Digital Signal Processing and Artificial Intelligence for Automatic Learning

DLINE JOURNALS

Online ISSN: 2583-5009

DSP 2024; 3 (3)

https://doi.org/10.6025/dspaial/2024/3/3/110-116

A Case Study of Traditional Chinese Martial Arts Inheritance Based on Deep Learning

Qian Ding
Xi'an Medical College,710021
Xi'an, Shaanxi,China
dingqian1031@163.com

ABSTRACT

With the progress of the times, the rapid development of socio-economy, and changes in human ideology, traditional martial arts have become more prominent based on ancient philosophy and ethics. However, due to modern science and technology advancements, this art is facing unprecedented challenges. This paper aims to explore the development process of traditional Chinese martial arts and achieve this goal using deep learning methods. To achieve this, we have taken a series of measures, including searching, integrating, summarizing, and classifying, better to grasp the skills of various martial arts masters and accurately measure their inheritance status. Through this study, we conclude that by using deep learning, we can accurately identify the characteristics of various traditional Chinese martial arts and thus better evaluate their value. This achievement brings a new perspective and approach to protecting and disseminating traditional Chinese martial arts.

Received: 30 May 2024
Revised: 21 July 2024
Accepted: 1 August 2024
Copyright: with Author(s)

Keywords: Traditional Chinese Martial Arts, Deep Learning,Martial Arts Techniques, Martial Arts Inheritance

1. Introduction

Traditional Chinese martial arts have a long history and culture, containing rich national essence. It has always been regarded as a valuable asset and has played an important role. Its origin can be traced back to ancient agricultural society and has been further promoted due to external threats in modern times. With the advancement of modernization, traditional martial arts have evolved from simple exercise and entertainment to an important cultural activity [1]. This also indicates that with the evolution of Chinese history, its social value and function have undergone significant changes. Therefore, we should combine the current practical situation, delve into and explore its historical origins, and provide strong support for its future development. With the progress of the times, traditional Chinese martial arts are facing unprecedented challenges [2]. Therefore, to adapt to the development of modern society, traditional Chinese martial arts should recognize their future development trends [3] and strive to find multiple feasible development models to contribute to

the prosperity of Chinese culture. To make a new skill lastingly passed down, we must carefully consider the environment in which it exists. Traditional Chinese martial arts are an ancient sports project, a valuable treasure of the Chinese nation, and a precious intangible cultural heritage [4]. Therefore, we must carefully consider the environment in which it exists to enable it to achieve sustainable development on the international stage [5]. Ancient Chinese martial arts have a long history. It can be used to exercise muscles and cultivate wisdom. It represents not only ancient skills but also a profound cultural heritage... This paper focuses on exploring how to use deep learning to protect traditional Chinese martial arts [6]. We will sample, integrate, and analyze dynamic data from various martial arts schools and use this information to create a model to identify better and protect the essence of these schools' arts [7].

2. Related Work

In the Web 6.0 era context, there has been a qualitative leap in the development of information technology, with closed-loop systems represented by artificial intelligence gradually expanding and the full integration of the Internet and the Internet of Things [8]. Combining traditional martial arts with artificial intelligence through salvage mining and organization can enrich and make the dissemination of content more authentic. The core framework of the AI+ traditional martial arts communication ecosystem is built using artificial intelligence technologies such as expert systems, deep learning, intelligent design, and human-computer interaction [9]. For example, the LeNet-5 model, now a requirement for many AI beginners, is used to achieve the MNIST data recognition task [10]. Despite the significant progress in developing neural networks, they still fail to receive sufficient attention, possibly due to the gap between technological advancement and theoretical practice and the lack of rigorous mathematical theory support, resulting in the under-recognition of their development achievements. Since AlexNet defeated the second-place SVM in the ImageNet image classification competition in 2012, the importance of deep learning has gradually become prominent. To achieve effective deep learning, many well-known architectures, including Caffe, TensorFlow, Pytorch, Keras, and MXNet, have been continuously improved and perfected and widely applied in computer programming technology to meet changing demands. With the rapid development of Python technology, it has become a benchmark in many fields, especially in gaming, e-commerce, entertainment, etc. [11,12]. Its powerful parallel processing capabilities make it the preferred choice for many companies, and deep learning applications, such as cost-effective training and large-scale clusters, are receiving widespread attention. With the development of ResNet and DenseNet, people can use them as a solid foundation for building more complex network architectures [13,14]. This promotes the development of many new network architectures and provides great convenience for many new applications.

