Rock evaluation of NATM tunnel face using deep learning

Masanari Nakata
Kansai University Graduate School of Science and Engineering, Osaka, Japan

Karnallisa Desmy Halim
Kansai University Graduate School of Science and Engineering, Osaka, Japan

Yeboon Yun
Kansai University, Osaka, Japan

Harushige Kusumi
Kansai University, Osaka, Japan

Akinobu Nishio
Construction Services in Kinki Region, Kinki Regional Development Bureau, Osaka, Japan

ABSTRACT: Because of the complexity of the geological features, when the NATM method is used in Japan, the rock mass is evaluated in nine categories (A. condition of tunnel face, B. condition of excavation face, C. compressive rock strength, D. weathering and alteration, E. spacing of discontinuities, F. condition of discontinuities, G. direction of discontinuities, H. presence of water inflow, I. deterioration due to water) The evaluation is graded on four levels. The objective of this study is to use deep learning to quantitatively evaluate the frequency, condition, and morphology of fractures, as well as weathering and alteration of the tunnel faces; CNN was used to grade the three criteria regarding fractures. Furthermore, ratio of weathering area was detected by HSV color space for categories regarding weathering and alteration. We also applied Grad-CAM to verify whether the CNN model could actually evaluate rock fractures as a decision criterion.

Keywords: New Austrian Tunneling Method, Rock mass rating, Convolutional neural network, Gradient-weighted class activation map.

1 INTRODUCTION

In Japan, the New Austrian Tunneling Method (NATM) is a very popular tunnel construction method because it can adapt to the complex geological formation of Japan. This construction method relies on the surrounding rock mass to ensure the stability of the structure. Furthermore, to maximize safety and minimize costs, support structure (determined through support patterns) could be change from the original design based on the observed rock mass as stated in the "Index for Road Tunnels Observation and Measurement (2009)". Rock mass on tunnel face are evaluated and graded based on a set of criteria, and labelled with a support pattern. Rock mass evaluation are typically done by onsite engineers, but since these decisions are based on their individual experiences, there is a discrepancy in judgement resulting in differing evaluations. This study will apply a type of deep learning method, the Convolutional Neural Network (CNN), to the process of rock mass evaluation. This study will focus on evaluating the visually observable features of rock mass, namely the rock fractures. Finally, Gradient-weighted Class Activation Map (Grad-CAM) is implemented to visualize

CNN. In addition to the evaluation of rock fractures, the weathering area of rocks is quantitatively extracted by using HSV color space.

2 TUNNEL FACE EVALUATION IN JAPAN

Companies and government agencies often have their own indexes for tunnel face evaluation. In this study, the index defined by the Ministry of Land, Infrastructure, Transport and Tourism (MLIT) was used. The indexes defined by MLIT for tunnel face observation is shown in Table 1. The nine evaluation items that indicate the observation results of the face are (A) condition of tunnel face, (B) condition of excavation face, (C) compressive rock strength, (D) change in rock mass condition due to weathering, (E) spacing of discontinuities, (F) condition of discontinuities, (G) direction of discontinuities, (H) presence of water inflow, and (I) deterioration due to water. Each of them is evaluated on a 4-point scale. In this study, we focus on four criteria: (E) spacing of discontinuities, (F) condition of discontinuities, (G) direction of discontinuities, and (D) change in rock mass condition due to weathering.

Table 1. Criteria of tunnel face evaluation.

