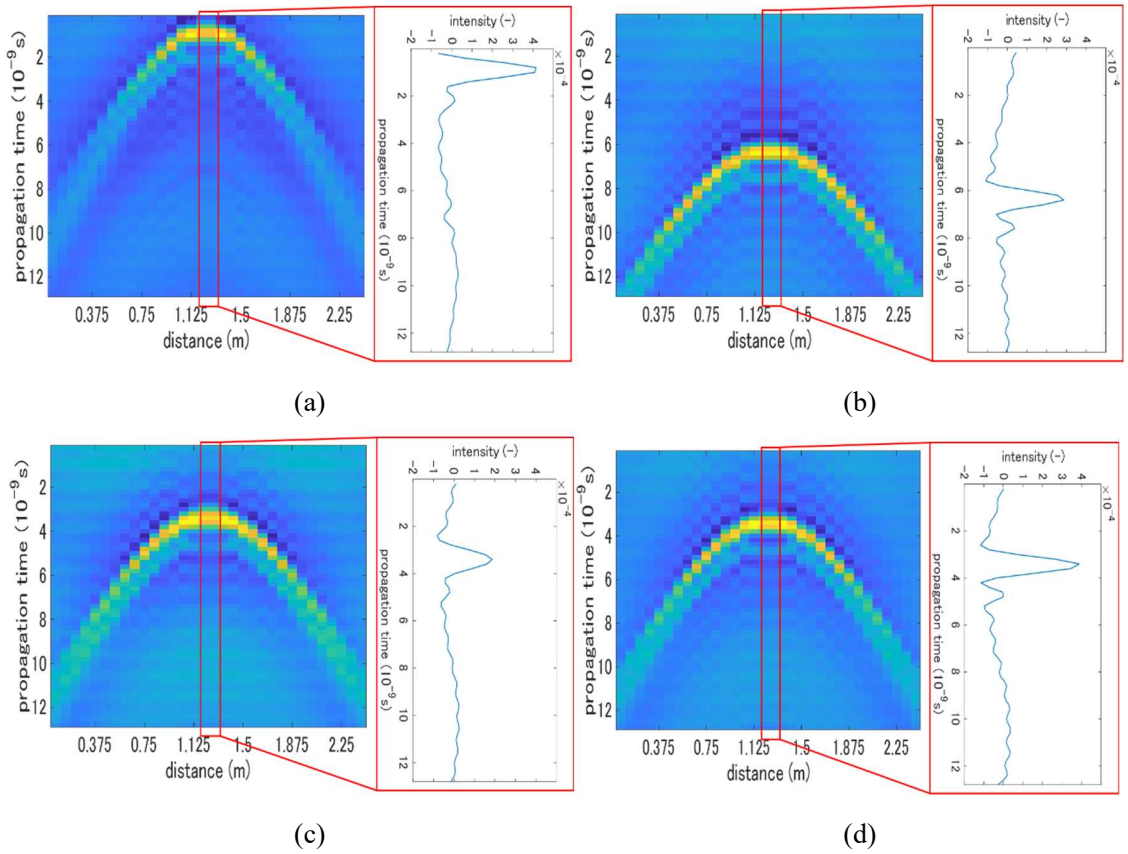


Fig. 6 Effects of depth and diameter on reflection patterns (a) (b) depth 0.3 m, 0.7 m (c) (d) diameter 5 cm, 10 cm

Section 8.2. The material and diameter of pipes are also important. In the case of plastic and concrete pipes, condition inside the pipe is also important. Large diameter with water inside the pipe is the easiest case. In the discussion of the experimental field results, the effects of material and diameter of pipes can be considered. Because of the limitation of measurement, only dried condition was considered in the experimental field case, which is the safer side case.

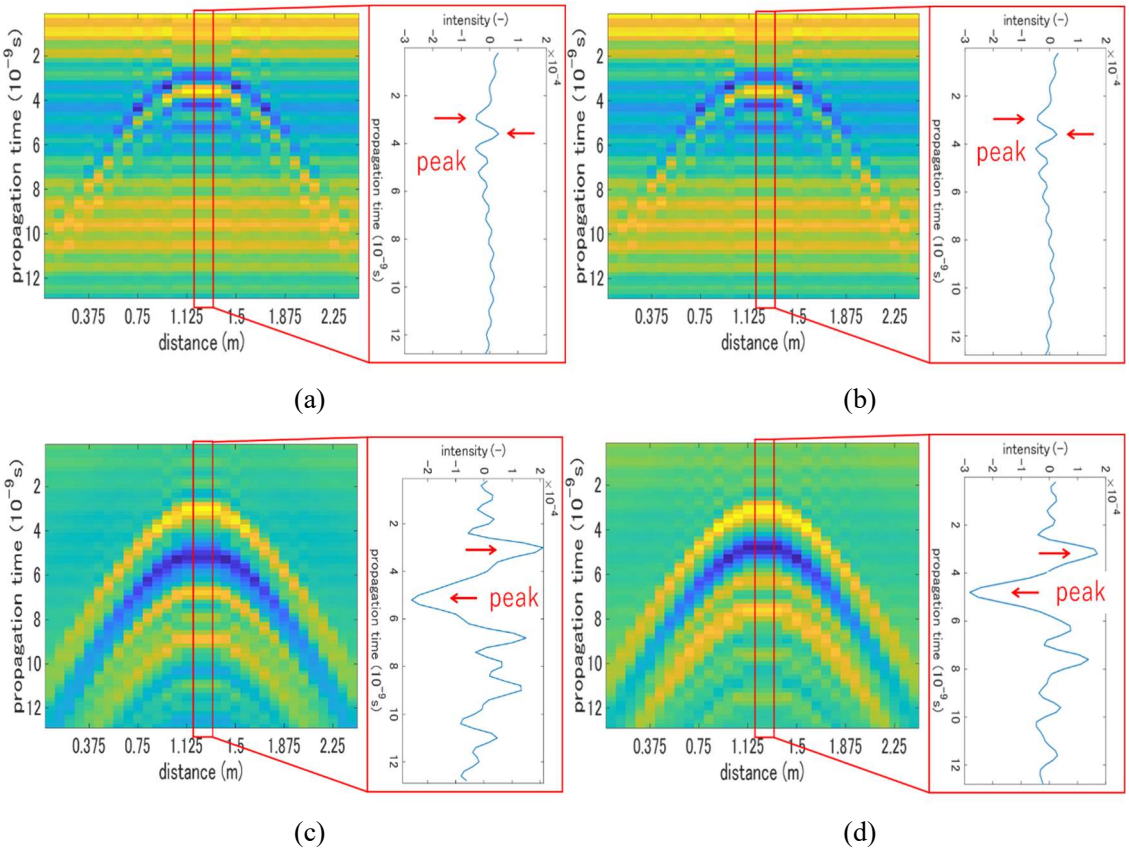


Fig. 7 Effects of thickness and condition (plastic pipe) (a) (b) dried, thickness 5 mm, 10 mm (c) (d) with water, thickness 5 mm, 10 mm

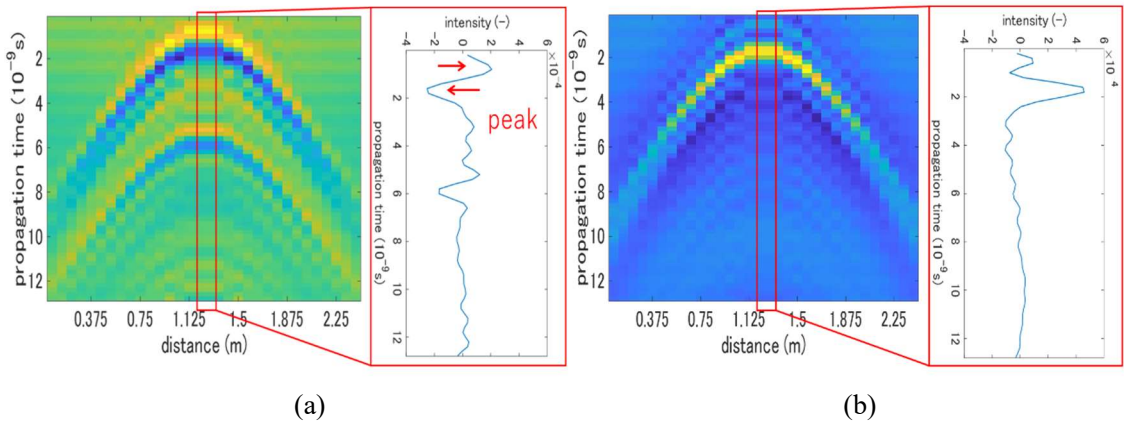


Fig. 8 Effect of condition (concrete pipe, diameter 50 cm, thickness 5 cm) (a) dried (b) with water

Appendix D: CNN parameters

An activation function is a transform function applied to the outputs of convolution and fully connected layers. An activation function is designed to mimic the activity of neurons. The derivative of an activation function should be large enough to efficiently update the parameters preventing the vanishing gradient problem. The simplest activation function is an identity function. Generally, to increase the classification accuracy of deep layer models, a nonlinear function is utilized. A conventional activation function is a sigmoid function F_{sig} defined below.

$$F_{sig}(x) = \frac{1}{1+e^{-x}} \tag{1}$$

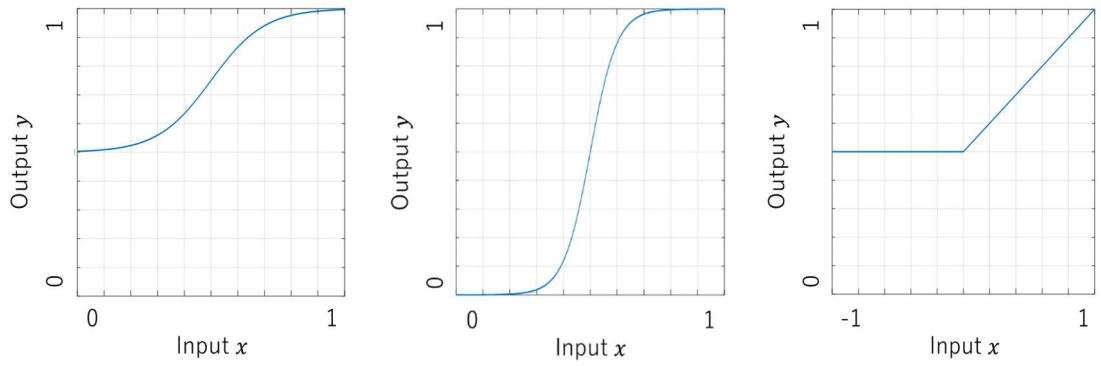
A Softmax function F_{Soft} is a multi-class version of F_{sig}. A n class Softmax function $F_{Soft}(x_n)$ is defined as below.

