2023 4th International Conference on Computers and Artificial Intelligence Technology

Weekly Earthquake Prediction Algorithm Based on Dense Net

Zhang Heng
School of Computer and Informatics
Hefei University of Technology
Hefei, China
2021111068@mail.hfut.edu.cn

Zhang Guofu*
School of Computer and Informatics
Hefei University of Technology
Hefei, China
zgf@hfut.edu.cn

Zhou Dongrui
Anhui Earthquake Agency
Hefei, China
332086433@qq.com

Su Zhaopin
School of Computer and Informatics
Hefei University of Technology
Hefei, China
szp@hfut.edu.cn

Yue Feng
School of Computer and Informatics
Hefei University of Technology
Hefei, China
yuefeng@hfut.edu.cn

2023 4th International Conference on Computers and Artificial Intelligence Technology (CAIT) | 979-8-3503-2671-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/CAIT59945.2023.10469381

***Abstract*—Earthquakes pose significant risks to life and property, making their prediction a crucial challenge. Traditional approaches to earthquake prediction suffer from limited generalization capabilities due to the influence of specific seismic zone characteristics and struggle with imbalanced seismic data management. Previous research has primarily relied on explicit seismic indicators designed by geologists or implicitly extracted feature vectors using deep learning techniques. However, effective fusion methods for utilizing the abundant information within earthquake data are lacking. To address these limitations, this study proposes a model decoupling methodology that separately trains a backbone network and a classification network. This approach specifically tackles the challenges posed by class imbalance in seismic data. Our research utilizes explicit seismic features derived from precursor signals and employs DenseNet as the backbone network to extract implicit features from seismic data. The fusion of explicit and implicit features enables accurate representation and comprehensive utilization of earthquake data. Evaluation metrics such as accuracy, recall rate, and F1-Score are employed to assess the performance of the proposed method. Experimental results validate its efficacy in earthquake prediction, with high accuracy rate (95.8%), recall rate (95.8%), and F1-Score (0.95). This research substantiates the effectiveness of our approach and its potential impact in the field of earthquake prediction.**

Keywords—earthquake precursors, earthquake prediction, Dense Net, imbalanced data, feature extraction

I. Introduction

Earthquakes are natural disasters that can instantaneously destroy buildings, roads, bridges, and other infrastructure, causing enormous loss of human lives and property. Therefore, predicting earthquakes has been a long-standing goal for scientists. Earthquake prediction involves analyzing various factors such as earthquake precursors, regularity of seismic activity, crustal deformation, and anomalous phenomena before earthquakes to infer information about future earthquake occurrences, including timing, location, and intensity. Research on earthquake prediction helps improve our understanding of seismic hazards and enhances our ability to prevent earthquake disasters, ultimately reducing casualties and property losses.

The Sichuan-Yunnan region is located in southwestern China and holds a prominent position within the internal tectonic structure of the Chinese mainland. It lies on the southeastern edge of the Qinghai-Tibet Plateau. Since the Cenozoic era, this region has been subjected to the northeastward thrusting of the Indian Plate, southwestward gravitational compression from the Qinghai-Tibet Plateau, and obstruction from the South China Block, leading to the formation of a highly active crustal movement and tectonic zone. Over time, a variety of faults with different scales, orientations, and activity rates have developed in the area. The compressional and shearing movements along these fault zones have led to intense seismic activity in the region. The complexity and multiple interpretations of long-term tectonic features have also made this area highly susceptible to sudden geological disasters, making it a vulnerable and heavily affected zone. Consequently, the Sichuan- Yunnan region has become an ideal location for geological scientists to study crustal deformation models, earthquake generation and migration patterns, as well as the formation and evolution mechanisms of the Qinghai-Tibet Plateau.