3. The Inheritance of Traditional Chinese Martial Arts Based on Deep Learning

3.1. Introduction to Neural Networks

With technological advances, the new generation of artificial intelligence has been applied in various aspects and widely recognized as a research milestone. Through research and exploration, it has been found that by utilizing the interaction and mechanisms of animal neurons, a new generation of artificial intelligence technology networks with high efficiency and accuracy can be created [15,16]. Due to technological progress, artificial intelligence technology has become a major means to help people change the data stored in human thinking, enabling better recognition, analysis, prediction, and decision-making. In the early stages, the new generation of artificial intelligence technology used a three-layer architecture, namely the input layer, hidden layer, and output layer, to transform the data stored in human thinking into something recognizable, analyzable, and predictable. In this section, we will explore a new approach to disseminating martial arts trends [17]. The basis of this approach is twelve convolutional layers, with 32 in the first and second layers and 64 in the rest. These layers have the same kernel size and stride and are continuously batch-processed for normalization. Based on the twelve convolutional layers, we merge two outputs and inject them into the next network model. By optimizing the two convolutional layers, the complexity and accuracy of the model are effectively controlled to obtain the final reconstruction result. Specifically, each layer is independent, as shown in Formula 1.

$$C_i(X) = W_i * X + B_i \quad (1)$$

where X is the input, i represents the *i-th* layer, id{1,2,…11}; C_i(5ØKÜ) denotes the convolution operation, and W and B represent the filters and biases, respectively; 5Ø5Ü5ØVÜ(5ØKÜ) indicates batch normalization, and the parameters to be learned are 5ØþÞ and 5Øýÿ; 5ØZÜ5ØNÜ5ØeÜ(0,) denotes the *ReLU* activation function. Originally, this study considered using MPJPE (Mean Per Joint Position Error) as the evaluation metric for the experiments. This metric is commonly used in 3D pose estimation tasks, where the goal is to estimate the 3D pose (*x, y, z* coordinates) of each joint from RGBD images. MPJPE calculates the Euclidean distance between the ground truth and the predictions and then takes the average of the position errors for all joints. The new task proposed in this study is also related to 3D pose estimation and requires measuring the error between the generated results and the ground truth. Therefore, evaluation metrics from 3D pose estimation tasks can be explored. However, considering the temporal dimension of the generated results in this task, a time dimension is added to the calculation based on MPJPE, as shown in formula 2:

$$tMPJPE = \frac{d(w, v)}{k \times t} \quad (2)$$

where 5ØdÜ is the output, 5ØcÜ is the target, J is the sample dimension, 5ØQÜ(5ØdÜ,5ØcÜ) represents the Euclidean distance, 5ØXÜ is the number of key points, and t is the time parameter. The *tMPJPE* metric in this section can be understood as the MPJPE per unit time, facilitating the comparison of results over the same time duration. It should be noted that the evaluation metric proposed in this section focuses on the accuracy of joint positions and is only applicable to short time intervals. The gap between different samples and the generated results can be accurately evaluated using mean squared error to determine if the experimental results meet the expected outcomes.

