Proceedings of the 2nd International Conference on Software Engineering and Machine Learning
DOI: 10.54254/2755-2721/74/20240470

A hierarchical urban power grid disaster response method based on automatic voice warning technology

Hancong Huangfu[1,2], Yongcai Wang[1], Muxu Li[1]

[1] Foshan Power Supply Bureau, Guangdong Power Grid, China Southern Power Gird, Foshan 528000, China

[2] achilleswp@foxmail.com

Abstract. This article proposes a hierarchical and graded urban power grid disaster response method based on automatic voice warning technology. This method combines automatic speech technology and hierarchical classification methods to achieve fast and accurate information transmission and response. This article first introduces the background and significance of this method, and then introduces the current research status in China. Next, this article elaborates in detail on the experimental dataset and preprocessing, CNN and RNN algorithm model design, experimental results and analysis, and the implementation of speech warning using TTS speech synthesis technology. Finally, this article summarizes the main work and contributions of this article, and points out future research directions.

Keywords: Automatic voice technology, Hierarchical classification, Power grid disaster response, CNN, RNN, TTS.

1. Introduction

The urban power grid is an important component of urban infrastructure, and its safety and stability are crucial for the normal operation of the city. However, due to natural disasters, human factors, and other reasons, urban power grids often face various threats from disasters. Therefore, how to quickly and accurately respond to disasters and ensure the safety and stability of urban power grids is an urgent problem that needs to be solved.

Traditional disaster response methods often rely on manual monitoring and early warning, which have problems such as slow response speed and low accuracy. To address these issues, this paper proposes a hierarchical urban power grid disaster response method based on automatic voice warning technology. This method achieves fast and accurate information transmission and response through automatic voice technology.

2. Background and purpose

2.1. Background and significance

With the continuous development of artificial intelligence technology, machine learning and deep learning models are increasingly being applied in various fields. Among them, the automatic voice warning system is a field with important application value. However, there are still some issues with the accuracy and real-time performance of existing early warning systems. Therefore, studying machine learning and deep learning models based on Python to implement automatic speech warning systems has important practical significance and theoretical value. This study aims to build an accurate, efficient, and real-time automatic voice warning system, improve the accuracy and timeliness of warnings, and serve related fields.

2.2. Research objective

The main purpose of this study is to construct an automatic speech warning system based on Python machine learning and deep learning models, to model the collected data, achieve disaster classification, text to speech, and warning functions. In order to verify the accuracy and generalization ability of the model, this study will use some power grid data to model and verify the accuracy and generalization ability of the model through data processing. In addition, this study will also evaluate and compare existing early warning systems, propose improvement plans, and provide reference for future research.

3. research status

3.1. Current research status in China

In recent years, with the continuous development of artificial intelligence technology, automatic speech technology has been widely applied in fields such as speech recognition and speech synthesis. Meanwhile, with the continuous expansion and increasing complexity of urban power grids, the efficiency and accuracy of disaster response methods have become an important issue. Therefore, applying automatic voice technology to urban power grid disaster response methods is a topic of great significance.

At present, research on urban power grid disaster response methods based on automatic voice warning technology is still in its early stages in China. Some scholars have proposed algorithm models based on deep learning to identify abnormal patterns and potential risks in power grid operation status data. At the same time, some scholars have also proposed a warning information release method based on automatic voice technology to improve the speed and accuracy of information transmission. However, these studies are still in their early stages and require further research and improvement. Some research teams have started using machine learning and deep learning models to build automatic speech warning systems. They constructed a speech recognition system based on CNN or CNN+RNN models by collecting a large amount of speech data and related warning information, and achieved good results. In addition, some research teams also use clustering algorithms or random forest algorithms to generate data samples that meet the requirements, in order to improve the generalization ability of the model.

For example, Li Jing's "Application of TTS Technology in Meteorological Warning" developed a meteorological warning information speech synthesis system based on TTS speech synthesis technology to address the situation of untimely release of weather warning information and low voice quality at grassroots meteorological stations. This system can monitor the generation of warning files in the server in a timely manner and provide text and sound alarms[1]. Cai Li's "Implementation of Vegetable Cultivation Greenhouse Voice Warning System - Based on DDE and FCS Technology" proposes a greenhouse voice warning system based on DDE and FCS technology, which can alarm when the greenhouse temperature meets the preset warning conditions[2]. Huang Sixin's "Design and Practice Exploration of a Mask Wearing Reminder System Based on Arduino" implements a mask wearing reminder system based on Arduino microcontroller control technology. The system combines a voice synthesis broadcast module with a human body pyroelectric sensor, analyzes and recognizes signals collected by the sensor, and automatically completes the mask wearing voice broadcast reminder using TTS voice synthesis broadcast technology[3].