II. Related Work

In the past decades, many researchers have approached earthquake prediction as a purely geological and physical problem. With advancements in physics and geology, they have sought more effective features and earthquake precursors to forecast future seismic events [1], [2]. For instance, Lei Zhang et al.[2] proposed a precursor pattern-based earthquake prediction feature extraction method, generating eight mathematical statistical features as seismic indicators. In comparison to different models such as SVM, BP, and PNN, experimental results on two historical earthquake records

Authorized licensed use limited to: Chaitanya Bharathi Institute of Tech - HYDERABAD. Downloaded on August 23,2024 at 08:56:16 UTC from IEEE Xplore. Restrictions apply.

demonstrated the effectiveness of the chosen CART algorithm for precursor pattern features in earthquake prediction. Kaftan et al.[3] utilized a combination of Multilayer Perceptron Neural Network (MLPNN), Radial Basis Function Neural Network (RBFNN), and Adaptive Neuro-Fuzzy Inference System (ANFIS) for the assessment of the western region of Turkey. The outcomes of their study revealed that, in comparison to actual data, the proposed RBFNN exhibited a higher correlation coefficient and lower error in predicting monthly earthquake frequencies. Külahcı et al.[4] employed a three-layer Levenberg-Marquardt feedforward learning algorithm to investigate the East Anatolian Fault System. Eight radon-related features were utilized in their artificial neural network to ultimately generate earthquake magnitude predictions. Upon comparing the predicted magnitudes with actual observed results, they encountered a favorable outcome, with an average relative error of approximately 2.3%. Building upon Külahcı et al.'s work, Niksarlıoğlu et al.[5] omitted earthquake depth, introduced wet bulb temperature and dry bulb temperature, and adjusted soil temperature depths to 10 cm, 20 cm, and 50 cm. The resultant trained model exhibited earthquake magnitude errors ranging from 0% to 6.25%.

Pan Xiong et al.[6] employed machine learning techniques to identify pre-earthquake electromagnetic interferences. The experimental findings indicated the superiority of the LightGBM algorithm over alternative methods. Across 13 training datasets, the Area Under the Curve (AUC) consistently exceeded 0.88, with a peak of 0.986, highlighting the precision of pre-earthquake electromagnetic interference recognition. Fangzhou Xu et al.[7] employed the backpropagation neural algorithm and utilized derived features from DEMETER satellite data. After training on earthquakes with magnitudes equal to or greater than 6.0, the results indicated that this method performed well in earthquake area prediction, achieving an 81.2% cumulative rate in seismic zones and 57.5% in non-seismic areas. Unfortunately, the performance of these methods is often constrained by seismic belt characteristics. For instance, Florido et al.[1] predicted earthquakes with magnitude above 4.4 in Chile, Lei Zhang et al.[2] focused on only two regions in China, and Kaftan et al.[3] solely considered a specific area in western Turkey. Previous methods frequently require adjustments or even modifications to prediction algorithms when applied to seismic data with distinct attributes. Moreover, utilizing the Gutenberg-Richter inverse power law to calculate seismic activity parameters for earthquake prediction remains effective only within the current region; if applied to a different region, parameter recalculations are necessary. In summary, these human-designed earthquake indicators (explicit features) possess strong interpretability within theoretical frameworks. However, they might not fully harness the information contained in seismic sequences. As a result, there is a desire to uncover hidden rich features within seismic data.

In pursuit of this, some researchers have attempted to establish neural network models that can directly learn implicit features from data without the need for explicit feature modeling. The theoretical framework of deep learning provides another avenue for earthquake prediction. Deep learning methods possess the capability to automatically extract features and have found widespread application in academic domains such as pattern recognition and natural language processing. In the context of earthquake prediction, training neural networks allows weight vectors to capture features of input data to a certain extent. DeVries et al.[8] introduced a deep learning approach to identify standard based on static stress, aiming to predict aftershock locations without assuming fault orientations beforehand. ROC curves revealed that the performance of the neural network significantly outperformed the classical Coulomb failure stress criterion. However, Mignan et al.[9] expressed skepticism about these findings from DeVries et al.[8] In their experiments, even with a relatively simple neural network employing a two-parameter logistic regression, based on measured distance and mainshock average slip parameter, the same level of performance as the DNN model could be achieved.