	Criteria	**Description**
(A)	Condition of tunnel face	State of rock mass at tunnel face
(B)	Condition of excavated surface	State of rock fall at tunnel face
(C)	Compressive rock strength	Hardness of rock
(D)	Change in rock mass condition due to weathering	Degradation by weathering
(E)	Spacing of discontinuities	Interval of discontinuity
(F)	Condition of discontinuities	State of discontinuity
(G)	Direction of discontinuities	Shape and direction of discontinuity
(H)	Presence of water inflow	Wetness of rock mass
(I)	Deterioration due to water	Degradation caused by spring water

3 CNN(CONVOLUTIONAL NEURAL NETWORK) ・ GRAD-CAM ・ HSV COLOR SPACE

In recent years, CNN is a type of deep learning method widely used in image recognition. By inputting images into several featured layers, the CNN model is able to analyze the images. A CNN model is typically made up of a convolutional layer, a pooling layer, and a fully connected layer. Firstly, features of the image is being selected and recognized in the convolutional layer through several filters. Next, the image goes through the pooling layer from the convolutional layer and gets spatially reduced to decrease calculation time. This process is repeated and the features selected from the layers so far are combined in the fully connected layer to predict outputs. Finally, the margin of error from the forward propagation (from the first layer to the final layer) output and the prepared supervised data is calculated. This is then processed through back propagation (from the final layer to the first layer) and by implementing the stochastic gradient descent, the bias and weight values of each layer is regulated, minimizing the error and optimizing the model.

Grad-CAM is a visualization technique for CNN output and is proposed by Selvaraju et al. (2017). The feature map and output results are extracted during forward propagation, and these results are then used in backward propagation to calculate the gradient of each attribute map. The bigger the pixel, the bigger the influence of the gradient on the prediction results. Therefore, by creating a heatmap of the calculated gradient, the parts of the image CNN observed to make the predictions can be visualized.

HSV color space is a system that expresses colors in terms of three elements: hue, saturation, and value/brightness. Hue is an element that defines a specific color such as red or green, saturation is an element that expresses the vividness or intensity of a color defined by hue, and brightness is an element that expresses the lightness or darkness of a color defined by hue. The advantages of the HSV color space include the possibility of fine color adjustment and the ease of color adjustment.

4 EVALUATION OF ROCK FRACTURE USING CNN

In this study, the CNN model is constructed, trained, and validated to evaluate rock fractures, using the tunnel face image as input in the following evaluation indexes; (E) spacing of discontinuities, (F) condition of discontinuities, (G) direction of discontinuities. The output is the 4-level evaluation points of the tunnel face. Grad-CAM is applied to image data of criteria (E), (F), and (G) to confirm the recognizability of the rock mass. The process of rock fracture evaluation is shown below. The programming language used to implement the CNN is Python, and the environment is built in Google Colaboratory. Keras, a deep learning library, is used to build the CNN model, and OpenCV is used for image processing.

4.1 Image data used in the study

The input image is pre-processed as shown in Figure 1. Pictures of tunnel faces photographed at construction sites were trimmed to remove the surrounding rock mass, and segmented into 1m^2 (400px*400px) images. Thus, the crest is divided into 14 images and the left and right shoulders are divided into 10 images respectively. Images without sprayed concrete and bolts are selected to be the input data. The labeled data for the input images are evaluation points used at the actual construction site, and revised evaluation points based on each divided image.

4.2 Composition of CNN model

Table 2 shows the structure of the CNN model used in this study. The CNN consists of four convolutional layers, three pooling layers, and three fully connected layers. The pooling layer uses Max pooling, which takes the maximum value in a filter generally applied to CNNs, the ReLU function for the CNN and NN parts, and the Softmax function for the output layer, which is applied to classification problems. In addition, several Dropout layers are added to the NN part to prevent overfitting, which is likely to occur in multi-layer learning.

4.3 Training and validation of the CNN model

The images used to train and validate the CNN model and the evaluation criteria for the grading scores are listed in Table 3. A total of 6064 images are used for evaluation items E, F, and G. The k-Fold cross-validation (k=5) is applied as the validation method. This method divides all image data into k datasets (5 in this study), trains k-1 datasets, and validates the remaining 1 dataset, as shown in Figure 2. The number of epochs for training is 300, batch size is 16, and Adagrad is used for the loss function to compute the error between the predicted and correct values and for the gradient descent algorithm to update the weights and biases.

