A Study on Interpretation of Ground Penetrating Radar by Deep Learning

ISO Shinichiro

Fukada Geological Institute

深層学習を用いた地中レーダデータ解釈の研究

磯　真一郎

深田地質研究所

Abstract: The experts analyze and interpret GPR (Ground Penetrating Rader) data to identify the buried objects and sub-surface structure, basing the characteristic difference in reflected wave shape caused by the relative permittivity of the underground. Progress of acquisition system technology and an increase of the infrastructure to be maintained have made a remarkable massive amount of acquired GPR data these years, and they have been accumulated. There is a strong demand to save labor efforts and increase the efficiency of analysis and interpretation. This decade, the ability of image object recognition has dramatically improved by Deep Learning technology, which is one of the Machine Learning algorithms. AlexNet, one of Deep Learning models, demonstrated the superiority of Deep Learning and CNN (Convolutional Neural Networks) comparing to the conventional Machine Learning algorithms. AlexNet equipped CNNs to imitate the real-living visual cortex and was open to dramatically improving Deep Learning accuracy. The author developed the Deep Learning system, which bases the AlexNet model architecture, to identify the hyperbolic reflection detection on the GPR images. This system gave fairly good performance scores, the F score was 0.9819, and the accuracy was 0.9875 to classify the hyperbolic reflections. With this system, the author also visualized the location of the hyperbolic reflections on the pseudo-3D image, which is supposed to be helpful in identifying the possible buried objects efficiently. This report mainly bases on the author's two works in 2019 (Iso et al., 2019; Iso, 2019).

Keywords GPR, Deep Learning, Object recognition, Convolution neural network, AlexNet

要旨：地中レーダの解析は，地下構造の比誘電率の違いによる生じる特徴的な反射波の形状に対して，熟練技術者が目視により走時断面画像上で解釈することで行われている。近年，データ取得システムの技術の発達と、保守の必要なインフラが増加にともない、解釈すべき大量の地中レーダデータが著しく増大している。このため，その解釈の自動化，省力化が期待されている。機械学習の一つである深層学習（ディープラーニング）による物体認識能力は近年大幅に向上し，多くの学習モデル・アーキテクチャが提唱され研究が進んでいる。とくに，AlexNet は生体の視覚野を模した畳み込みニューラルネットワーク（CNN: Convolution Neural Network）などの手法を用い，深層学習精度向上の端緒を開き、深層学習の優位性を示した学習モデルである。本報告では，著者が AlexNet に基づき開発した地中レーダデータでの双曲線反射面を判別するシステムにより、双曲線状の反射を F 値で 0.9819，精度で 0.9875 と，良い結果を得たことを示した。さらに，地中埋設物による双曲線状反射の検出を多チャンネルの現場データに適用し、擬似 3 次元的にその存在を示した。この報告は、主に著者の 2019 年の 2 つの論文に基づいている（磯ほか 2019; 磯 2019）。

キーワード：地中レーダ・深層学習・物体認識・畳み込みニューラルネットワーク・AlexNet

1. Introduction

As the progress of data acquisition system technology and increase of social infrastructure to be maintained, a massive amount of GPR data has been acquired these years remarkably. GPR experts need significant working efforts to analyze and interpret the acquired data to identify the buried object and sub-surface structure. There is a strong demand to improve the efficiency of analysis and interpretation of the system.

1.1 Theory of Ground Penetrating Radar

GPR survey is a method of geophysical exploration. The system has antennas on the ground surface that emit electromagnetic waves and receive the reflected electromagnetic waves due to the subsurface boundary and buried objects (including cavities). Electromagnetic waves are reflected by the difference in relative permittivity between the underground soil and the buried object. Recording (profile measurement) of moving the antenna along the survey line while keeping the constant distance between the transmitting and receiving antennas obtains a profile that represents the underground structure. The velocity at which the electromagnetic wave is traveling in the ground (propagation velocity) is determined by the relative dielectric constant of the underground soil. In order to determine the propagation velocity of electromagnetic waves, the wide-angle measurement method is generally used (The Society of Exploration Geophysicists of Japan, 2016). If the propagation velocity is known, the depth of the buried object can be estimated by the period to the arrival of the reflected wave from the electromagnetic wave transmission. GPR is a high-resolution exploration method compared to other geophysical exploration methods and can acquire a large amount of data per exploration area. The data processing is applied to the acquired data to emphasize the buried object to be detected for the interpretation, such as amplitude recovery and offset removal. The many expert engineers estimate and interpret buried objects by comprehensively judging the shape of reflected waves, the magnitude, and polarity of reflected waves in a two-dimensional travel cross-section (The Society of Exploration Geophysicists of Japan, 2016; Sato, 2017).

