

Article

Noise Annoyance Prediction of Urban Substation Based on Transfer Learning and Convolutional Neural Network

Shengping Fan ¹, Jun Li ², Linyong Li ¹ and Zhigang Chu ^{2,*}

¹ Electric Power Research Institute of Guangdong Power Grid Co., Ltd., Guangzhou 510080, China; 13926035445@163.com (S.F.); lilyong@foxmail.com (L.L.)

² College of Mechanical and Vehicle Engineering, Chongqing University, Chongqing 400044, China; lijun0702@cqu.edu.cn

* Correspondence: zgchu@cqu.edu.cn

Abstract: The noise pollution caused by urban substations is an increasingly serious problem, as is the issue of local residents being disturbed by substation noise. To accurately assess the degree of noise annoyance caused by substations to surrounding residents, we established a noise annoyance prediction model based on transfer learning and a convolution neural network. Using the model, we took the noise spectrum as the input, the subjective evaluation result as the target output, and the AlexNet network model with a modified output layer and corresponding parameters as the pre-training model. In a fixed learning rate and epoch setting, the influence of different mini-batch size values on the prediction accuracy of the model was compared and analyzed. The results showed that when the mini-batch size was set to 4, 8, 16, and 32, all the data sets had convergence after 90 iterations. The root mean square error (RMSE) of all validation sets was lower than 0.355, and the loss of all validation sets was lower than 0.067. As the mini-batch size increased, the RMSE, loss, and mean absolute error (MAE) of the verification set gradually increased, while the number of iterations and the training duration decreased gradually. In this test, a mini-batch size value of four was appropriate. The resultant convolutional neural network model showed high accuracy and robustness, and the error between the prediction result and the subjective evaluation result was between 2% and 7%. The model comprehensively reflects the objective metrics affecting subjective perception, and accurately describes the subjective perception of urban substation noise on human ears.

Keywords: urban substation; noise; annoyance; convolutional neural network; transfer learning



Citation: Fan, S.; Li, J.; Li, L.; Chu, Z. Noise Annoyance Prediction of Urban Substation Based on Transfer Learning and Convolutional Neural Network. *Energies* **2022**, *15*, 749. <https://doi.org/10.3390/en15030749>

Academic Editor:
Djaffar Ould-Abdeslam

Received: 22 December 2021

Accepted: 18 January 2022

Published: 20 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the continued popularization of electric vehicles, more charging piles are being installed, creating a corresponding increase in the demand for power supply, which has led to the construction of more urban substations and other basic power grid facilities. However, the location of the new urban substations is increasingly problematic due to the associated noise pollution. With the acceleration of urbanization in China, residential areas have grown up around existing substations, resulting in more residents being disturbed by the noise produced by these substations. The local residents frequently complain of noise pollution, and the social problems caused by residents who are adversely affected by the noise from the substations have become more pronounced. Therefore, the accurate evaluation and prediction of the degree of substation noise disturbance to residents is important in the construction of new substations and with respect to implementing noise control measures for existing substations, which is a common problem in the power grid industry which needs to be solved urgently.

Urban substation noise is usually composed of low-frequency noise caused by transformers and reactors and high-frequency noise caused by transformer cooling systems,

although the former is dominant [1–3]. Low-frequency noise is characterized by strong penetration, slow attenuation, and long transmission distance, and it is easily perceived by local residents. In addition, the most typical low-frequency substation noise occurs at 100 Hz and its harmonic frequency components are prominent [4], which leads to a high level of subjective annoyance. When residents are exposed to substation noise for a long time, their physical and mental health is seriously affected [5–8]. Therefore, it is important to explore methods for evaluating noise annoyance and to establish a noise annoyance prediction model that conforms to the noise characteristics of urban substations.