Explicit features have strong interpretability but may not fully utilize the latent information in seismic data. In contrast, implicit features can better utilize information in seismic data but have weaker interpretability. Each of these approaches has its own advantages and limitations, and achieving the best performance solely with one approach is challenging. To address these limitations, this study combines explicit and implicit features to leverage their respective strengths. By connecting the implicit features extracted by DenseNet with the original explicit features, the model can benefit from both types of features. This approach allows the model to not only extract the latent information embedded in seismic data but also retain the interpretability of explicit features.

In seismic data analysis, feature selection is a common preprocessing step that helps us choose the most relevant features to describe seismic events. However, relying solely on explicit features may not fully harness all the information embedded in seismic data. To better explore this data, we can employ backbone networks to compute implicit features.

Backbone networks are capable of extracting abstract and high-level feature representations from input data. By inputting the raw seismic data into a backbone network, we can obtain these implicit features, which may encompass hidden correlations and patterns present in the original data.

Subsequently, we concatenate the explicit features with the implicit features, combining the strengths of both types of features to describe seismic data in a more comprehensive manner. Such feature fusion aids in capturing a broader range of seismic data characteristics, thereby enhancing classification accuracy and prediction capabilities.

Finally, the fused features are input into a classifier network to perform classification tasks. Through the classifier network, we can infer specific outcomes from the seismic data.

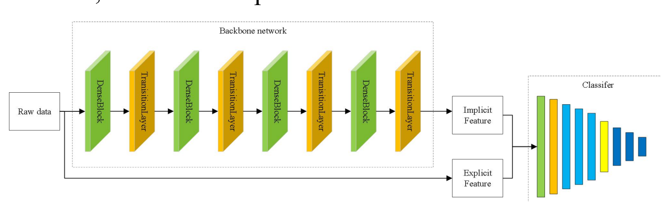

Fig. 1. Complete structure of proposed algorithm.

Authorized licensed use limited to: Chaitanya Bharathi Institute of Tech - HYDERABAD. Downloaded on August 23,2024 at 08:56:16 UTC from IEEE Xplore. Restrictions apply.

A. *Imbalanced Data*

In the dataset of 1105 instances, there are 771 instances of earthquakes below magnitude 3.5 and 38 instances of earthquakes with magnitude 5 and above. This imbalance in data distribution poses challenges in training models, as models tend to favor the majority class, resulting in poorer performance on the minority class with limited samples. To address this issue and enhance the performance across all classes, various approaches can be employed, including resampling the data[10] or designing specific loss functions[11, 12] to better facilitate learning from imbalanced data. Another direction involves transferring knowledge from the majority class to improve the recognition performance of the minority class[13, 14]. However, these methods share the common goal of designing appropriate sampling strategies, losses, and potentially more complex models, while emphasizing joint learning of representations and classifiers.

A study by Bingyi K et al.[15] found that the issue of data imbalance does not necessarily hinder the learning of high-quality representations. Using the simplest form of instance-balanced (natural) sampling to learn features, followed by adjusting the learning of the classifier alone, can yield good results. Based on this insight, Bingyi K et al. divided the learning process of the long-tail classification model into two steps. In the first step, no sampling is performed, and an initial classification model (consisting of a feature-extracting backbone network and a fully connected classifier) is learned directly from the raw data. In the second step, the parameters of the backbone network learned in the first step are fixed, and an individual classifier is added, which is trained using instance-balanced sampling. Additionally, Bingyi K et al. observed a positive correlation between the magnitudes of the weight parameters of the fully connected classifier and the number of corresponding class samples. This implies that classes with larger sample sizes have larger weight parameter magnitudes, leading to higher scores for the majority class during final classification. Therefore, the classifier in the second step is designed as a normalized classifier.