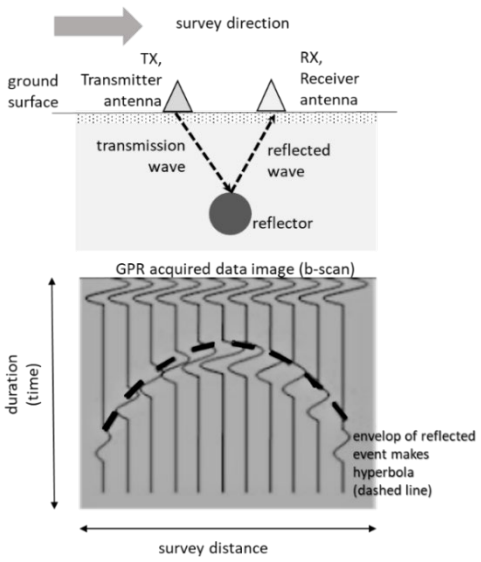

Fig. 1 Principle of Ground Penetrating Rader method and schematic images (B-Scan) of the reflection by embedded objects (Iso et al, 2019).

1.2 Machine Learning

A Neural Network (or called Artificial Neural Network) is one of Machine Learning algorithms for

solving classification and regression problems such as image and speech recognition by altering the connection among nodes in the network through the training process. A multilayered Neural Network is called Deep Learning, which is an algorithm of Machine Learning (Okatani, 2015). Many researchers, including Hinton et al. (2006) showed that multilayer Neural Networks called Deep Belief Networks could learn efficiently, and more complex features were able to extract from the data automatically. Various kind of learning model architectures has been developed and improved for general image object recognition. Especially ILSVRC (ImageNet Large Scale Visual Recognition Challenge) is one of the renowned image recognition competitions. Today, the accuracy of image recognition exceeds humans one, and it is supposed to be made by the recent Deep Learning technologies. CNN (Convolutional Neural Networks), which emulates the feature extraction mechanism of the visual cortex of a living body, contributed significantly to its improvement.

2. Deep Learning Model Architecture

The Deep Learning Model, which was developed by Krizhevsky et al. (2012) (Fig. 1) is an eight-layer Deep Learning model with CNNs, and called AlexNet. The input takes 227 × 227 pixels with three color channels. The first layer convolutes the input image with 11 × 11 for each channel with every four pixels strides and generates 96 outputs. These are activated by ReLU, which is a ramp function. Then the maximum-pooling layer picks the maximum value in the range of 3 × 3 with sliding by two pixels. Convolution is performed in the second, third, fourth, and fifth layers. The fifth layer is the last convolution-layer, which has is 3 × 3 × 192 filters and 256 outputs. The sixth and seventh layers are the fully-connected-layers, and the eighth output layer is the Softmax-layer (Okatani, 2015) to set the output range from 0 to 1. The Deep Learning system contains about 60 million parameters. AlexNet model architecture was designed for the competition of the 2012 ILSVRC, and it often uses some of the object recognition application studies. These days, various Deep Learning frameworks provided the AlexNet model to implement affordably.

This study developed a Deep Learning system that based the AlexNet model architecture, to classify the presence or absence of the hyperbolic reflection waveforms of GPR cross-section image data.

Deep Learning system computes the optimum solution by updating the weighting parameter iteratively during a training process, and some of the parameters, such as learning rate, mini-batch, must be determined manually. If the learning rate is too small, training may take a longer time, and if the learning rate is too large, it may not converge to the optimum value. Therefore it is essential to define the hyperparameters for efficient training.