$$F_{Soft}(x_n) = \frac{e^{-x_n}}{\sum_n e^{-x_n}} \tag{2}$$

A Softmax function is typically added to the final layer to derive the probability of each class. The summation of all the classes of $F_{soft}(x_n)$ is 1 by definition. Fig. 1 (a) shows a normalized sigmoid function. As shown in Fig. 1 (a), at large x the derivative of the function decreases to 0, causing the vanishing gradient problem. Therefore, other nonlinear functions such as a tangent hyperbolic (tanh) function and Rectified Linear Unit (ReLU) function as shown in Fig. 1 (b), (c) is utilized. A ReLU function is 0 in a negative range and an identity function in a positive range. A ReLU function is one of the most common activation functions.

$$F(x) = 0 \quad (x \leq 0) \tag{3}$$

$$= x \quad (x \geq 0) \tag{4}$$



(a) (b) (c)

Fig. 1 Activation functions (a) Sigmoid (b) tanh (c) ReLU

Because the derivative of a ReLU function is constant at a positive range, the vanishing gradient problem cannot occur. In the research, the classification accuracy of each activation function, identity, sigmoid, tanh and ReLU functions, was evaluated.

An activation function has a small effect on classification accuracy. Fig. 2 and Fig. 3 show the effect of activation functions in the case of transverse and longitudinal pipes respectively. From Fig. 2, in the case of transverse pipes, the classification accuracy of a sigmoid function decreased to 52 %. The vanishing gradient problem may have occurred. From Fig. 3, in the case of longitudinal pipes, the classification accuracy of a sigmoid function was 75 %. It was about 7 % lower than the classification accuracy of a ReLU function, about 82 %. However, except for a sigmoid function, classification accuracy did not change regardless of

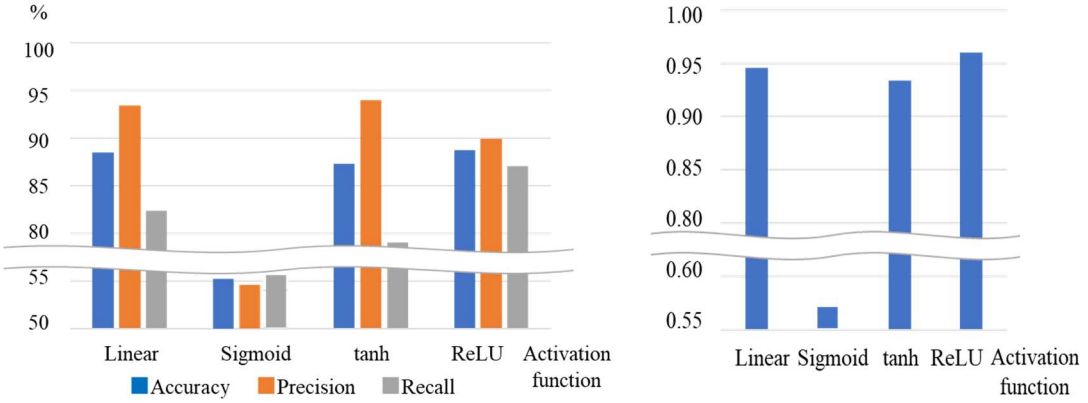


Fig. 2 Effect of activation functions (transverse pipes) (a) Accuracy, Precision, Recall (b) AUC

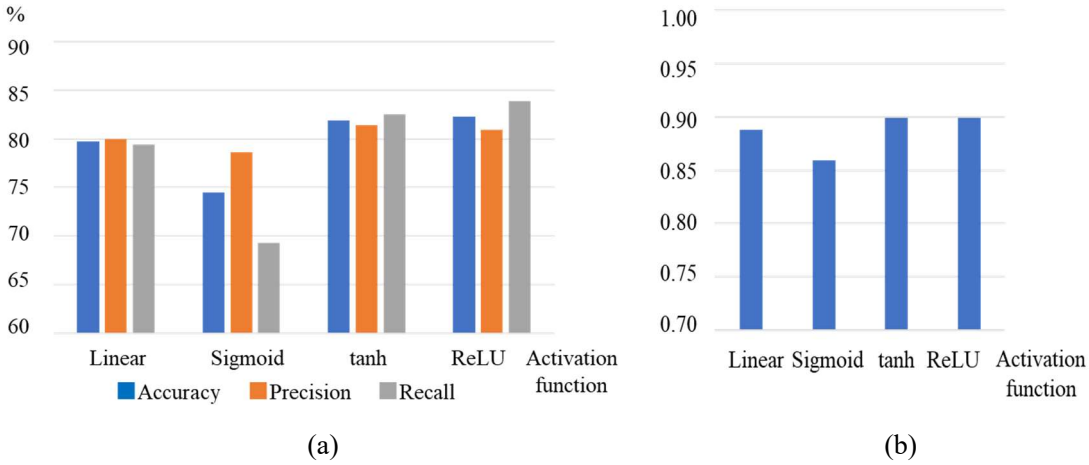


Fig. 3 Effect of activation functions (longitudinal pipes) (a) Accuracy, Precision, Recall (b) AUC

activation functions. In the research, ReLU and Softmax functions were adopted. A Softmax function was only applied to the outputs of the final fully connected layer.

As stated in Section 6.2.2, dropout has a possibility to increase accuracy by the ensemble learning effect. In the research, dropout was added to the convolution and fully connected layers to evaluate the effect. From the results, dropout has no effect on classification accuracy. Fig. 4 and Fig. 5 show the effect of dropout with probability p in the case of transverse and longitudinal pipes respectively. $p = 0$ corresponds to no dropout. In all the cases, classification accuracy did not change regardless of p. The effect of dropout depends on the target problem. In the research, dropout was not utilized.

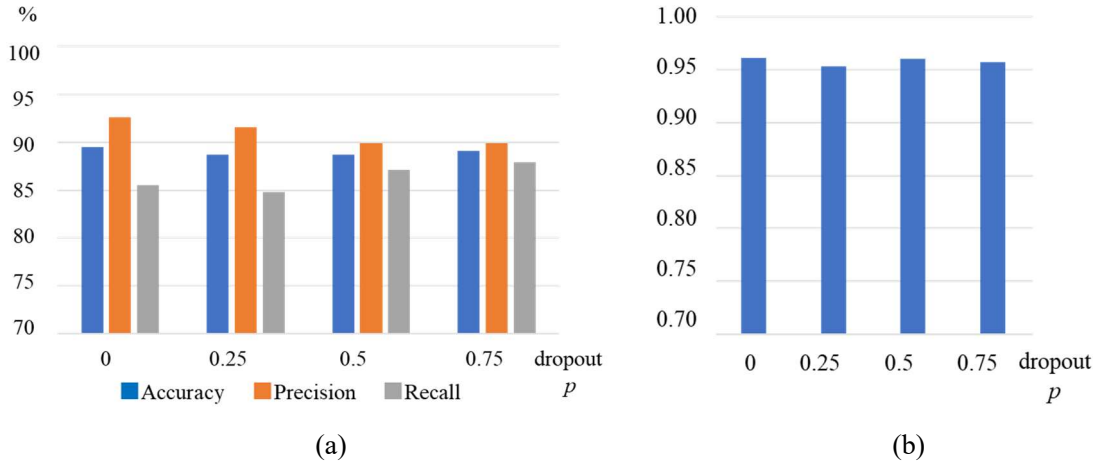


Fig. 4 Effect of dropout (transverse pipes) (a) Accuracy, Precision, Recall (b) AUC

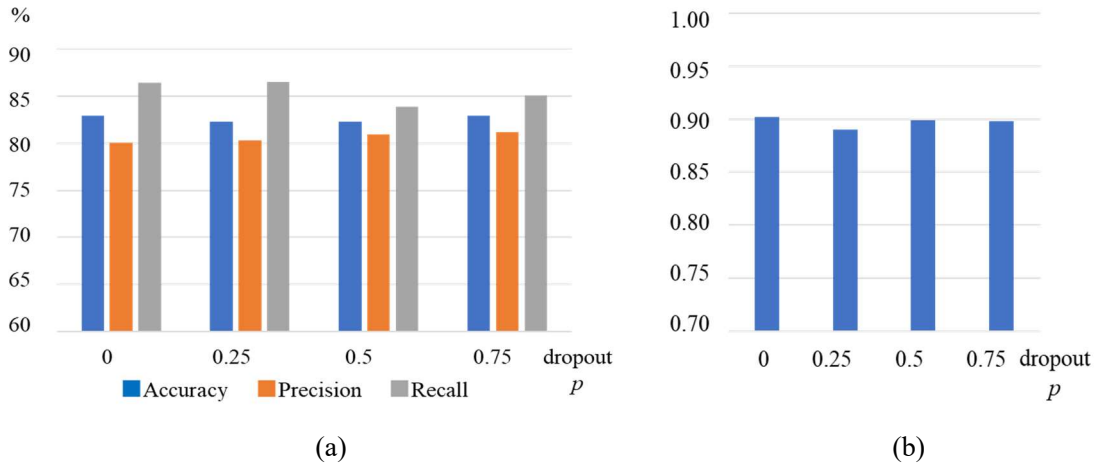


Fig. 5 Effect of dropout (longitudinal pipes) (a) Accuracy, Precision, Recall (b) AUC

Appendix E: Visualization of CNN features

It is difficult to interpret learned features simply by showing the patterns of coefficients of convolution filters and fully connected layers. Fig. 1 and Fig. 2 display the two examples of learned convolution filters and coefficients of fully connected layer perceptrons. The transverse and longitudinal sections of convolution filters are shown. The initial values of the layers were random numbers. From Fig. 1 and Fig. 2, the learned patterns were not clear.

