Seismic-ionospheric Precursor Prediction Using Deep Learning

Bach-Tung Pham, Pao-Chi Chang and Jia-Ching Wang
National Central University, Taiwan
E-mail: pbtung@g.ncu.edu.tw
E-mail: pcchang@ce.ncu.edu.tw
E-mail: jcw@csie.ncu.edu.tw

***Abstract*- Numerous studies have indicated the potential involvement of seismic-ionospheric precursors (SIPs) in catastrophic earthquakes. These SIPs have been extensively analyzed in terms of characteristics such as the degree of decrease or increase, timing of occurrence, and duration, with a specific focus on electron density or total electron content (TEC). While deep neural networks have demonstrated remarkable accuracy in various applications, their capability to estimate SIPs has been severely limited. In this study, we developed a novel AI prediction model for forecasting SIPs based on a six-channel data sequence: TEC, median TEC, lower bound TEC, upper bound TEC, positive polarity, and negative polarity. To leverage the power of deep learning image classification methods, we treated this data as a six-dimensional image. This approach yielded a maximum accuracy of 67.1 percent.**

I. Introduction

Seismo-Ionospheric precursors (SIPs), occurring before large earthquakes, have been well-documented in previous studies [1],[2],[3]. These notable anomalies typically appear a few days prior to major earthquakes [2],[4],[5]. In this study, we employ a deep learning technique to detect SIPs by analyzing temporal variations in the ionosphere. The experiment is focused on the Japan region due to its high frequency of earthquakes (over 1000 events with magnitudes larger than 5.0) between 1999 and 2021.

II. Data analysis

The Center for Orbit Determination in Europe (CODE) utilizes spherical harmonics expansion and VTEC (vertical total electron content) observations from several hundred ground-based GNSS receiving stations to generate the global ionospheric map (GIM) [6]. The GIM offers a comprehensive global distribution of ionospheric VTEC, covering latitudes from 87.5 degrees to -87.5 degrees and longitudes from -180 degrees to 180 degrees. It provides a latitude resolution of 2.5 degrees and a longitude resolution of 5 degrees. With hourly updates, the GIM enables monitoring of global ionospheric variations and analysis of the ionospheric horizontal structure. For our analysis, we have selected the location at 36.5N and 142E for GIM TEC assessment.

For deep learning training, we utilize six categories: TEC, median TEC, lower bound TEC, upper bound TEC, positive polarity, and negative polarity. Each category is structured as a matrix with dimensions of 24 hours * 30 days. The 24-hour resolution data represents the TEC values for each hour of the day, while the 30-day duration signifies the data collected in the 30-day period leading up to the earthquake. The TEC category represents the GIM TEC at the selected location, providing the ionospheric conditions prior to the earthquake. The median TEC represents the median TEC value calculated from the previous 15 days, serving as a reference state for the ionosphere. The upper bound and lower bound TEC correspond to the boundaries of ionospheric anomaly TEC values, as defined by the following equations,

$$UB=M+k(UQ-M).$$

$$LB=M-k(M-LQ).$$

, which UB is TEC upper boundary, LB is TEC lower boundary, M is the median TEC of the previous 15 days, UQ is the third quartile TEC of the previous 15 days, LQ is the first quartile TEC of the previous 15 days, and k is threshold constant. The positive and negative polarities represent ionospheric states that exceed or fall below the upper or lower

boundary, respectively, indicating positive and negative ionospheric anomalies. The six categories utilized in this study consist of the ionospheric state (TEC), ionospheric reference (median TEC, lower bound TEC, upper bound TEC), and ionospheric anomaly polarity (positive polarity, negative polarity). This comprehensive set of categories is well-suited for training a deep learning model to accurately classify SIAs features.

III. Experiment and Methods:

1. Proposed Method:

This study explores the application of EfficientNetV2 [7], a highly advanced image classification method, for a binary classification task. The original EfficientNetV2 model was slightly modified to suit the specific requirements of this task. The modifications included adjusting the "strides" parameter in the first stem convolution and the first block to one. This modification allows the model to capture more detailed information in the initial stages before down sampling, enhancing its ability to extract relevant features. To mitigate overfitting, the model was equipped with high dropout regularization layers. For binary classification, the sigmoid function was employed as the activation function in the final output layer.

