ISSN 2616-5775 Vol. 7, Issue 3: 63-67, DOI: 10.25236/AJCIS.2024.070309

Research on traditional music score conversion algorithm based on image recognition technology

Dong Li

Pingdingshan University, Pingdingshan, 467000, Henan, China

Abstract: *Based on image recognition technology, this paper studies the traditional music score conversion algorithm, aiming at realizing the automatic conversion from paper score to digital score. In this paper, a music score image recognition method based on convolutional neural network is proposed, and its effectiveness is verified by experiments. On this basis, the related technologies and problems of music score conversion are further discussed, including how to preprocess music score images, how to extract music score features and how to realize music score conversion algorithms. The research results show that the traditional music score conversion algorithm based on image recognition technology has high accuracy and robustness, and can effectively realize the digital conversion of music score.*

Keywords: *image recognition technology, traditional music score, automatic conversion, convolutional neural network, score feature extraction, digital conversion*

1. Introduction

Traditional music score is an important carrier of music inheritance, but in the digital age, paper score can no longer meet people's needs for music education, performance and communication. Therefore, it is of great practical significance to realize the automatic conversion from paper music score to digital music score. [1]The traditional music score conversion method is mainly based on manual annotation or rule design, which has some problems such as low efficiency, poor accuracy and poor expansibility. The music score conversion algorithm based on image recognition technology can effectively solve these problems and become one of the research hotspots in the field of music score conversion.

2. Music score conversion algorithm based on image recognition

(1) Pretreatment of music score image

Before music score image recognition, it is necessary to preprocess the music score image in order to improve the accuracy and robustness of recognition. The preprocessing of music score image mainly includes the following steps:

De-noising: Noise and stains often exist in music score images, so de-noising is needed. Common denoising methods include median filtering and Gaussian filtering.

Binary processing: the score image is converted into a black-and-white binary image for subsequent processing. The binarization methods include global threshold method and adaptive threshold method.

Removing interference lines: There are often horizontal lines, vertical lines and other interference lines in music score images, which need to be removed. Common methods to remove interference lines include projection and morphological processing.

Segmentation processing: the score image is divided into single note regions for subsequent feature extraction and recognition. Common segmentation methods include projection-based segmentation method and edge detection-based segmentation method.

Normalization: Normalize the note regions to make them have the same size and shape, so as to facilitate subsequent feature extraction and recognition.

In the process of music score image preprocessing, it is necessary to consider the characteristics and differences of music scores of different instruments and adopt corresponding preprocessing methods. At the same time, it is necessary to pay attention not to lose the important information of the score image in the preprocessing process.

ISSN 2616-5775 Vol. 7, Issue 3: 63-67, DOI: 10.25236/AJCIS.2024.070309

(2) Music score feature extraction

Music score feature extraction is to transform the preprocessed music score image into a feature vector that can be processed by the computer, and provide input for the subsequent recognition algorithm. The purpose of music score feature extraction is to extract useful information from the image to distinguish different notes, while minimizing the interference of useless information.

Common methods for extracting musical score features include:

Geometric features: Note areas in music score images usually have a certain geometric shape, which can be expressed by extracting geometric features such as area, perimeter, width-height ratio and rectangle.

Texture features: Note areas in music score images usually have certain texture information, which can be expressed by extracting texture features such as gray level co-occurrence matrix and local binary pattern.

Frequency domain features: Note regions in music score images usually have certain frequency domain features, which can be expressed by extracting their Fourier transform coefficients, wavelet transform coefficients and other frequency domain features.

Symbolic features: Note areas in music score images usually have certain symbolic information, which can be expressed by extracting the symbolic features such as stroke number, stroke width and symbol shape.

The above feature extraction methods can be used alone or in combination to obtain more abundant and differentiated feature vectors. In the process of feature extraction, it is necessary to adopt corresponding feature extraction methods according to the characteristics and differences of scores of different musical instruments in order to improve the recognition accuracy and robustness.

(3) Music score image recognition based on convolutional neural network.

Music score image recognition based on Convolutional Neural Network (CNN) is one of the most widely used and effective music score recognition methods in recent years. CNN is a deep learning model, and its structure mainly includes convolution layer, pooling layer and fully connected layer.

The main advantage of CNN is that it can automatically learn the characteristics of images, and it has strong robustness and generalization ability. In order to improve the recognition accuracy and robustness, a large number of training data and suitable network structure should be used in music score image recognition, and reasonable data enhancement and model optimization should be carried out.

In recent years, many scholars have improved and optimized on the basis of CNN, such as introducing attention mechanism and increasing residual connection, which further improves the effect of music score image recognition.

(4) Realization of music score conversion algorithm

The realization of music score conversion algorithm is the process of inputting the pre-processed and feature-extracted music score images into the trained model for music score recognition and conversion. When realizing the music score conversion algorithm, it mainly includes the following steps:

Data set preparation: a large number of music score images are obtained from the network, and data cleaning, preprocessing, feature extraction and other operations are carried out to construct a data set suitable for training.

