Optimization of DL Technology for Auxiliary Painting System Construction Based on FST Algorithm

Pengpeng Xu*, Guo Chen
Engineering Training Center, Shandong Huayu University of Technology, Ezhou, 253000, China

***Abstract*—The continuous development of computers has brought about the emergence of many image processing software, but these software have relatively limited functions and cannot learn and create works according to the prescribed style. To make it easier for ordinary people to create artistic style paintings, this study proposes the construction of an auxiliary painting system based on finite state transducer algorithm-optimized deep learning technology. The results demonstrated that when there were 12 images, the accuracy of the optimized convolutional neural network model in extracting image features increased by 1.1% compared to before optimization. When the number of images was 1, the optimized model reduced the image feature extraction time by 15.1s compared to before optimization. Compared with other algorithms, the accuracy of extracting image style information based on a convolutional neural network was the highest at 80% under different iteration times. The research algorithm has improved the accuracy and time of extracting image style information.**

Keywords*—Finite state transducer; deep learning; CNN; auxiliary painting; style transfer

I. INTRODUCTION

As a visual art, painting can stimulate people's imagination and creativity. Through the observation and practice of painting, people can cultivate their creativity, form a unique way of thinking and problem-solving ability [1]. In recent years, with the rapid development of industrial automation and intelligent technology, spraying, as a key link in the field of industrial manufacturing, has gradually received widespread attention. In the field of modern industrial manufacturing, spraying technology as a key surface treatment process, its accuracy and efficiency are directly related to the final quality of products and market competitiveness [2-3]. However, the traditional spraying system has many problems in the construction process, such as low spraying accuracy, serious material waste and high manual operation cost. In recent years, the development of intelligent spraying robots and automated spraying systems has provided a new direction to solve these problems. In the automatic development of spraying system, Fuzzy SynST Technology (FST) can effectively optimize the spraying process through state conversion, avoid repetitive programming, and reduce human error [4-5]. And automatically diagnose and adjust faults such as nozzle blockage to improve construction efficiency. In the spraying system, FST is prone to lack of control accuracy due to complex environment, and it is difficult to cope with changeable construction scenarios [6]. Deep learning (DL) optimizes spraying paths and parameters through big data analysis and adaptive learning to improve construction efficiency and quality. At the same time, DL can adjust the spraying strategy in real time to better respond to environmental changes and reduce human intervention [7]. In view of this situation, in order to enable most ordinary people to integrate the style of painting art into their works, this study proposed a deep learning technology based on FST algorithm optimization, and innovatively constructed an auxiliary spraying system (APS) based on FST and DL, aiming to improve the intelligence level of the auxiliary spraying system. This study innovatively optimized the Convolutional Neural Network (CNN) model based on the FST, using FST as the output layer of the model and constructing APS based on the TensorFlow framework. The main contribution is the proposal of APS construction based on FST-optimized DL technology, which has significant implications for the successful transfer of painting styles in images.

The research structure has six sections. Section II is a review of the current research status of DL and FST algorithms. Section III is a study on APS based on FST-optimized DL technology. Section IV is the experimental verification of the proposed research method. Discussion is given in Section V. Section VI is a summary of the research content.

II. RELATED WORKS

DL is a ML method based on artificial neural networks that can perform nonlinear transformations and feature extraction on complex data [8]. Yu K et al. proposed a DL-based auxiliary diagnosis scheme for breast cancer to solve the problem of low diagnostic efficiency of breast cancer due to the lack of high-quality medical resources in remote areas. This scheme was based on DL model for transfer learning to gain a diagnostic model. This method improved the diagnostic efficiency of breast cancer to 98.19%, and could be used for auxiliary diagnosis in hospitals in remote areas [9]. Zhou X et al. proposed a smart automatic tagging scheme based on deep Q-network to improve the accuracy of Human Activity Recognition (HAR) in healthcare Internet of Things. This scheme identified fine-grained patterns by extracting advanced features from sequential motion data. This method better solved the problem of insufficient sample labeling [10]. Wan S et al. proposed a real-time HAR method based on DL model to solve the problem that traditional methods cannot recognize complex and real-time human activities. This method utilized CNN for local feature extraction and preprocessed the data through denoising, normalization, and other methods. This method was superior to other traditional methods and improved the accuracy of HAR [11]. Chohan M et al. proposed a DL-based plant disease detection model to promptly detect various diseases caused by bacteria, fungi, and viruses in plants. This model could use images of plant leaves to detect plant diseases, improving the accuracy of image detection of diseased plants, with an accuracy rate of up to 98.3% [12]. Chowdhury M E H et al. proposed a DL architecture based on EfficientNet to classify tomato diseases to lower down the adverse influences of diseases on plants and weaken the drawbacks of continuous human monitoring. This method could improve the segmentation accuracy of leaf images and has shown excellent performance in the ten class classification of images, with an accuracy of 99.89% [13].