To compare the performance of the modified EfficientNetV2 with other methods, a range of models was selected, varying from lightweight to highly effective. The lightweight model chosen was MobileNetV3-Large [8], while popular methods such as VGG16 [9], DenseNet121 [10], and ResNet50 [11] were also included. Additionally, ConvNext [12] and ResNet-RS [13], which are state-of-the-art convolutional neural networks for classification tasks, were incorporated. By including these diverse models, the study aimed to provide a comprehensive assessment of the performance of the modified EfficientNetV2 in comparison to established and state-of-the-art approaches for classification tasks.

2. Experiments:

2.1. Data Preprocessing:

For this study, we conducted a random selection of 2356 samples. These samples were divided equally into two groups: normal events and SIPs events. Each sample is represented by a 30x24x6 array, indicating a time series with six channels. To approach the problem as an image classification task, we treated the time series data as a 6-channel image array. This means that each channel of the time series corresponds to a particular aspect or feature, much like the different color channels in an image.

To input the data into the model, we resized the 30x24 array to 32x32. When resizing, we added zero padding to maintain the aspect ratio between the width and height of the data. By resizing the data to 32x32, we transformed the time series to a format compatible with image-based models so that they could take advantage of the image classification capabilities of convolutional neural networks (CNNs).

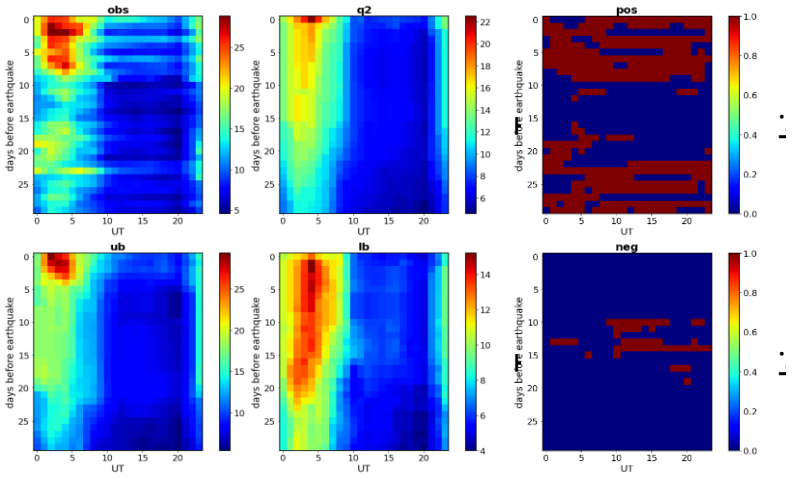

Figure 1. A sample of the dataset (30x24x6)

2.2. Data augmentation:

CutMix [14] and Mixup [15] are widely employed image augmentation techniques in deep learning for enhancing model performance and generalization. CutMix combines cutout [16] and mixup methods by replacing a portion of an image with a patch from another, introducing diversity and perturbations. On the other hand, Mixup blends pairs of images and their corresponding labels. These techniques effectively improve model regularization and generalization, and they have found extensive usage in various computer vision tasks.

2.3. Data division:

We divided the data into three groups: Test, Validation, and Training. Initially, we randomly assigned 20 percent of the data to the test set. The next 10 percent were allocated to the validation set, while the remaining data were designated for the training set.