Currently, the noise level of substation equipment is evaluated using the A-weighted sound pressure level (AWSPL) of a sound signal [9–12]. Chen et al. [10] established a noise annoyance prediction model for urban substations based on commonly used psychoacoustic metrics utilizing the multiple linear regression (MLR) method and concluded that annoyance depends mainly on the AWSPL. Liu [11] established a substation noise annoyance prediction model based on logical regression and stepwise regression analysis, concluding that in addition to AWSPL, loudness and loudness levels also have a significant impact. Another study [12] predicted the noise annoyance of urban substations by establishing a multiple linear regression model, concluding that pure tones of typical frequencies have a significant effect on annoyance [13]. The MLR model is simple and less time-consuming, but its prediction accuracy needs to be improved to better reflect the subjective perception of the human ear [14]. In this paper, a convolutional neural network model is presented as a means of evaluating the degree of urban substation noise disturbance to residents. Compared to traditional sound quality prediction methods that are based on objective acoustic metrics, the convolutional neural network automatically extracts and learns features from images or signals without tedious manual feature extraction and has achieved satisfactory results in many applications [15]. In 2012, the AlexNet convolutional neural network was proposed and won first place with only a 15.3% error rate in the ImageNet image classification competition [16]. In 2013, Wan et al. [17] established an upgraded version of Dropout technology, DropConnect, which converts the fully connected layer into a sparsely connected layer to make the network more generalized and robust. In 2015, He et al. [18] built the ResNet network, and the error rate on the ImageNet dataset set a new record of only 4.49%, which is lower than the human error rate (5.1%). In 2016, Yuan et al. [19] constructed a convolutional neural network model for rolling bearing fault diagnosis. The input data were the grayscale time-frequency map after continuous wavelet transform of the vibration signals of rolling bearings. The training results showed that the model has a strong recognition ability. In 2019, Liang et al. [20] established a sound quality prediction model based on a convolutional neural network after processing the input sound signals of internal combustion engines with the auditory spectrum and short-time average energy. The results showed that the model has a high prediction accuracy. In 2020, Huang et al. [21] used a self-constructed convolutional neural network to evaluate the sound quality of interior noise. They used two-dimensional noise time-frequency images and one-dimensional time and frequency vector data as the input. They found that the network performance was better when using two-dimensional images as the input. In 2021, the same team proposed a convolutional neural network model based on an adaptive learning rate tree (ALRT) to predict the interior sound quality of a pure electric vehicle (PEV) under non-steady state conditions. It adjusts the learning rate adaptively according to the training loss. It was proved that the ALRT convolutional neural network has a better parameter update effect than traditional methods, and its use in future technology is promising [22]. Based on transfer learning, we developed a convolutional neural network model for the degree of noise annoyance under small data sets. Compared to traditional convolutional neural networks, the transfer learning method can make the model achieve the expected effect in a shorter training time, significantly saving the training costs such as time and computing resources, and considerably improving the generalization ability of the model.

The remainder of this paper is structured as follows: Section 2 briefly introduces the basic principles of the convolutional neural network and the transfer learning methods. Section 3 develops a convolutional neural network model for the prediction of substation noise annoyance. This section also compares and analyzes the effects of different mini-batch size values on the prediction accuracy of the model when the learning rate and epoch remain unchanged. Section 4 discusses the advantages of the proposed model by comparing the prediction results with a well-established multiple linear regression model. Section 5 sets out our conclusions.

2. Materials and Methods

A convolutional neural network (CNN) is a feedforward neural network [23], and its basic structure includes mainly convolutional layers, pooling layers, and fully connected layers. These layers are stacked hierarchically, as shown in Figure 1 [21]. A convolutional neural network usually takes multi-dimensional data as input, and it automatically extracts input features and performs recognition and classification. It is essentially a non-linear mapping from input to output, which can automatically learn the mapping relationship between input and output without specific mathematical expressions between input and output.

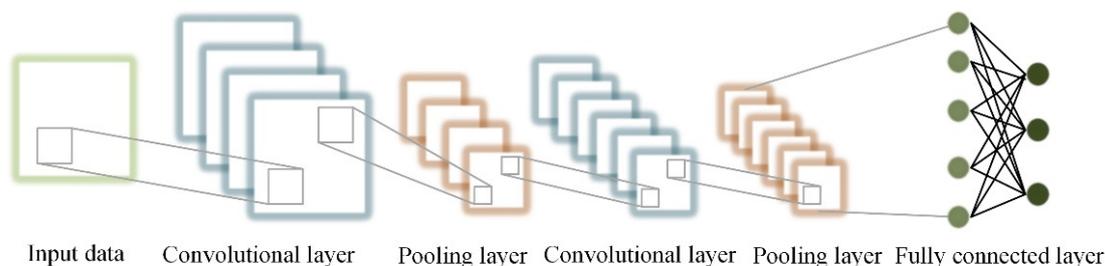


Figure 1. A typical convolutional neural network architecture.