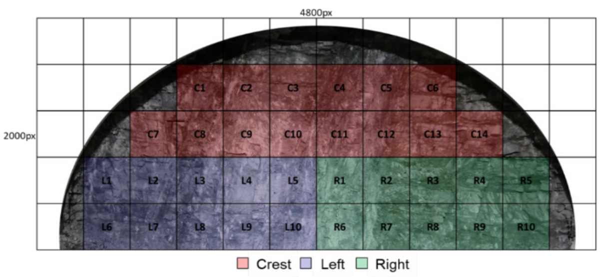

Figure 1. Revision to tunnel face segmentation size.

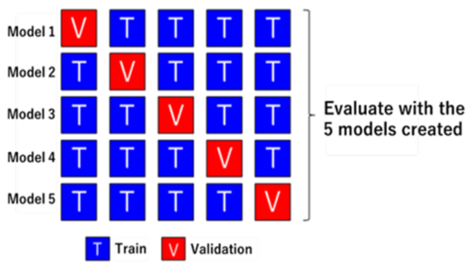

Figure 2. k-Fold cross-validation (k=5).

Table 2. Composition of CNN model.

Layer	Filter size	Number of Filter	Strides	Output size
Conv_1	(3, 3)	16	(1, 1)	(None, 400, 400, 16)
Max_pool_1	(2, 2)		(2, 2)	(None, 200, 200, 16)
Conv_2	(3, 3)	32	(1, 1)	(None, 200, 200, 32)
Max_pool_2	(2, 2)		(2, 2)	(None, 100, 100, 32)
Conv_3	(3, 3)	64	(1, 1)	(None, 100, 100, 64)
Conv_4	(3, 3)	64	(1, 1)	(None, 100, 100, 64)
Max_pool_4	(2, 2)		(2, 2)	(None, 50, 50, 16)
Dense1				(None, 64)
Dense2				(None, 32)
Dense3				(None, 4)

Table 3. Number of data in each evaluation criteria.

Evaluation Criteria	All Data	Grade1	Grade2	Grade3	Grade4
(E)Spacing of discontinuity	6064	Space d >= 1m, no cracks	1m > d >20cm	20cm > d >5cm	5cm > d Smashed unconsolidated
		646	3232	1152	1034
(F)Condition of discontinuity		Joint is closed	Some opening in joint	Open joint	Clay-like, unconsolidated
		537	4570	14	943
(G)Direction of discontinuity		random	column	Stratified, fragmented	Sand-like, unconsolidated
		1058	19	4156	831

5 CHANGE IN ROCK MASS CONDITION DUE TO WEATHERING EVALUATION IN HSV COLOR SPACE

As weathering and alteration progresses, the color of the bedrock tends to change. Therefore, considering that criteria (D) Change in rock mass condition due to weathering can be evaluated by detecting the percentage of weathered areas. In this study, we used deep learning to detect the fraction of weathered and altered areas.

The common weathering colors for many rock types are reddish brown. Therefore, we set the hue of the HSV color space to 0-90 and 140-179 to detect the weathering areas. These values were determined based on the HSV color space color range shown in Figure 3. The saturation was set from 30 to 255. Since each region of Japan has its own unique rock types, we added the region-specific color range values to the common weathering area values to detect the weathered area by region. All values of brightness are detected.

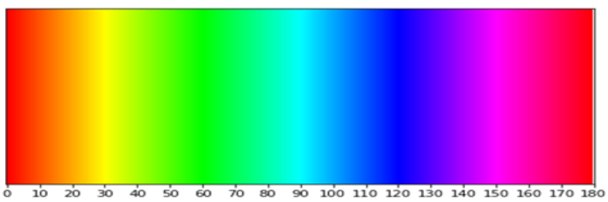

Figure 3. HSV color space range.