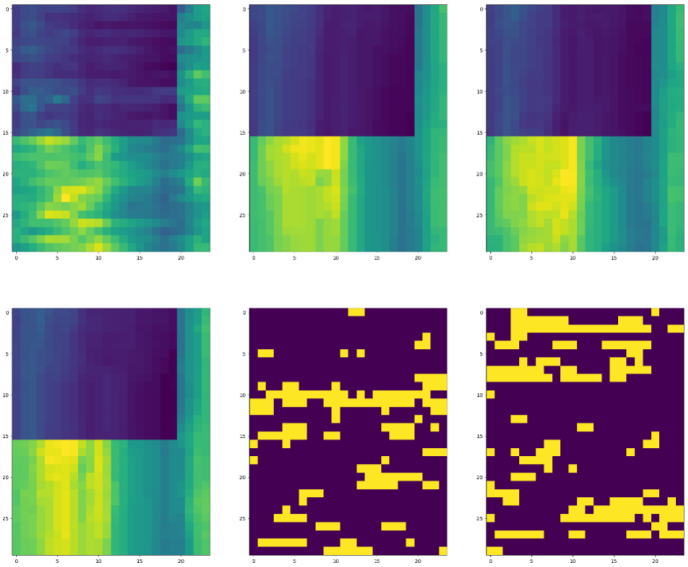

Figure 2. Cutmix augmentation

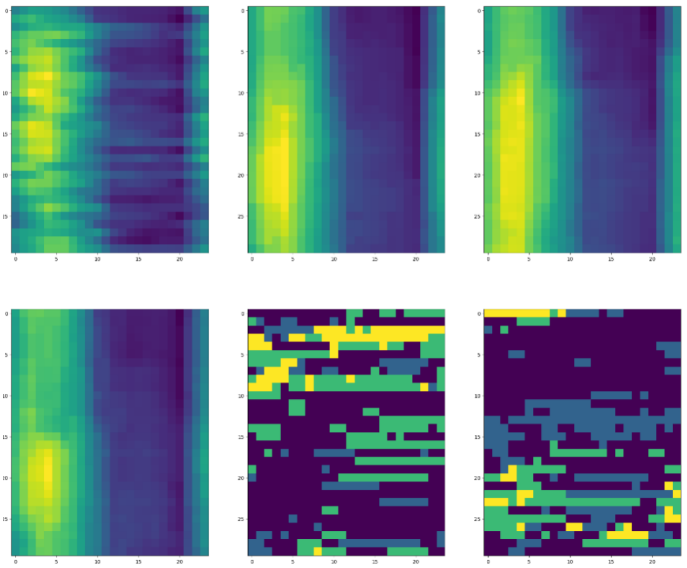

Figure 3. Mixup augmentation

2.4. Experiment setting:

In this study, we implemented the following approach: We initialized all models with pre-trained ImageNet weights and set the input size to 32x32x6. Each model underwent ten training iterations, with each iteration comprising 80 epochs. Throughout the training process, we utilized the Adam optimizer with a learning rate of 1e-4 and employed binary cross-entropy loss, which is commonly used for binary classification tasks. Following the ten training iterations, we selected the best-performing model and conducted a comparative analysis using the test data.

In the second experiment, we utilized the same settings and compared the utilization of CutMix and Mixup augmentations in combination with our method.

3. Results:

Table 1. Comparison with state-of-the art methods on Test Set

Models	ResnetRS 50	ResNet50	VGG16	DenseNet 121	Mobinet_ v3	Efficient Net-v2	ConvNext
Accuracy	0,6208	0,6186	0,5826	0,6123	0,5530	**0,6716**	0,5000

Among the models evaluated in our study, the modified EfficientNet-v2 demonstrated the highest performance with an accuracy of 0.6716. It outperformed the other models, including Resnet-RS50 (accuracy of 0.6208) and its predecessor ResNet50 (accuracy of 0.6186). DenseNet121 achieved an accuracy of 0.6123, VGG16 had an accuracy of 0.5826, and MobinetV3-Large had an accuracy of 0.5530. However, it is worth noting that the ConvNext model did not effectively converge during the experiment.

Overall, our study demonstrates the superior performance of the modified EfficientNet-v2, followed by Resnet-RS50 and ResNet50, while noting the convergence issues encountered by the ConvNext model.

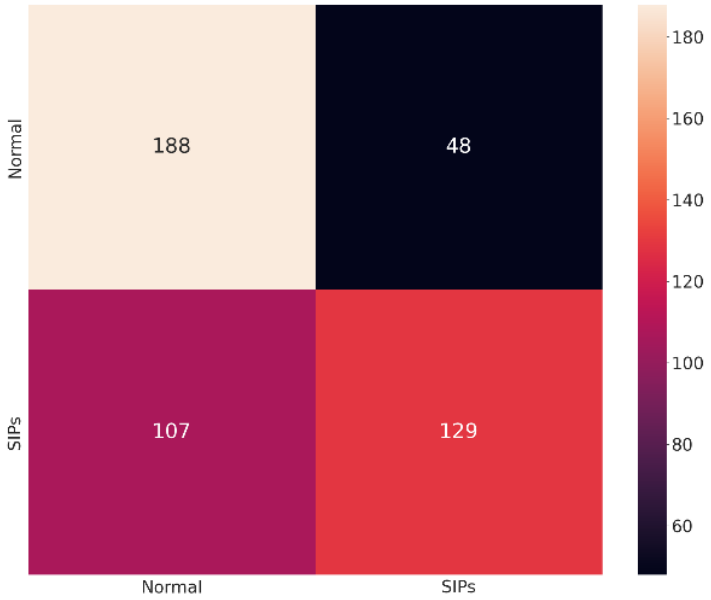

Figure 4. The confusion matrix

The Figure 4 present the confusion matrix of the efficientNet-v2.