B. *Densely Connected Block*

As convolutional neural networks (CNNs) become deeper, they can become challenging to train. When deeper networks start to converge, a degradation problem can arise: with increasing network depth, accuracy saturates (which is not surprising), and then rapidly deteriorates. Surprisingly, this degradation is not due to overfitting, and adding more layers to an appropriately deep model can lead to higher training errors. Kaiming He et al.[16] proposed the use of skip connections between residual structures, which can be represented using (1).

$$x_{l+1}=x_l+W'*(\sigma(B*(W*(\sigma(B(x_l)))))) \tag{1}$$

The *l-th* residual block's input features are denoted as, and represent the weight matrix, * represents the convolution operation, *B* represents normalization[17], andσrepresents the activation function[18]. The skip connection is achieved by adding the input of the block to the output of a set of convolutional layers, making the network resemble a stack of small networks. However, as the network deepens, the number of parameters becomes significantly large. Furthermore, due to the presence of the vanishing gradient problem, deepening the network does not necessarily lead to improvement. Huang G et al. introduced DenseNet[19], an improvement over ResNet. The *l+1-th* layer receives input from all previous layers' feature maps, as shown in (2).

$$x_{l+1}=H_{l+1}([x_1, x_2, x_3, ..., x_n]) \tag{2}$$

In this context, represents the connection of feature maps from layer 0 to layer *l*, and is a composite function typically involving regularization, activation functions, pooling, and convolution, among others.

In a Densely Connected Block, the output of each layer is connected to the input of the deeper layers within the same block, effectively leveraging the advantages of skip connections. Additionally, the features from shallower layers are reused in deeper layers, resulting in a reduction of the overall number of parameters. This architecture promotes information flow, facilitates feature reuse, and contributes to the efficient training of deep neural networks.

According to the information provided in TableI, the proposed model in this study employs a DenseNet architecture as its backbone network. As the input consists of 1D time-series data, all convolutional operations in the network are 1D convolutions. The backbone network comprises four DenseBlocks and three TransitionLayers. Fig. 2 illustrates a DenseBlock, where the data passes through two DenseLayers and a Dropout layer before being channel-concatenated with the original input data. Each DenseLayer consists of a regularization layer, an activation function, and a 1D convolutional layer. Different DenseBlocks are connected through TransitionLayers, which include average pooling operations for downsampling the feature maps.

The classifier is composed of a DenseBlock, a TransitionLayer, three 1D convolutions, a global average pooling layer, and three fully connected layers. The DenseBlock, TransitionLayer, and 1D convolutions process the features extracted by the backbone network. The global average pooling layer calculates the average value of all features and flattens them. The fully connected layers are responsible not only for downsampling the data but also for classification.

TABLE I. STRUCTURE DIAGRAM OF EARTHQUAKE PREDICTION NETWORK BACKBONE NETWORK BASED ON DENSENET.

Layer	Output
Conv1D	(None, None, 24)
BatchNorm	(None, None, 24)
Relu	(None, None, 24)
DenseBlock1	(None, None, 48)
TransitionLayer1	(None, None, 24)
DenseBlock2	(None, None, 48)
TransitionLayer2	(None, None, 24)
DenseBlock3	(None, None, 48)
TransitionLayer3	(None, None, 24)
DenseBlock4	(None, None, 48)
BatchNorm	(None, None, 48)
GlobalAveragePooling	(None,48)
Dense	(None,5)

Authorized licensed use limited to: Chaitanya Bharathi Institute of Tech - HYDERABAD. Downloaded on August 23,2024 at 08:56:16 UTC from IEEE Xplore. Restrictions apply.

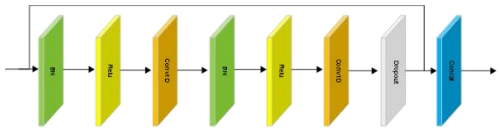

Fig. 2. Dense Layer.

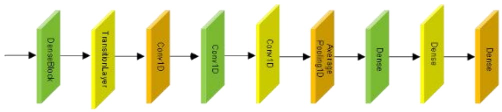

Fig. 3. Classifier.