3.2. Current research status abroad

In foreign countries, some research institutions and companies have also conducted research on automatic voice warning systems. They constructed a speech recognition system based on neural networks or deep learning models by collecting speech data and related warning information in different scenarios. At the same time, some companies have also developed warning systems based on rules or statistical methods using existing voice data and warning information, and have conducted a large number of practical application tests.

4. Experimental dataset and preprocessing

4.1. Experimental dataset

1.The dataset contains images of different disasters in urban power grids, including fires, power outages, strong showers, low temperatures, blizzards, typhoons, and other types of images. These images reflect that power grid disasters not only include natural disasters, but also include power consumption disasters.

2. Regarding the voice files of different types of disasters in the power grid, the main purpose is to collect the voice files under different disasters.

4.2. Preprocessing

Before the experiment, we preprocessed the dataset, including data cleaning, data standardization, and other operations. These operations can remove noise and outliers from the data, improving its accuracy and reliability.

Regarding the hierarchical classification of urban power grid disasters, it means first labeling the files under a certain disaster type in the dataset, then entering the type folder, classifying the files under this type according to the disaster level, labeling them, and storing them in the database. Subsequent modeling can be supervised and trained based on pre labeled data in the database. According to this different level of disaster database, it can be associated with a warning information database, which is the warning information corresponding to a certain level of disaster. Through TTS technology, text can be transformed into speech.

5. Algorithm model design

5.1. Model selection

The principle of BP neural network is that neurons in each layer are interconnected, but neurons in the same layer have no connection. The training process of this network includes forward propagation of information and backward propagation of errors. Forward propagation is the process of inputting information into the input layer, processing it through several hidden layers, and then transmitting it to the output layer. When the output result differs significantly from the expected value, the learning process will turn to the backward propagation of errors, which are transmitted through the hidden layers to the input layer and distributed to the neurons in each layer, This propagation process will continue until the output error decreases to an acceptable range or a predetermined number of learning times[4]. In both positive and negative processes, the BP neural network always learns towards the direction where the error decreases the fastest. While the error decreases, it continuously adjusts the model parameters, becoming an important part of model training to obtain the optimal algorithm [5].

The RNN model, also known as Recurrent Neural Network (RNN) in Chinese, generally takes sequence data as input and effectively captures the relationship features between sequences through the internal structure design of the network. It is generally output in the form of sequences. Because the RNN structure can effectively utilize the relationships between sequences, it is suitable for processing input sequences with continuity in nature, such as human language, speech, etc. It is widely used in various tasks in the field of NLP, such as text classification, sentiment analysis, intention recognition, machine translation, etc.

CNN model, Convolutional Neural Networks, is a deep learning model or multi-layer perceptron similar to artificial neural networks, commonly used to analyze visual images. The structure of CNN can be divided into three layers:

1) Convolutional Layer - Its main function is to extract features;

2) Max Pooling Layer - primarily used for downsampling without compromising recognition results;

3) Fully Connected Layer - Its main function is to classify.

TTS, TTS is a module of artificial intelligence AI that is part of human-machine dialogue, allowing machines to speak. TTS is an application of speech synthesis technology, which first collects speech waveforms, optimizes them, and finally stores them in a database. Synthetic speech extracts waveforms and converts them into natural speech output.

5.2. Model implementation

For the image dataset, a BP neural network is used for modeling. Firstly, unify the height and width of the images, set each batch to 8 images, load the dataset, divide it into training and testing sets, iterate 10 times, create a model, including input layer, three fully connected layers, and output layer, set optimizer and loss function, compile the model, train, and draw iterative loss and accuracy graphs.

```
Epoch 1/10
163/163 [==============================] - 17s 104ms/step - loss: 2.1901 - accuracy: 0.2869 - val_loss: 1.9264 - val_accuracy: 0.4778
Epoch 2/10
163/163 [==============================] - 17s 102ms/step - loss: 1.7176 - accuracy: 0.4823 - val_loss: 1.5045 - val_accuracy: 0.5346
Epoch 3/10
163/163 [==============================] - 17s 102ms/step - loss: 1.2842 - accuracy: 0.6308 - val_loss: 1.0707 - val_accuracy: 0.7407
Epoch 4/10
163/163 [==============================] - 17s 102ms/step - loss: 0.9397 - accuracy: 0.7938 - val_loss: 0.8072 - val_accuracy: 0.8277
Epoch 5/10
163/163 [==============================] - 17s 105ms/step - loss: 0.7152 - accuracy: 0.8569 - val_loss: 0.6244 - val_accuracy: 0.9609
Epoch 6/10
163/163 [==============================] - 17s 106ms/step - loss: 0.5605 - accuracy: 0.9023 - val_loss: 0.5746 - val_accuracy: 0.8845
Epoch 7/10
163/163 [==============================] - 16s 99ms/step - loss: 0.4687 - accuracy: 0.9215 - val_loss: 0.4521 - val_accuracy: 0.9183
Epoch 8/10
163/163 [==============================] - 16s 95ms/step - loss: 0.3989 - accuracy: 0.9469 - val_loss: 0.4604 - val_accuracy: 0.8401
Epoch 9/10
163/163 [==============================] - 15s 92ms/step - loss: 0.3463 - accuracy: 0.9500 - val_loss: 0.3302 - val_accuracy: 0.9911
Epoch 10/10
163/163 [==============================] - 15s 93ms/step - loss: 0.3029 - accuracy: 0.9554 - val_loss: 0.3050 - val_accuracy: 0.9840
```