The traditional convolutional neural network requires tens of thousands of training data, but it is time- and labor-consuming to obtain sufficient tagged data. Transfer learning uses the principle of analogy to transfer the feature information learned from the source domain to the target domain [24]. The two domains are different but somewhat related. The higher the correlation, the easier it is to obtain ideal migration results. By directly calling the trained network framework and migrating it to its own data set, the model can achieve the desired effect in a relatively short training time, fundamentally solving the problem of insufficient training data. At the same time, the software and hardware environment required for transfer learning is more relaxed. The pre-training model is a deep learning framework trained by predecessors using massive data, and pre-training is a way of transfer learning. Instead of training the network from scratch, a scene-similar pre-trained model is chosen to solve the problem, which greatly reduces the training time. In this study, the AlexNet classical network was selected as the pre-training model based on transfer learning. Since the data set used to train the model was small and the data had relatively simple shallow features such as prominent spikes, the network architecture was completely preserved, save for the output layer. The last SoftMax layer and the classified output layer were modified as the fully connected layer with response number 1 and the regression layer, thereby converting the classification network into a regression network. The AlexNet network architecture is shown in Figure 2 [22].

Convolutional Layer filter size 11×11 , filter number 96, stride 4/ReLU
Local Response Normalization Layer
Max Pooling Layer filter size 3×3 , stride 2
Convolutional Layer filter size 5×5 , filter number 256, stride 1/ReLU
Local Response Normalization Layer
Max Pooling Layer filter size 3×3 , stride 2
Convolutional Layer filter size 3×3 , filter number 384, stride 1/ReLU
Convolutional Layer filter size 3×3 , filter number 384, stride 1/ReLU
Convolutional Layer filter size 3×3 , filter number 256, stride 1/ReLU
Max Pooling Layer filter size 3×3 , stride 2
Fully Connected Layer filter number 4096/ReLU
Fully Connected Layer filter number 4096/ReLU
Fully Connected Layer filter number 1000

Figure 2. The AlexNet network architecture.

3. Results

The performance of the method used in this study was verified by the urban substation noise recorded and the corresponding subjective evaluation results of a previous study [12]. These urban substation noise samples were recorded using a PULSE data acquisition system (Bruel & Kjaer type 3050-A-060, Nærum, Denmark) and microphones (Bruel & Kjaer type 4189-A-021, Nærum, Denmark). Noise measurement points were arranged mainly near substation plant boundaries and residential areas to reflect the impact on residents and pedestrians. In order to make the data more diverse and to reflect the impact of noise on the staff in the substation, a number of measurement points were also arranged within the substation plant. A total of 17 measurement points were arranged, and three time periods were selected: 2:00 to 3:00, 9:00 to 10:00, and 20:00 to 22:00. Across the measurement points and time periods, a total of 51 noise samples with a duration of 5 s were collected. From these samples, 43 samples were randomly selected to participate in the training process of the network model and the remaining 8 samples were used to verify the reliability of the network model. Given that the number of samples was too small, each noise sample was divided into multiple segments of 1 s, so that each recorded noise signal was divided into five non-overlapping noise samples; a total of 215 noise sample segments were obtained. Each sample had the same subjective evaluation score (due to the substation noise being steady-state). Three segments of each noise sample were classified as the training set, and the remaining two segments were classified as the test set, so the ratio of the training set to the test set in the dataset was 3:2. Each noise sample was converted into the corresponding frequency spectrum to obtain its frequency characteristics using a MATLAB 2020b (MathWorks, USA). All the calculated frequency spectra were saved as color pictures of the same size. The pictures were used as the input of the convolutional neural network (the pixels of each sample were set to 227×227 , which is consistent with the input layer size of the AlexNet network). Part of the input spectra is shown in Figure 3. As can be seen in Figure 3, the noise energy of these samples was concentrated at the integer multiple frequency of 100 Hz within 1000 Hz, and pure tones were very prominent. The subjective evaluation score of each sound sample was used as the output of the convolutional neural network.