C. *Train*

In the article, all convolutional layers' kernel weight matrices are initialized using Xavier initialization[20], and bias vectors are initialized to 0. The optimizer chosen is the Adam optimizer, with an initial learning rate of 0.001. A cosine annealing learning rate schedule is employed[21], and the loss function utilized is sparse categorical cross-entropy. The model is built using the TensorFlow framework and trained on an NVIDIA RTX 2070s graphics card.

The training process is divided into two parts: the backbone network and the classifier. In the first step, the backbone network is trained with just one fully connected layer as the classifier. Training is stopped when the loss no longer decreases. In the second step, the classifier is trained. The parameters of the backbone network are fixed, the global average pooling layer and fully connected layers are removed, and the output of the backbone network is combined with the original features as inputs to the classifier. Instance-balanced sampling is used to train the classifier. Training is concluded when the validation loss no longer decreases within 5 epochs.

III. Experiment And Results

To validate the effectiveness of the features, the parameters of the classifier are fixed, and the implicit features, explicit features, and the fusion of implicit and explicit features are separately input. In order to assess the effectiveness of the classifier network used in this study, implicit features and explicit features are input into multiple different network architectures. For performance comparison, the model proposed in this article is compared with two machine learning methods (Weighted-SVM[22], Weighted- Decision Tree[22])and a deep learning method (CNN[23]).

A. *Datasets*

This study utilizes the AETA (Array Observation of Broadband Electromagnetic and Acoustic Disturbances for Earthquake Prediction) seismic dataset for network training. The AETA system consists of 159 stations in the Sichuan-Yunnan region, covering areas such as the Seismic Bureau of Leshan City, Ganhaizi in Jiuzhaigou, and Yajiang in Ganzi. Among these, only 131 stations are capable of recording data. The AETAsystem integrates ground-based processing terminals, electromagnetic sensors, acoustic sensors, and a cloud platform for data storage and analysis. It monitors underground electromagnetic and acoustic signals in real time, transmitting collected data to the cloud platform via wired or wireless networks. After sampling at predefined intervals, 95 feature values related to electromagnetic and acoustic signals are obtained, including variance, power, mean, fourth-level detail energy, sixth-level approximation energy, and more.

The training dataset spans from January 1, 2017, to June 30, 2021, and the testing dataset covers July 1, 2021, to May 30, 2022. The labels are calculated using seismic data from the China Seismic Network. Specifically, data from the two provinces within the time range (January 1, 2017, to May 30, 2023) are selected, encompassing all seismic events within the latitude range of 22°N to 34°N and the longitude range of 98°E to 107°E. This results in a total of 1105 data points. Among these, there are 771 instances of earthquakes below 3.5 magnitude and 38 instances of earthquakes with magnitude above 5. The earthquake catalog data comprises six features, including time, longitude, latitude, magnitude, depth, and reference location. For this research, only four features—time, longitude, latitude, and magnitude—are selected, and the remaining features are discarded.

B. *Evaluation Index*

In this paper, the three indicators of accuracy (Acc), F1-SCORE, and recall (Recall) are used as the measurement standard. Acc is usually used to measure the classification performance of a classifier, and its definition is shown in (3).

$$Acc = (TP+TN)/(TP+TN+FP+FN) \tag{3}$$

Recall is used to assess the coverage of the classifier on the detected targets, and its calculation method is shown in (4).

$$Recall = TP/(TP+FN)$$

The F1-Score considers both precision and recall and serves as a measure of testing accuracy. Its calculation method is shown in (5).

$$F1\text{-}Score = 2*Precision*Recall/(\mathrm{Precision} + Recall)$$

C. *Feature Validity*

TABLE II. Comparison of Feature Effects

Feature	Acc	Recall	F1-Score
Implicit features	84.34%	87.64%	0.86
Explicit features	46.95%	55.80%	0.59
Proposed featuers	95.80%	95.80%	0.95

From Table II, it can be observed that using only explicit features achieves an accuracy of 46.95%, a recall of 55.80%, and an F1-Score of 0.59. Since explicit features can only utilize a portion of the information in the earthquake data and ignore hidden relationships, their performance is relatively poor. Meanwhile, using only implicit features results in an accuracy of 84.34%, a recall of 87.64%, and an F1-Score of 0.86. Although implicit features can leverage the hidden information in earthquake data more effectively compared to explicit features, they may not capture certain explicit characteristics. Hence, by fusing explicit and implicit features and inputting them into the classifier, the performance is significantly enhanced, achieving an accuracy of 95.80%, a recall of 95.80%, and an F1-Score of 0.95.