Model training: select the appropriate CNN network structure, use data sets for training, and optimize model parameters through back propagation algorithm to obtain a model that can accurately identify music scores.

Music score conversion: input the music score image that needs to be converted into the trained model, and perform identification and conversion operations to obtain the converted music score.

Post-processing: Post-processing the converted music score, including music score formatting, error removal, MIDI file generation and other operations.

When implementing the music score conversion algorithm, we need to pay attention to the following aspects:

Selection of data sets: the quality and quantity of data sets have great influence on the effect of model

ISSN 2616-5775 Vol. 7, Issue 3: 63-67, DOI: 10.25236/AJCIS.2024.070309

training, so it is necessary to select high-quality, large-quantity and representative data sets.

Choice of CNN network structure: Different CNN network structures will affect the effect of music score recognition and conversion, so it is necessary to choose the appropriate network structure according to the actual situation.

Data enhancement and model optimization: Data enhancement and model optimization can effectively improve the recognition accuracy and robustness of the model and need reasonable operation.

Importance of post-processing: Post-processing operation has an important influence on the final score conversion result, which needs careful processing and verification.

Real-time consideration: the score conversion algorithm needs to be used in scenes with high real-time requirements, so it needs to consider the real-time performance of the algorithm.

Scalability of the algorithm: With the increase of data and the continuous advancement of algorithm optimization, the score conversion algorithm needs to maintain good scalability to meet the future needs.

In order to realize the music score conversion algorithm, we need to adopt a variety of technologies and tools, such as Python programming language, deep learning frameworks such as PyTorch, image processing libraries such as OpenCV, and commonly used computer vision algorithms and music analysis algorithms. By reasonably combining and applying these technologies and tools, an efficient and accurate music score conversion algorithm can be realized.

3. Experiment and result analysis

(1) Experimental setup

In the experiment and result analysis of music score conversion algorithm, the following experimental settings are needed:

Data set: Select the appropriate music score image data set, such as MUSCIMA+++data set.

Data preprocessing: preprocessing the score image, such as binarization, resizing and image enhancement.

Feature extraction: feature extraction is carried out on the processed music score image, such as using CNN network to extract convolution features.

Model training: using the score image data set after feature extraction for model training, such as using cross entropy loss function and Adam optimizer for training.

Model evaluation: Use the test set to evaluate the trained model, and calculate the accuracy, recall, F1-score and other indicators.

Analysis of experimental results: analyze the experimental results, compare the effects of different models and different parameter settings, and show them visually.

In the process of setting up the experiment, it is necessary to reasonably select the experimental parameters, such as CNN network structure, training rounds, learning rate, batch size, etc., in order to obtain better experimental results. At the same time, it is necessary to pay attention to the repeatability and comparability of the experiment, follow scientific experimental methods and norms, and ensure the reliability and credibility of the experimental results.

(2) Analysis of experimental results

When analyzing the experimental results of music score conversion algorithm, it is necessary to evaluate and compare the performance of the model on the test set in order to get the effect and advantages and disadvantages of the model. The specific steps are as follows:

Data set division: the original data set is divided into training set, verification set and test set, usually with a ratio of 8:1:1.

Model training: use the training set for model training, and train according to the preset parameters such as training rounds and learning rate.

Model evaluation: the trained model is evaluated by using the verification set, and the indexes such as accuracy, recall and F1-score are calculated, and the model is adjusted according to the evaluation results.

ISSN 2616-5775 Vol. 7, Issue 3: 63-67, DOI: 10.25236/AJCIS.2024.070309

Model test: Use test set to test the adjusted model, and calculate the accuracy, recall, F1-score and other indicators to evaluate the performance of the model.

In the process of analyzing the experimental results, it is necessary to pay attention to the credibility and reliability of the results to avoid data errors and interference [2]. At the same time, it needs to be evaluated in combination with practical application scenarios, and the actual effect and feasibility of the music score conversion algorithm should be considered. Finally, the experimental results should be scientific and practical, and provide reference for the optimization and improvement of music score conversion algorithm.

4. Problems and Prospects

(1) Problems existing in the research

When studying the traditional music score conversion algorithm based on image recognition technology, there may be the following problems:

Data set problem: the scale and quality of data set have great influence on the effect of the model. The existing music score data set may be too limited to cover all music types and styles, resulting in insufficient generalization ability of the model.

Feature extraction: The effect of music score conversion algorithm is closely related to the ability of feature extraction. How to extract meaningful music score features is one of the research difficulties. The existing music score feature extraction methods may be too simple or too complicated.

Model training problem: the score recognition model based on convolutional neural network needs to spend a lot of time and computing resources to train, and how to improve training efficiency and reduce resource consumption is a problem that needs to be solved.