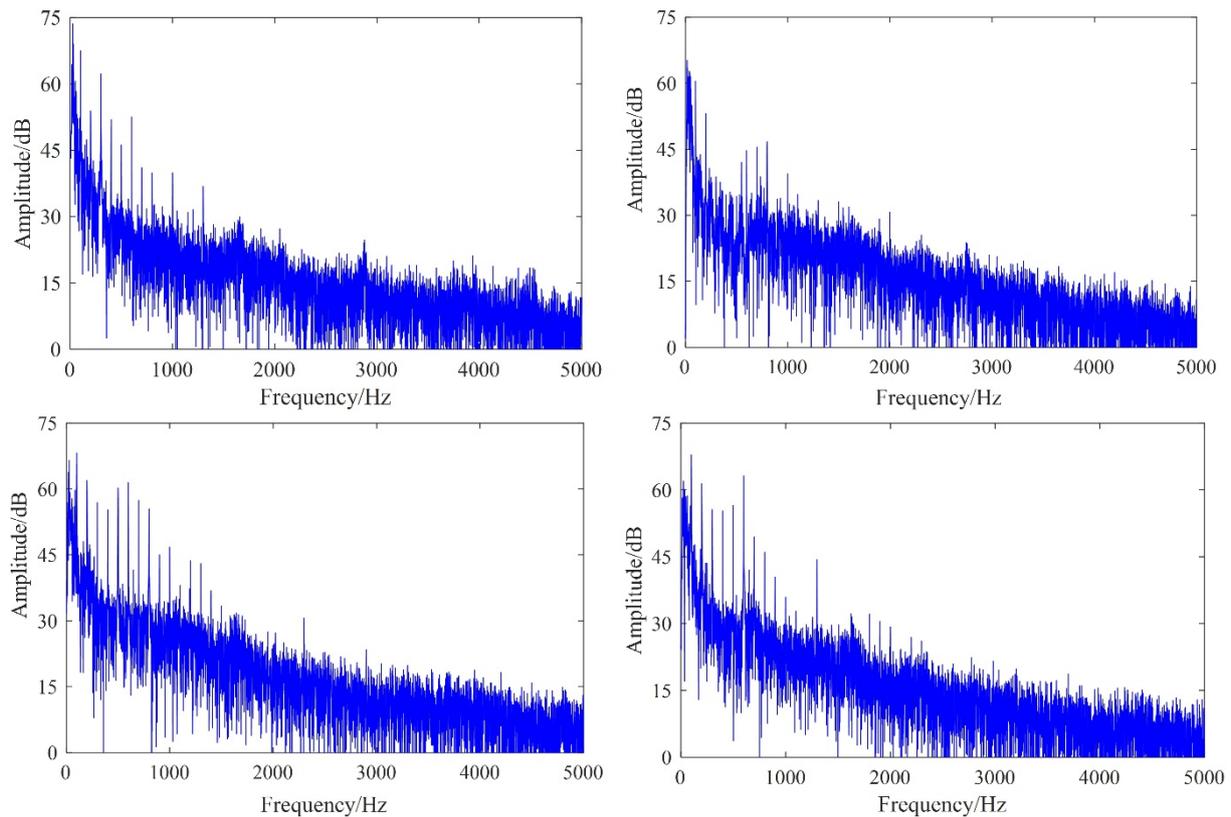


Figure 3. Part of the input spectra.

In this study, a single GPU of Nvidia GeForce 940MX was used to run the MATLAB 2020b environment, and the built-in stochastic gradient descent method (SGDM) with a first-order momentum solver was utilized. For the learning rate, it is usually necessary to set a learning rate lower than that of the general training model when performing fine-tuning training. In this model, it was more appropriate to set the initial learning rate to 0.0001. The mini-batch size is usually set to a power of two, which can make the GPU perform better [25]. In this study, the mini-batch size was set to 4, 8, 16, and 32, separately. The total number of iterations was set to 90 epochs (the process of training the complete data set once on the neural network was one epoch).

In this study, the root mean square error (RMSE) and loss were selected to evaluate the quality of the prediction results, and loss is the square of RMSE. The smaller the RMSE or the loss function, the better the quality of the model. The calculation formula of RMSE was

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{k=1}^n (t_k - y_k)^2} \quad (1)$$

where n denotes the number of training samples, and t_k and y_k denote the target output and the forecast output of the training samples, respectively.

Based on the aforementioned settings of the network model parameters, the spectrum diagrams of urban substation noise were input into the model for training. Ten trials were conducted for each experiment to reduce the impact of randomness. During the training period, the changing curves of the average RMSE and loss of the training set and the test set in each epoch, with different mini-batch size values, are shown in Figure 4 (taken as the average of ten trials). As shown in Figure 4, as the number of iterations increased, the RMSE and loss of the training set and the test set generally decreased. After 90 epochs, all data sets had reached convergence, and the RMSE of all test sets was not higher than 0.355, and the loss was not higher than 0.067, which indicates that the convolutional neural network

based on AlexNet transfer learning can achieve high accuracy when used to predict the noise annoyance of urban substations.

The errors between the subjective score values and the predicted values of the eight verification samples not involved in the modeling, when the mini-batch size was 4, 8, 16, and 32, are shown in Figure 5. As shown in Figure 5, when the mini-batch size values were different, although the convolutional neural network model had various errors for verification samples, the trend of prediction errors was the same. At the same time, when the mini-batch size was four, the prediction error was the smallest, and the accuracy of the convolutional neural network model was the highest, so the mini-batch size selected was four.