Authorized licensed use limited to: Chaitanya Bharathi Institute of Tech - HYDERABAD. Downloaded on August 23,2024 at 08:56:16 UTC from IEEE Xplore. Restrictions apply.

D. Network Validity

TABLE III. COMPARISON OF NETWORK EFFECTS

Model	Acc	Recall	F1-Score
Flatten	85.91%	92.81%	0.85
CNN	83.33%	92.92%	0.83
Decoupling DenseNet	95.80%	95.80%	0.95

From the table III, it can be observed that using only CNN and Flatten as the classifier results in accuracies and F1-Scores that do not surpass 90%. This suggests that both approaches are inadequate in effectively integrating explicit and implicit features, as well as in acquiring information effectively. In contrast, the classifier employed in this study first inputs features into DenseNet for feature fusion and extraction, then feeds the extracted information into a convolutional network to further extract information, and finally inputs the data into a Flatten layer to obtain earthquake information. As a result, the proposed algorithm in this study achieves an accuracy of 95.8% and an F1-Score of 0.95, which outperforms the alternatives.

E. Compared with Existing Methods

The performance of the proposed model in this study, along with other deep learning models and traditional methods, was evaluated and compared using the AETA station data from the Sichuan-Yunnan region for the years 2021-2023. The deep learning models used for comparison were retrained based on the AETA data. The test dataset comprised electromagnetic and ground sound signals collected at ten-minute intervals from each station. All tests were conducted without any preprocessing or filtering applied to the test data.

Table IV presents the experimental results of the proposed method and the comparative methods on the test set. In terms of Accuracy (Acc), Weighted-DT and Decoupling DenseNet achieved the best results, showcasing the best classification outcomes. Decoupling DenseNet's excellent performance can be attributed to its effective integration of explicit and implicit features. Weighted-DT's multiple branches compared to Weighted-SVM contribute to better feature handling. CNN and DenseNet both rely on implicit features and, due to the imbalanced samples that tend to favor the majority class, their accuracy is relatively lower. Recall represents the ratio of correctly predicted positives to the total positives in the dataset. Despite sampling being applied to Weighted-SVM and CNN, due to the scarcity of data in the large earthquake class compared to no-earthquake and small earthquake classes, the models tend to favor the latter. This leads to reasonable accuracy but low recall. On the other hand, Weighted-DT, DenseNet, and Decoupling DenseNet perform better in recognizing tail-end class data samples. F1-Score balances precision and recall, making it an effective measure of a model's recognition ability. The decrease in recognition rates for earthquake samples in the CNN model on the test set could be attributed to its limited parameters and susceptibility to overfitting. Weighted-SVM's nature as a multi-classification is essentially a combination of multiple binary classifications, and since SVM is a linear classifier, it struggles to perfectly separate nonlinear geomagnetic and geosonic data using feature hyperplanes.

Compared to CNN, DenseNet exhibits a similar level of accuracy, with the main differences being in recall and F1-score. Specifically, DenseNet outperforms CNN in terms of recall and F1-score, indicating its superior recognition performance for tail-end classes. This can be attributed to DenseNet's skip-connection and feature reuse mechanisms, which enable better information propagation and utilization, thereby enhancing feature extraction capabilities. However, it's important to note that although DenseNet performs well in terms of recall and F1-score, it still falls short of reaching the optimal level, implying potential limitations in recognizing tail-end classes.