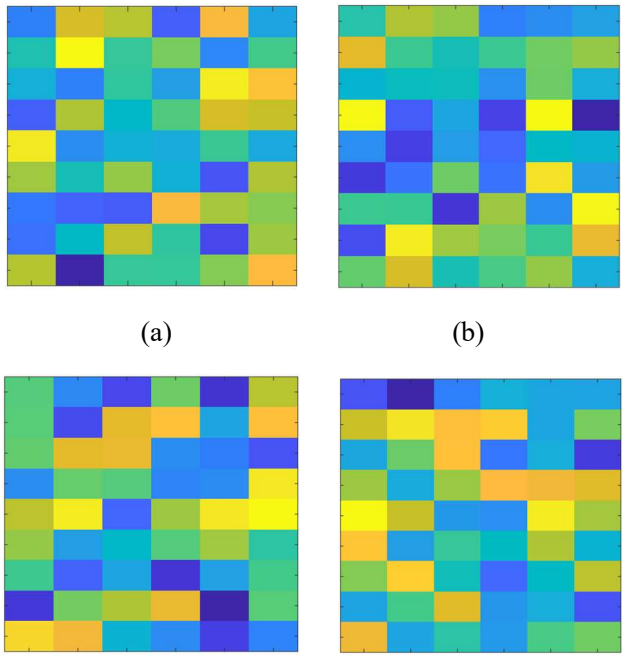


Fig.1 Convolution filters (transverse and longitudinal sections) (a) (b) filter I (c) (d) filter II

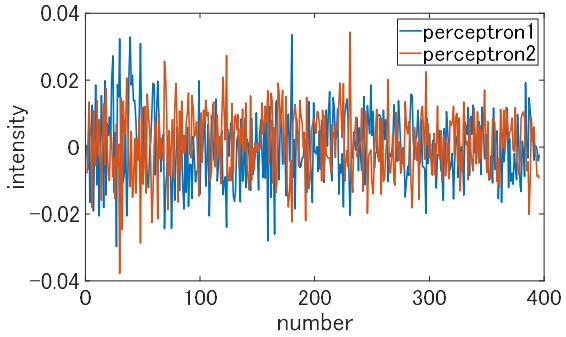


Fig. 2 Fully connected layer coefficients

Feature maps after convolution filters show clear patterns. Fig. 3 and Fig. 4 show the feature maps of a transverse and longitudinal pipe after applying filters I and II shown in Fig. 1. From Fig. 3 (c) - (f), the transverse pipe showed a hyperbolic reflection pattern in a longitudinal section and a straight line in a transverse section. The filter I successfully extracted the feature of the transverse pipe, though the filter II failed. Conversely, from Fig. 4 (c) - (f), the filter I failed to obtain the feature of the longitudinal pipe. The filter II extracted the reflection pattern of the longitudinal pipe. Therefore, the filter I served as a filter for transverse pipes and filter II for longitudinal pipes. In three-category classification by 3D-CNN, each convolution filter corresponds to each direction pipe.

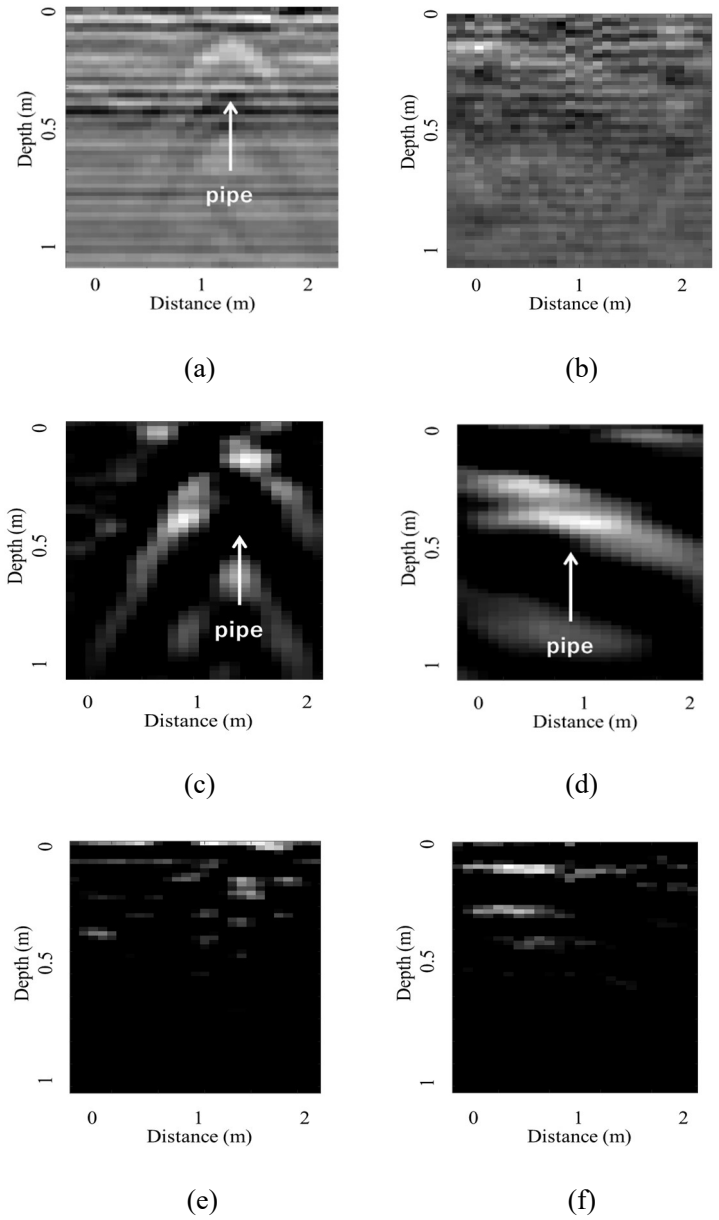


Fig. 3 Transverse pipe (transverse and longitudinal sections) (a) (b) raw images (c) (d) after filter I (e) (f) after filter II

Feature maps show clear patterns also in the case of void. Fig. 5 shows the feature maps of void after applying two convolution filters. From Fig. 5 (c) - (f), the two filters enhanced the hyperbolic reflection patterns of void in the both transverse and longitudinal sections. The 3D convolution filters utilized 3D hyperboloid reflection patterns to detect void accurately.

As a visualization method of learned features of CNN, other than simply displaying filters and feature maps, baseline method, Ultrametric Contour Map (UCM), Class Activation Mapping (CAM), Gradient-

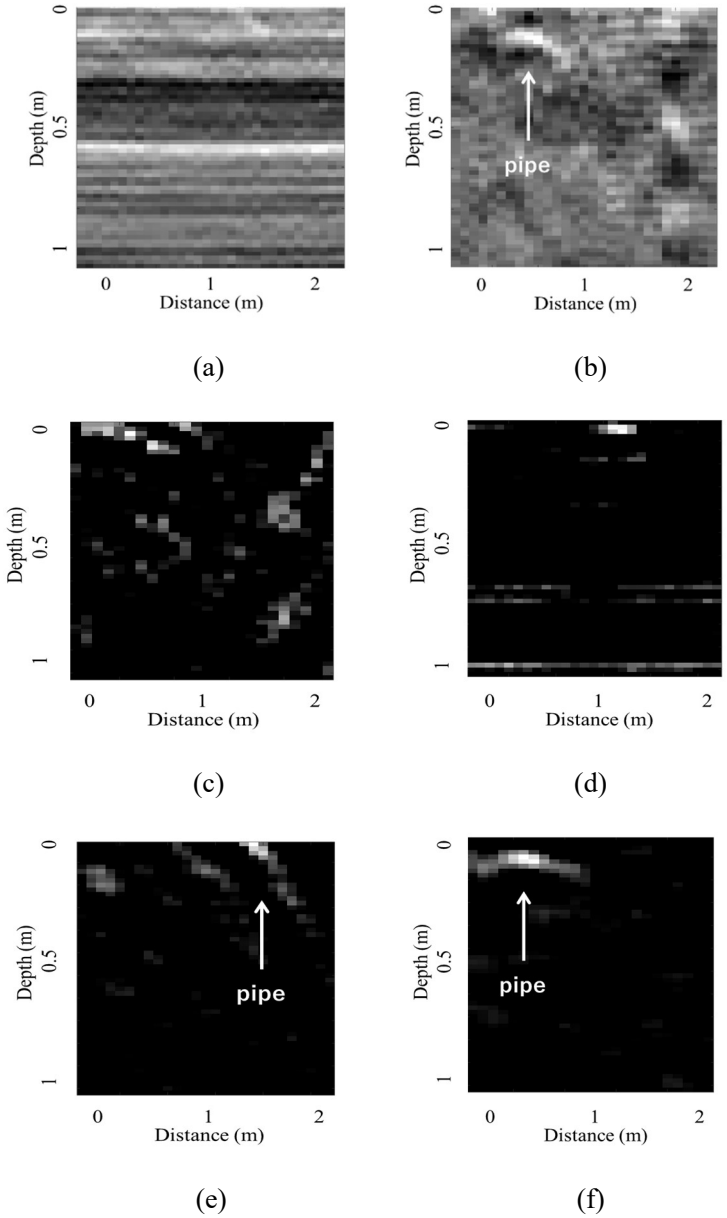


Fig. 4 Longitudinal pipe (transverse and longitudinal sections) (a) (b) raw images (c) (d) filter I (e) (f) filter II

Weighted Class Activation Mapping (Grad-CAM) are proposed [148]. Baseline method and UCM are types of sensitivity analysis. On the other hand, CAM and Grad-CAM analytically calculate a heat map by backpropagation. The problem of baseline method is it is difficult to consider the context of the whole image, which decreases the accuracy. In the research, the author applied baseline method to the data of Fig. 3 - Fig. 5, resulting in unreasonable heat maps. In terms of CAM and Grad-CAM, application to 3D data is discussed in previous research though it still remains a target of research [148]. The development of a visualization method is left as future work.

