

RESEARCH

Open Access



The construction of Bert fusion model of speech recognition and sensing for South China electricity charge service scenario

Guangcai Wu^{1,2} and Yinglong Zheng^{1,2*}

*Correspondence:
yinglong_64@163.com;
zhengyinglong@stu.cpu.edu.cn

¹ Division of R & D and Construction, Guangdong Branch of China Southern Power Grid Digital Grid Group Co., Ltd, Guangzhou 510000, Guangdong, China

² South China University of Technology, Guangzhou 510000, Guangdong, China

Abstract

Electric charge service and management is an important part of electric power work. The effective recovery of the electric charge relates to the smooth development of daily work and continuous improvement of the operation and management of power supply enterprises. With the large-scale implementation of the card prepayment system, the problem of electricity customers defaulting on electricity charges has been solved to a large extent, but some large electricity users still fail to pay electricity charges on time. Therefore, under the current situation of power grid development, it is still necessary to strengthen the service and management of electricity charges to promote efficient recovery of electricity charges. Speech recognition technology has increasingly become the focus of research institutions at home and abroad. People are committed to enabling machines to understand human speech instructions and hope to control the machine through speech. The research and development of speech recognition will greatly facilitate people's lives shortly. The development of 5G technology and the proposal of 6G technology make the interconnection of all things not only a hope but also a reality. To realize the interconnection of all things, one of the key technical breakthroughs is the development of a new human-computer interaction sensing system. Under the guidance of relevant theories and methods, this paper systematically analyzes the user structure, electricity charge recovery management and service system, existing problems and causes in South China, and clarifies the necessity of design and application of electricity charge service system in South China power supply companies. The experimental data and empirical analysis results show that the optimized Bert fusion model can provide more digital support for the power supply companies in South China in terms of electricity charge recovery efficiency, management level system improvement, and electricity charge service.

Keywords: Electricity service, Speech recognition, Sensing, Optimization of Bert fusion model

1 Introduction

Electricity charge recovery is the most important work of power supply enterprises, which is directly related to the survival and development of power enterprises. "Difficulty in electricity charge recovery" has always been the focus and difficulty of power

supply enterprises [1]. For electric power companies, after the monthly electricity bill is issued, the recovery of electricity bills is very important to work. To complete this work, a lot of manpower and material resources are often required. At present, the individual users of power supply companies in South China have realized the card prepaid electricity meter system and the recovery of electricity charges has become easier for prepaid users to deal with; However, for enterprise users and key male users, simple and fast reminder tools and means are still needed to accelerate the speed and efficiency of electricity charge recovery. According to relevant data, the total arrears of major customers accounting for all customers account for about 68% of the total arrears of power supply enterprises [2].

In recent years, the State Grid Corporation of China has vigorously promoted the construction of a smart grid within the power supply system nationwide, conducted business transformation on the traditional power system and power supply network, actively introduced the advanced experience and technology of the power industry in developed countries, and combined the development situation and market demand of the domestic power industry to create a new type of smart power supply system [3]. This requires local power supply companies to build an information management platform covering multiple business links such as power and energy production, allocation, supply, and communication, under the unified planning and deployment of the State Grid Corporation of China. Based on upgrading the basic physical facilities of the power supply system, they should also adopt a variety of modern high-tech means to improve the management level of related businesses [4]. At present, a new power supply system management mode with an intelligent power supply and distribution network as the core and business information management platform as the auxiliary has been preliminarily implemented in power supply companies around the country. Based on the innovation and transformation of the physical facilities of the previous power supply system, the intelligent power supply network system in the new era has been fully implemented [5, 6].

Among them, electricity, as an important frontier business for power supply companies to achieve the economic transformation of power energy and social benefits, is an important guarantee for local power supply companies to finally achieve their strategic development goals. It expands the meaning of text vectors from multiple dimensions so that the subsequent model can learn more deep semantic representations and obtain better detection efficiency. Language is by far the most important communication tool for human development [7].

Since the computer was invented, the computer has played an increasingly important role in human daily life. Therefore, any product technology that helps to fill the huge gap between humans and computers can be developed rapidly, and voice recognition technology is one of them. As the main technology of human–computer interaction communication, speech recognition has been widely concerned by domestic and foreign research institutions [8]. The rapid development of speech recognition technology makes speech recognition more and more important in various fields [9]. The world has entered the information age, with a sharp increase in scientific and technological materials. International cooperation and exchanges have become increasingly extensive and in-depth. However, language differences have become a serious obstacle for people to obtain information, enhance understanding, and expand exchanges and cooperation. Therefore, the

development of speech recognition technology is not only an important fulcrum for the development of each country but also an important technology to strengthen national and international competitiveness [10, 11]. The ultimate goal is to develop a machine that can understand human speech. On the one hand, it can convert the speaker's voice into text information, and on the other hand, it can make correct responses to the speaker's voice, not just adhere to the accurate conversion of all words into written text.

Speech recognition technology is a comprehensive technology integrating acoustics, phonetics, computer, information processing, and artificial intelligence. It has a relatively broad application prospect in computer and communication fields [12]. Although robots have not yet begun to appear in the human world in a large area, this trend will be irresistible and cannot be ignored. However, the large-scale application of robots still faces some problems, such as how to make robots have the same sensing and sensing functions as human skin, imitate human muscle movement, and how make robots have the thinking ability similar to the human brain. Among them, how to make robots have the same sensing and perception function as human skin can imitate various information of human sensing and perception of the external environment through the relevant knowledge and methods of electronics, and human-computer interaction related fields. This paper focuses on the research of visual sensing and perception characteristics and the corresponding image processing process, focusing on the three aspects of visual attention characteristics in the early stage of visual sensing and perception, visual discrimination in the process of sensing and perception, and the comprehensive experience of image quality in the later stage of visual sensing and perception.

In the research work of this paper, the electric charge service subsystem in the electric charge service platform of South China Power Supply Company is taken as the specific research object. Based on the actual project participation experience, through the investigation and analysis of the relevant technical basis of the electric charge service subsystem in the design and development process, according to the relevant management mode and theory of software engineering, the design and development work of the system are analyzed and described in detail; In the trial run, computer software technology and database technology are used to conduct the centralized online implementation of platform charging management, daily settlement of charges, data query and retrieval, and liquidated damages management in the electric charge service business of power users of power supply companies. Under the unified deployment and management mode of the power digital management platform, automatic maintenance and implementation of the electric charge service business are realized [13].

The main contribution of the paper is discussed as follows:

The problem of electricity customers defaulting on electricity charges has been solved to a large extent with the large-scale implementation of the card prepayment system, but some large electricity users still fail to pay electricity charges on time. Based on this issue, the paper analyzes the user structure, electricity charge recovery management and service system, existing problems and causes in South China, and clarifies the necessity of design and application of electricity charge service system in South China power supply companies. The optimized Bert fusion model can provide more digital support for the power supply companies in South China in terms of electricity charge recovery efficiency, management level system improvement, and electricity charge service.

2 Related works

The literature [14] mentioned that the informatization of power business has long been an important task in the development of domestic and foreign power industry, especially after the gradual popularization of computer technology, power business management tools and platforms based on automatic network communication technology and computer software technology have come out one after another. While popularizing and improving the overall business capability of the power industry, it also has a significant impact on the business implementation mode and management process of the power industry. According to the literature [15], since the development of the power industry in various countries is closely related to its economic development model, the domestic industrial sector's demand for power energy and national macro strategy, a relatively unified power information management model and technical system have not been formed in the world at present. The national electric power management department mainly implements and plans according to local conditions based on the actual national conditions and the development needs of the electric power industry.