TABLE IV. EXPERIMENTAL RESULTS

Model	Acc	Recall	F1-Score
Weighted-DT	89.96%	89.83%	0.90
Weighted-SVM	75.56%	55.00%	0.70
CNN	78.85%	31.44%	0.30
DenseNet	76.70%	79.70%	0.83
Proposed model	95.80%	95.80%	0.95

The primary advantage of Decoupling DenseNet over DenseNet lies in its two-step training process. Initially, a pre-trained DenseNet is used as the backbone network, leveraging its powerful feature extraction capabilities to extract implicit features from geomagnetic and geosonic data. Subsequently, the implicit features extracted from DenseNet are combined with the original explicit features, and a retrained classifier is used to fuse these features. This approach not only maximizes the feature extraction capability of DenseNet but also effectively addresses its limitations in classifying tail-end classes. Decoupling DenseNet demonstrates significant improvements in accuracy, recall, and F1-score compared to DenseNet alone.

This indicates that by fusing DenseNet features with explicit features, Decoupling DenseNet can better recognize and classify tail-end classes, thereby enhancing overall performance.

IV. CONCLUSION

The reliable prediction of major earthquakes presents one of the most significant challenges in urban disaster management. Earthquake catalogs often exhibit a severe class imbalance, with non-seismic events and minor earthquakes comprising the vast majority of entries. As a result, addressing class imbalance becomes crucial. Concurrently, with the widespread adoption of deep learning algorithms, the demand for utilizing deep learning models for earthquake prediction is increasing.

In this study, DenseNet was employed as the backbone network for feature extraction, and seismic events in the Sichuan-Yunnan region were predicted in weekly intervals. By utilizing DenseNet for feature extraction, the fusion of shallow and deep features was achieved through dense connections,

Authorized licensed use limited to: Chaitanya Bharathi Institute of Tech - HYDERABAD. Downloaded on August 23,2024 at 08:56:16 UTC from IEEE Xplore. Restrictions apply.

facilitating feature reuse. Moreover, the dense connectivity structure enables smoother backpropagation of gradients, making the network more trainable. To tackle class imbalance, a feature decoupling approach was adopted, involving separate training phases for the backbone network and the classifier. During classifier training, the implicit features extracted by DenseNet were combined with the original explicit features and underwent data augmentation to prevent feature loss.

Nonetheless, this study also has certain limitations. Firstly, it exclusively used data from the Sichuan-Yunnan region and did not extend testing to other areas, potentially raising concerns about the model's generalizability, as seismic characteristics may differ across regions. Secondly, the earthquake prediction method, using weekly intervals, may lack timeliness since seismic activity could change over shorter time scales. Additionally, recent research indicates that ultra-low-frequency geomagnetic signals exhibit a stronger correlation with earthquakes and are less susceptible to external disturbances, making them better suited for earthquake prediction.

In conclusion, for future research, it would be beneficial to consider cross-regional studies to enhance the model's generalization capabilities. Simultaneously, exploration of more real-time and accurate earthquake prediction methods, combined with investigations into the features of ultra-low-frequency geomagnetic signals, should be pursued to elevate the effectiveness and reliability of earthquake prediction.

REFERENCES

[1] E. Florido, F. Martínez-Álvarez, A. Morales-Esteban, J. Reyes, and J. L. Aznarte-Mellado, "Detecting precursory patterns to enhance earthquake prediction in Chile," (in English), Computers and Geosciences, vol. 76, pp. 112-120, March 01, 2015 2015.

[2] L. Zhang, L. Si, H. Yang, Y. Hu, and J. Qiu, "Precursory Pattern Based Feature Extraction Techniques for Earthquake Prediction," (in English), IEEE Access, vol. 7, pp. 30991-31001, 2019.

[3] I. Kaftan, M. Şalk, and Y. Şenol, "Processing of earthquake catalog data of Western Turkey with artificial neural networks and adaptive neuro-fuzzy inference system," (in English), Arabian Journal of Geosciences, vol. 10, no. 11, p. 243, 2017/06/02 2017.

[4] F. Külahcı, M. İnceöz, M. Doğru, E. Aksoy, and O. Baykara, "Artificial neural network model for earthquake prediction with radon monitoring," (in English), Applied Radiation and Isotopes, vol. 67, no. 1, pp. 212-219, 2009/01/01/ 2009.