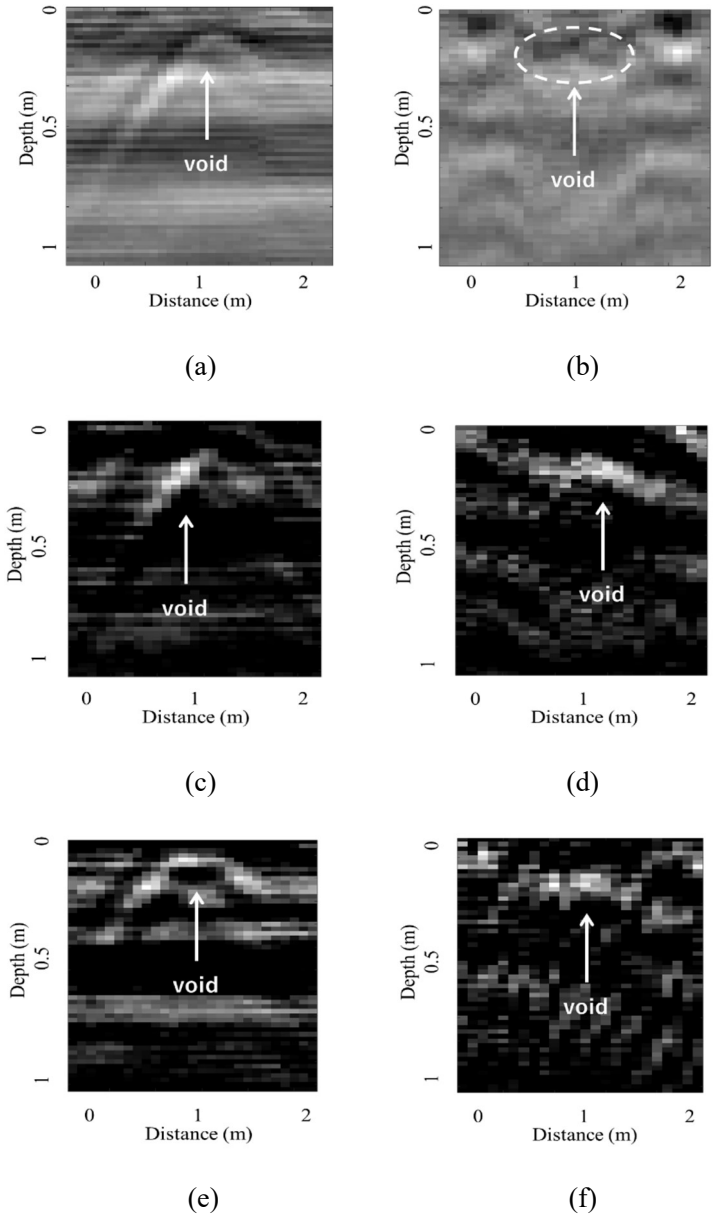


Fig. 5 Void (transverse and longitudinal sections) (a) (b) raw images (c) (d) filter I (e) (f) filter II

Appendix F: Measurement results

This appendix introduced other measurement data results of pipes and void. The parameters of the algorithm were the same as Chapter 7. From the results, the positions of transverse, longitudinal pipes and void were successfully visualized in 3D maps. Therefore, the optimized parameters were applicable to other data. The parameters may depend only on the characteristics of the measurement system. In some cases, the points were missing, varied or falsely detected possibly because of noise and clutters of other buried objects. As a whole, the variation of the points from the lines was several to several tens cm, indicating the estimation accuracy of the algorithm was high. The migration method was validated by the experimental field results as discussed in Section 7.4.