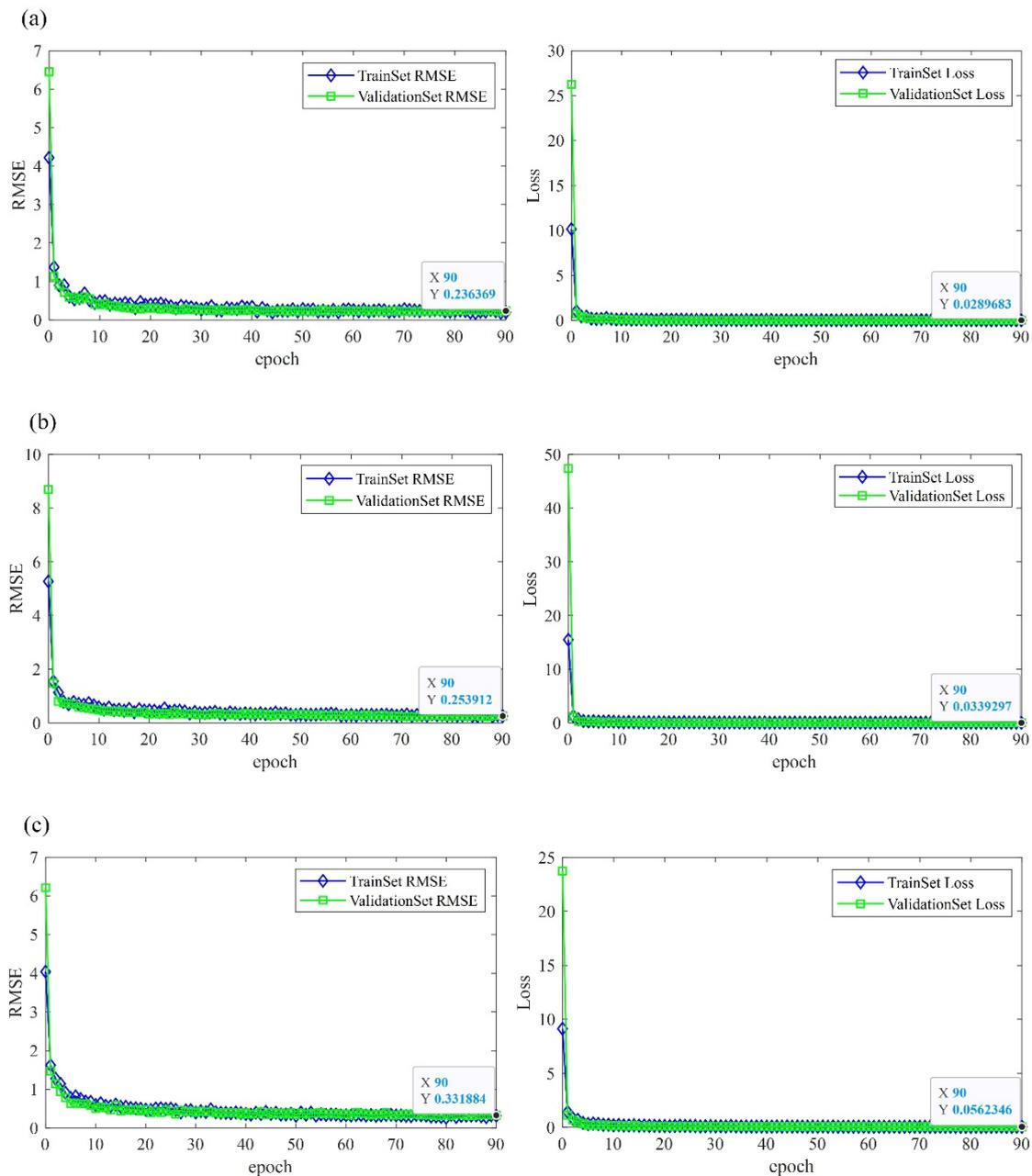


Figure 4. Cont.