3.2. Convolutional Neural Networks

LeCun Y and other sub-investigators made a breakthrough by applying convolutional neural networks to more complex scenarios [18]. By adjusting the network structure and parameters, they achieved globalized neural network training from input to output, greatly improving the effectiveness of deep learning. By dividing complex hierarchical information, such as convolutional neural networks, pooling layers, or fully connected layers, into different layers, the differences between these layers can be reflected through comparison. Although these differences can be represented through comparison, for example, in the case of perceptrons, their hidden layers are represented through comparison, while in the case of convolutional neural networks, their results are not entirely the same. Through convolutional operations, information can be collected and processed from multiple receptive fields to obtain more valuable information. This operation may need to be repeated numerous times and require various convolution kernels to capture and process the diversity of information better. The parameters and values can be calculated using mathematical methods by processing the receptive field of the convolution kernel. By inputting the extracted short video data into the human pose recognition network, the body key points of the martial arts instructor can be accurately recognized and recorded. Furthermore, we need to collect coordinate information from the original images to better estimate the human pose. To ensure the reliability of the data, the obtained human skeleton data needs to be processed. Firstly, all keypoint data is reduced to 1280*720 centered around the head joint, and the coordinate points are standardized and normalized.

4. Experimental Design and Analysis

4.1 Experimental Design

In the history of computer vision technology, especially in image processing, evaluating the performance of a neural network often requires considering its feature extraction capabilities. Therefore, building a model with good feature extraction performance and effectively utilizing this information is crucial. The development of deep learning technology has become an important tool that helps us better predict and recognize complex physical phenomena. For example, DeepPose, OpenPose, Hourglass, and HRNet are all typical representatives that use deep learning to predict physical phenomena. Researchers are actively exploring new techniques, such as expanding the

boundaries of models, combining different models, and leveraging the characteristics of different models to create more complex human pose models. However, the success of these techniques depends on their effective application in practical scenarios to achieve optimal model accuracy.

After years of in-depth exploration, we ultimately decided to use Simple Baselines as the base network and adopt it as the new feature extraction unit, replacing ResNet, to better evaluate recent human pose estimation algorithms. Based on the foundation of Involution, we constructed an IPN (InvolutionPose Estimation Net) for human identity recognition, as shown in Figure 1. *InNet* is a method based on convolutional neural networks that extracts 17 key information. In this method, we first expand the convolutional layers, then process the convolutional layers, then process three deconvolution layers, and finally, apply L2Loss processing. By using *InNet* as the base grid, the complexity of parameters can be greatly reduced, and the advantages of Involution can significantly improve accuracy.