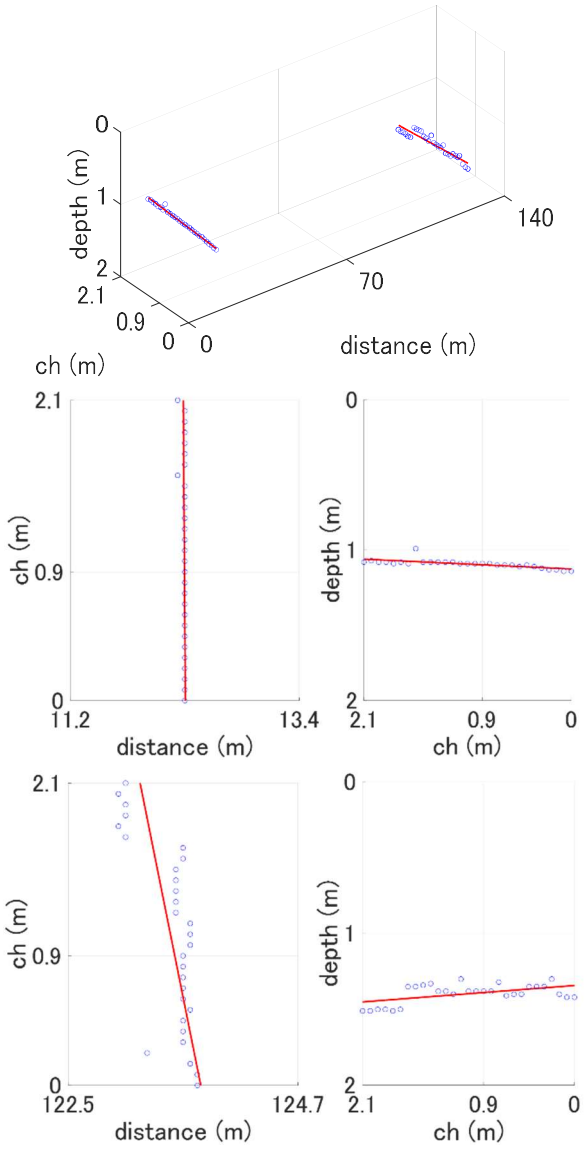


Fig. 1 3D map of transverse pipe I

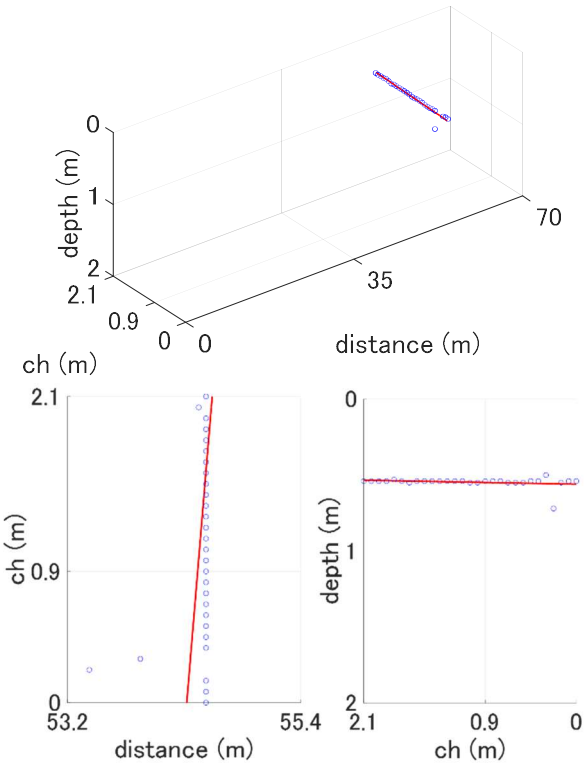


Fig. 2 3D map of transverse pipe II

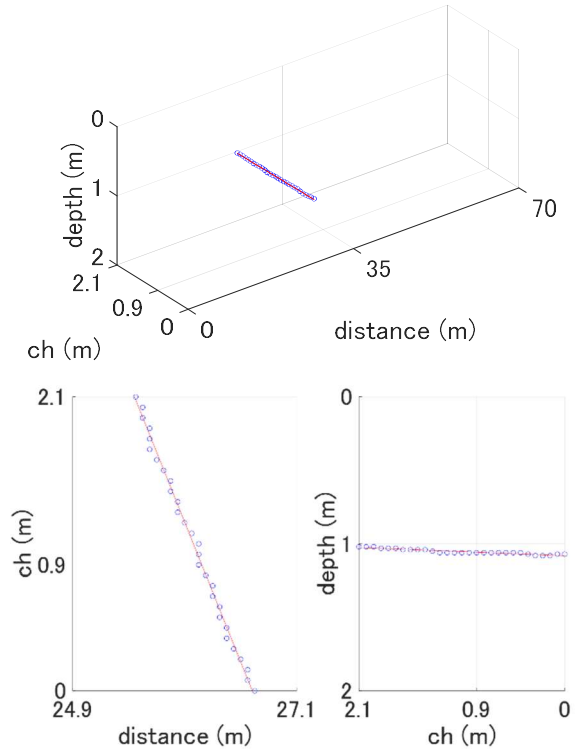


Fig. 3 3D map of transverse pipe III

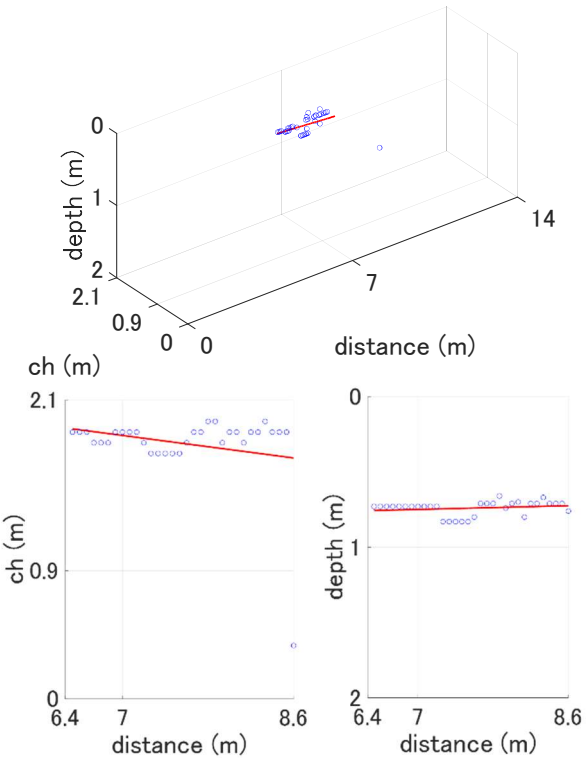


Fig. 4　3D map of longitudinal pipe I

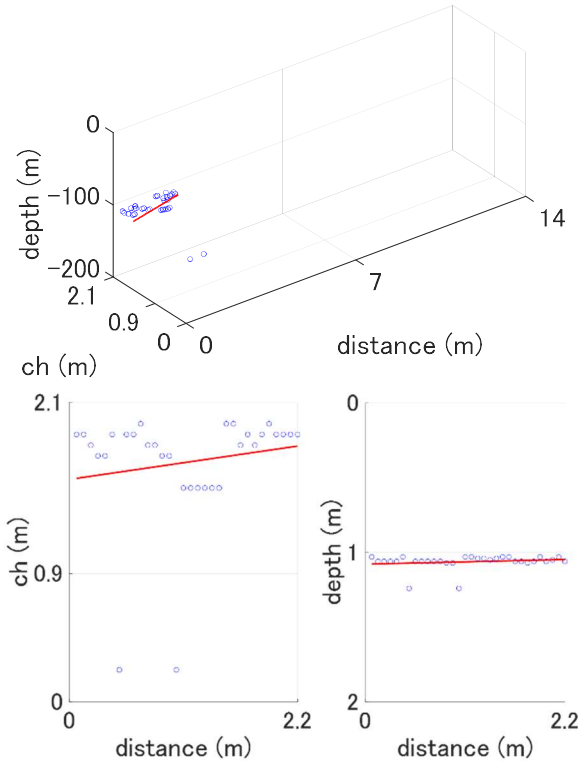


Fig. 5　3D map of longitudinal pipe II

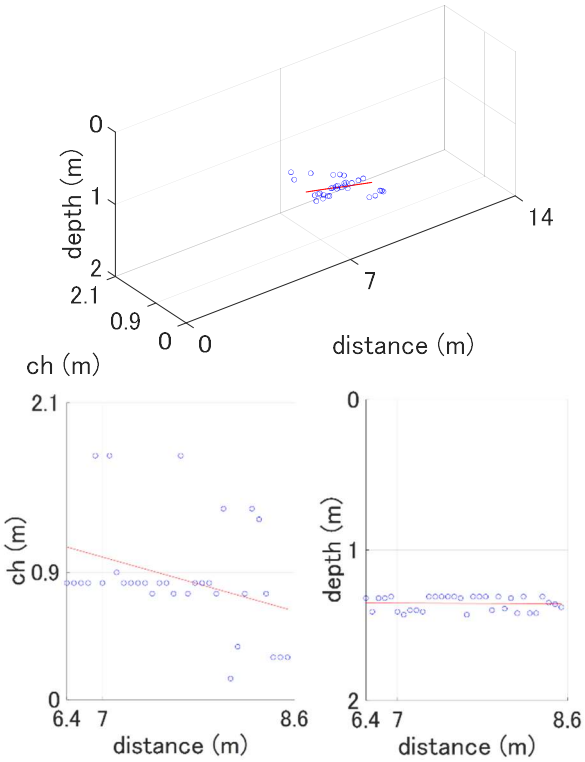


Fig. 6 3D map of longitudinal pipe III

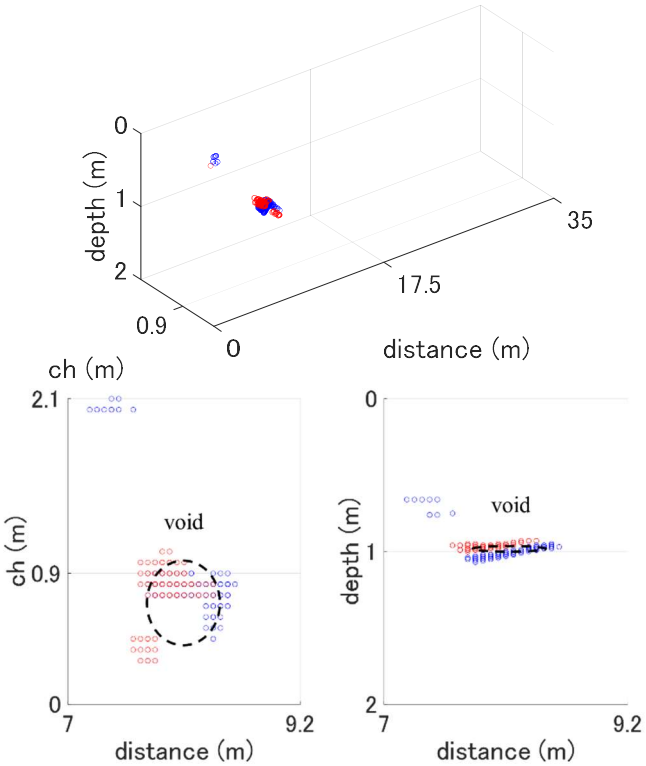


Fig. 7 3D map of void I

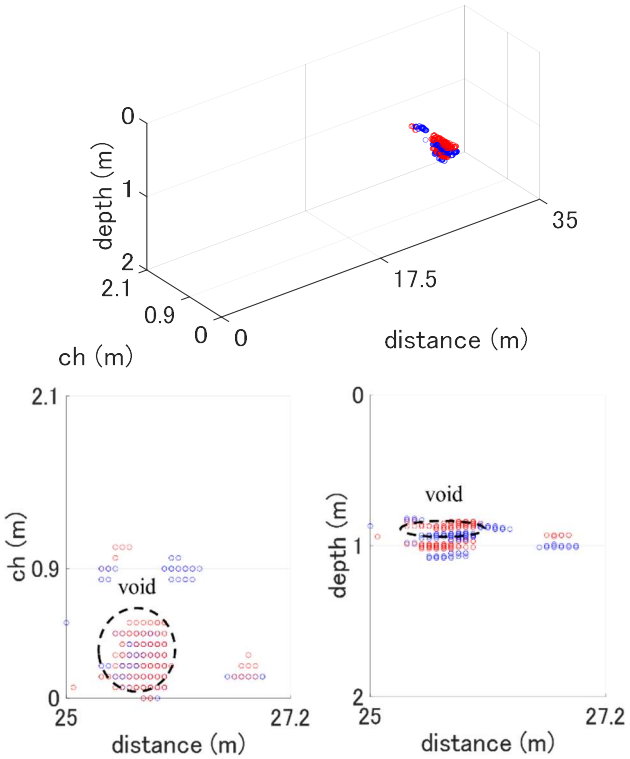


Fig. 8　3D map of void II

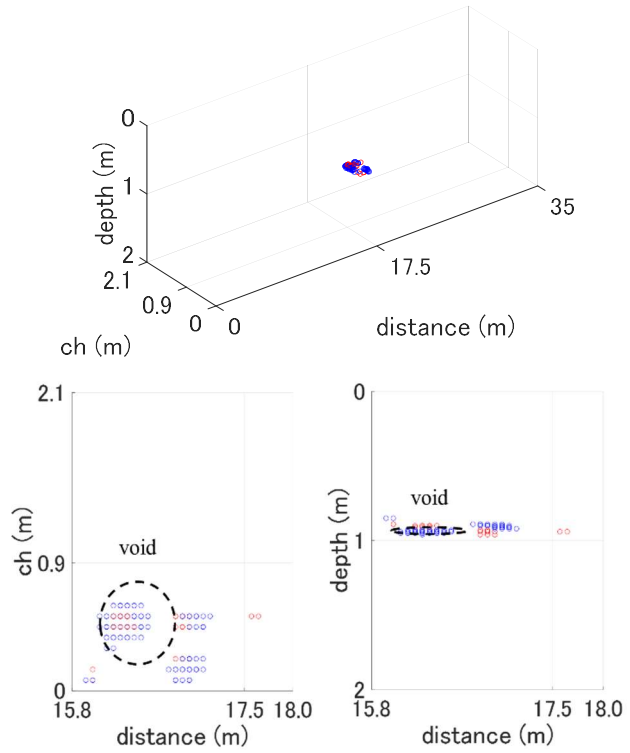


Fig. 9　3D map of void III

References

1) Ministry of Land, Infrastructure, Transport and Tourism: Maintenance of sewage pipes (in Japanese) (http://www.mlit.go.jp/mizukokudo/sewerage/crd_sewerage_tk_000135.html accessed on 2020.6.30)
2) Ministry of Economy, Trade and Industry: Manual for subsurface pipes maintenance (in Japanese) (https://www.lpgpro.go.jp/guest/learning/h19_01/ accessed on 2020.6.30)
3) Ministry of Land, Infrastructure, Transport and Tourism: Road collapse and causes (in Japanese) (http://www.mlit.go.jp/road/sisaku/ijikanri/pdf/h27-29kanbotu.pdf accessed on 2020.6.30)
4) R. Kuwano: Subsurface void and countermeasures (in Japanese) (http://www.elgard.com/html/kouenkai/2018/180416_soukai-kouenkai-siryou.pdf accessed on 2020.3.4)
5) T. Yokota, W. Fukaya and T. Miyamoto: The present situation of the road cave-in sinkholes caused by sewer systems (in Japanese), Report of National Institute for Land and Infrastructure Management, No. 688, 2012.
6) Japan Federation of Construction Contractors: Buried objects and overhead line accidents under construction in 2016 (in Japanese) (https://www.nikkenren.com/anzen/file_chika/H28chika_jiko.pdf accessed on 2020.6.30)
7) Cabinet Office: Examples of risks in PFI projects (in Japanese) (https://www8.cao.go.jp/pfi/iinkai/kaisai/sougou/kentou/3kai/pdf/shiryo_sbk_33.pdf accessed on 2020.6.30)
8) Shueisha: Amendment of water act - collapse of water pipes system (in Japanese) (https://wpb.shueisha.co.jp/news/society/2019/01/15/107973/ accessed on 2020.6.30)
9) Asahi Shimbun: 30 m road collapse in front of Hakata station (in Japanese) (https://www.asahi.com/articles/ASJC82Q11JC8TIPE002.html accessed on 2020.6.30）
10) Asahi Shimbun: Road collapse, motorcycle falls in (in Japanese), 2008.8.6 Morning Edition.
11) Ryoutan Nichinichi Shimbun: 3 m road collapse, damages to sewage pipe (in Japanese) (https://www.ryoutan.co.jp/news/2015/05/27/009190.html accessed on 2020.6.30)
12) M. Sato: Monitoring by ground-coupling aperture synthesis radar and array radar (in Japanese) (https://www.jst.go.jp/sip/dl/k07/booklet/15_sato.pdf accessed on 2020.6.30)
13) S. Yamada: Development of chirp radar for detection of void and subsidence of backfilling (in Japanese) (https://www.jst.go.jp/sip/dl/k07/booklet/13_yamada.pdf accessed on 2020.6.30)
14) Ministry of Internal Affairs and Communications: Report on GPR techniques (in Japanese) (http://www.soumu.go.jp/main_content/000477180.pdf accessed on 2020.6.30)
15) Consortium for Quality Assurance of Analysis Technique of Subsurface Sensing Vehicle: Manual of subsurface sensing (in Japanese) (https://www.zenchiren.or.jp/news/download/140 accessed on 2020.6.30)