Mini-batch training is a technique to group the sample data to a small number instead of every single sample of training data in order to update the weights in units efficiently. Mini-batch training, a specified number of samples are randomly selected from whole training data once for the training. An epoch is a state in which a single weight update is completed for all samples to be trained (Okatani, 2015). In other words, assuming the number of data is 1000, and the mini-batch is 10, the state that the weight is updated

100 times per group is one epoch. Based on the preliminary experiments, the learning rate was set at 0.0001, the mini-batch was four, and the epoch was eight times for the training experiments. The object recognition of images is a classification problem. Deep Learning utilizes the probability distribution between the target value (a state that the classification result is properly performed), i.e., true value and the predicted value (network output) are defined by the Softmax function. In order to quantify the difference between these two probability distributions, the loss (Loss) is calculated using a cross-entropy function (Saito, 2016). If the distributions of the target value and the predicted value resemble, the value of the loss should be small. Training execution alters weight of the parameters of neural nodes so that the difference of probability distribution between the true value and the predicted value should be minimizing Two data sets, training data set and the validation data set, are used to calculate both of the training loss and validation loss as well as accuracy value, to monitor the training progress. The accuracy calculated using the training data set is called "training accuracy," and the accuracy calculated using the verification data set is called "verification accuracy," respectively. The profile of loss value, in training execution, is evaluated to judge whether training is being appropriately performed (Okatani, 2015; Goodfellow et al., 2016).

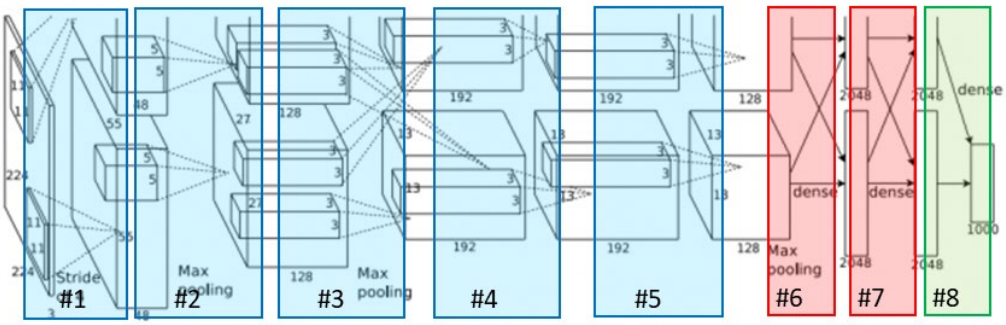

Fig. 2 AlexNet model architecture quoted from Krizhevsky et al. (2012) Figure2. The blue boxes are added to illustrate the five convolution layers, and the red boxes are for the fully connected layers (Iso et al., 2019).

3. Data Set

In general, Deep Learning requires a large amount of data. This study provided the data (labeled images) for three of survey lines acquired data by the identical GPR system. This GPR system acquired the data with 20 channels every 1.0 centimeter in the driving direction. Although the number of samples varies on the survey lines, they were recorded as about 330 samples in one channel (one receiver antenna) and about 27,000 samples along the driving direction. The sample intervals were 0.098 nanoseconds and 1.0 centimeters, respectively. The value of each sample represents the amplitude of the reflection of electromagnetic waves, and they were recorded as a 16-bit unsigned integer type. The GPR system emitted electromagnetic waves with a step frequency method, that is, the frequency from 60 MHz to 2,990 MHz in 294 steps. The acquired nascent data is converted into the time domain by inverse Fourier Transform and

taking the real part so that a waveform equivalent to the pulse type GPR system data can be obtained. Basic preprocessing such as noise reduction according to depth (frequency), a high-pass filter for removing horizontal waves, and signal gain recovery were applied. Since the limitation of computer resources, the amount of data handled was reduced, and only 10 out of 20 channels were processed. Images were generated for the amplitude intensity of the GPR two-dimensional cross-section data as training data.