Literature [16] proposed that the initial research on speech recognition originated in 1930 when the research focus was mainly on speaker recognition. At first, Bell Laboratories conducted speech recognition only by observing the voiceprint spectrum and then introduced the method of probability and statistics for voiceprint recognition [17]. After 1950, the basic theoretical research of speech recognition became the focus of this period, including feature extraction and cepstrum analysis technology, which developed rapidly during this period, and the earliest voice-controlled typewriter was developed during this period. Then, speech recognition entered the era of rapid development and greatly promoted the future development trend. According to the literature [18], due to the difference between Chinese pronunciation and English pronunciation, Chinese speech recognition technology is more difficult than English. China also attaches great importance to the development of speech recognition, but due to the late start compared with Western countries, there are few papers published in domestic academic journals. Although it started relatively late, the research of speech recognition has been following the international development step. According to literature [19], in the 1980s, the International 863 Program was implemented and scientific research institutions such as the State Key Laboratory of Pattern Recognition and the Institute of Acoustics of the Chinese Academy of Sciences were established, which greatly promoted the research and development of Chinese speech recognition. In the past two years, with the popularity of smartphones, iFLYTEK has made outstanding contributions to the application of Chinese speech recognition in the market and launched products related to Chinese continuous speech recognition with a good recognition rate. In general, domestic speech recognition technology has developed rapidly with outstanding achievements and has gradually moved from the experimental development stage to market application.

At the initial stage of sensing and perception, the human visual system does not process all image regions equally but filters out important regions through a visual attention mechanism for further priority interpretation [20]. The image saliency detection algorithm based on visual attention can effectively reduce the content of the image to be processed, thus improving image processing efficiency. Literature [21] proposed that in the process of sensing and sensing input, due to the limited resolution of the visual system,

it is impossible to detect changes in signal content below a certain threshold. Distortion threshold can be just identified to represent the sensing and perception ability of the visual system, which can effectively remove the redundant information of the image, thus improving the image compression performance. With the introduction of Bidirectional Encoder Representations from Transformers (BERT) and other algorithms, deep learning has further improved its performance in the natural language processing (NLP) field.

Due to the leading developer of the social service industry in foreign countries, especially in developed countries, a relatively reasonable electricity service and management model has also appeared in the field of electricity service and management. According to the literature [22], the practice of Japan and the USA is outstanding, which provides guidance and reference for the practice of electricity tariff service and management in other countries. As an important part of electric power, both Japan and the USA attach great importance to the management and service of electric charge and have well reflected the mode of electric charge service and management in their electric power systems. Tokyo Electric Power Company of Japan and Edison Electric Power Company of the USA are two representative enterprises that have developed well in the field of electric charge service and management. This Tokyo Electric Power Company, they have established a relatively sound customer service system; At present, there are customer service centers that provide advisory services to customers, including electricity fee service business, mainly including inquiry of electricity fee composition, electricity fee settlement, and early settlement of electricity fees. According to the literature [23], after applying for payment of electricity charges through the customer center, network, or telephone, users can choose to pay through financial institutions, post offices or business departments of power supply companies, or they can choose the automatic bank transfer function to pay electricity charges, which provides convenience for users to pay electricity charges, and also makes it easier for power supply enterprises to recover electricity charges. In Literature [24] proposed approach, called "Fusion-ConvBERT," utilizes deep learning techniques to effectively extract significant features associated with speaker emotions from speech signals. Through extensive experiments, the model was found to surpass existing state-of-the-art methods in a majority of test scenarios, demonstrating its superior performance. This innovative approach maximizes the utilization of available information in the given speech signals, leading to improved results compared to previous techniques. In Literature [25], two competitive fusion approaches are proposed to address the challenge of covering diverse linguistic content in streaming recurrent neural network transducer (RNNT) models for speech recognition: Competitive Shallow Fusion and Competitive Cold Fusion. These approaches dynamically select the most suitable language model during streaming processing, allowing us to handle a broader range of linguistic content and improve accuracy.

According to the literature [26], as the main operating income of power supply enterprises, electricity charges have become a common concern of power supply enterprises in various countries, including China. The literature [27] shows that, especially with the arrival of the information age, there is more and more research on the design and development of electricity fee services and service systems, which strive to create an effective and reasonable electricity fee service mode for power supply enterprises.

Comparison of prior works

Reference	Method	Results	Limitations
[11]	A research study was conducted to create a speech recognition system specifically designed for classifying nine Thai syllables. The system utilized surface electromyography (sEMG) signals from the articulatory muscles, obtained through five different channels	The average classification accuracies achieved were 94.5% for healthy volunteers and 89.4% for dysarthric participants	The recognition system used for communication in dysarthric patients is not much effective
[14]	Preoperative visual assessments of working memory and inhibition can predict postoperative speech recognition measures six months after cochlear implantation. Patients with lower baseline cognitive abilities showed the greatest improvements in cognitive measures following the procedure	Patients with lower baseline cognitive abilities improved the most after CI use, with visual WM, concentration, and inhibition tasks improving the most	The study had a small sample size of 19 patients with good preoperative cognition, the correlations between effect sizes were significant, indicating promising results. While the risk of test–retest learning bias is unknown, Spearman correlations suggest that it would not significantly impact cognitive improvements
[17]	Convolutional neural networks (CNNs) have proven to be highly effective in various machine learning applications, including computer vision, speech recognition, board game playing, and medical diagnosis	The proposed accelerator efficiently processes images with an impressive resolution of 250,000 pixels, enabling successful recognition of handwritten digit images with an impressive accuracy of 88%	Even though optical neuromorphic processing is a relatively new subject, ONNs have now entered the TOPS regime, with the potential to reach the petaops per second regime
[19]	This paper presents three robust architectures for automatic speech emotion recognition. These architectures, based on hybrid CNN and FDN models, outperform state-of-the-art models on the RAVDESS dataset	The proposed models achieve an overall precision ranging from 81.5 to 85.5% and an overall accuracy between 80.6 and 84.5%	The proposed architecture can analyze certain languages with particular datasets
[23]	The paper introduces a novel Convolutional Neural Networks (CNNs) architecture for Speech Emotion Recognition. The researchers conducted extensive evaluations using the Acted Emotional Speech Dynamic Database (AESDD) as the training and testing dataset	The results demonstrate that the proposed CNN architecture surpasses the performance of previous baseline models by a significant margin of 8.4% in terms of accuracy	The investigation of language and speaker-dependent and independent approaches are not possible in this model

To sum up, while studying electricity charge recovery, experts and scholars are also constantly exploring the new mode of electricity charge service and management and putting forward the design and development of electricity charge collection systems from different technical levels. No matter what kind of system is developed and designed, the purpose is to provide modern management tools for electric charge service and electric charge recovery, to improve the level of electric charge service and the efficiency of electric charge recovery, which is worthy of the reference.