Compared with other data structures, FST has higher spatial efficiency and query performance when processing large-scale string collections. Abro WA et al. proposed a model combining weighted FST and BERT architecture to reduce the cost and time of collecting high-quality labeled data. This model utilized weighted FST to enhance the fine-tuning of BERT-like architecture. Compared with other models, this model had a higher recall rate and F1 score, which can reduce the need for massive supervised data [14]. Mohammadghuliha M et al. constructed a frequency controllable acoustic sensor for guided wave detection in order to simplify the hardware of phased array systems and reduce the cost of guided wave based systems. FSAT was made by patterning spiral electrodes on a piezoelectric plate. The generation of directional guided waves in the main structure of the converter has been significantly improved, reducing the cost of the system [15]. Dolatian et al. designed a type of finite state machine with two deterministic FSTs to address the issue of limited state processing not fully capturing the productivity of unbounded replication. This sensor was easy to design and debug in practice, and had linguistic motivation in the origin semantics of the transducer. This sensor could capture almost all processes, which can be found in online repetitive databases [16]. Martin K et al. proposed a bidirectional motion device based on FST to eliminate the descriptive costs of bidirectional motion, non-determinism, and scanning, especially the cost of converting to deterministic or non-deterministic finite automata. This method was more powerful in bidirectional motion than unidirectional motion under deterministic conditions [17]. Somerset W. E. et al. proposed a novel bending ultrasonic transducer for high-pressure environments to handle the issue of pressure imbalance caused by the internal air cavity of traditional transducers on the vibrating membrane of the transducer. The internal chamber of the transducer was filled with an incompressible fluid in the form of non-volatile oil. This method could achieve stable ultrasound measurement [18].

In summary, literature in [8] and [9] have made some progress in the field of medical image recognition, but they are still limited in real-time adaptability and accuracy in dealing with complex industrial scenes. Literature in [10] and [11] improve the recognition accuracy through human activity recognition, but it is difficult to cope with the high environmental variability in the spraying system. The plant disease recognition model proposed in literatures in [12] and [13] performs better under high noise data, but it is insufficient in extracting complex spray features. Literature in [14] combined FST and BERT to improve the recall rate, but the acquisition cost of large-scale data is high. Literatures [15] to [18] has optimized the sensor design, but it still needs to be improved in the stability and diversity of industrial spraying. DL reduces manual intervention through automatic feature extraction of neural networks and is suitable for diverse data scenarios. FST has advantages in optimizing state transition and processing big data, and can effectively improve the spatial efficiency of the algorithm. Therefore, this paper proposes the research of APS construction based on FST optimization DL technology.

III. APS Construction based on FST-Optimized DL Technology

A. *Optimizing DL Technology based on FST Algorithm*

A painting mainly consists of two parts, namely the content and style of the painting. Among them, painting styles have diversity, and each painter has their own unique style, expressing emotions through elements such as color and lines. Even if the content of the painting is the same, the effects displayed by different painting styles are also different [19]. Compared with traditional ML methods, the most significant advantage of DL is its ability to automatically extract and learn meaningful feature hierarchies from raw data. This hierarchical structure simulates the hierarchical organization of the human brain's neural network, where each layer can capture features of different complexity and abstraction levels [20-21]. CNN is a DL model that can automatically learn the features of images without the need for manual design or selection of feature extractors. CNN uses convolutional and pooling layers to extract local features of images and perform dimensionality reduction, thereby reducing the number of parameters and computational complexity. CNN has advantages such as high efficiency in image recognition [22-23]. Therefore, to integrate the painting styles of famous masters into daily painting works, this study extracts the style information features of images based on CNN. The structure of CNN mainly consists of convolutional layers, activation functions, pooling layers, fully connected layers, output layers, etc. The process of extracting image style features using CNN is shown in Fig. 1.