The reflection event showing the hyperbola in the cross-sectional is supposed to due to the buried object. This study compared the originally acquired cross-section image to the migration processed image (Ishitsuka et al., 2018) in order to find ROIs (regions of the interests). A hyperbolic reflection waveform was called an "event," and the ROI which had an "event" was called "event presence" image.

If the subjected image has an "event," that state is defined as "event presence," If not, it is defined as "event absence." The presence or absence of an event is clarified visually and manually, and each image is labeled.

The area extracted as an image this time was a 65 × 65 square to capture the entire reflected hyperbolic wave, and corresponded to 6.4 nanoseconds in the depth direction and 65 centimeters in the moving direction. This image was enlarged to 227 × 227 pixels by bicubic interpolation to meet the developed Deep Learning system as the training data sets.

Data Augmentation is a technique to increase the number of image data by hiding part of the image, adding noise, enlarging, shrinking, inverting, and rotating the images, in order to improve the training performance (Goodfellow et al. 2016; Okatani, 2015). Considering the data acquisition principle of the GPR, the image "translation" and "stretching and shrinking" are supposed to be reasonable for data augmentation. "Translation" is the case when the hyperbolic reflection waveform is not acquired in the exact center of the target image. The 6-point data is shifted up and down can increase the number of training data nine times to the original one. "stretching and shrinking" was considered as the case that is caused by the dielectric constant of the underground medium local variation. The depth direction in GPR cross-section data was visualized as the arrival time of the reflected wave. Even in the same types of soil, the relative permittivity can change from 2.5 to 5 times due to changes in water content (Daniels, 2004). Since the velocity of the electromagnetic wave in the ground is inversely proportional to the square root of the relative permittivity in the medium (Sato, 2017), the variation of 40% in the relative permittivity of the underground soil makes about 18% variation in depth. The Data Augmentation process set 70% of the number of total images for training, and the rest of them for verification randomly picked at the beginning of training. (Table 1). The verification data was used for monitoring the progress of training.

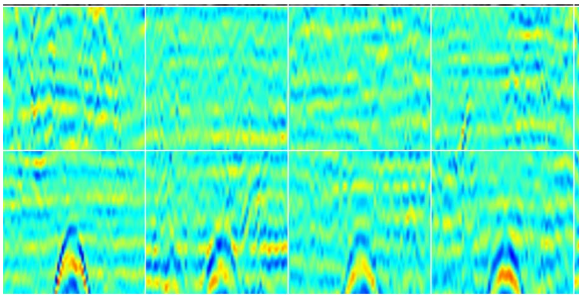

Fig. 3 Example of the target images Upper row shows four event-absence images are in the upper row, and event-presence images are shown in the bottom row.

Table 1 Number of dataset images for the experiments (Iso et al. 2019).

Label	Total images	Original target images	Augmented images by shift	Augmented images by frame extension	Augmented images by frame shrunk
Event presence	4857	440	3537	440	440
Event absence	9064	824	6592	824	824

4. Evaluation method

Fig. 4 shows the confusion matrix indicating the result of classification. TP is True-Positive, the number of cases where the "event" is correctly judged, FN is what can not be judged as an "event," FP is mistakenly judged as an "event," TN is correctly judged "event absence." The training results were evaluated based on the instruction of Goodfellow et al. (2016). It used the F-score in addition to accuracy and presented regarding TPR and PPV (Positive Predictive Value). According to Chinchor (1992), the F-score shows a trade-off between accuracy and recall, and it is better the value closer to one. The scores were calculated as follows.