2.1 Optimization of Bert fusion model

The pre-training model uses a lot of data to train in advance, to obtain a basic model with strong generalization ability. When encountering a specific task, you can fine-tune the task on the model and get good results. This advantage is that after using a large

amount of data for training in advance, you can complete the training without too much data for a specific task, saving time and cost. According to the principle of the exponential smoothing method, the formula for calculating the exponential smoothing value is:

$$S_t^{(1)} = aX_t + (1 - a)S_{t-1}^{(1)} \tag{1}$$

$$S_t^{(2)} = aS_t^{(1)} + (1 - a)S_{t-1}^{(2)} \tag{2}$$

$$S_t^{(3)} = aS_t^{(2)} + (1 - a)S_{t-1}^{(3)} \tag{3}$$

The multi-dimensional vector set is:

$$u'_i = (u_{1i}, u_{2i}, \dots, u_{pi}) \tag{4}$$

Then:

$$F_i = u_{1i}X_1 + u_{2i}X_2 + \dots + u_{pi}X_p = u'_iX \tag{5}$$

Then, the mathematical model of the principal component is:

$$\begin{cases} F_1 = u_{11}X_1 + u_{21}X_2 + \dots + u_{p1}X_p \\ F_2 = u_{12}X_1 + u_{22}X_2 + \dots + u_{p2}X_p \\ \dots \\ F_r = u_{1r}X_1 + u_{2r}X_2 + \dots + u_{pr}X_p \end{cases} \tag{6}$$

The formula for processing the directional indicators of the original data of the sample observation matrix is:

$$X_{ij} = \frac{x_{ij} - x_j}{\sqrt{\text{var}(x_i)}} \quad i = 1, 2, \dots, n; j = 1, 2, \dots, p \tag{7}$$

The iteration process of free parameters is as follows:

$$\frac{\partial \varepsilon(n)}{\partial w_i(n)} = \sum_{j=1}^n e_j(n)g(x_j - c_i(n)) \tag{8}$$

$$w_i(n + 1) = w_i(n) - \eta_1 \frac{\partial \varepsilon(n)}{\partial w_i(n)} \tag{9}$$

The formulas for these three steps are as follows:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{10}$$

$$\tilde{c}_t = \tan d(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{11}$$

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \tag{12}$$

Center of hidden unit:

$$\frac{\partial \varepsilon(n)}{\partial c_i(n)} = 2w_i(n) \sum_{j=1}^n e_j(n)g(x_j - c_i(n)) \sum_i^{-1} [x_j - c_i(n)] \tag{13}$$

$$c_i(n + 1) = c_i(n) - \eta_2 \frac{\partial \varepsilon(n)}{\partial c_i(n)} \tag{14}$$

3 Methods

3.1 Data storage organization

Based on the types and demand characteristics of power customers in the areas under the jurisdiction of power supply companies in South China, combined with the current status of power charge service and management of power supply companies in South China, this paper studies the development and design of the power charge service system. In terms of power supply capacity, power supply companies in South China currently have jurisdiction over substations and substations; the main transformer shall be installed accumulatively, and the capacity of the main transformer shall reach mean subtraction, variance normalization, and ARMA filtering (MVA). In addition, it also has jurisdiction over substations, sub-district distribution rooms, sub-box transformers, sub-distribution transformers, overhead lines, cables and low-voltage cables. The power supply companies in South China have a total transformer and voltage transformer capacity of, km of overhead lines, km of low-voltage overhead lines, km of cables and km of low-voltage cables.

In terms of organization and management, the organization and management level of power supply companies in South China has been continuously improved, with a manager’s office, a political work office, a supervision office, a production command center, a development planning office, a labor and personnel office, a financial office, a safety supervision office, a production technology office, a project construction office, and an administrative security office; the multi-business office has 8 functional management departments and grassroots work units including 2 productions, operation, and maintenance departments, 1 department and multi-business companies. In terms of electricity charge recovery service, the power supply companies in South China have adopted the service mode shown in Fig. 1. They have achieved differentiated management and services for customers who pay their electricity bills on time, while for customers who fail to pay their electricity bills on time, they have to implement power outages and resumption following relevant laws and regulations.

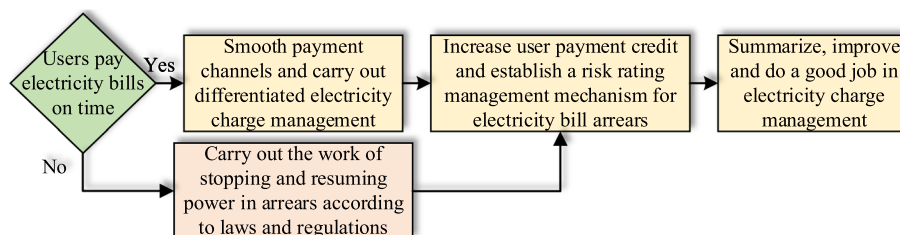


Fig. 1 Electric charge service mode of power supply companies in South China

Figure 1 represents the electric charge service mode of power supply companies in South China. Users who pay their electricity bills on time are used as input for the smooth payment channels and process differentiated electricity charge management. Electric charge service and management are critical components of power generation. The efficient recovery of electric charges is linked to the smooth expansion of daily work and continuous enhancement of power supply solid operation and management. The issue of power consumers not paying their electricity bills on time has been mitigated to a large extent by the widespread implementation of the card prepayment system, though some major electricity users continue to fail to do so. The smooth payment channels then help in the improvement of user payment credit and the establishment of a risk rating management mechanism for electricity bill arrears. If a user pays electricity bills on time, the process does a good job in electricity charge management; otherwise, if the user does not pay electricity bills on time, the process performs the work of stopping and resuming power in arrears in accordance with laws and regulations.

The distribution network model is an abstraction of the actual distribution network. The first purpose of modeling is to provide the global basic data for the calculation of electricity and charge and to define the measurement point where the calculation is located and the distribution network elements used. In addition, it can also make the business personnel and management personnel intuitively understand the basic structure of the calculation of the electric charge, to facilitate the power consumption analysis of the electric energy system and the summary of the data reports of each metering point. On this basis, there are both objective and subjective reasons for the generation of electricity charge security risks, including the reasons of power supply companies and customers. The objective reason is the security risk of electricity charges caused by the on-site environment, abnormal metering devices, and other objective conditions. If the metering device is abnormal, the reading power of the power customer is far less than the actual power consumption, which reduces the user's payment and leads to property losses for the power supply company. The subjective reason is the safety risk of electricity charges caused by the factors such as users, power supply companies, etc. For example, the user refuses to pay when using the electricity, resulting in property losses for the power supply company, as shown in Fig. 2.

From the current electricity charge service and management mode of power supply companies in South China, their electricity charge service mode has been quite mature, providing good guarantees and support for more efficient power development. However, the current model pays more attention to electricity service, while the implementation of electricity service is poor. When collecting electricity charges from key customers or important customers, the staff mostly regard the collection of electricity charges as mechanical work, neglecting the implementation of the concept of electricity charge service in this process, believing that the work content is completed when the collection of electricity charges is completed. There are two typical ways to deal with meters: combination and aggregation of the relationship between meters and distribution network. The whole and part cannot exist independently. The meters, substation lines, switching stations, and transformers are regarded as elements of the distribution network. At the same time, meters are divided into two meters: purchase and sale meters and user

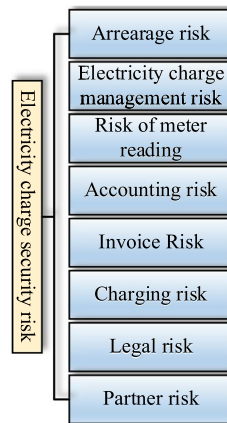


Fig. 2 Classification of electricity charge security risks

meters. In the purchase and sale meters, attribute fields are used to indicate whether they are user meters. As shown in Figs. 3 and 4.