Real-time problem: the existing music score conversion algorithm may have some delay, which is difficult to meet the real-time requirements. How to improve the conversion speed without reducing the accuracy of the algorithm is a problem to be solved.

(2) Improvement direction of the algorithm

Real-time optimization: at present, the traditional music score conversion algorithm based on image recognition technology usually takes a long time to calculate, which can not meet the real-time demand. In the future, we can explore the use of more efficient algorithms or technologies to improve the calculation speed of algorithms, such as lightweight convolutional neural networks and hardware acceleration.

Multi-modal fusion: At present, the traditional music score conversion algorithm based on image recognition technology only considers the image as the input source. In the future, more input sources can be integrated into the algorithm, such as video and audio, to improve the accuracy and efficiency of score recognition. At the same time, the performance of the algorithm can be further improved by combining various modal information, such as semantic information and musical instrument sound information.

The improvement direction of the algorithm can start with improving robustness, realizing real-time and multi-modal fusion, so as to further improve the performance and application value of the traditional music score conversion algorithm based on image recognition technology.

(3) Future research direction.

Image processing and recognition algorithm based on deep learning: The development of deep learning technology provides more powerful tools and methods for image processing and recognition. In the future, more advanced deep learning algorithms can be explored, such as Generative Confrontation Network (GAN) and Variational Self-Encoder (VAE) to improve the accuracy and robustness of the score conversion algorithm.

Music score recognition with multiple input sources: The current music score conversion algorithm based on image recognition only considers the image as the input source. In the future, more input sources can be integrated into the algorithm, such as video and audio, to further improve the accuracy and efficiency of music score recognition.

Recognition of non-traditional music score: At present, the music score conversion algorithm based

ISSN 2616-5775 Vol. 7, Issue 3: 63-67, DOI: 10.25236/AJCIS.2024.070309

on image recognition can only recognize traditional music score forms, such as staff and notation. In the future, we can explore the application of this algorithm to more non-traditional music forms, such as guitar music, drum music and so on.

Music style conversion: In addition to music score conversion, music style conversion algorithm based on image recognition technology can be explored in the future. This algorithm can automatically convert different styles of music into other styles of music, providing more tools and ideas for music production and creation.

The traditional music score conversion algorithm based on image recognition technology still has a lot of room for expansion and application prospects in the future research [3]. These research directions can bring more innovation and value to music production, education, entertainment and other fields.

(4) Future Prospect The traditional music score conversion algorithm based on image recognition technology may be improved and developed as follows:

Construction of large-scale data set: more music type and style score data are included in the data set to improve the generalization ability of the model.

Improvement of deep learning model: introduce more advanced deep learning model, such as attention mechanism, generating confrontation network, etc., to improve the accuracy and real-time performance of score conversion algorithm.

Combination of artificial intelligence and music: combine artificial intelligence technology with music creation, performance and other fields to explore more intelligent and innovative music score conversion algorithm.

Cross-disciplinary cooperation: strengthen cross-disciplinary cooperation with musicology, computer vision and other fields to achieve technological innovation and the expansion of application scenarios.

Real-time optimization: use more efficient algorithms and hardware resources, such as GPU and FPGA, to improve the real-time and response speed of music score conversion algorithm.

User experience optimization: On the premise of achieving high precision and high efficiency, pay more attention to the optimization of user experience, and improve the usability and usability of music score conversion algorithm.

Through the above improvement and development, the traditional music score conversion algorithm based on image recognition technology will have a wider application and better performance.

5. Conclusion

In this paper, the problem of traditional music score conversion based on image recognition technology is deeply studied, and a score conversion algorithm based on convolutional neural network is proposed, and each link of the algorithm is described and analyzed in detail. The experimental results show that the algorithm has achieved good results in music score image recognition and conversion.

In addition, this paper also discusses the existing problems of the algorithm and puts forward the future research direction. Aiming at the robustness and real-time performance of the algorithm, we explore the use of more robust and efficient algorithms or technologies for optimization. At the same time, the fusion of various input sources and modal information can be considered to further improve the performance and application value of the algorithm.

In a word, the research results of this paper have certain reference value for further improving the performance and application of traditional music score conversion algorithm based on image recognition technology. Future research can continue to explore the optimization direction of the algorithm, and verify and improve it with more practical application scenarios.

References

[1] Li Chun, Zheng Shuquan, Jin Tao. Research on the extraction method of Guqin score based on image processing [J]. Musicology Research, 2019, 37(3): 119-124.
[2] Gu Ping, Gong Jing, Yu Dongmei, et al. Image recognition algorithm based on improved HOG feature and SVM classifier [J]. computer applications and software, 2018, 35(8): 12-17.
[3] Guo Weihong, Yang Lin, Dong Hui, et al. Research on Chinese character recognition method based on deep learning [J]. Computer Science, 2018, 45(2): 225-230.