ACC (Accuracy)

$$\mathrm{ACC} = \frac{TP + FN}{TP + FP + FN + TN}$$

PPV (Precision)

$$\mathrm{PPV} = \frac{TP}{TP + FP}$$

TPR (Recall, or Sensitivity)

$$\mathrm{TPR} = \frac{TP}{TP + FN}$$

F-Score is the harmonic mean of PPV and TPR,

$$\mathrm{F} = \frac{2PPV}{PPV + TPR}$$

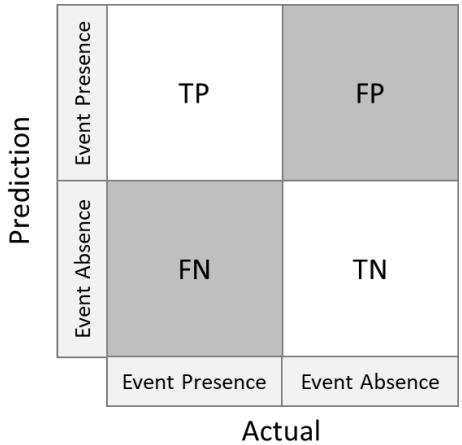

Fig. 4 Confusion plot to illustrate the event presence vs. absence (Iso et al., 2019).

5. Experimental results

Experiments were conducted with the developed Deep Learning system. The number of images, which were applied for training was 9,745, which was 70% of all 13,939 images, as shown in table 1. Each training used the images which were randomly picked. The training execution was monitored by learning curves based on the accuracy and loss of the training data set and the validation data set, not to do over-fitting and lack of iteration. The loss was the cross-entropy error, as described in the previous section, and the likelihood was maximized when the value was minimized.

The experiment was repeated three times, and it was evaluated whether the prediction result was true or false. Table 2 shows the total results for each event of the three experiments. It was observed that "event absence" and "event presence" were determined correctly. This result was summarized as the scores, aka Performance Indicators. The F score was 0.9819, and the accuracy was 0.9875, both of which were nearly close to 1.0.

Fig. 5 shows the learning curve. As learning progresses, Accuracy was improving steadily, while Loss was decreasing. After approximately 180,000 iterations, the loss reduction had a peak. Therefore it was considered that this learning was neither over-fitting nor insufficient fitting.

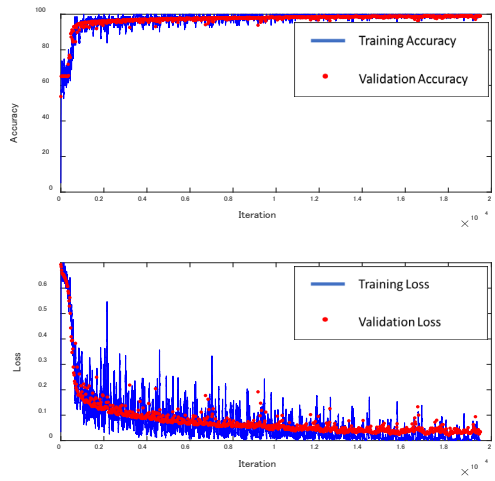

Fig. 5 The learning curves. It shows the validation and training accuracy is gradually increased and mostly saturated middle of the training. The training value is averaged for every 20 iterations (Iso et al., 2019).

This study visualized the predicted buried objects with the GPR cross-section image, as shown in Fig 7 (b). The red-colored clouds consist of the dots, which are the center of predicted images, represent possible buried objects. Fig .7 (a) and (b) show the same GPR image data that have one "void" on the top side of each image, which was identified by an expert. The red cloud in Fig. 7 (b) shows as hyperbolic reflected wave location, and it agreed to include the void, although the red cloud also were shown in non-void reflectors. The prediction was also applied to the multi-channel field data, as shown in Fig. 6.

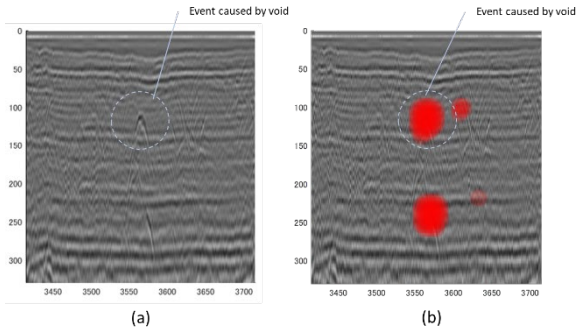

Fig. 6 The Deep learning system identifies the event location which is caused by void #1. Figure (a) shows the scanned event presence images position as red. Figure (b) is the GPR 2D image at the same place as a reference (Iso, 2019).