For example, word vector feature extraction is based on the BERT algorithm and word vector feature extraction using the Word2Vec method. The BERT model is a pre-training model for natural language processing. The pre-training model uses a lot of data to train in advance, to obtain a basic model with strong generalization ability.

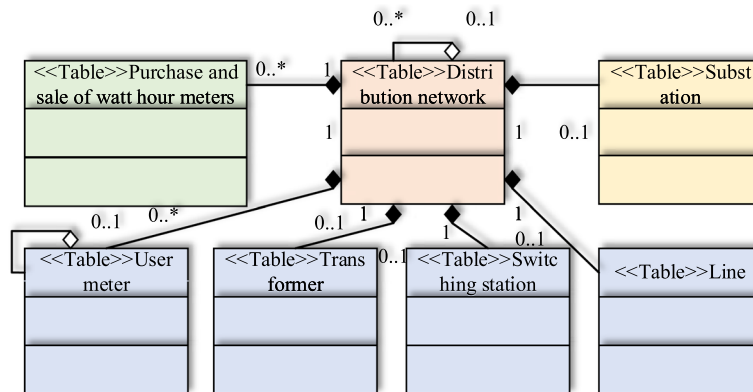


Fig. 3 Distribution network data organization model optimization diagram—overall mode

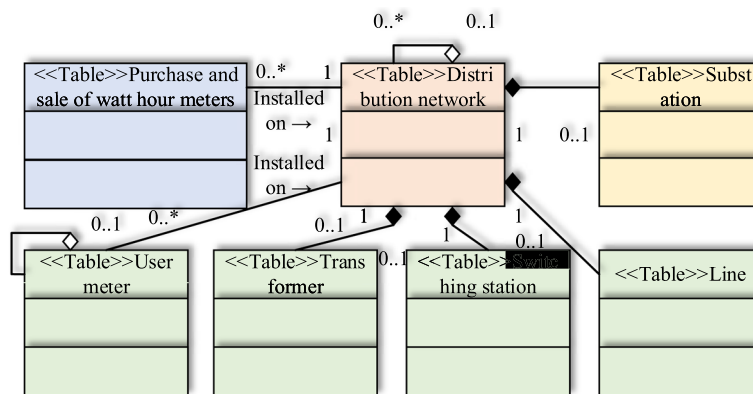


Fig. 4 Distribution network data organization model optimization diagram—expansion mode

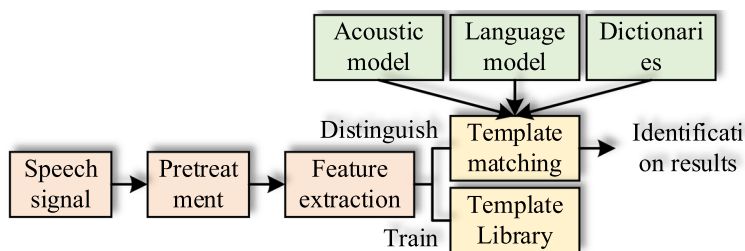


Fig. 5 Optimization diagram of calculation activity algorithm of flow framework ratings table of speech recognition system

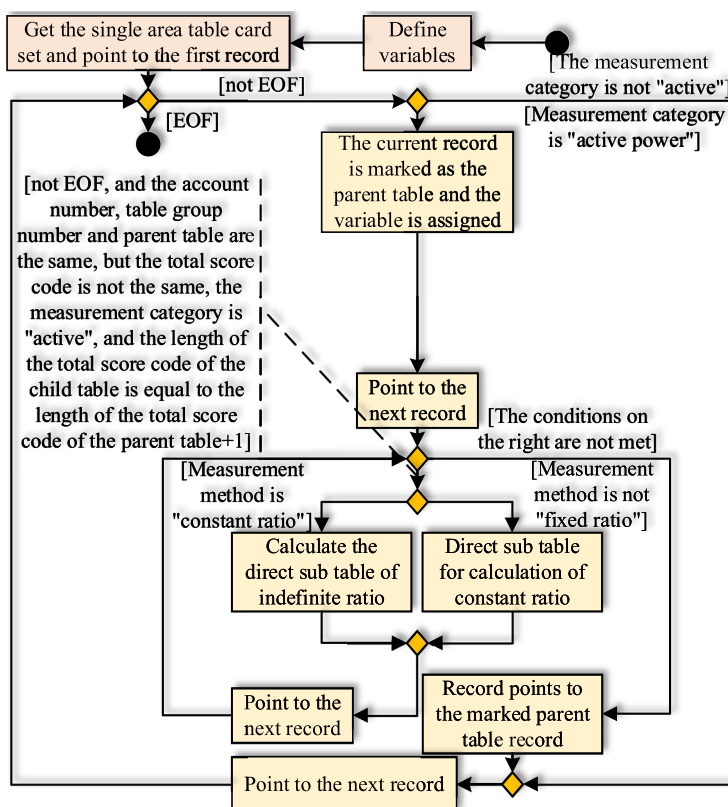


Fig. 6 Optimization diagram of calculation activity algorithm of constant ratio table

3.2 Data addition, deletion, and search

The optimized voice recognition system of South China electric charge service belongs to the category of pattern recognition system in essence, and its overall framework is shown in Fig. 5.

The algorithm shown in Fig. 6 is described in detail as follows.

From the above analysis, it is concluded that the basic calculation process for unlocking the transfer relationship is shown in Fig. 7.

Of course, before analyzing and processing the voice signal, the first thing to do is endpoint detection, which can find the signal to be analyzed from the original voice input signal. The calculation process for solving the total score relationship is shown

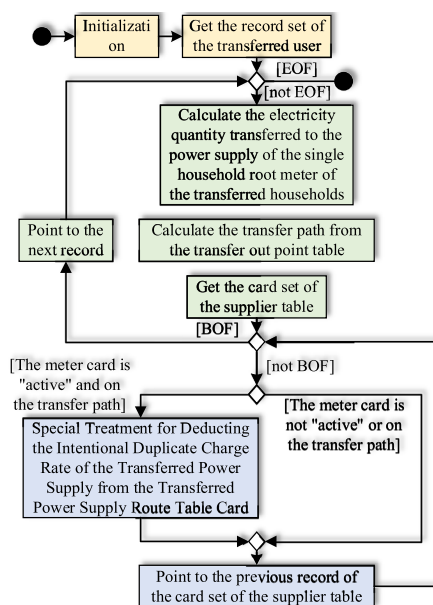


Fig. 7 Optimization of algorithm process for unlocking the transfer relationship

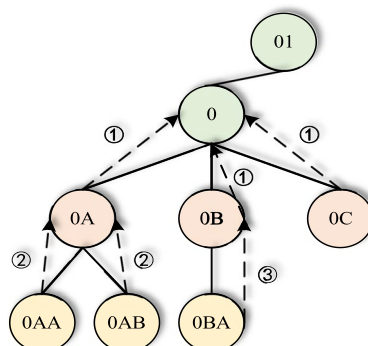


Fig. 8 Flow optimization diagram of total score relationship decomposition

in Fig. 8. The detailed processing of all users’ cards in the form reading area is as follows.