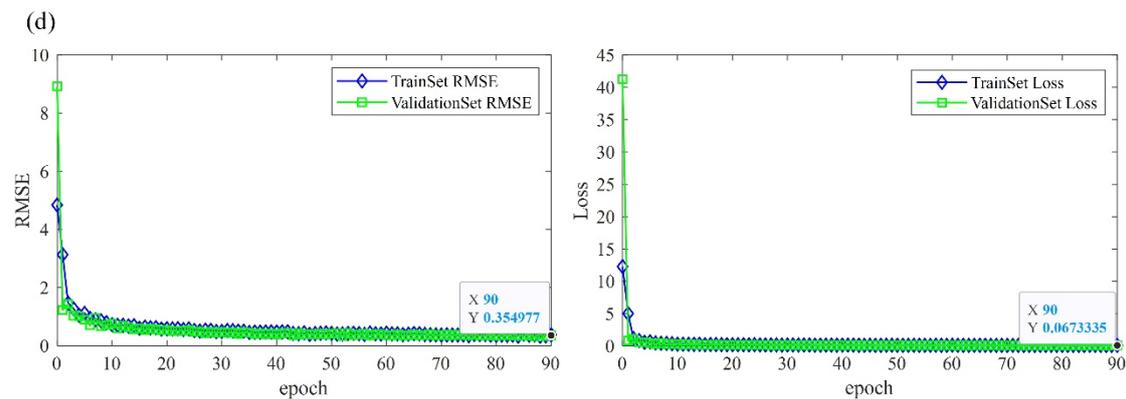


Figure 4. Transfer learning training process based on AlexNet: (a) mini-batch size of 4; (b) mini-batch size of 8; (c) mini-batch size of 16; (d) mini-batch size of 32.

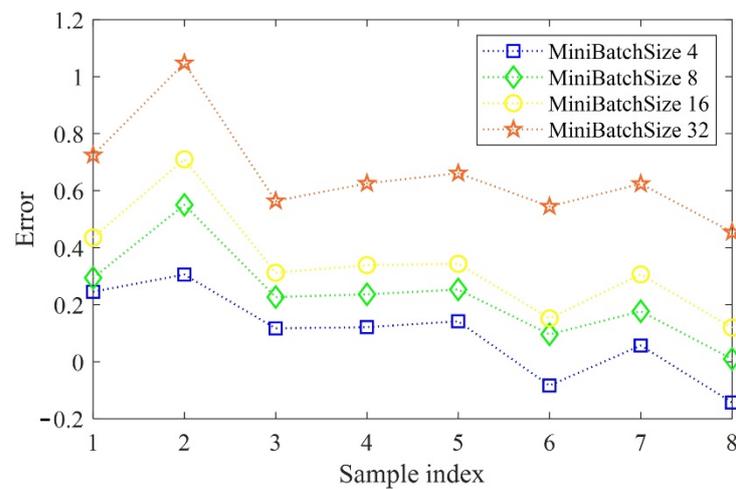


Figure 5. The relative error of different mini-batch size values.

4. Discussion

The multiple linear regression model established in a prior study [12] was used to predict the annoyance level of the eight verification noise samples, and the results were compared with the subjective evaluation results. The multiple linear regression model is based on the least squares method, which takes the objective parameters that are highly correlated with the subjective evaluation results as independent variables, and the subjective annoyance level as dependent variables. Using multiple linear stepwise regression, the independent variables are introduced into the regression model one by one according to the significance of their influence on the dependent variables, and the original independent variables that become irrelevant to the model due to the introduction of new variables are deleted. Therefore, the final regression model only includes variables that have a significant impact on the dependent variable. The multiple linear regression model is shown in Formula (2).

$$P = -3.354 + 0.035L_{600} + 0.089L_A \quad (2)$$

where P , L_{600} , and L_A denote subjective annoyance, the 600 Hz AWSPL, and the total AWSPL, respectively.

The comparison of the prediction errors of the two models is shown in Figure 6. As shown in Figure 6, the prediction errors of the multiple linear regression model for the noise annoyance of urban substations are between 1% and 9%, but the prediction errors of

the convolutional neural network model are between 2% and 7%. Using the latter model, a relatively ideal and robust result was obtained.

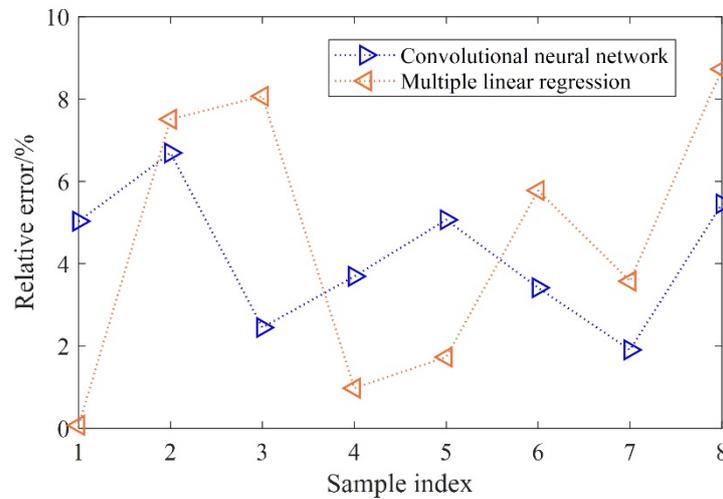


Figure 6. Comparison of prediction errors of the two models.