Table 2 The result (Number of the images)of the prediction (Iso et al., 2019).

	TN (True Negative)	FN (False Negative)	FP (False Positive)	TP (True Positive)
Number of images	2697	30	22	1427

Table 3 The performance indicators (Iso et al., 2019).

Accuracy	TPR	PPV	F1-Score
0.9875	0.9794	0.9849	0.9819

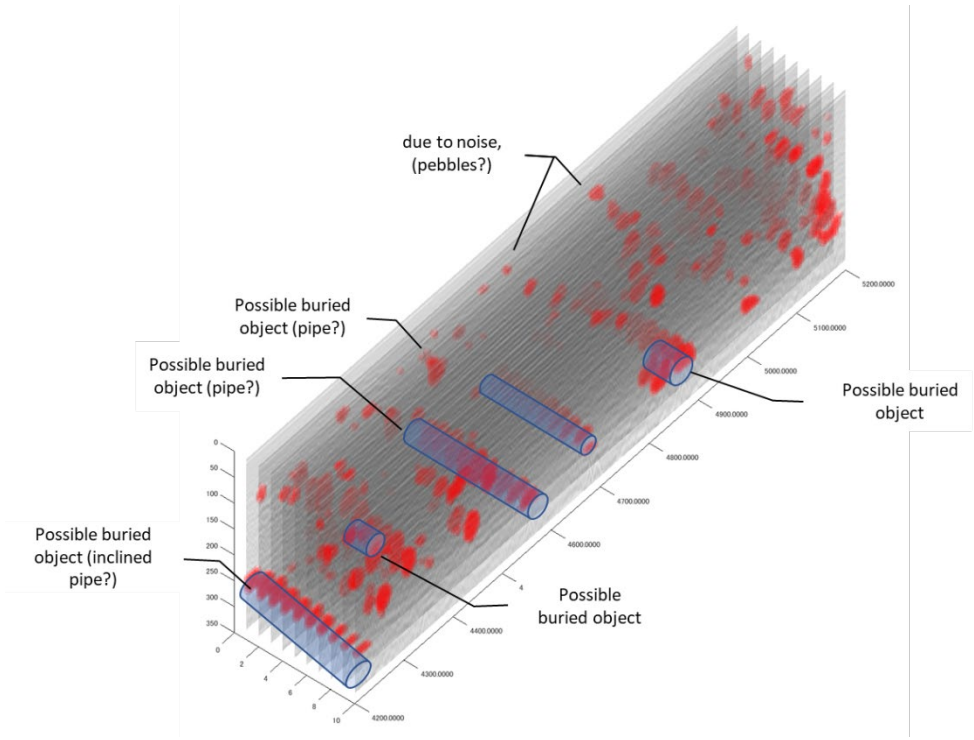

Fig. 7 An example of event detection in multi-channel 2D GPR profile. Dense red color crowd is the group of dots, which shows the center of center of the predicted event image (Iso, 2019).

5. Discussion

These experiments showed that the developed Deep Learning system was able to discriminate with high accuracy that an event contained a hyperbolic reflection waveform (event). It is interesting what features derived by the Deep Learning system and if it is understandable for interpretation by the experts. The degree of weight activation was visualized to analyze it.

Images of the output value at each stage in the neural network layers are called feature maps. In feature maps, portions that react to specific features of the input image makes the significant output as extracted features. The entire image was normalized so that the maximum value was white, and the minimum value s black. With the test data set, "event presence" and "event absence" were examined at each layer.

Fig. 8 is a feature map at the 5th layer, the last maximum pooling layer, and only 9 of the 256 channels are shown. These are the feature map which has the most activated output from the maximum pooling. In the image, where an "event" was present, the hyperbolic reflection waveform was activated and did not react to the background. As observing the significant output in the images, it was able to be seen that "event presence" was determined from the characteristics of the center of the image, and "event absence" was determined by characteristics at the entire image. The pattern of the activation location at the feature map agreed with experts' visual basic check, and it could be realized that the developed Deep Learning system was performed appropriately.