According to the data simulation calculation, the existing just noticeable difference (JND) calculation model is mainly based on the sensing and perception characteristics of the visual system on image brightness and edges. However, there are complex and nonlinear interactions between them, so it is necessary to consider the structural characteristics of image content to further analyze the resolution of HVS. In addition, it has limited knowledge of the specific process of human visual system (HVS) sensing perception, so it needs to further analyze the attenuation effect of noise, to propose an objective quality evaluation algorithm that is more consistent with subjective sensing perception. The last three recorded warning signals are shown in Table 1.

From the overall perspective, the power business information platform of the domestic power industry still needs to be further improved and enhanced in the development process. The power service system platform is an information-based business management

Table 1 Early warning signal of electric charge

Date	The total amount of arrears	Early warning signal	Accounts in arrears	Early warning signal
20,110,716	3,238,998.86	Red	155	Yellow
20,110,720	77,051.52	Green	5	Green
20,110,812	1,139,297.74	Orange	28	Blue

platform developed by the power center of Dayi Power Supply Company. Among them, as one of the main subsystems of the platform, the electric charge service subsystem needs to strictly follow the overall planning of the power business of the power supply company, manage and maintain the business processes related to the power management activities of the power supply company and the electric charge collection management business of power users in a unified manner, and provide an efficient online business management and implementation platform for relevant users.

To sum up, for power supply enterprises, establishing and maintaining good relations with power customers can not only smoothly carry out power sales but also play a role in smooth electricity charge recovery, making power become easy and smooth work. State Grid Corporation of China and power grid enterprises at all levels have begun to explore the application of advanced technology in electricity charge service and management, and have continuously introduced advanced electricity charge service and management models. The extensive application of advanced technology will become the most important feature in the future development of electricity charge service and management models.

3.3 Case study

3.3.1 Regression analysis test

Under the current electric charge service and management system, the electric charge service work of power supply companies in South China is often carried out only when the electric charge is found to be in arrears, which is very purposeful and completely aimed at the recovery of electric charge. However, this kind of service work will be weakened or even terminated to varying degrees after the charge recovery, which makes the charge service on the surface become a passive behaviour under the condition of overdue charge, which is not only detrimental to the establishment of a perfect charge service system within power supply enterprises but also will damage the relationship. The work of electricity charge recovery should be based on the electricity charge service, which should be a service rather than a passive service. Only when the service level reaches the degree of user satisfaction can it play a positive role in promoting electricity charge recovery and virtually reduce the difficulty of electricity charge recovery. This kind of service should include not only the service work in daily work but also the warm reminder and holiday greetings provided by the electricity charge service system. These services will become the bridge between power supply enterprises and electricity customers, and the key to forming good customer relations.

Table 2 KMO inspection of electricity service

KMO sampling moderate inspection value		0.614
Bartlett sphericity test value	X square	831.715
	Unit root test	156
	Level of significance	0.000

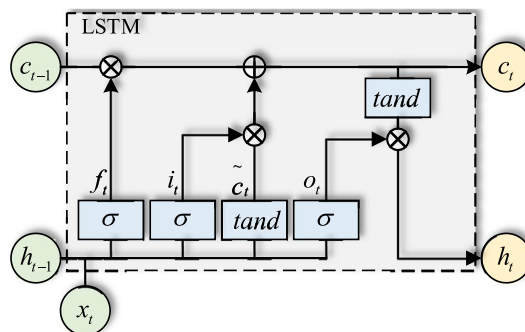


Fig. 9 LSTM model optimization structure

Under the leadership of the State Grid, the power supply companies in South China have established a system of one department and three centers, but the service practices are insufficient. They should strengthen the targeted electricity charge service and management in the future to form a good online and offline coordination situation. The KMO test, also known as the Kaiser–Meyer–Olkin test, is used to determine whether a dataset is appropriate for factor analysis. It assesses the strength of correlations between variables and converts a value ranging from 0 to 1. Higher values (closer to one) indicate a better fit for factor analysis. The KMO statistic is used by researchers to determine whether the variables are sufficiently correlated. If the KMO value is low, it indicates that the dataset may not be suitable for factor analysis and that alternative approaches should be considered. The results of the KMO test on their indicators system are represented in Table 2, in which the KMO sampling moderate inspection value is taken as 0.614 and the Barlett sphericity test value of x square is 831.715, the unit root test value is 156 and the significance level is 0.000, respectively. The results of the KMO test on the indicator system are shown in Table 2.

The frequency domain of digitized speech signals still changes with time. To use traditional signal analysis methods for subsequent processing, it is necessary to assume that the signal has short-time stationary characteristics in a certain short time. The following analysis and processing are feasible only on the premise that the short-time stationarity of speech signal is established. To solve the long-distance dependency problem, long short-term memory (LSTM) introduces three gates to implement its structure, as shown in Fig. 9.

The initial state of the BERT model used in this experiment is the parameters in the pre-training model of the BERT model. BERT structure is shown in Fig. 10.

The AdaBoost method adds the idea of iterative updating. When using the basic classifier for training, the weight of the samples that are wrongly classified will increase. At the

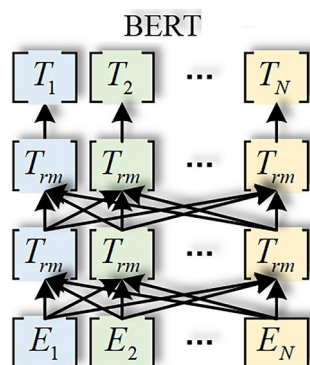


Fig. 10 Structure of BERT model optimization

same time, the weighted training set will be used to train the next basic trainer until the whole model reaches the set error rate or the maximum number of iterations. However, using the AdaBoost method as a classifier can avoid overfitting and its results are understandable. The process is shown in Fig. 11.

People cannot receive and process voice signals. The computer cannot process the received voice signal and distinguish the semantics of the voice signal at the same time.

3.3.2 Baseline model comparison

For the confusion matrix of the voice recognition system for electric charge service in South China, it can be classified as a second-order text classification, as shown in Table 3.

After feature extraction, the three methods uniformly access Text CNN for downstream classification tasks. The comparison results of the final classification accuracy are shown in Table 4 and Fig. 12, in which three different models such as BERT feature extraction, Word2Vec feature extraction and Word feature combination model are

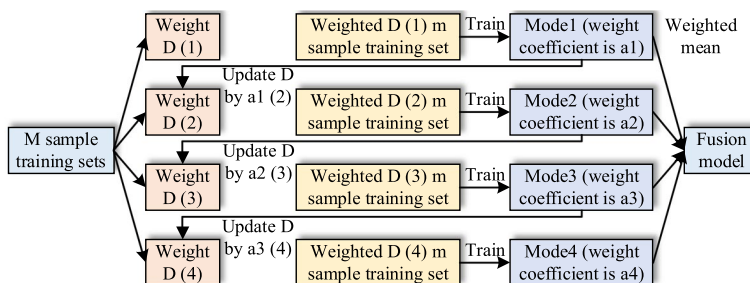


Fig. 11 Optimization of AdaBoost training process

Table 3 Power charge service confusion matrix

	Actual value: true	Actual value: false
Predicted value: true	TN	FN
Predicted value: false	TP	FP

Table 4 Comparison of classification accuracy of different feature extraction methods

Model Selection	Accuracy	Recall	F1 value
BERT feature extraction	0.8714	0.8708	0.8713
Word2Vec feature extraction	0.8643	0.8905	0.8818
Word feature combination model	0.8833	0.8826	0.8826