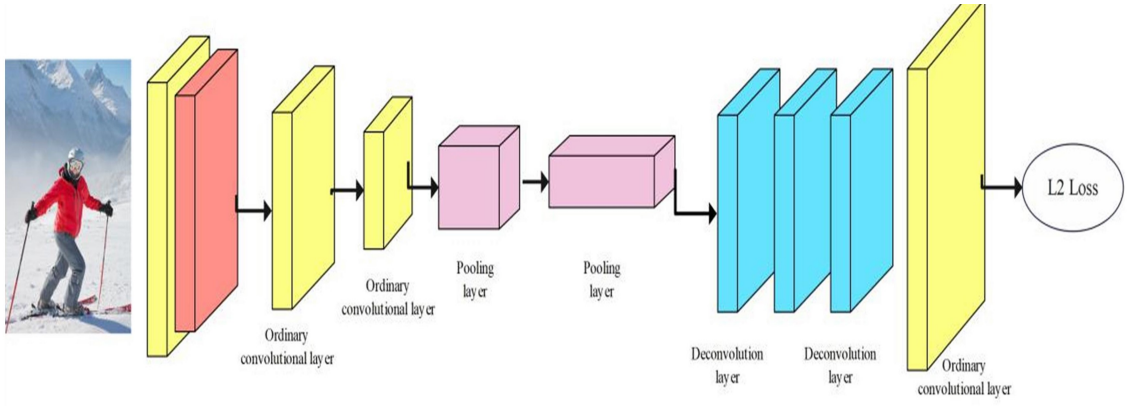

Figure 1. IPN Network Architecture

4.2. Experimental Testing and Result Analysis

To better recognize the martial arts movements in "少年连环拳" (Youth Serial Boxing), we need to use two of the most common and authoritative datasets, namely NTU-RGB+D and UTD-MHAD. Currently, these datasets have not been widely utilized, so we need to obtain the information from them in an automated way. The NTU-RGB+D dataset is a large-scale publicly available dataset worldwide, containing over 60 action categories with more than 56,880 samples. Among these, 40 are daily activities such as drinking water, sitting down, standing, etc., 9 are health-related actions like dancing, waving, etc., and 11 are interactive actions like handshaking, high-fiving, etc. 40 subjects of different age groups performed the experiments, and their behavior data were collected using Kinect v2 sensors. In this experiment, participants needed to perform two actions captured by three cameras from -45 degrees, 0 degrees, and 45 degrees angles. The data collected by these cameras can be categorized into four types: depth, 3D skeletal, RGB frames, and infrared sequences. The RGB frames have a resolution of up to 1920*1080, and the 3D skeletal frames have 25 body joints, all visible to the participants. This computer has rich observational abilities, capturing various dynamic changes. These two datasets cover interactive actions in daily life, but these movements are typically unique in martial arts. Most of these datasets are also from foreign countries, so further research is needed. Although there is currently no specialized action database for a specific martial art in China, it is still necessary to establish a dedicated martial arts action database.

Based on the data in Figure 2, Model B1 shows excellent training performance, with a much higher convergence rate than B2, but it remains relatively low. However, B1's performance is still outstanding, especially in the 1s and 3s experiments, where it performs even better. B2 relies on residual blocks, while B0 uses only one channel. We found that this model demonstrates remarkable efficiency through systematic training and experiments. After 200 epochs of training, while some network parameters have changed, other parameters remain unchanged, and the final training results are shown in Figure 3.

