Comparative detection of rock cracks based on two convolutional neural networks

Xuegang Zhang[a], Tao Wang[*b,c], Shanji Chen[b,c], Yubo Wang[a]
[a]School of Physics and Electronic Information Engineering, Qinghai Minzu University, Xining 810007, Qinghai, China; [b]National Demonstration Center for Experimental Communication Engineering Education, Xining 810007, Qinghai, China; [c]Key Laboratory of Communication Engineering, Xining 810007, Qinghai, China

ABSTRACT

Rock mass fractures are one of the main factors leading to slope instability, and the detection of rock mass fractures can predict slope instability to a certain extent. The rapid development of deep learning has provided a low-cost and efficient method for the detection of rock mass fractures. This paper employs the DenseNet121 model and the InceptionV2 model for the detection of rock mass fractures, and improves the models by incorporating an attention mechanism. The dataset consists of rock masses with fractures from various regions to enhance the model's applicability in different scenarios. Experiments have revealed that the InceptionV2 family of models exhibits overall better performance than the DenseNet121 family of models. Among them, the InceptionV2-ECA model performs the best with an F1 score of 0.9850 and an accuracy rate of 98.73%. Compared to the original InceptionV2 model, the accuracy rate has increased by 9.57%.

Keywords: Rock cracks, DenseNet121, InceptionV2, attention mechanism

1. INTRODUCTION

Slope disasters occurring worldwide constantly threaten people's property and safety, and the demand for underground engineering and geological disaster prevention is continuously increasing[1]. The prevention of slope disasters is a critical issue in the fields of underground engineering and geological disaster prevention. Slope instability and landslides pose threats not only to personal and property safety but also have significant socio-economic impacts. Rock fractures are one of the primary factors causing slope instability. With the rapid development of infrastructure in China, including highways, railways, and urban construction, extensive tunnel excavations and mountain road projects are underway[2]. These engineering activities alter the natural state of the original rock mass, leading to stress redistribution, which may cause the formation and propagation of rock fractures. Additionally, natural factors such as rainfall and earthquakes can exacerbate the development of rock fractures, further triggering slope disasters. Therefore, researching rock fracture detection technologies is of significant practical importance for preventing slope disasters. Traditional rock fracture detection methods primarily rely on manual observation and expert judgment, which are time-consuming and labor-intensive. These methods are also susceptible to subjective influences, leading to inconsistencies and reduced accuracy in detection results. In recent years, with the rapid advancement of computer vision and machine learning technologies, automated rock fracture detection methods based on deep learning have gained increasing attention. Deep learning, especially convolutional neural network (CNN), has demonstrated exceptional performance in image recognition and processing. CNNs automatically extract latent features of images through multi-layer convolution and pooling operations, effectively recognizing complex image patterns. Therefore, applying deep learning to rock fracture detection can improve detection efficiency and accuracy, reducing human error. Existing research indicates that deep learning models can achieve automated detection of rock fractures[3-7]. For example, Li et al. (2023) proposed an improved YOLOv7 algorithm based on the SimAM attention mechanism for rock fracture detection, achieving favorable detection results[8]. Ali et al. (2020) proposed using models such as ResNet101 and MobileNetV2 for concrete crack detection, achieving significant results[9].

This study aims to develop a general rock fracture detection model based on convolutional neural networks in deep learning, utilizing real rock fracture images from various regions for training, to enhance the model's applicability and detection accuracy in different scenarios. Additionally, by incorporating techniques such as data augmentation, the study

[*]wt_2004331@163.com; phone 13997069090

International Conference on Optics, Electronics, and Communication Engineering (OECE 2024), edited by Yang Yue, Proc. of SPIE Vol. 13395, 133952Q · © 2024 SPIE · 0277-786X · Published under a Creative Commons Attribution CC-BY 3.0 License · doi: 10.1117/12.3049238

aims to maintain high detection performance even with limited data availability. Furthermore, this study employs both DenseNet121 and InceptionV2 for fracture detection, and through comparative experiments, selects the better-performing model as the target model. Simultaneously, different attention mechanisms are added to the base models for comparative experiments to identify the one that most significantly enhances performance, thereby improving the model. The study found that the InceptionV2 series models outperform the DenseNet121 series models overall. Among them, the InceptionV2-ECA model exhibits the best performance, while the DenseNet121-SE model shows the worst performance.