[5] S. Niksarlıoğlu and F. Kulahci, "An Artificial Neural Network Model for Earthquake Prediction and Relations between Environmental Parameters and Earthquakes," (in English), vol. 7, pp. 87-90, 2013.

[6] P. Xiong, C. Long, H. Zhou, R. Battiston, X. Zhang, and X. Shen, "Identification of Electromagnetic Pre-Earthquake Perturbations from the DEMETER Data by Machine Learning," (in English), vol. 12, no. 21, p. 3643, 2020.

[7] F. Xu, X. Song, X. Wang, and J. Su, "Neural Network Model for Earthquake Prediction Using DMETER Data and Seismic Belt Information," presented at the Second WRI Global Congress on Intelligent Systems Dec 16-17 2010. Available: https://doi.org/10.1109/GCIS.2010.237

[8] P. M. R. DeVries, F. Viégas, M. Wattenberg, and B. J. Meade, "Deep learning of aftershock patterns following large earthquakes," (in English), Nature, vol. 560, pp. 632-634, August 01, 2018 2018.

[9] A. Mignan and M. Broccardo, "One neuron versus deep learning in aftershock prediction," (in English), Nature, vol. 574, no. 7776, pp. E1-E3, 2019/10/01 2019.

[10] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," (in English), Journal of Artificial Intelligence Research, vol. 16, no. 1, pp. 321–57, 2002.

[11] K. Cao, C. Wei, A. Gaidon, N. Arechiga, and T. Ma, "Learning Imbalanced Datasets with Label-Distribution-Aware Margin Loss," in 33rd International Conference on Neural Information Processing Systems, 2019, vol. 33, pp. 1567–78.

[12] Y. Cui, M. Jia, T.-Y. Lin, Y. Song, and S. Belongie, "Class-Balanced Loss Based on Effective Number of Samples," presented at the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), January 01, 2019. Available: https://ui.adsabs.harvard.edu/abs/2019arXiv190105555C

[13] Y.-X. Wang, D. Ramanan, and M. Hebert, "Learning to model the tail," presented at the Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, California, USA, 2017.

[14] X. Yin, X. Yu, K. Sohn, X. Liu, and M. Chandraker, "Feature Transfer Learning for Face Recognition With Under-Represented Data," in IEEE Conference on Computer Vision and Pattern Recognition 2019, pp. 5697-5706.

[15] B. Kang et al., "Decoupling Representation and Classifier for Long-Tailed Recognition," in 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2019, pp. 1875–1879.

[16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition 2016, pp. 770-778.

[17] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," presented at the 32nd International Conference on International Conference on Machine Learning, Lille, France, July 06 2015.

[18] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," presented at the 27th International Conference on International Conference on Machine Learning, Haifa, Israel, 2010.

[19] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in 2017 IEEE Conference on Computer Vision and Pattern Recognition 2017, pp. 2261-2269.

[20] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in International Conference on Artificial Intelligence and Statistics, 2010, vol. 9, pp. 249-256.

[21] I. Loshchilov and F. Hutter, "SGDR: Stochastic Gradient Descent with Warm Restarts," in 5th International Conference on Learning Representations, Toulon, France, 2016.

[22] A. Sharma, A. Ahuja, S. Devi, and S. Pasari, "Use of Spatio-temporal Features for Earthquake Forecasting of imbalanced Data," in International Conference on Intelligent Innovations in Engineering and Technology, 2022, pp. 178-182.

[23] J. Huang, X. a. Wang, Y. Zhao, C. Xin, and H. Xiang, "Large Earthquake Magnitude Prediction in Taiwan Based on Deep Learning Neural Network," (in English), Neural Network World, vol. 28, pp. 149-160, 2018.

Authorized licensed use limited to: Chaitanya Bharathi Institute of Tech - HYDERABAD. Downloaded on August 23,2024 at 08:56:16 UTC from IEEE Xplore. Restrictions apply.