The mean absolute error and the sum of squared errors of the prediction errors of the two models are shown in Table 1. The calculation formulas for the mean absolute error (MAE) and the sum of squared errors (SSE) are, respectively:

$$\text{MAE} = \frac{1}{n} \sum_{k=1}^n |(t_k - y_k)| \quad (3)$$

$$\text{SSE} = \sum_{k=1}^n (t_k - y_k)^2 \quad (4)$$

where n denotes the number of verification noise samples and t_k and y_k denote the target output and the forecast output of the verification noise samples, respectively.

Table 1. The statistical values of the prediction error of the two models.

Statistic	Convolutional Neural Network	Multiple Linear Regression
MAE	0.132	0.299
SSE	0.978	4.960

As shown in Table 1, the MAE and the SSE of the convolutional neural network model are significantly lower than those of the multiple linear regression model, indicating that the convolutional neural network model has higher prediction accuracy.

5. Conclusions

This paper presents a convolutional neural network model for noise annoyance prediction, based on AlexNet transfer learning, to predict the noise annoyance level of urban substations. The frequency spectra of the noise samples were taken as the input for the model and the sample features were automatically extracted during the training process. The subjective evaluation results were the output of the model. The epoch was fixed, and the mini-batch size was set to 4, 8, 16, and 32, separately. Finally, all the data sets were converged. The root mean square errors of all test sets were not higher than 0.355, and the losses were not higher than 0.067. Compared to the prediction results of the well-established multiple linear regression model, the convolutional neural network model has higher prediction accuracy and robustness, with the prediction error falling between 2% and 7%. Thus, the network comprehensively reflects the objective metrics affecting

subjective perception, and more accurately describes the subjective perception of urban substation noise on human ears. Nevertheless, in future studies, it would be worthwhile to explore different methods that exhibit high prediction accuracy in other application fields, such as multilayer perceptron neural networks [26,27], to predict the noise annoyance level of urban substations.

Author Contributions: Conceptualization, S.F., J.L. and Z.C.; methodology, S.F., J.L. and Z.C.; software, Z.C.; validation, L.L. and Z.C.; formal analysis, L.L.; investigation, J.L. and L.L.; resources, S.F.; data curation, J.L.; writing—original draft preparation, S.F. and J.L.; writing—review and editing, J.L., L.L. and Z.C.; visualization, J.L.; supervision, Z.C.; project administration, Z.C.; funding acquisition, S.F. All authors have read and agreed to the published version of the manuscript.

Funding: The authors disclose receipt of the following financial support for the research, authorship, and/or publication of this article: this work was supported by the Science and Technology Project of China Southern Power Grid (No. GDKJXM20200377).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no potential conflict of interest with respect to the research, authorship, and/or publication of this article.