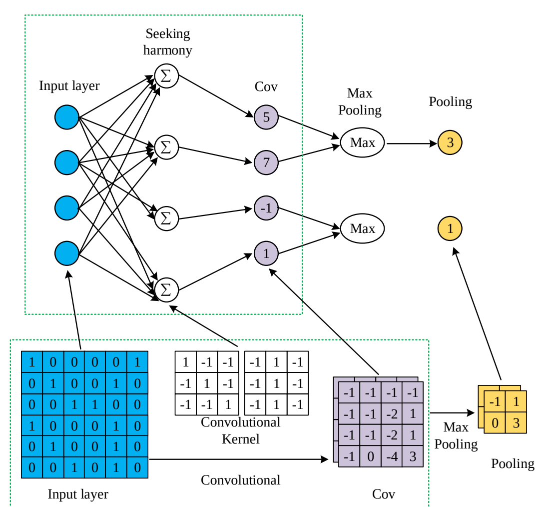


Fig. 1. Flow chart of CNN extracting image style features.

Among them, the input layer receives raw image data. The convolutional layer convolves the input image with the convolutional kernel, introducing nonlinearity by applying activation functions, enabling the network to learn complex features. The pooling layer reduces computational complexity by reducing the feature map size, by selecting the maximum or average value within the pooling window. The fully connected layer transforms the extracted feature maps into the final output. The first step in extracting the style information features of an image is to perform convolution operations, where attention should be paid to the position step size of each slide. The basic idea of the Adam algorithm is to keep an adaptive Learning Rate (LR) for each model parameter during the training process, so that parameter updates can more accurately control the step size [24]. Therefore, this study uses the Adam parameter update method combined with the First-Order Moment Estimation (1OME) and Second-Order Moment Estimation (2OME) of the loss function for calculation. The Adam parameter update steps are shown in Fig. 2.

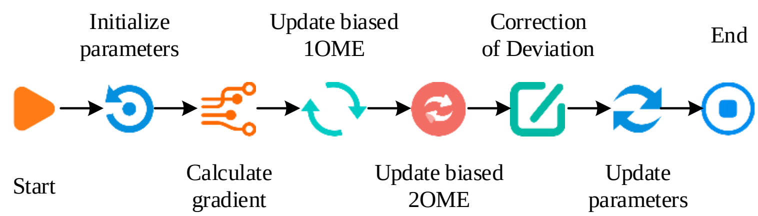


Fig. 2. Adam parameter update steps.

Firstly, the 1OME m_t is corrected using the correction formula shown in Eq. (1).

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t} \tag{1}$$

In Eq. (1), β_1 is the 1OME attenuation coefficient. $\hat{m}_t$ is the corrected 1OME. t is the number of updated steps. Secondly, the 2OME v_t is corrected, and its correction expression is Eq. (2).

$$\hat{v}_t = \frac{v_t}{1-\beta_2^t} \tag{2}$$

In Eq. (2), β_2 is the 2OME attenuation coefficient. $\hat{v}_t$ is the corrected 2OME. Finally, the model parameters were updated based on the corrected $\hat{m}_t$ and $\hat{v}_t$. The formula for updating model parameters θ_{t+1} is Eq. (3).

$$\theta_{t+1} = \theta_t - \frac{\ell \hat{m}_t}{\sqrt{\hat{v}_t} + \varepsilon} \tag{3}$$

In Eq. (3), ℓ is the LR. θ_t is the parameter model before the update. ε is a hyperparameter. When performing convolution processing on each convolutional channel, different filters are used, and the calculation process is Eq. (4).