2. MODEL INTRODUCTION

2.1 DenseNet121 model architecture

DenseNet (Densely Connected Convolutional Network) is a deep learning model primarily used for image classification tasks. Through dense connections, this model ensures that the output of each layer is directly inputted to every subsequent layer. This facilitates efficient information flow and effective gradient propagation. Not only does this reduce the number of parameters in the model, but it also mitigates the issue of vanishing gradients, making the entire training process more stable and efficient. DenseNet consists of multiple Dense Blocks and Transition Layers, with each layer within a Dense Block interconnected. The DenseBlock also employs a structure of BN (Batch Normalization)+ReLU (Rectified Linear Unit)+Conv (Convolution). This design differs from the conventional Conv+BN+ReLU, specifically to handle the situation where the input to the convolutional layer includes the output features from all previous layers. Since these features come from different layers, their numerical distributions may vary significantly. To address this issue, we first pass the input features through a BN (Batch Normalization) layer for standardization before performing the convolution operation. Transition Layers are used to connect adjacent Dense Blocks and also to adjust the size and number of channels of the feature maps. DenseNet121 is a common variant of the DenseNet model, where 121 denotes the total number of convolutional layers and fully connected layers in the model.

2.2 InceptionV2 model architecture

Inception, also known as InceptionNet or GoogLeNet, is a deep learning model used for image classification. The core component of the model is a convolutional structure called the Inception Module. This structure combines various convolutional layers and pooling layers into a single module. The entire model assembles these modules into a complete network architecture. The model consists of multiple such modules and includes auxiliary classifiers to aid in gradient propagation. InceptionV2 is an improved version of InceptionV1. The main improvement is the introduction of BN layers, inserting a Conv-BN-ReLU structure between convolutional layers and activation functions. Additionally, the 5×5 convolutional kernel in InceptionV1 is replaced with two 3×3 convolutional kernels, increasing depth while significantly reducing the number of parameters. The initial layer consists of three convolutional layers and two max-pooling layers. This is followed by ten layers of Inception Modules, which include two different types of Inception Modules, denoted as IM1 and IM2 in the figure. The output layer consists of average pooling and fully connected layers. Since the model outputs two classes--images without fractures and images with fractures--the fully connected layer has two nodes.

2.3 Model improvements

The experiment involved adding different attention mechanisms to two models and selecting the one with better performance for model improvement. Attention mechanisms can be seen as intelligent filtering tools that allow the model to focus on the most relevant information while ignoring less important details. This enables the model to more accurately identify the features of images. Additionally, attention mechanisms can enhance the interpretability of the model's decisions by highlighting the information the model focuses on when making decisions.

This experiment employed SE (Squeeze-and-Excitation) attention mechanism, CAM (Channel Attention Mechanism), SAM (Spatial Attention Mechanism), CBAM (Convolutional Block Attention Module), and ECA (Efficient Channel Attention) attention mechanism. In the DenseNet121 model, the attention mechanisms were added after the convolutional layers in the Dense Block and Transition Layer. In the InceptionV2 model, the attention mechanisms were added after the convolutional layers in each Inception module[10-12].

3. EXPERIMENTS

3.1 Data set design

The training dataset consists of rock mass images from different slope areas, with formats including JPG and PNG, totaling 6,689 images. Of these, 3,834 images depict rock masses with cracks, and 2,855 images show rock masses without cracks. The training set contains 2,230 images without cracks and 3,195 images with cracks. The test set includes 625 images without cracks and 639 images with cracks.

3.2 Model training

The model used the Adam optimizer, with the cross-entropy loss function. During training, the initial learning rate (lr) was set to 0.01, the batch size (batch_size) to 32, and the number of epochs to 40. By continuously adjusting the hyperparameters and observing the changes in the loss function and accuracy over the epochs during training, the best-performing hyperparameters were determined. The final hyperparameter settings are shown in Table 1.

Table 1. Over parameter setting.

Model	lr	Batch_size	Epochs
DenseNet121	0.000001	12	14
DenseNet121-SE	0.000001	12	14
DenseNet121-ECA	0.00001	12	14
DenseNet121-SAM	0.000001	12	14
DenseNet121-CAM	0.000001	12	14
DenseNet121-CBAM	0.000001	12	14
InceptionV2	0.000001	32	15
InceptionV2-SE	0.000001	32	15
InceptionV2-ECA	0.00001	32	15
InceptionV2-SAM	0.00001	32	15
InceptionV2-CAM	0.00001	32	15
InceptionV2-CBAM	0.0001	24	15

After setting the final hyperparameters, the comparison graphs of the loss values and accuracy over the epochs for the DenseNet121 model and its improved model are shown in Figures 1 and 2, respectively. The comparison graphs of the loss values and accuracy over the epochs for the InceptionV2 model and its improved model are shown in Figures 3 and 4, respectively.