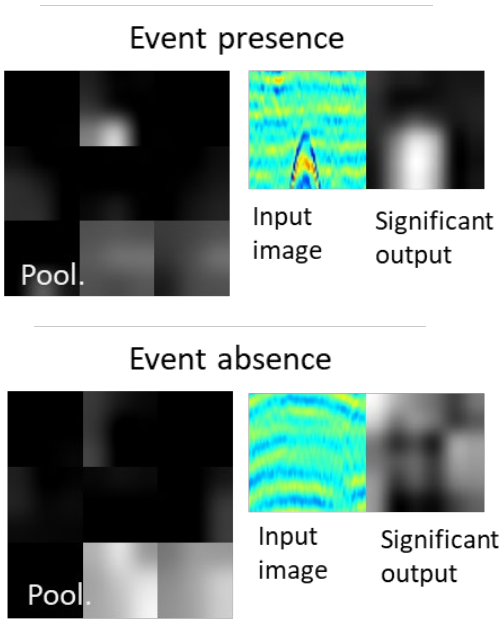

Fig. 8 Output at the fifth layer. The nine images out of the entire 256 of 2D images (Feature map) are shown(left) The middle image is the input image as a reference and the right image is the most significant output of the 256 2D images (Iso et al., 2019).

6. Conclusion

The developed Deep Learning system based on the AlexNet model demonstrated to recognize the hyperbolic reflection waveforms in GPR images, and it was useful to assist the interpretation, precisely a large amount of data, such as multi-channel system. It was examined by the activation pattern of the feature maps that the system extracted the reasonable features automatically. In the case of "event presence," the feature extraction was activated only for the whole image of the hyperbolic shape by the reflected wave.

Besides the efficiency of interpretation, the developed system can prevent the personal psychological deviation from data interpretation, and the system has reproducibility. The system is supposed to be useful for standardization and quantification of interpretation.

Acknowledgements

I would particularly like to thank Dr. Ishitsuka and Dr. Onishi, especially they gave the basics about GPR data processing and theories. I received generous support from Dr. Takahashi and the FGI colleagues. Finally, I am deeply grateful to Prof. Matsuoka.

I also greatly thank Cannan GeoResearch, for accepting to use the filed GPR data.

References cited

Chinchor, N. (1992): MUC-4 Evaluation Metrics, in Proc. of the Fourth Message Understanding Conference, 22–29.

Daniels, J., D. (2004): Ground Penetrating Radar 2nd edition, The Institute of Engineering and Technology, Hertfordshire, UK, Chapter 4, Properties of materials, 74.

Goodfellow, I., Bengio, Y., Courville, A. (2016) : , Deep Learning (Adaptive Computation and Machine Learning), The MIT Press, Cambridge US-MA, Chapter 11 Practical Methodology, 417-419.

Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A Fast Learning Algorithm for Deep Belief Nets. Neural Computation, 18(7), 1527–1554.

Ishitsuka, K., S. Iso, K. Onishi, and T. Matsuoka, (2018): Object Detection in Ground-Penetrating Radar Images Using a Deep Convolutional Neural Network and Image Set Preparation by Migration, International Journal of Geophysics, Article ID 9365184, 8.

Iso. S., (2019): A study of interpretation of ground

penetrating radar data by deep learning, doctoral thesis, Waseda University, http://hdl.handle.net/2065/00063317.

Iso, S., K. Ishitsuka, K. Onishi, and T. Matsuoka (2019): GPR data interpretation by the deep learning with coloring data, Geophysical Exploration, 72, 8-77, https://doi.org/10.3124/segj.72 .

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012) : ImageNet Classification with Deep Convolutional Neural Networks. Advances In Neural Information Processing Systems, 1–9.

Okatani, T., (2015): Deep Learning, Machine Learning Professional Series, Kodansha Scientific, Tokyo, 1-181.

Sato, M., (2017): Basics of Radar, Corona Publishing Co.,Ltd., Tokyo , 9, 233-259.

The Society of Exploration Geophysics of Japan, (2016): Geophysical Exploration Handbook, Tokyo , 7, 233-259.