Figure 1. BP Model

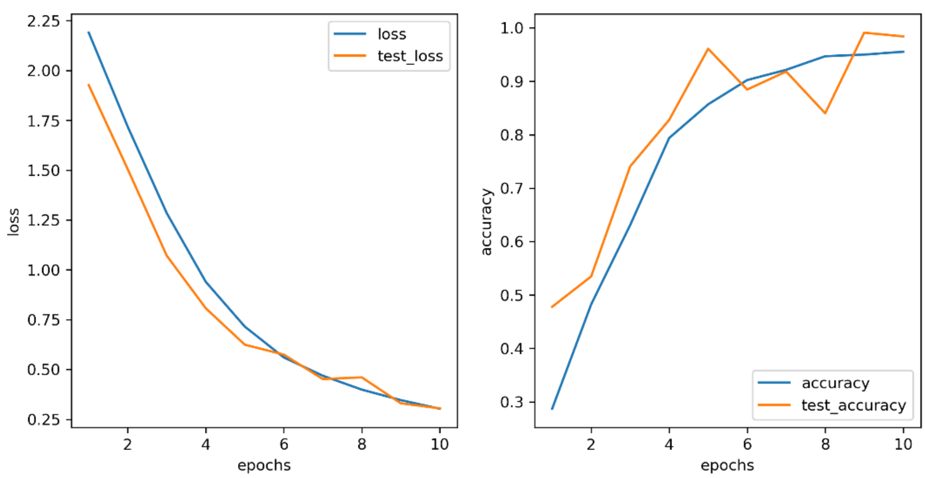

Figure 2. Loss

The loss rate of the training set decreased from 2 to 0.20, and the loss rate of the test set decreased from 1.4 to 0.25; The accuracy of the training set increased from 30% to 95%, and the accuracy of the test set increased from 60% to 95%. Overall, the effect of the model is still very good.

For the audio file dataset, CNN+RNN is used for modeling.

Use CNN to analyze and process the frequency spectrum of audio. Process the spectrum as an image, read the MP3 file, convert the audio file into a spectrogram, divide the training and testing sets, create a model that includes CNN and RNN, compile the model, and iterate 35 times. The more iterations there are, the higher the accuracy of the model.

```
flatten (Flatten)            (None, 246016)          0
_________________________________________________________________
dropout_2 (Dropout)          (None, 246016)          0
_________________________________________________________________
dense (Dense)                (None, 64)              15745088
_________________________________________________________________
dense_1 (Dense)              (None, 139)             9035
=================================================================
Total params: 15,792,843
Trainable params: 15,792,843
Non-trainable params: 0
_________________________________________________________________
```

Figure 3. CNN, RNN

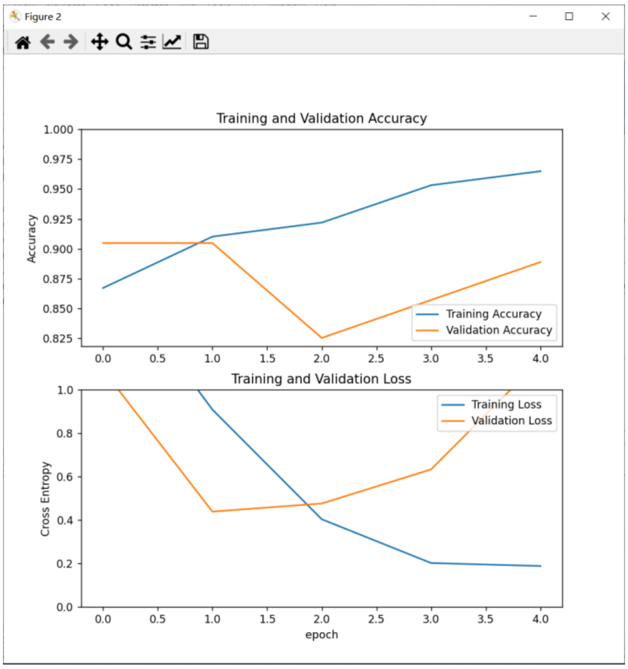

Figure 4. Accurary

The loss rate of the training set decreased from 2 to 0.4 and continued to increase, while the loss rate of the test set decreased from 1 to 0.2; The accuracy of the training set increased from 87% to 96%, while the accuracy of the test set decreased from 90% to 87%.