Table 2. Comparison with the use of Cutmix and Mixup Augmentation

Num of Experiment	EfficientNet-V2 Cutmix+Mixup	EfficientNet-V2
1	**0.6504**	0,6398
2	0.6292	0,6335
3	0.6186	0,6356
4	**0.6504**	0,6102
5	0.6377	0,6398
6	0.6186	0,6229
7	0.6462	0,6483
8	0.6462	0,6144
9	0.6335	**0,6716**
10	0.6271	0,6314
Avg ± std	**0.6358 ± 0.012**	0.6348 ± 0.017

Table 2 demonstrates the significance of the CutMix and Mixup augmentation methods. While they did not directly improve the performance to achieve the new best result, they contributed to enhancing the stability of the performance. This improvement was observed not only in the average results but also in reducing the standard deviation, indicating a more consistent performance.

IV. Conclusions:

In this study, we propose an image representation approach for SIPs time sequence data, achieving a promising performance of 0.67 despite the limited data availability. To the further enhance our methodology; we intend to explore enhancement techniques, self-supervised and unsupervised methods, as well as meta-learning approaches. These strategies aim to improve the model's capability to work with limited data and enable its application in real-world domains.

Reference

[1] Liu J Y, Chen Y I, Pulinets S A, et al. (2000) Seismo-ionospheric signatures prior to M>6.0 Taiwan earthquakes. Geophysical Research Letters, 27:3113-3116

[2] Liu J Y, Chen Y I, Chuo Y J, Tsai H F. (2001) Variations of ionospheric total electron content during the Chi-Chi earthquake. Geophys. Res. Lett. , 28:1383-1386

[3] Hayakawa M, Molchanov O A. (2002) Seismo Electromagnetics: Lithosphere Atmosphere Ionosphere Coupling. Tokyo: TERRAPUB, 477

[4] Chen Y I, Liu J Y, Tsai Y B, Chen C S (2004) Statistical tests for pre-earthquake ionospheric anomaly. Terr. Atmos. Ocean. Sci., 15:385-396

[5] Liu J Y, Chen Y I, Chuo Y J, Chen C S (2006) A statistical investigation of pre-earthquake ionospheric anomaly. J. Geophys. Res., 2006, 111, A05304, doi:10.1029/2005JA011333

[6] Schaer, S. (1999) Mapping and predicting the Earth's ionosphere using theGlobal Positioniong System,Geod. Geophys. Arb. Schweiz,59, 205 pp.

[7] Tan, Mingxing and Quoc V. Le. "EfficientNetV2: Smaller Models and Faster Training." ArXiv abs/2104.00298 (2021): n. pag.

[8] Howard, Andrew G. et al. "Searching for MobileNetV3." 2019 IEEE/CVF International Conference on Computer Vision (ICCV) (2019): 1314-1324.

[9] Simonyan, Karen and Andrew Zisserman. "Very Deep Convolutional Networks for Large-Scale Image Recognition." CoRR abs/1409.1556 (2014).

[10] Huang, Gao et al. "Densely Connected Convolutional Networks." 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016): 2261-2269.

[11] He, Kaiming et al. "Deep Residual Learning for Image Recognition." 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2015): 770-778.

[12] Liu, Zhuang et al. "A ConvNet for the 2020s." 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2022): 11966-11976.

[13] Bello, Irwan et al. "Revisiting ResNets: Improved Training and Scaling Strategies." ArXiv abs/2103.07579 (2021)

[14] Yun, Sangdoo et al. "CutMix: Regularization Strategy to Train Strong Classifiers With Localizable Features." 2019 IEEE/CVF International Conference on Computer Vision (ICCV) (2019): 6022-6031.

[15] Zhang, Hongyi et al. "mixup: Beyond Empirical Risk Minimization." ArXiv abs/1710.09412 (2017)

[16] Devries, Terrance and Graham W. Taylor. "Improved Regularization of Convolutional Neural Networks with Cutout." ArXiv abs/1708.04552 (2017)