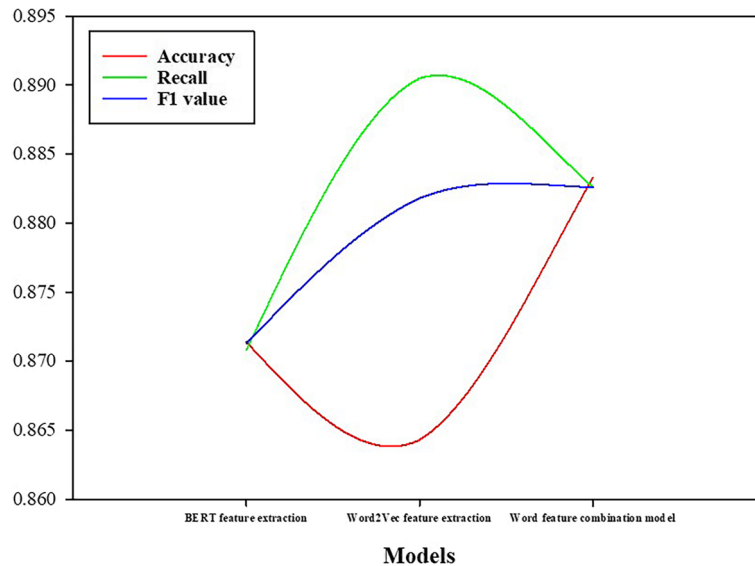


Fig. 12 Comparison of classification accuracy of different feature extraction methods

compared based on their accuracy, recall and F1 value. BERT, a transformer-based language model, is used for feature extraction in natural language processing (NLP) tasks. By tokenizing and inputting text into BERT, contextualized word and sentence embeddings are generated. These embeddings can be used to extract features such as word embeddings, sentence embeddings, or pooled representations. These features are then utilized for downstream NLP tasks like text classification or sentiment analysis. Feature extraction with BERT enables leveraging its pre-trained contextual representations to enhance NLP model performance. Word2Vec generates word embeddings that capture semantic meaning. Features can be extracted by training the model on text data and retrieving embeddings for words or averaging them for sentences/documents. These features enhance NLP tasks like classification or clustering. Word2Vec leverages semantic information for effective NLP applications.

At the same time, to ensure the reliability of the speech recognition experiment of the South China electric charge service, the test is conducted by repeating the experiment and took the average set as the monitoring point. The model loss changes are shown in Figs. 13, 14 and Table 5:

From this figure, the model loss of different models in training varies with the number of training iterations. Since the epoch of model training is 5 and the batch sent each time is 128, the total number of iterations is about 2000. It can be seen from the figure that all models tend to converge after 2000 iterations. Then, the comparison

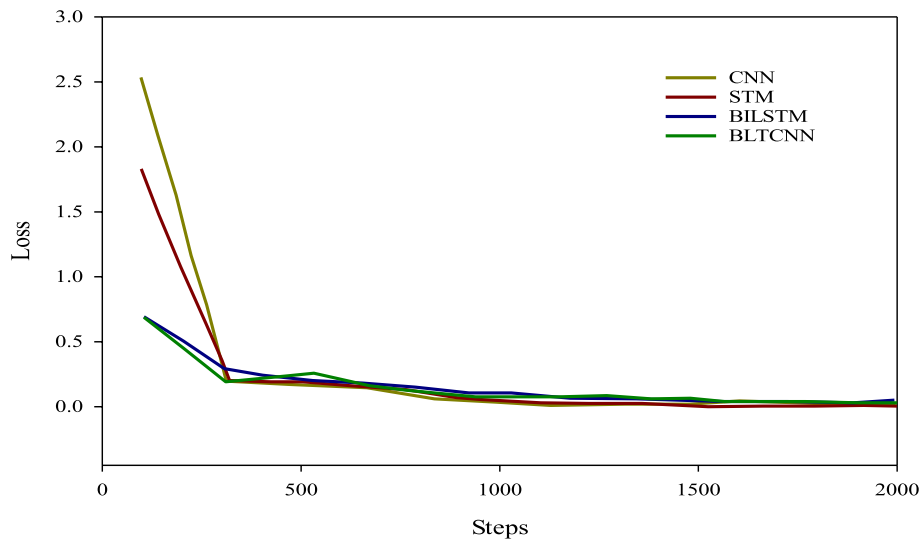


Fig. 13 Change of loss of each model with iteration times

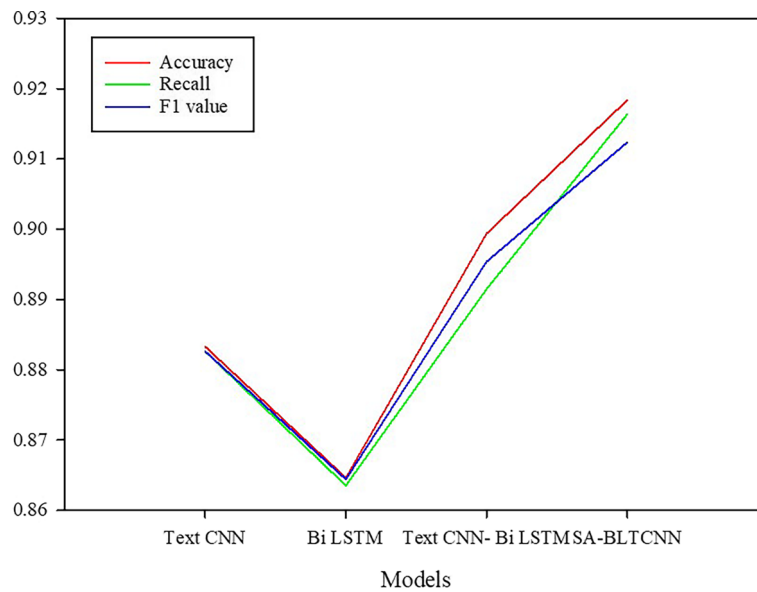


Fig. 14 Comparison of classification results of different models

Table 5 Change of Loss of Each Model with Iteration Times

Steps	CNN	STM	BiLSTM	BLTCNN
0	2.519	1.808	0.671	0.671
500	0.15	0.185	0.193	0.228
1000	0.011	0.028	0.098	0.046
1500	-0.006	-0.015	0.028	0.037
2000	0.002	-0.024	0.02	0.002

Table 6 Comparison of classification results of different models

Model selection	Accuracy	Recall	F1 value
Text CNN	0.8833	0.8826	0.8826
Bi LSTM	0.8646	0.8635	0.8644
Text CNN-Bi LSTM	0.8994	0.8915	0.8954
SA-BLTCNN	0.9184	0.9164	0.9124

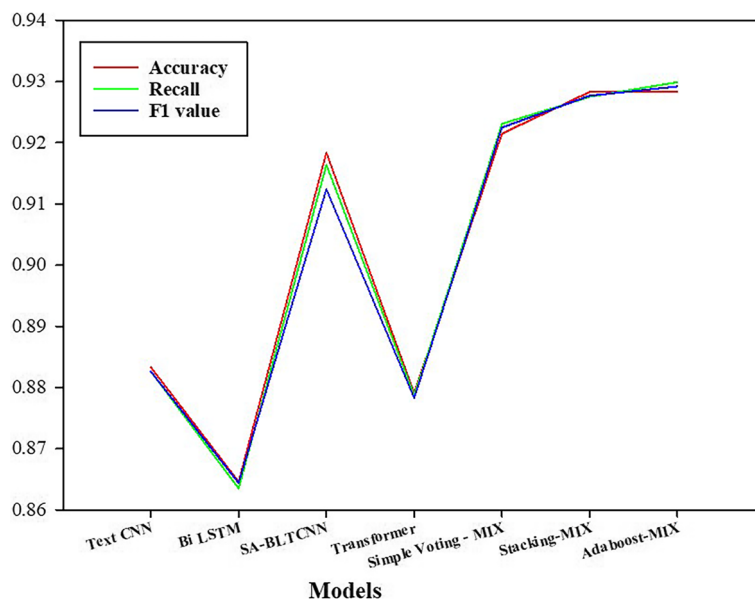


Fig. 15 Comparison of different fusion strategies

results of the results obtained by each model on the same data set after the test are obtained, and the results are shown in Table 6 and Fig. 15.