$$c_i = F\left(K_i * X + b_i\right) \tag{4}$$

In Eq. (4), K_i and b_i are the convolution kernels and biases of the i-th channel. X is the input data. c_i is the convolution result of the i-th channel. $F(\cdot)$ is a non-linear activation operation in the convolution process. The activation function of CNN is ReLU. ReLU trains in the negative area. When a neuron with a gradient of 0 appears, the gradient of this neuron and subsequent neurons will always be 0 and will no longer respond to any data, causing the correlated parameters never being updated [25]. To solve the problem of neuronal necrosis in ReLU function, this study improves ReLU using exponential linear units, and the improvement process is Eq. (5).

$$ELU(x) = \begin{cases} x, & x>0 \\ \alpha\cdot(\exp(x)-1), & x\le 0 \end{cases} \tag{5}$$

In Eq. (5), x is the input value, α is the predefined hyperparameter, and $\exp(\cdot)$ is the exponential function operation. The composition of CNN includes a max pooling layer, which discards non max information and may lose some detail information. Average Pooling (AveP) considers the mean value of all values in the pooling window, which is smoother than Maximum Pooling (MaxP) and can preserve more detailed information. The calculation of mixed pooling and AveP is shown in Fig. 3.

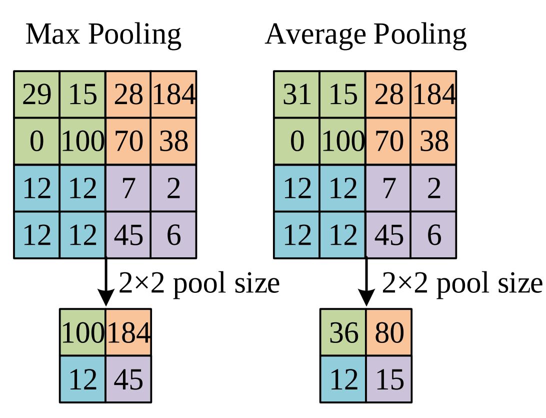


Fig. 3. Mixed pooling and AveP calculation.

The calculation of AveP G_{ij} is Eq. (6).

$$G_{ij} = \frac{1}{p\times p}\cdot\left(\sum_{i=1}^{p}\sum_{j=1}^{p} M_{ij}\right) + b \tag{6}$$

In Eq. (6), M denotes the feature map of the pooling operation, b is the bias of each channel, and $p\times p$ is the size of the pooling region. To solve the problem of feature weakening caused by using only the max pooling layer for feature processing, this study adopts a mixed pooling method to pool the features. The calculation of mixed pooling is Eq. (7).

$$G_{ij} = \frac{1}{2}\left(\frac{1}{p \times p} \cdot \left(\sum_{i=1}^{p}\sum_{j=1}^{p} M_{ij}\right) + \max_{(p\times p)i=1,\, j=1} M_{ij}\right) + b \qquad (7)$$

In Eq. (7), $\max_{(p\times p)i=1,\, j=1} M_{ij} + b$ represents the MaxP operation. To transform the image style information features extracted by CNN into a series of string states containing semantic information, this study optimizes DL technology based on FST and uses FST as the output layer of the model. The formula for the time complexity O of FST is Eq. (8).

$$O = n \times m \qquad (8)$$

In Eq. (8), n is the text length and m is the pattern length. The optimization steps are as follows: Firstly, the number of states in FST and the transition relationship between states are determined. Secondly, in the output layer of the CNN model, a node layer with an equal number of FST states is created. Finally, the image feature information extracted based on the FST algorithm is transformed into a string.

B. Construction of APS

After optimizing DL technology based on FST, this study constructs APS based on the optimized model. TensorFlow is an open-source ML framework dedicated to automatic differentiation of various data flow graphs and computation of deep neural networks. This framework provides multiple advanced Application Programming Interfaces (APIs), making building and training DL models relatively simple. At the same time, it also maintains great depth and flexibility to meet the needs of researchers exploring complex models. Therefore, this study constructs APS based on the TensorFlow framework. The APS constructed mainly consists of two parts. Part 1 is the front-end page: The main functions of the system's front-end include uploading images that require style fusion, selecting the desired style, style conversion buttons, and outputting images after style conversion is completed. The process of requesting front-end image style conversion is shown in Fig. 4.