6 ANALYSIS RESULTS

6.1 Results of rock fracture evaluation by CNN

$$accuracy = \frac{Number\ of\ matches\ between\ predicted\ and\ correct\ values}{Number\ of\ all\ data} \times 100 \qquad (1)$$

Figure 4 shows the decision accuracy of the CNN model for each crack evaluation item(E,F,G). Accuracy in this study refers to the percentage of correctly predicted evaluation points (1-4) of each criterion (A-I) and can be expressed by Equation (1).

While accuracy is the main focus of model performance, precision and recall are also provided for each evaluation item and grade. The precision is the percentage of data predicted to be true positive (e.g. percentage of the data predicted by the CNN to be grade 1 is actually grade 1). The recall is the percentage of correct predictions out of those that are actually positive. (e.g. percentage of the data that the CNN was able to evaluate as grade 1 out of those that were actually grade 1).

The results are shown in Figure 4. The average accuracy of evaluation items (E), (F), and (G) was 93.8%, 96.8%, and 94.4%, respectively, indicating that deep learning is feasible to evaluate the criteria regarding cracks. The reason for the poor recall and precision of the evaluation point 3 of criteria F (condition of discontinuity) is thought to be due to the small number of images. The same factor applies to criteria G (direction of discontinuity).

Next, the Grad-CAM results are shown. It can be seen that Grad-CAM focuses on cracks. However, the results for E (spacing of discontinuity) and F (condition of discontinuity) show that the color is used to determine whether the rock is whole or sediment-like, which may indicate a risk of judging the originally brown rock as earthly debris. From Figures 5-8, it can be seen that for G (direction of discontinuity), the CNN judges according to the evaluation criteria of grades 1 to 4.

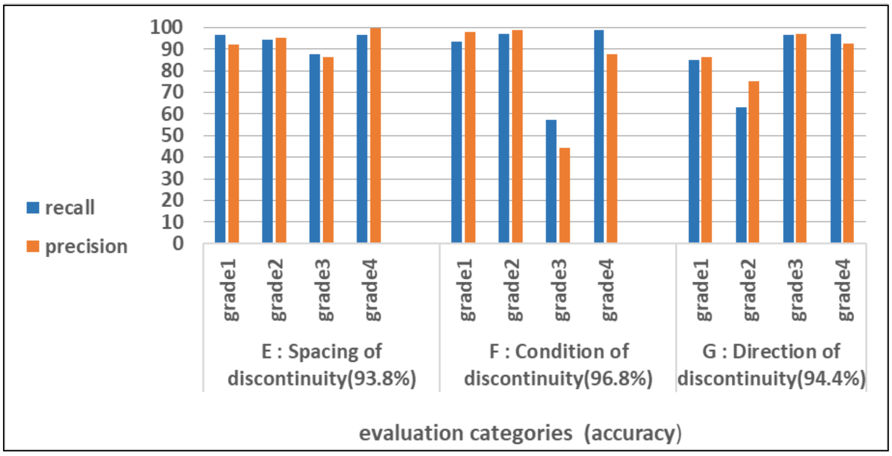

Figure 4. Results of rock fracture evaluation by CNN.

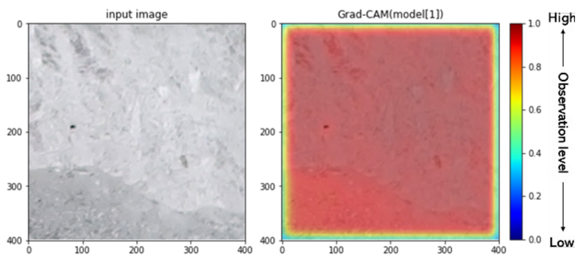

Figure 5. Heatmap of rock mass image by Grad-CAM (Grade 1).

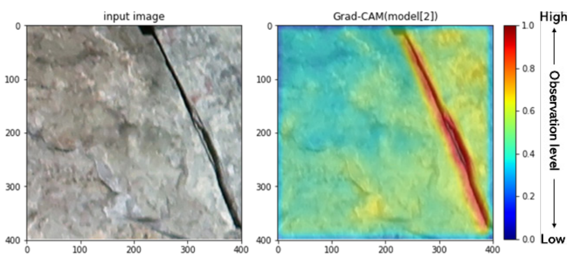

Figure 6. Heatmap of rock mass image by Grad-CAM (Grade 1).