16) Nikkei Bizgate: Accelerating removal of electric pole by “Sukeruka” (in Japanese) (https://bizgate.nikkei.co.jp/article/DGXMZO35509020190920180000000 accessed on 2020.6.30)
17) Nikkan Kensetsu Kougyou Shimbun: Kawasaki Geological Engineering, AI for subsurface void detection in service (in Japanese) (https://www.decn.co.jp/?p=93776 accessed on 2020.6.30)
18) Shutoko Engineering Company: Daily inspection (https://www.shutoko-eng.jp/business/daily.php accessed on 2020.6.30)
19) K. Sato: Study on simple estimation of cavity below road surface considering soundness of pavement (in Japanese), Doctoral dissertation, Nippon University, 2016.
20) K. Sato, S. Akiba, Y. Kano and K. Akatsu: Estimation method of cavity below the road surface considering soundness of pavement of pavement for the heavy traffic road (in Japanese), Journal of JSCE E1, Vol. 71, No. 3, pp. I_135 - I_143, 2015.
21) N. Kawamura and M. Tsubokawa: Influence of void beneath asphalt pavement on FWD deflection (in Japanese), Journal of JSCE E1, Vol. 73, No. 3, pp. I_1 - I_8, 2017.
22) S. Jomoto, M. Aoki and Y. Takeuchi: Examination on road surface lower cavity investigation method using portable FWD and GPR (in Japanese), JSCE E1, Vol. 69, No. 3, pp. 167 - 173, 2013.
23) Tokyo Measuring Instruments Lab.: Portable FWD test (https://tml.jp/knowledge/special_ins/hfwd.html accessed on 2020.6.30)
24) NIPPO: Evaluation system of integrality of pavement by FWD (https://www.nippo-c.co.jp/tech_info/general/SG03002_g.html accessed on 2020.6.30)
25) K. Hayano, R. Maeda, T. Suzuki and H. Hashimoto: Application of infrared thermography scanning technique to airport asphalt pavement for field investigation of detached layers (in Japanese), Japanese Geotechnical Journal, Vol. 3, No. 1, pp. 13 - 23, 2007.
26) F. Osada, Y. Yamada, N. Mushiake and Y. Akamatsu: Development of inspecting method for a defect inside of concrete railway viaduct using thermal image (in Japanese), Journal of JSCE, Vol. 63, No. 760, pp. 121 - 133, 2004.
27) PASCO：Total management of road surface condition (https://www.pasco.co.jp/products/real/ accessed on 2020.6.30)
28) S. Miki, Y. Yoshida, K. Inoue and K. Nakagawa: A study on the evaluation of in-situ rock conditions by the tomographic method in tunnel geological investigations (in Japanese), Journal of JSCE, Vol. 707, No. VI - 55, pp. 111 - 124, 2002.
29) K. Yamada and M. Yamamoto: Experimental study on estimation of void and rebar inside concrete by seismic wave method (in Japanese), Proceedings of the Japan Concrete Institute, Vol. 13, No. 1, pp. 491 - 496, 1991.
30) M. Otsu: Theory and application of seismic method for non-destructive testing of concrete (in Japanese), Journal of the Japan Concrete Institute, Vol. 46, No. 2, pp. 5 - 11, 2008.
31) Nittetsu Mining Consultants: Geophysical exploration

(http://www.nmconsults.co.jp/construction/investigation/ accessed on 2020.6.30)
32) S. Yuyama, Z. W. Li, M. Yoshizawa and T. Uomoto: Evaluation of fatigue damage in reinforced concrete slab by acoustic emission (in Japanese), Journal of JSCE, Vol. 44, No. 627, pp. 251 - 259, 1999.
33) Y. Tsuyuguchi and A. Machida: Fundamental study on direct current resistivity method to presume the inner parts of RC structures (in Japanese), Journal of JSCE E, Vol. 62, No. 4, pp. 641 - 656, 2006.
34) D. J. Daniels: Ground penetrating radar, the Institution of Engineering and Technology, 2004.
35) H. M. Jol: Ground penetrating radar theory and applications, Elsevier, 2009.
36) L. B. Conyers: Ground-penetrating radar for archaeology, AltaMira Press, 2013.
37) B. Saleh: Introduction to subsurface imaging, Cambridge University Press, 2011.
38) A. S. Turk, A. K. Hocaoglu and A. A. Vertiy: Subsurface sensing, Wiley, 2011.
39) M. Sato: Subsurface sensing by ground penetrating radar - from basics to applications (from lecture note, in Japanese), 2015.
40) T. Yamaguchi, T. Mizutani, M. Tarumi and D. Su: Sensitive damage detection of reinforced concrete bridge slab by "time-variant deconvolution" of SHF-band radar signal, IEEE Transactions on Geoscience and Remote Sensing, Vol. 57, No. 3, pp. 1478 - 1488, 2019.
41) T. Mizutani, N. Nakamura, T. Yamaguchi, M. Tarumi, Y. Ando and I. Hara: Bridge slab damage detection by signal processing of UHF-band ground penetrating radar data, Journal of Disaster Research, Vol. 12, No. 3, pp. 415 - 421, 2017.
42) P. M. Parker: Ground-penetrating radar webster's timeline history 1979 - 2007, Icon Group International Inc., 2009.
43) S. Tanaka and M. Okamoto: Highly accurate non-destructive measurement of the covers of reinforcing bars in concrete using electromagnetic wave radar (in Japanese), Journal of the Japanese Society of Non-Destructive Inspection, Vol. 59, No. 4, pp. 183 - 188, 2010.
44) S. Tanaka, Y. Mayumi and T. Takenaka: Simultaneous estimation of rebar location, diameter and concrete permittivity by ground penetrating radar (in Japanese), Journal of the Japan Concrete Institute, Vol. 29, No. 2, pp. 775 - 780, 2007.
45) J. Hugenschmidt: Concrete bridge inspection with a mobile GPR system, Construction and Building Materials, Vol. 16, No. 3, pp. 147 - 154, 2002.
46) N. Martino, K. Maser, R. Birken and M. Wang: Determining ground penetrating radar amplitude thresholds for the corrosion state of reinforced concrete bridge decks, Journal of Environmental and Engineering Geophysics, Vol. 19, No. 3, 2014.
47) A. Loizos and C. Plati: Accuracy of pavement thicknesses estimation using different ground penetrating radar analysis approaches, NDT & E International, Vol. 40, No. 2, pp. 147 - 157, 2007.
48) S. Zhao and I. L. Al-Qadi: Development of regularization methods on simulated ground-penetrating radar signals to predict thin asphalt overlay thickness, Signal Processing, Vol. 132, pp. 261 - 271, 2015.
49) IEICE (Editor: T. Yoshida): Radar engineering (in Japanese), Corona Publishing, 1996.
50) IEICE (Editor: M. Sekine): Signal processing of radar engineering (in Japanese), Corona Publishing,

1991.
51) 3D-RADAR: High resolution step-frequency 3D GPR using wideband antenna arrays (http://3d-radar.com accessed on 2020.6.30)
52) Federal Highway Administration Research and Technology: RABIT (http://www.fhwa.dot.gov/research/tfhrc/programs/infrastructure/structures/ltbp/ltbpresearch/rabit/index.cfm accessed on 2020.6.30)
53) D. J. Daniels: A review of GPR for landmine detection, Sensing and Imaging, Vol. 7, No. 3, pp. 90 - 123, 2006.
54) M. Sato: Subsurface imaging by UWB radar: application to humanitarian demining in Cambodia (in Japanese), Journal of IEICE A, Vol. 132, No. 1, pp. 1 - 4, 2012.
55) Chigasaki City (Kanagawa Prefecture): Why it takes long to carry out construction work (https://www.city.chigasaki.kanagawa.jp/sisei_gesuido/koji/1008435.html accessed on 2020.6.30)
56) Keysight Technologies: Basics of measurement on dielectrics (http://literature.cdn.keysight.com/litweb/pdf/5989-2589JAJP.pdf accessed on 2020.6.30)
57) C. Ekes and B. Neducza: Pipe condition assessments using ground penetrating radar, the 14th International Conference on Ground Penetrating Radar, China, Jun. 2012.
58) M. H. Powers and G. R. Olhoeft: Modeling the GPR response of leaking, buried pipes, Symposium on the Application of Geophysics to Engineering and Environmental Problems, 1996.
59) M. Sato and R. Kuwano: Influence of location of subsurface structures on development of underground cavities induced by internal erosion, Soils and Foundations, Vol. 55, No. 4, pp. 829 - 840, 2015.
60) R. Kuwano, M. Sato and R. Sera: Study on the detection of underground cavity and ground loosening for the prevention of ground cave-in accident (In Japanese), Japanese Geotechnical Journal, Vol. 5, No. 2, pp. 219 - 229, 2010.
61) A. Benedetto, F. Tosti, L. B. Ciampoli and F. D'Amico: An overview of ground-penetrating radar signal processing techniques for road inspections, Signal Processing, No. 132, pp. 201 - 209, 2017.
62) O. Yilmaz: Seismic data analysis: processing, inversion, and interpretation of seismic data, Vol. I, 1987.
63) O. Yilmaz: Seismic data analysis: processing, inversion, and interpretation of seismic data, Vol. II, 1987.
64) R. Solimene, A. Cuccaro, A. Dell'Aversano, I. Catapano and F. Soldovieri: Ground clutter removal in GPR surveys, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 7, No. 3, pp. 792 - 798, 2014.
65) R. Persico and F. Soldovieri: Effects of background removal in linear inverse scattering, IEEE Transactions on Geoscience and Remote Sensing, Vol. 46, No. 4, pp. 1104 - 1114, 2008.
66) F. Soldovieri, A. Brancaccio, G. Prisco, G. Leone and R. Pierri: A Kirchhoff-based shape reconstruction algorithm for the multimonostatic configuration: the realistic case of buried pipes, IEEE Transactions on Geoscience and Remote Sensing, Vol. 46, No. 10, pp. 3031 - 3038, 2008.
67) E. Pettinelli, A. Di Matteo, E. Mattei, L. Crocco, F. Soldovieri, J. D. Redman and A. P. Annan: GPR response from buried pipes: measurement on field site and tomographic reconstructions, IEEE