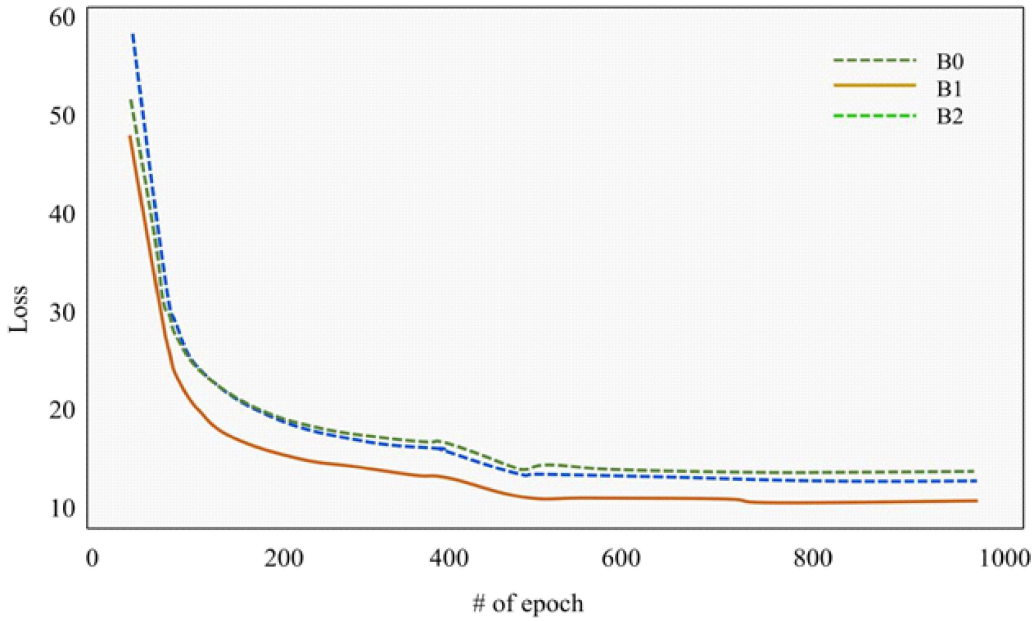

Figure 2. Model B1 Data

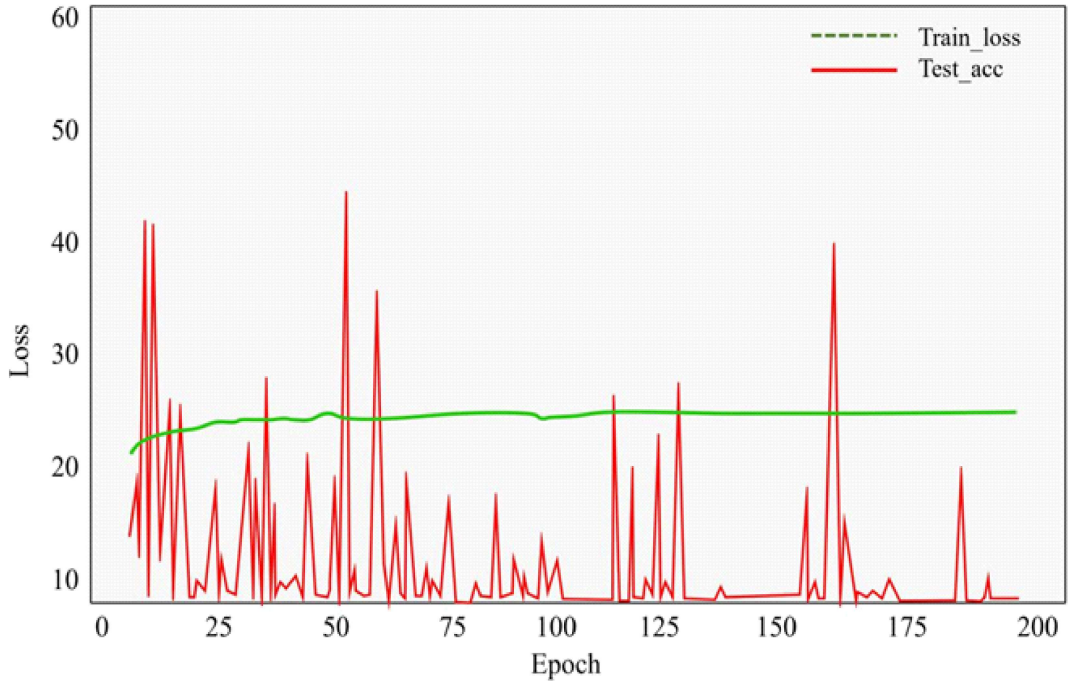

Figure 3. Training Results

According to Figure 3, through training, we have transformed the received information into recognizable martial arts actions. We used the videos from multiple schools, and each performance had different martial arts movements. By observing each performance, we could accurately identify which actions the audience saw and their positions. Through in-depth research, we found that the proposed martial arts action recognition method, which considers both human body postures and corresponding key information, achieves high accuracy and holds significant academic value.

5. Conclusion

In general, deep learning has provided new ideas and methods for inheriting and developing traditional Chinese martial arts. Conventional martial arts inheritance often relies on communication and teaching between masters and disciples, but this approach has become increasingly challenging with societal changes and population mobility. With the advent of the intelligent era, deep learning has become the darling of the times. It plays an indispensable role in the field of artificial intelligence and accelerates the development of AI. Deep learning enables machines to mimic human beings in analyzing things, especially in speech and image processing. Many research results based on deep learning have already been applied in various aspects of daily life, such as semantic analysis, face recognition, etc. Its performance in digital recognition is also excellent. The emergence of deep learning models offers new possibilities for martial arts inheritance. Through our research, we have discovered that the proposed martial arts action recognition method, which considers both human body postures and corresponding key information, achieves high accuracy and holds significant academic value. In the future, we can apply deep learning models to a wider range of martial arts fields to promote traditional Chinese martial arts inheritance.

References

[1] Chude, C. I and Schaaf, S. (2021). A unique cause of sciatica in a master's martial arts athlete: 1264. *Medicine & Science in Sports & Exercise, 53*(8S), 416–416.