References

- Ming, R.S.; Pan, J.; Norton, M.P.; Wende, S.; Huang, H. The sound-field characterization of a power transformer. *Appl. Acoust.* **1999**, *56*, 257–272. [[CrossRef](#)]
- Zhou, B.; Pei, C.M.; Ni, Y.; Zhang, J. Measurement and analysis of noise in UHV AC ausatation. *High Volt. Eng.* **2013**, *39*, 1447–1453.
- Piana, E.A.; Roozen, N.B. On the control of low-frequency audible noise from electrical substation: A case study. *Appl. Sci. Basel* **2020**, *10*, 637. [[CrossRef](#)]
- Di, G.Q.; Zhou, X.W.; Chen, X.W. Annoyance response to low frequency noise with tonal components: A case study on transformer noise. *Appl. Acoust.* **2015**, *91*, 40–46. [[CrossRef](#)]
- Sjödén, F.; Kjellberg, A.; Knutsson, A.; Landström, U.; Lindberg, L. Noise exposure and auditory effects on preschool personnel. *Noise Health* **2012**, *14*, 72–82. [[CrossRef](#)] [[PubMed](#)]
- Leventhall, G. Low Frequency Noise. What we know, what we do not know, and what we would like to know. *J. Low Freq. Noise Vib. Act. Control* **2009**, *28*, 79–104. [[CrossRef](#)]
- Abbasi, M.; Yazdanirad, S.; Mehri, A.; Fallah Madvari, R.; Alizadeh, A.; Ghaljahi, M.; Falahati, M. Noise exposure and job stress—A structural equation model in textile industries. *Arch. Acoust.* **2015**, *45*, 601–611.
- Alayrac, M.; Marquis-Favre, C.; Viollon, S.; Morel, J.-C.; Le Nost, G. Annoyance from industrial noise: Indicators for a wide variety of industrial sources. *J. Acoust. Soc. Am.* **2010**, *128*, 1128–1139. [[CrossRef](#)]
- Liu, H.; Yin, J.G.; Zhang, G.Y.; Zhang, Y.; Xie, L.K.; Ji, Y.F.; Liu, G.S.; An, S.H. Analysis of noise pollution of 110 kV outdoor substation and discussion of noise control. *Guangdong Electr. Power* **2018**, *31*, 12–19.
- Chen, Y.Y.; Wang, X.P.; Yang, W. Experimental research on evaluation of sound quality annoyance of urban substation noise. *Sci. Technol. Eng.* **2018**, *18*, 214–218.
- Liu, G. *Study on Evaluation Method of Sound Quality Improvement Effect of Noise Control Measures on Substation Noise*; Zhejiang University: Hangzhou, China, 2018.
- Li, J.; Li, L.; Zhang, Y.; Fan, X. Annoyance evaluation of noise emitted by urban substation. *J. Low Freq. Noise Vib. Act. Control* **2021**, *40*, 2106–2114.
- Di, G.-Q.; Chen, X.-W.; Song, K.; Zhou, B.; Pei, C.-M. Improvement of Zwicker’s psychoacoustic annoyance model aiming at tonal noises. *Appl. Acoust.* **2016**, *105*, 164–170. [[CrossRef](#)]
- Su, L.L.; Wang, D.F.; Wang, Q. Sound quality prediction of vehicle interior noise based on GRNN. *J. Jilin Univ.* **2011**, *41*, 82–86.
- Khan, A.; Sohail, A.; Zahoora, U.; Qureshi, A.S. A survey of the recent architectures of deep convolutional neural networks. *Artif. Intell. Rev.* **2020**, *53*, 5455–5516. [[CrossRef](#)]
- Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Commun. ACM* **2017**, *60*, 84–90. [[CrossRef](#)]
- Wan, L.; Zeiler, M.; Zhang, S.; Cun, Y.L.; Fergus, R. Regularization of neural networks using dropconnect. In Proceedings of the 30th International Conference on Machine Learning, Atlanta, GA, USA, 16–21 June 2013; pp. 1058–1066.
- He, K.M.; Zhang, X.Y.; Ren, S.Q.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

19. Yuan, J.H.; Han, T.; Tang, J.; An, L. An approach to intelligent fault diagnosis of rolling bearing using wavelet time-frequency representations and CNN. *Mech. Design Res.* **2017**, *33*, 93–97.
20. Liang, K.; Zhao, H.J.; Song, W.Z. Research on evaluation method of internal combustion engine sound quality based on convolutional neural network. *Chin. Intern. Combust. Engine Eng.* **2019**, *40*, 67–75.
21. Huang, X.; Huang, H.; Wu, J.; Yang, M.; Ding, W. Sound quality prediction and improving of vehicle interior noise based on deep convolutional neural networks. *Expert Syst. Appl.* **2020**, *160*, 113657. [[CrossRef](#)]
22. Huang, H.; Wu, J.; Lim, T.C.; Yang, M.; Ding, W. Pure electric vehicle nonstationary interior sound quality prediction based on deep CNNs with an adaptable learning rate tree. *Mech. Syst. Signal Process.* **2021**, *148*. [[CrossRef](#)]
23. Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [[CrossRef](#)]
24. Pan, S.J.; Yang, Q.A. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [[CrossRef](#)]
25. Zhao, L.X.; Hou, F.D.; Lu, Z.C.; Zhu, H.C.; Ding, X.L. Image recognition of cotton leaf diseases and pests based on transfer learning. *Trans. Chin. Soc. Agric. Eng.* **2020**, *36*, 184–191.
26. Roshani, G.H.; Nazemi, E.; Roshani, M.M. Flow regime independent volume fraction estimation in three-phase flows using dual-energy broad beam technique and artificial neural network. *Neural Comput. Appl.* **2016**, *28*, 1265–1274. [[CrossRef](#)]
27. Sattari, M.A.; Roshani, G.H.; Hanus, R.; Nazemi, E. Applicability of time-domain feature extraction methods and artificial intelligence in two-phase flow meters based on gamma-ray absorption technique. *Measurement* **2021**, *168*, 108474. [[CrossRef](#)]