Transactions on Geoscience and Remote Sensing, Vol. 47, No. 8, pp. 2639 - 2645, 2009.
68) H. Zhou, M. Sato and H. Liu: Migration velocity analysis and prestack migration of common-transmitter GPR data, IEEE Transactions on Geoscience and Remote Sensing, Vol. 43, No. 1, pp. 86 - 91, 2005.
69) L. Li: Sparsity-promoted blind deconvolution of Ground-Penetrating Radar (GPR) data, IEEE Geoscience and Remote Sensing Letters, Vol. 11, No. 8, pp. 1330 - 1334, 2014.
70) A. Simi, S. Bracciali and G. Manacorda: Hough transform based automatic pipe detection for array GPR: algorithm development and on-site tests, IEEE Radar Conference, Italy, May. 2008.
71) P. Gamba and S. Lossani: Neural detection of pipe signatures in ground penetrating radar images, IEEE Transactions on Geoscience and Remote Sensing, Vol. 38, No. 2, pp. 790 - 797, 2000.
72) W. Al-Nuaimy, Y. Huang, M. Nakhkash, M.T.C. Fang, V.T. Nguyen and A. Eriksen: Automatic detection of buried utilities and solid objects with GPR using neural networks and pattern recognition, Journal of Applied Geophysics, No. 43, pp. 157 - 165, 2000.
73) U. S. Khan: Interpretation of ground penetrating radar data for utilities, VDM Verlag Dr. Muller, 2011.
74) E. Pasolli, F. Melgani and M. Donelli: Automatic analysis of GPR images: a pattern-recognition approach, IEEE Transactions on Geoscience and Remote Sensing, Vol. 47, No. 7, pp. 2206 - 2217, 2009.
75) S. Shihab and W. Al-Nuaimy: Radius estimation for cylindrical objects detected by ground penetrating radar, Subsurface Sensing Technologies and Applications, Vol. 6, No. 2, pp. 151 - 166, 2005.
76) H. Kotaki, Y. Nomura, N. Kato and H. Matsui: Simultaneous estimation method of relative permittivity and pipe radius by ground penetrating radar (in Japanese), Journal of IEICE B, Vol. J92-B, No. 9, pp. 1538 - 1549, 2009.
77) Q. Hoarau, G. Ginolhac, A. M. Atto and J. M. Nicolas: Robust adaptive detection of buried pipes using GPR, Signal Processing, No. 132, pp. 293 - 305, 2017.
78) D. Ayala-Cabrera, M. Herrera, J. Izquierdo and R. Perez-Garcia: Location of buried plastic pipes using multi-agent support based on GPR images, Journal of Applied Geophysics, No. 75, pp. 679 - 686, 2011.
79) U. Boniger and J. Tronicke: Subsurface utility extraction and characterization: combining GPR symmetry and polarization attributes, IEEE Transactions on Geoscience and Remote Sensing, Vol. 50, No. 3, pp. 736 - 746, 2012.
80) A. Shaari, R. S. Ahmad and T. H. Chew: Effects of antenna-target polarization and target-medium dielectric contrast on GPR signal from non-metal pipes using FDTD simulation, NDT & E International, No. 43, pp. 403 - 408, 2010.
81) C. Warren, A. Giannopoulos and I. Giannakis: gprMax: open source software to simulate electromagnetic wave propagation for ground penetrating radar, Computer Physics Communications, No. 209, pp. 163 - 170, 2016.
82) T. Wang and M. L. Oristaglio: 3-D simulation of GPR surveys over pipes in dispersive soils, Geophysics, Vol. 65, No. 5, pp. 1560 - 1568, 2000.
83) J. Sonoda and T. Kimoto: Object identification and visualization for GPR images using deep learning

with convolutional neural network and generative adversarial network (in Japanese), Japanese Geotechnical Journal, Vol. 67, No. 6, pp. 12 - 15, 2019.
84) J. Sonoda and T. Kimoto: Experimental images identification with simulation images and finetuning for objects identification of ground penetrating radar using deep learning (in Japanese), the 32nd Annual Conference of the Japanese Society for Artificial Intelligence, Japan, Jun. 2018.
85) D. H. Chen and T. Scullion: Detecting subsurface voids using ground-coupled penetrating radar, Geotechnical Testing Journal, Vol. 31, No. 3, pp. 217 - 224, 2008.
86) S. K. Park, T. Uomoto and M. Yoshizawa: A study on a quantified evaluation of void under the concrete pavement by ground penetrating radar (in Japanese), Proceedings of the Japan Concrete Institute, Vol. 18, No. 1, pp. 1365 - 1370, 1996.
87) 3d-Radar (E. Eide, P. A. Valand and J. Sala): High resolution step-frequency 3D GPR using wideband antenna arrays (from presentation slides), 2012.
88) K. Iizuka, A. P. Freundorfer, K. H. Wu, H. Mori, H. Ogura and V. K. Nguyen: Step frequency radar, Journal of Applied Physics, Vol. 56, No. 9, pp. 2572 - 2583, 1984.
89) A. Langman and M. R. Inggs: Pulse versus stepped frequency continuous wave modulation for ground penetrating radar, IEEE Geoscience and Remote Sensing Symposium, Australia, Jul. 2001.
90) IEEE: IEEE Standard Letter Designations for Radar-Frequency Bands (521-2002)
91) 3D-RADAR: 3dr-Examiner (http://3d-radar.com/wp-content/uploads/2011/07/FS-3dR-Examiner-071411v1-US.pdf accessed on 2020.6.30)
92) Passmark Software: Videocard benchmarks (https://www.videocardbenchmark.net/high_end_gpus.html accessed on 2020.6.30)
93) NVIDIA: GeForce GTX 1080 Ti (https://www.nvidia.com/ja-jp/geforce/products/10series/geforce-gtx-1080-ti/ accessed on 2020.6.30)
94) Mouse Computer: MASTERPIECE i1640 (https://www.g-tune.jp/desktop/mp_iz/ accessed on 2020.6.30)
95) Panasonic: CF-SZ5 (https://panasonic.biz/cns/pc/prod/note/sz5p/ accessed on 2020.6.30)
96) MathWorks: MATLAB (https://jp.mathworks.com/ accessed on 2020.6.30)
97) I. Kobayashi: MATLAB handbook (in Japanese), Shuwa System, 2014.
98) grpMax: Open source Finite-Difference Time-Domain electromagnetic simulation software (https://www.gprmax.com/ accessed on 2020.6.30)
99) B. Lubanovic: Introducing python, O'reilly, 2014.
100) M. Jaworski and T. Ziade: Expert python programming, Packt Publishing, 2016.
101) Google: TensorFlow (https://www.tensorflow.org/ accessed on 2020.6.30)

102) K. Saito: Deep learning from scratch - theory and application by Python (in Japanese), O'reilly Japan, 2016.
103) Y. Sugomori: Details of deep learning - time series analysis by TensorFlow and Keras (in Japanese), Mynavi Publishing, 2017.
104) Google: TensorBoard
(https://www.tensorflow.org/tensorboard accessed on 2020.6.30)
105) kwikteam: npy-matlab
(https://github.com/kwikteam/npy-matlab accessed on 2020.6.30)
106) T. H. Friis: A note on a simple transmission formula, Proceedings of the IRE, Vol. 34, No. 5, pp. 254 - 256, 1946.
107) Institute of Electronics, Information and Communication Engineers: Antenna engineering handbook (in Japanese), Ohmsha, 2008.
108) J. L.Volakis: Antenna engineering handbook, McGrawHill, 2007.
109) A. V. Oppenheim and R. W. Schafer: Discrete-time signal processing, Prentice Hall, 2010.
110) S. Adachi: Digital signal processing and systems by MATLAB (in Japanese), Tokyo Denki University Publishing, 2002.
111) M. Ikehara, T. Shimamura and T. Sanada: MATLAB multimedia signal processing, the first volume (in Japanese), Baihukan, 2004.
112) T. Higuchi and M. Kawamata: Digital signal processing by MATLAB (in Japanese), Morikita Publishing, 2015.
113) The Mathematical Society of Japan: Iwanami mathematics dictionary (in Japanese), Iwanami Shoten, 2007.
114) S. Abe: Introduction to support vector machine for pattern recognition (in Japanese), Morikita Publishing, 2011.
115) N. Cristianini and J. Shawe-Taylor: An introduction to support vector machines and other kernel-based learning methods, Cambridge University Press, 2000.
116) C. M. Bishop: Pattern recognition and machine learning, Springer, 2011.
117) Y. Hirai: Introduction to pattern recognition (in Japanese), Morikita Publishing, 2012.
118) Y. Panga, Y. Yuan, X. Li and J. Pan: Efficient HOG human detection, Signal Processing, Vol. 91, No. 4, pp. 773 - 781, 2010.
119) D. F. Llorca, R. Arroyo and M. A. Sotelo: Vehicle logo recognition in traffic images using HOG features and SVM, Proceedings of the 16th International IEEE Annual Conference on Intelligent Transportation Systems, Netherlands, Oct. 2013.
120) J. Davis and M. Goadrich: The relationship between Precision-Recall and ROC curves, the 23rd International Conference on Machine Learning, USA, Jun. 2006.
121) J. M. Lobo, A. Jimenez-Valverde and R. Real: AUC: A misleading measure of the performance of predictive distribution models, Global Ecology and Biogeography, No. 17, pp. 145 - 151, 2008.