The results have a guiding significance for analyzing the sensitivity of human eyes to image content. In addition, according to the internal derivation mechanism, analyzing the impact of noise on image content sensing perception is of great significance for designing a more effective image quality evaluation algorithm. The experiment is mainly to verify whether the model fusion design can ultimately improve the detection performance. This experiment still carries out ten experiments for each fusion method and takes the average value of ten experiments as the final result, as shown in Table 7 and Fig. 15.

The detection results of the above models can be converted into more intuitive figures, as shown in Fig. 16.

Comparison Chart At present, under the overall framework of the State Grid Corporation of China, it has become a positive measure for regional power supply enterprises to actively develop the electricity charge service and management mode that combines with the actual regional economic and social development. With the

Table 7 Comparison of different fusion strategies

Model selection	Accuracy	Recall	F1 value
Text CNN	0.8833	0.8826	0.8826
Bi LSTM	0.8646	0.8635	0.8644
SA-BLTCNN	0.9184	0.9164	0.9124
Transformer	0.8791	0.8787	0.8783
Simple voting-MIX	0.9215	0.9231	0.9225
Stacking-MIX	0.9283	0.9275	0.9277
Adaboost-MIX	0.9284	0.9299	0.9292

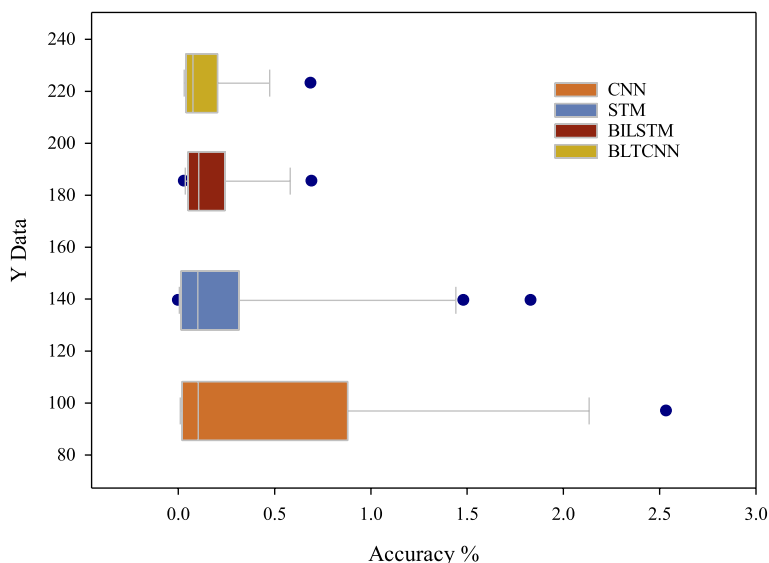


Fig. 16 Classification results of models

gradual appearance of the differences in economic development status and economic development level between regions, power grid enterprises at all levels will pay more attention to combining closely with the local actual situation, constantly innovate the power model, and build a regional electricity charge service and management system will still be an important link and development trend in the future power industry.

To sum up, after the implementation of the concept of "customer-centered" in the electric charge service and management model, the direction of the original electric charge service and recovery of electric power enterprises has been completely changed. The service has gradually changed from passive, and the electric charge recovery has also changed from forcing users to users. By providing satisfactory services to customers, the proposed model achieves the purpose of effective recovery of electricity charges based on strengthening customer relations and achieve a win-win situation for power supply enterprises and customers.

4 Conclusion

Based on the actual situation of electricity charge recovery of power supply companies in South China, this research designed, analyzed, and applied the electricity charge service system of power supply companies in South China using project management theory, electricity charge service theory, theory and method. The related algorithms in the speech recognition process, including preprocessing, are studied in depth in this paper, and the main endpoint detection and recognition algorithms involved in the process are simulated. Simultaneously, an improved feature parameter extraction algorithm is proposed, which reduces the computation required for Mel-frequency cepstral coefficients (MFCC) extraction by nearly 50% and significantly improves feature extraction efficiency. This paper made some theoretical developments as well as technical implementation innovations. It makes the way for new methods of objective image processing based on subjective visual perception. The electric charge service system is not only a collection system for electric charges but also a link and bridge between power supply enterprises and important customers. According to the development strategy of power supply enterprises, the concept of customer-centricity will be more deeply implemented in the service and management of electricity bills in the future.

Abbreviations

MFCC	Mel-frequency cepstral coefficients
Text CNN	Text convolution neural network
Bi LSTM	Bidirectional long short-term memory
SA-BLTCNN	Sentiment analysis with bi-layer temporal convolutional neural networks
BLTCNN	Binary latent tree convolutional neural network
TN	True negative
TP	True positive
FN	False negative
FP	False positive
BERT	Bidirectional Encoder Representations from Transformers
NLP	Natural language processing
LSTM	Long short-term memory
KMO	Kaiser–Meyer–Olkin test
HVS	Human visual system
JND	Just noticeable difference
MVA	Mean subtraction, variance normalization, and ARMA filtering
AESDD	Acted Emotional Speech Dynamic Database
CNN	Convolutional neural networks
RAVDESS	Ryerson Audio-Visual Database of Emotional Speech and Song
TOPS	Trillions of operations per second
sEMG	Surface electromyography
RNNT	Recurrent neural network transducer

Acknowledgements

The authors would like to sincerely thank those techniques who contributed to this research.

Funding

There is no specific funding to support this research.

Availability of data and materials

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Declarations

Consent for publication

All authors reviewed the results, approved the final version of the manuscript and agreed to publish it.

Competing interests

The authors declared that they have no conflicts of interest regarding this work.