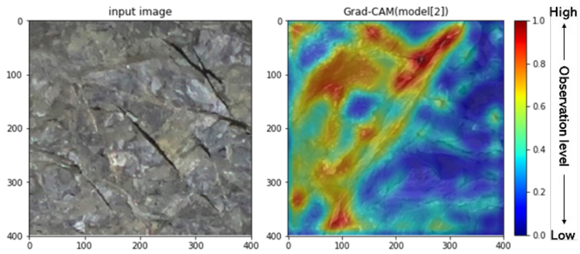

Figure 7. Heatmap of rock mass image by Grad-CAM (Grade 3).

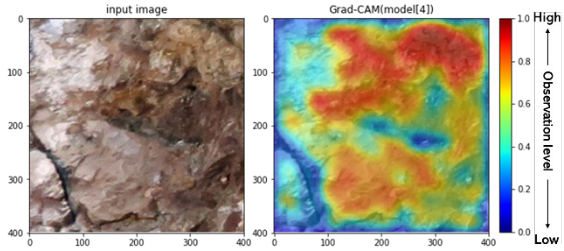

Figure 8. Heatmap of rock mass image by Grad-CAM (Grade 4).

6.2 Result of area of change in rock mass condition due to weathering using HSV color space

From Figures 9 and 10, the change in rock mass condition due to weathering area could be extracted by using the HSV color space.

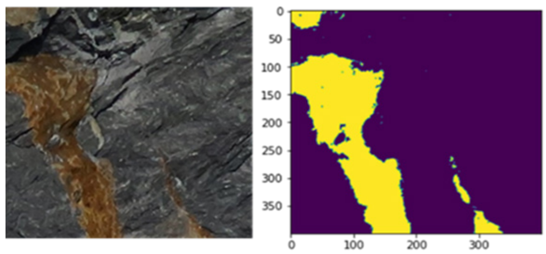

Figure 9. Change in rock mass condition due to weathering extraction image (19.2%).

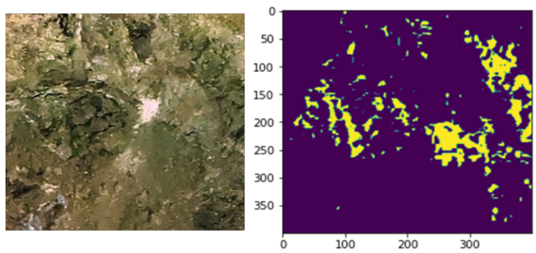

Figure 10. Change in rock mass condition due to weathering extraction image (10.1%).

7 CONCLUSION

In this study, one of the deep learning methods, CNN and its visualization method, Grad-CAM, is applied to the analysis of rock joints in tunnel faces. The weathering area was calculated using HSV color space. This study can be concluded as the following.

- In the analysis using CNN, accuracy was 93.8% for criteria (E), 96.8% for criteria (F), and 94.4% for criteria (G).
- In evaluation criteria (G), it is clear that the CNN model is able to evaluate rock fractures as a basis for decision making.
- HSV color space allowed the weathered area ratio based on discoloration of the rock surface to be correctly calculated.

Based on the results achieved in this study, CNN is effective in the quantitative evaluation of rock joints, and further study is still required for the quantitative evaluation of weathering alteration.

REFERENCES

Hino, Tomohiro. 2017. *Manual for Rock Mass Evaluation during Tunnel Construction*(in Japanese). Proceedings of the Japan Road Conference (CD-ROM). 32nd. ROMBUNNO.6001

Ramprasaath R. Selvaraju *et al.* 2017. *Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization.* In: 2017 IEEE International Conference on Computer Vision (ICCV). DOI: 10.1109/ICCV.2017.74