122) T. Harada: Image recognition (in Japanese), Kodansha, 2017.
123) CG-ARTS Society: Digital image processing (in Japanese), CG-ARTS Society, 2015.
124) M. Ikehara, T. Shimamura and T. Sanada: MATLAB multimedia signal processing, the second volume (in Japanese), Baihukan, 2004.
125) H. Bay, T. Tuytelaars and L. Van Gool: SURF: Speeded Up Robust Features, the 9th European Conference on Computer Vision, Austria, May. 2006.
126) N. Dalal and B. Triggs: Histograms of Oriented Gradients for human detection, the 17th IEEE Conference of Computer Vision and Pattern Recognition, USA, Jun. 2005.
127) T. Uno: FDTD method for numerical electromagnetic analysis (in Japanese), Corona Publishing, 2016.
128) O. Hashimoto and T. Abe: Introduction to FDTD time domain difference method (in Japanese), Morikita Publishing, 1996.
129) S. Sunakawa: Theory of electromagnetics (in Japanese), Kinokuniya Shoten, 1999.
130) S. Adachi: Engineering electromagnetics (in Japanese), Corona Publishing, 1983.
131) W. H. Hayt: Engineering electromagnetics, McGraw-Hill, 1979.
132) K. S. Lee: Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media, IEEE Transactions on Antennas and Propagation. Vol. 14, No. 3, pp. 302 - 307, 1966.
133) J. P. Berenger: A perfectly matched layer for the absorption of electromagnetic waves, Journal of Computational Physics, No. 114, pp. 185 - 200, 1994.
134) H. Takayasu: Fractal (in Japanese), Asakura Shoten, 2010.
135) N. R. Peplinski, F. T. Ulaby and M. C. Dobson: Dielectric properties of soils in the 0.3 - 1.3 GHz range, IEEE Transactions on Geoscience and Remote Sensing, Vol. 33, No. 3, pp. 803 - 807, 1995.
136) K. Belli, C. M. Rappaport, H. Zhan and S. Wadia-Fascetti: Effectiveness of 2-D and 2.5-D FDTD ground-penetrating radar modeling for bridge-deck deterioration evaluated by 3-D FDTD, IEEE Transactions on Geoscience and Remote Sensing, Vol. 47, No. 11, pp. 3656 - 3663, 2009.
137) H. Zhan, K. Belli, S. Wadia-Fascetti and C. Rappaport: Effectiveness of 2D FDTD ground penetrating radar modeling for bridge deck deterioration evaluated by 3D FDTD, International Geoscience and Remote Sensing Symposium, USA, Jul. 2008.
138) Japan Society for Artificial Intelligence (Editor: T. Kamishima): Deep learning (in Japanese), Kindaikagakusha, 2017.
139) T. Okatani: Deep learning - machine learning professional series - (in Japanese), Kodansha, 2015.
140) I. Goodfellow, Y. Bengio and A. Courville: Deep learning, the MIT press, 2016.
141) Y. LeCun, Y. Bengio and G. Hinton: Deep learning, Nature, No. 521, pp. 436 - 444, 2015.
142) G. E. Hinton and R. R. Salakhutdinov: Reducing the dimensionality of data with neural networks, Science, No. 313, pp. 504 - 507, 2006.
143) Y. LeCun, L. Bottou, Y. Bengio and P. Haffner: Gradient-based learning applied to document recognition, Proceedings of the IEEE, Vol. 86, No. 11, pp. 2278 - 2324, 1998.
144) K. Kamnitsas, C. Ledig, V. F. J. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert and

B. Glocker: Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation, Medical Image Analysis, No. 36, pp. 61 - 78, 2017.

145) S. Ji, W. Xu, M. Yang and K. Yu: 3D convolutional neural networks for human action recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 35, No. 1, pp. 221 - 231, 2013.

146) C. Yang, A. Rangarajan and S. Ranka: Visual explanations from deep 3D convolutional neural networks for Alzheimer's disease classification, the 30th IEEE Conference on Computer Vision and Pattern Recognition, USA, Jul. 2018.

147) D. E. Rumelhart, G. E. Hinton and R. J. Williams: Learning representations by back-propagating errors, Nature, No. 323, pp. 533 - 536, 1986.

148) N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov: Dropout: a simple way to prevent neural networks from overfitting, Journal of Machine Learning Research, No. 15, pp. 1929 - 1958, 2014.

149) K. Simonyan and A. Zisserman: Very deep convolutional networks for large-scale image recognition, the 3rd International Conference on Learning Representations, USA, May. 2015.

150) A. Krizhevsky, I. Sutskever and G. E. Hinton: ImageNet classification with deep convolutional neural networks, the 25th International Conference on Neural Information Processing Systems, USA, Dec. 2012.

151) C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich: Going deeper with convolutions, the 26th IEEE Conference on Computer Vision and Pattern Recognition, USA, Jun. 2014.

152) C. Szegedy, V. Vanhoucke, S. Ioffe and J. Shlens: Rethinking the inception architecture for computer vision, the 28th IEEE Conference on Computer Vision and Pattern Recognition, USA, Jun. 2016.

153) Y. Shao and R. S. Lunetta: Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points, ISPRS Journal of Photogrammetry and Remote Sensing, No. 70, pp. 78 - 87, 2012.

154) S. Mukkamala, G. Janoski and A. Sung: Intrusion detection using neural networks and support vector machines, the International Joint Conference on Neural Networks, USA, Aug. 2002.

155) J.R.R. Uijlings, K.E.A. van de Sande, T. Gevers and A.W.M. Smeulders: Selective search for object recognition, International Journal of Computer Vision, No. 104, pp. 154 - 171, 2013.

156) R. Girshick, J. Donahue, T. Darrell and J. Malik: Rich feature hierarchies for accurate object detection and semantic segmentation, the 26th IEEE Conference on Computer Vision and Pattern Recognition, USA, Jun. 2014.

157) R. Girshick: Fast R-CNN, the 27th IEEE Conference on Computer Vision and Pattern Recognition, USA, Submitted Apr. 2015.

158) S. Ren, K. He, R. Girshick and J. Sun: Faster R-CNN: towards real-time object detection with region proposal networks, the 27th IEEE Conference on Computer Vision and Pattern Recognition, USA, Submitted Jun. 2015.

159) J. Redmon, S. Divvala, R. Girshick and A. Farhadi: You Only Look Once: unified, real-time object detection, the 27th IEEE Conference on Computer Vision and Pattern Recognition, USA, Jun. 2016.
160) W. Liu, D. Anguelov, D. Erhan and C. Szegedy: SSD: Single Shot MultiBox Detector, the 28th IEEE Conference on Computer Vision and Pattern Recognition, USA, Jun. 2016.
161) P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus and Y. LeCun: OverFeat: integrated recognition, localization and detection using convolutional networks, the 26th IEEE Conference on Computer Vision and Pattern Recognition, USA, Jun. 2014.
162) D. Garcia, L. L. Tarnec, S. Muth, E. Montagnon, J. Poree and G. Cloutier: Stolt's f-k migration for plane wave ultrasound imaging, IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol. 60, No. 9, pp. 1853 - 1867, 2013.
163) T. Sakamoto, T. Sato, P. J. Aubry and A. G. Yarovoy: Ultra-wideband radar imaging using a hybrid of Kirchhoff migration and Stolt F-K migration with an inverse boundary scattering transform, IEEE Transactions on Antennas and Propagation, Vol. 63, No. 8, pp. 3502 - 3512, 2015.
164) M. Sato: Subsurface imaging by ground penetrating radar (in Japanese), Journal of IEICE C, Vol. 85, No. 7, pp. 520 - 530, 2002.

Acknowledgements

The author would like to thank all the professors, staffs and laboratory members. This research cannot be done without the cooperation of the persons concerned. Measurement data was collected and experimental field was newly constructed in the research. The author is grateful for helpful comments and advices on the research. In my private life, the three years of the doctoral course was sometimes hard and sometimes enjoyable. The author could interact many junior students in the laboratories and lectures to maintain the motivations to study and improve the research. The research was satisfying to write this thesis as a compilation of the study of the three years. The author hopes this experience will be useful in the future.

Most of the research is based on deep learning and other machine learning techniques. The author was studying about Ground Penetrating Radar (GPR) and has basic knowledges on civil infrastructures. However, three years is not enough to follow all of the up-to-date deep learning and machine learning algorithms, which are being developed even while writing this acknowledgement. Deep learning is popular in many research fields, which is also being introduced for civil engineering purposes in recent years. Luckily, some acquaintances studying in abroad are doing research about deep learning, who provided some hints. The author hopes this thesis will make an impact on the practice of subsurface sensing and be a stimulus to the future developments of other deep learning and machine learning algorithms for the monitoring and assessment of civil infrastructures.

The author would like to thank my advisor, Prof. Mizutani for his support for my academic life. Without him, of course, the thesis could not have been finished. My research interest is very close to him. We had many discussions about radar, signal processing, deep and machine learning. The author could study in good research environment.

The author would also like to thank my committee members, Prof. Meguro, Prof. Honda, Prof. Nagayama and Prof. Hirano. The professors provided helpful comments and advices. The author thanks Prof. Nagayama for his kind support for my academic life. Prof. Hirano in Tokyo City University provided us some advices based on his outstanding knowledge of electromagnetics.

The author appreciates the related persons of Fujimura Crest Co. Ltd. for their support for the research. As stated above, the research could not have been finished without their support for collecting the measurement data and construction of the experimental field. Training data of the research is based on the inspection results of the skilled inspectors.

The author would like to thank Prof. Di Su for his comments for the research. The author thanks his kind support.

The author would like to thank Prof. Ishihara, Prof. Yamaguchi, Prof. Kikuchi and Prof. Chen for their helpful comments and advices. Prof. Kikuchi and Prof. Chen are very young and kind professors, who gave me useful advices not only about academic matters but also about trivial private things from the view point of their experiences as doctoral course students.

The author would like to thank some graduate students for their support for my research. Ms. Sakaguchi and Mr. Harada helped my analysis and had many discussions about the research. Dr. Takeda was a doctoral course student when I was a bachelor and master course student, always being a good senior researcher for me. I would like to thank secretary Ms. Takasaki and Ms. Takahashi on behalf of the laboratory staffs.

The author remembers many senior and junior members graduated during my bachelor, master and doctor courses. Especially, the author had many chances to talk with the junior members in the same laboratory room. The author felt relaxed through interactions even when the research was not going well. The author hopes their success in the future.

The research activities of the author are financially supported by Japan Society for the Promotion of Science research fellowship for young scientists. The author was also adopted as a research assistant of graduate school of engineering, the university of Tokyo. The author was engaged in some lectures and projects as a teaching assistant.

Finally, the author would like to express my gratitude to my family. Without their understanding, the research could not have been completed.

August 31, 2020

Takahiro Yamaguchi