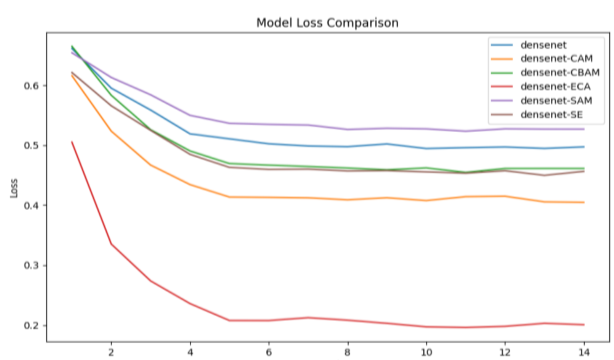

Figure 1. Comparison of loss values across models for DenseNet121.

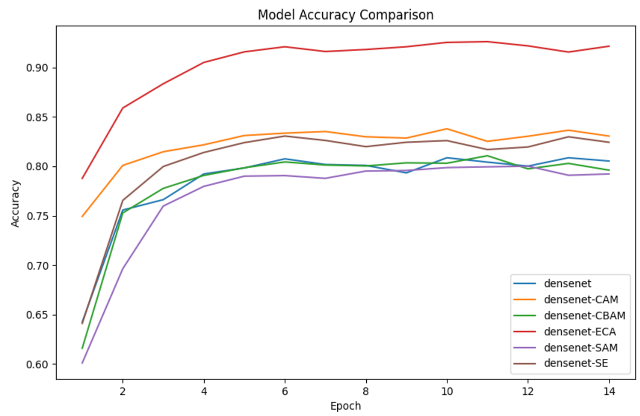

Figure 2. Comparison of the accuracy of each model of DenseNet121.

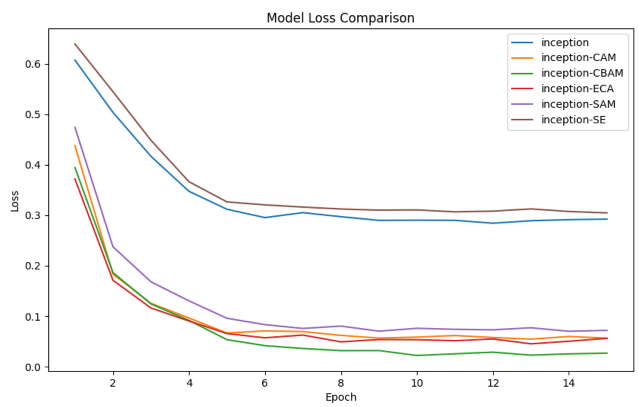

Figure 3. Comparison of loss values across models for InceptionV2.

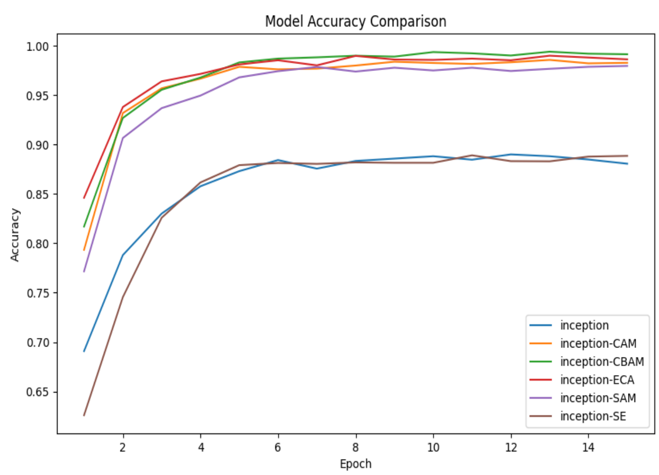

Figure 4. Comparison of accuracy across models for InceptionV2.

3.3 Model validation

After training the model, its performance was evaluated using the test dataset. In this experiment, the performance of the trained model was evaluated using four evaluation metrics: F1 score (F1), recall (Rec), precision (Pre), and accuracy (Acc). The specific experimental results are shown in Table 2 below.

Table 2. Model performance metrics on validation set.

Model	Rec	Pre	F1	Acc
DenseNet121	0.98	0.83	0.8988	0.8916
DenseNet121-SE	1.00	0.81	0.8950	0.8797
DenseNet121-ECA	1.00	0.87	0.9305	0.9280
DenseNet121-SAM	0.97	0.90	0.9337	0.9312
DenseNet121-CAM	0.96	0.91	0.9343	0.9100
DenseNet121-CBAM	0.85	0.92	0.8836	0.8892
InceptionV2	0.96	0.92	0.9396	0.9359
InceptionV2-SE	0.97	0.90	0.9337	0.9359
InceptionV2-ECA	0.99	0.98	0.9850	0.9873
InceptionV2-SAM	0.99	0.98	0.9850	0.9858
InceptionV2-CAM	0.99	0.97	0.9799	0.9771
InceptionV2-CBAM	1.00	0.95	0.9744	0.9731