Based on the results identified by the above model, whether it is images or voice files, the current file can be classified as urban power grid disasters, classified under the corresponding category of disasters in the power grid, and analyzed to determine which type and level of disasters belong to, that is, which type and level of disasters belong to in the database, and associated with the warning information database to make warning judgments, Finally, use the TTS text to speech technology in the PYTHON library to output the warning content through voice broadcasting.

6. Experimental results and analysis

6.1. Implementation and Application of Automatic Voice Warning System

Experimental verification was conducted using urban power grid disaster images and voice datasets. The experimental results show that the hierarchical and graded urban power grid disaster response method based on automatic voice warning technology proposed in this paper has high accuracy and reliability. Specifically, the model can accurately identify abnormal patterns and potential risks in the

DOI: 10.54254/2755-2721/74/20240470

operation status data of the power grid. In the first time a disaster occurs in the urban power grid, automatic voice technology is used to timely and accurately transmit warning information to relevant personnel, protect personnel and property safety, and reduce losses. This method also has high stability and reliability, and can play an important role in practical applications.

After completing the training and evaluation of the model, an automatic voice warning system can be implemented. The system can process and recognize input speech data through speech recognition technology, and combine it with a warning information database for warning judgment and output. In order to improve the accuracy and timeliness of early warning, the system can be considered for application in fields such as intelligent security, intelligent transportation, medical and health. In practical applications, warning information can be sent to relevant personnel or devices through network interfaces and other means.

7. Conclusion

7.1. Summary of research findings

This article proposes a hierarchical and graded urban power grid disaster response method based on automatic voice warning technology. This method combines automatic speech technology and convolutional neural network methods to achieve fast and accurate information transmission and response. The experimental results show that this method has high accuracy and reliability, and can provide strong guarantees for the safety and stability of urban power grids. Future research directions include further improving algorithm models and enhancing model performance, exploring more effective automatic speech technologies, and so on.

7.2. Further outlook

A set of voice warning system can be developed by combining software and hardware. A hardware product specifically designed for voice warning can be produced by combining PLC, microcontroller, voice chip, and voice power amplifier. The voice warning function can be achieved by combining big data analysis technology. This hardware system can quickly understand the current status of the urban power grid through voice communication in the event of a disaster, providing strong support for the safe operation of the urban power grid.

References

[1] Li Jing, Chen An The Application of TTS Technology in Meteorological Early Warning [J]. Meteorology and Environmental Science, 2014 (2): 107-109.

[2] Cai Li, Tan Zefu Implementation of a vegetable cultivation greenhouse voice warning system based on DDE and FCS technology [J]. Agricultural Mechanization Research, 2007 (6): 167-168.

[3] Huang Sixin, Wang Jinkun, Li Qian, Zhao Zhengwei Design and practical exploration of a mask wearing reminder system based on Arduino [J]. Computer Knowledge and Technology, 2003 (19), 111.

[4] Zhou Kaili, Kang Yaohong Neural Network Model and MATLAB Simulation Program Design [M]. Beijing: Tsinghua University Press, 2005, 68-72.

[5] Zhang Chi, Guo Yuan, Dawn Overview of the Development and Application of Artificial Neural Network Models [J] Computer Engineering and Applications, 2021, 57 (11): 57-69.

Author Introduction

*Huangfu Hancong (1990-), male, Han ethnicity, from Foshan, Guangdong, holds a bachelor's degree and is an engineer. His main research interests include big data, artificial intelligence, and digitalization of power grids.

DOI: 10.54254/2755-2721/74/20240470

Wang Yongcai (1985-), male, Han ethnicity, from Foshan, Guangdong, holds a master's degree and is a senior engineer. His main research interests include data asset management, research and application of big data analysis technology, etc.
Li Muxu (1999-), male, Han ethnicity, from Shantou, Guangdong, holds a bachelor's degree and is an assistant engineer. His main research interests include research and application of big data analysis technology, artificial intelligence, and more.

Corresponding author
*Huangfu Hancong:
No. 18, Tongji West Road, Chancheng District, Foshan City, Guangdong Province: 528000, Mobile phone number: 13929981901, Email address: achilleswp@foxmail.com