[2] Selitrenikova, T., Ageev, E and Kolokoltsev, M. et al. (2022). Transcranial electrical stimulation to increase psychophysiological stability, technical and tactical readiness of MMA fighters. *Journal of Physical Education and Sport, 22*(6), 1419–1425.

[3] Monterrosa Quintero, A., De La Rosa, A., Arc Chagnaud, C., et al. (2023). Morphology, lower limbs performance and baropodometric characteristics of elite Brazilian Jiu-Jitsu athletes. *Ido Movement for Culture: Journal of Martial Arts Anthropology, 23*(2), 58–69.

[4] Starke, S., Zhao, Y., Zinno, F., et al. (2021). Neural animation layering for synthesizing martial arts movements. *ACM Transactions on Graphics, 40*(4), 1–16.

[5] Liu, M and Zhang, J. (2022). Gesture estimation for 3D martial arts based on neural network. *Displays: Technology and Applications, 72*, 72.

[6] Envelope, V. L. A. A., Envelope, Y. F. F. B. P., Envelope, P. B. A. C., et al. (2022). Validity and reliability of a specific anaerobic test for mixed martial arts. *Science & Sports, 37*(5–6).

[7] Zhao, C., Yang, H and Li, X. et al. (2021). Analysis and application of martial arts video image based on fuzzy clustering algorithm. *Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology, 40*(4).

[8] Peng, X. B., Abbeel, P., Levine, S., et al. (2018). Deep Mimic: Example-guided deep reinforcement learning of physics-based character skills. *ACM Transactions on Graphics, 37*(4CD), 143.1–143.

[9] Theeboom, M., Dong, Z and Vertonghen, J. (2012). Traditional Asian martial arts and youth: Experiences of young Chinese wushu athletes. *Archives of Budo, 8*(1), 27–35.

[10] Mawardi, S and Firdaus, F. (2019). The character values in Minangkabau traditional martial arts. *International Journal of Scientific & Technology Research, 8*(10), 846–850.

[11] Jain, A., Awan, A. A., Subramoni, H., et al. (2019). Scaling TensorFlow, PyTorch, and MXNet using MVAPICH2 for high-performance deep learning on Frontera. In *2019 IEEE/ACM Third Workshop on Deep Learning on Supercomputers (DLS)* (pp. 76–83). IEEE.

[12] Tabani, H., Pujol, R., Abella, J., et al. (2020). A cross-layer review of deep learning frameworks to ease their optimization and reuse. In *2020 IEEE 23rd International Symposium on*

Real-TimeDistributed Computing (ISORC) (pp. 144–145). IEEE.

[13] Zhang, J., Zhang, Y and Lu, Y. (2021). A comparative study of deep learning frameworks based on short-term power load forecasting experiments. *Journal of Physics: Conference Series, 2005*(1), 012070.

[14] Haixia, Y., Chuandong, X., Wenbo, F., et al. (2022). Intelligent mask recognition and voice dialogue system based on Raspberry Pi control. In *2022 IEEE 2nd International Conference on Data Science and Computer Application (ICDSCA)* (pp. 103–105). IEEE.

[15] Barba-Guaman, L., Eugenio Naranjo, J and Ortiz, A. (2020). Deep learning framework for vehicle and pedestrian detection in rural roads on an embedded GPU. *Electronics, 9*(4), 589.

[16] Demidovskij, A., Tugaryov, A., Suvorov, A., et al. (2020). OpenVINO deep learning workbench: A platform for model optimization, analysis, and deployment. In *2020 IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI)* (pp. 661–668). IEEE.

[17] Long, G and Chen, T. (2022). On reporting performance and accuracy bugs for deep learning frameworks: An exploratory study from GitHub. In *Proceedings of the 26th International Conference on Evaluation and Assessment in Software Engineering* (p. 90–99).

[18] LeCun, Y., Bottou, L., Bengio, Y., et al. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE, 86*(11), 2278–2324.