Analysis of the data in Table 2 reveals that the overall performance of the InceptionV2 series models is superior to that of the DenseNet121 series models. Among various improved models of InceptionV2, InceptionV2-ECA and InceptionV2-SAM demonstrate the best performance, with F1 scores and accuracies of 0.9850 and 0.9873 for InceptionV2-ECA, and 0.9850 and 0.9858 for InceptionV2-SAM, respectively. In contrast, the performance of the DenseNet121-SE model is the poorest, despite its recall rate reaching 1.00. However, its precision is low, at only 0.81, with F1 scores and accuracies being the lowest among all models, at 0.8950 and 0.8797, respectively. Therefore, this experiment concludes that the InceptionV2-ECA model performs the best, while the DenseNet121-SE model performs the worst.

4. CONCLUSION

This study employs DenseNet121 and InceptionV2 models as base models and incorporates five attention mechanisms, namely ECA, SE, CBAM, SAM, and CAM, to conduct rock mass fracture detection. The experiments reveal that the InceptionV2 model with the added ECA attention mechanism performs the best, with F1 score and accuracy reaching 0.9850 and 0.9873, respectively, the highest among all models. However, adding attention mechanisms may also decrease the model’s performance, as evidenced by the decreased precision, accuracy, and F1 score of the DenseNet121 model after incorporating the SE attention mechanism compared to the original model.

The improved InceptionV2 model, although achieving satisfactory results in rock mass fracture detection, still presents areas for improvement in the experiment. For example, expanding the dataset and enlarging the data collection area are necessary. The limited scope of dataset collection and the scarcity of data samples may lead to the model performing well in some environments while exhibiting decreased recognition performance in other regions.

ACKNOWLEDGMENTS

In the fulcrum of academic pursuit, I extend my sincerest appreciation to the cohort of mentors and collaborators whose unwavering support and sagacious guidance were instrumental in navigating the intricate journey of thesis composition. Additionally, I am profoundly grateful for the receipt of financial endowment from the 2022 Qinghai Provincial Key R&D Program for Scientific and Technological Assistance and Cooperation Special Project (2022-Qinghai Cooperation Special Project (2022-QY-205)), which furnished the essential resources for this scholarly endeavor.

REFERENCE

[1] Liu, T., Chen, T., Niu, R., et al., "Landslide detection mapping employing CNN, ResNet, and DenseNet in the three gorges reservoir, China," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14, 11417-11428 (2021).

[2] Ge, Y., Liu, G., Tang, H., et al., "Comparative analysis of five convolutional neural networks for landslide susceptibility assessment," Bulletin of Engineering Geology and the Environment 82(10), 377 (2023).

[3] Yang, Y. and Mei, G., "Deep transfer learning approach for identifying slope surface cracks," Applied Sciences 11(23), 11193 (2021).

[4] Peng, G., Tuo, X., Shen, T., et al., "Recognition of rock micro-fracture signal based on deep convolution neural network inception algorithm," IEEE Access 9, 89390-89399 (2021).

[5] Chen, X., Lian, Q., Chen, X., et al., "Surface crack detection method for coal rock based on improved YOLOv5," Applied Sciences 12(19), 9695 (2022).

[6] Ai, D., Jiang, G., Lam, S. K., et al., "Automatic pixel-wise detection of evolving cracks on rock surface in video data," Automation in Construction 119, 103378 (2020).

[7] Duan, Y., Sun, L. and Wang, Y., "Se-densenet: Attention-based network for detecting pathological images of metastatic breast cancer," Proceedings of the 2019 8th International Conference on Computing and Pattern Recognition, 240-245 (2019).

[8] Li, J., Tian, Y., Chen, J., et al., "Rock crack recognition technology based on deep learning," Sensors 23(12), 5421 (2023).

[9] Ali, S. B., Wate, R., Kujur, S., et al., "Wall crack detection using transfer learning-based CNN models," 2020 IEEE 17th India Council International Conference (INDICON), 1-7 (2020).

[10] Niu, J., Li, H., Zhang, C., et al., "Multi-scale attention-based convolutional neural network for classification of breast masses in mammograms," Medical Physics 48(7), 3878-3892 (2021).

[11] Du, H., Cui, X., Wang, Y., et al., "Food image classification based on CBAM-Inception V3 transfer learning," Science & Technology of Cereals, Oils & Foods 32(1), 91-98 (2024).

[12] Wu, K. Q., Duan, S. B., Wang, Y. J., et al., "Convolutional neural network-based automatic classification for incomplete antibody reaction intensity in solid phase anti-human globulin test image," Medical & Biological Engineering & Computing 60(4), 1211-1222 (2022).