Received: 17 April 2023 Accepted: 17 October 2023

Published online: 06 November 2023

References

1. L. Peng, S. Liu, R. Liu, L. Wang, Effective long short-term memory with differential evolution algorithm for electricity price prediction. *Energy* **162**, 1301–1314 (2018). <https://doi.org/10.1016/j.energy.2018.05.052>
2. L. Wen, K. Zhou, S. Yang, A shape-based clustering method for pattern recognition of residential electricity consumption. *J. Clean. Prod.* **212**, 475–488 (2019). <https://doi.org/10.1016/j.jclepro.2018.12.067>
3. A.B. Nassif, I. Shahin, I. Attili, M. Azzeh, K. Shaalan, Speech recognition using deep neural networks: a systematic review. *IEEE Access.* **7**, 19143–19165 (2019). <https://doi.org/10.1109/ACCESS.2019.2896880>
4. T. Kawase, M. Okamoto, T. Fukutomi, Y. Takahashi, Speech enhancement parameter adjustment to maximize the accuracy of automatic speech recognition. *IEEE Trans. Consum. Electron.* **66**(2), 125–133 (2020). <https://doi.org/10.1109/TCE.2020.2986003>
5. A. Jacks, K.L. Haley, G. Bishop, T.G. Harmon, Automated speech recognition in adult stroke survivors: comparing human and computer transcriptions. *Folia Phoniatr. Logop.* **71**(5–6), 286–296 (2019). <https://doi.org/10.1159/000499156>
6. H. Kwon, H. Yoon, K.W. Park, Acoustic-decoy: detection of adversarial examples through audio modification on speech recognition system. *Neurocomputing* **417**, 357–370 (2020). <https://doi.org/10.1016/j.neucom.2020.07.101>
7. X. Cui, W. Zhang, U. Finkler, G. Saon, M. Picheny, D. Kung, Distributed training of deep neural network acoustic models for automatic speech recognition: a comparison of current training strategies. *IEEE Signal Process. Mag.* **37**(3), 39–49 (2020). <https://doi.org/10.1109/MSP.2020.2969859>
8. C. Tong, J. Li, C. Lang, F. Kong, J. Niu, J.J. Rodrigues, An efficient deep model for day-ahead electricity load forecasting with stacked denoising auto-encoders. *J. Parallel Distrib. Comput.* **117**, 267–273 (2018). <https://doi.org/10.1016/j.jpdc.2017.06.007>
9. K. Kanagarathinam, K. Sekar, Text detection and recognition in raw image dataset of seven-segment digital energy meter display. *Energy Rep.* **5**, 842–852 (2019). <https://doi.org/10.1016/j.egy.2019.07.004>
10. K.M. Rashid, J. Louis, K.K. Fiawoyife, Wireless electric appliance control for smart buildings using indoor location tracking and BIM-based virtual environments. *Autom. Constr.* **101**, 48–58 (2019). <https://doi.org/10.1016/j.autcon.2019.01.005>
11. N.S. Jong, P. Phukpattaranont, A speech recognition system based on electromyography for the rehabilitation of dysarthric patients: a Thai syllable study. *Biocybern. Biomed. Eng.* **39**(1), 234–245 (2019). <https://doi.org/10.1016/j.bbe.2018.11.010>
12. D. Yongda, L. Fang, X. Huang, Research on multimodal human-robot interaction based on speech and gesture. *Comput. Electr. Eng.* **72**, 443–454 (2018). <https://doi.org/10.1016/j.compeleceng.2018.09.014>
13. S. Helbig, Y. Adel, M. Leinung, T. Stöver, U. Baumann, T. Weissgerber, Hearing preservation outcomes after cochlear implantation depending on the angle of insertion: indication for electric or electric-acoustic stimulation. *Otol. Neurotol.* **39**(7), 834–841 (2018). <https://doi.org/10.1097/MAO.0000000000001862>
14. K.Y. Zhan, J.H. Lewis, K.J. Vasil, T.N. Tamati, M.S. Harris, D.B. Pisoni, W.G. Kronenberger, R. Christin, A.C. Moberly, Cognitive functions in adults receiving cochlear implants: predictors of speech recognition and changes after implantation. *Otol. Neurotol.* **41**(3), e322–e329 (2020). <https://doi.org/10.1097/MAO.0000000000002544>
15. L. Tan, K. Yu, L. Lin, X. Cheng, G. Srivastava, J.C.W. Lin, W. Wei, Speech emotion recognition enhanced traffic efficiency solution for autonomous vehicles in a 5G-enabled space-air-ground integrated intelligent transportation system. *IEEE Trans. Intell. Transp. Syst.* **23**(3), 2830–2842 (2021). <https://doi.org/10.1109/TITS.2021.3119921>
16. J. Huang, T. Lu, B. Sheffield, F.G. Zeng, Electro-tactile stimulation enhances cochlear-implant melody recognition: effects of rhythm and musical training. *Ear Hear.* **41**(1), 106–113 (2020). <https://doi.org/10.1097/AUD.0000000000000749>
17. X. Xu, M. Tan, B. Corcoran, J. Wu, A. Boes, T.G. Nguyen, T.C. Sai, B.E. Little, D.G. Hicks, R. Morandotti, A. Mitchell, D.J. Moss, 11 TOPS photonic convolutional accelerator for optical neural networks. *Nature* **589**(7840), 44–51 (2021). <https://doi.org/10.1038/s41586-020-03063-0>
18. J. Berg, S. Lu, Review of interfaces for industrial human-robot interaction. *Current Robot. Rep.* **1**(2), 27–34 (2020). <https://doi.org/10.1007/s43154-020-00005-6>
19. M. Ezz-Eldin, A.A. Khalaf, H.F. Hamed, A.I. Hussein, An efficient feature-aware hybrid model of deep learning architectures for speech emotion recognition. *IEEE Access.* **9**, 19999–20011 (2021). <https://doi.org/10.1109/ACCESS.2021.3054345>
20. M. Dua, R.K. Aggarwal, M. Biswas, Performance evaluation of Hindi speech recognition system using optimized filter banks. *Eng. Sci. Technol. Int. J.* **21**(3), 389–398 (2018). <https://doi.org/10.1016/j.jestch.2018.04.005>
21. R. Haeb-Umbach, S. Watanabe, T. Nakatani, M. Bacchiani, B. Hoffmeister, M.L. Seltzer, H. Zen, M. Souden, Speech processing for digital home assistants: combining signal processing with deep-learning techniques. *IEEE Signal Process. Mag.* **36**(6), 111–124 (2019). <https://doi.org/10.1109/MSP.2019.2918706>
22. N. Apergis, G. Gozgor, C.K.M. Lau, S. Wang, Decoding the Australian electricity market: new evidence from three-regime hidden semi-Markov model. *Energy Econ.* **78**, 129–142 (2019)
23. N. Vryzas, L. Vrysis, M. Matsiola, R. Kotsakis, C. Dimoulas, G. Kalliris, Continuous speech emotion recognition with convolutional neural networks. *J. Audio Eng. Soc.* **68**(1/2), 14–24 (2020). <https://doi.org/10.17743/jaes.2019.0043>
24. S. Lee, D.K. Han, H. Ko, Fusion-ConvBERT: parallel convolution and BERT fusion for speech emotion recognition. *Sensors.* **20**(22), 6688 (2020). <https://doi.org/10.3390/s20226688>
25. R. Cabrera, X. Liu, M. Ghodsi, Z. Matteson, E. Weinstein, A. Kannan, Language model fusion for streaming end-to-end speech recognition. *arXiv preprint* (2021). <https://doi.org/10.48550/arXiv.2104.04487>

26. J. Jiang, R. Chen, M. Chen, W. Wang, C. Zhang, Dynamic fault prediction of power transformers based on hidden Markov model of dissolved gases analysis. *IEEE Trans. Power Deliv.* **34**(4), 1393–1400 (2019). <https://doi.org/10.1109/TPWRD.2019.2900543>
27. Q. Chen, D. Xiang, L. Wang, Y. Tang, E. Harkin-Jones, C. Zhao, Y. Li, Facile fabrication and performance of robust polymer/carbon nanotube-coated spandex fibres for strain sensing. *Compos. A Appl. Sci. Manuf.* **112**, 186–196 (2018). <https://doi.org/10.1016/j.compositesa.2018.06.009>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