The other part is the backend of the system. The backend of APS mainly has three modules: Uniform Resource Locator (URL) request forwarding, logical functionality, and style conversion. The task of Module 1 is to pass the URL request link transmitted by the frontend to a specific function for execution. The task of Module 2 is to preprocess, transcode, encode, and store uploaded images. The task of Module 3 is to convert the style of uploaded images and return the converted images to the user. The relationship between the three modules in the APS backend is shown in Fig. 5.

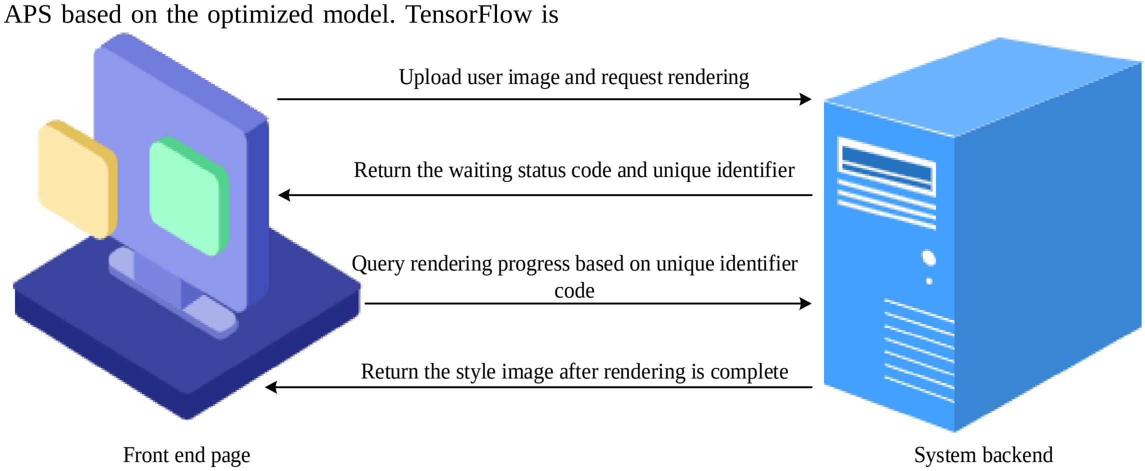


Fig. 4. Process of front-end image style conversion request.

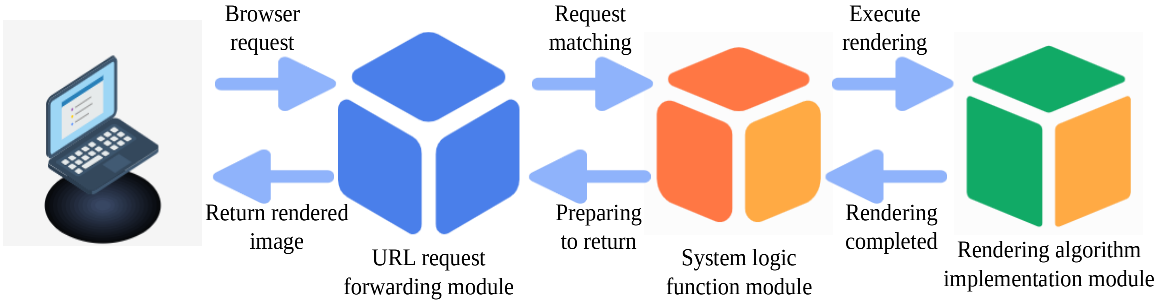


Fig. 5. The relationship between the three modules in the background of the APS.

The core module of the system backend is to extract image feature information. Due to various noises, interferences, and deformations that may exist in the original image, directly using the original image for DL model training may lead to a decrease in model accuracy. As a result, this study requires preprocessing of the original images to enhance the training effectiveness and recognition accuracy. The image data preprocessing steps are shown in Fig. 6.

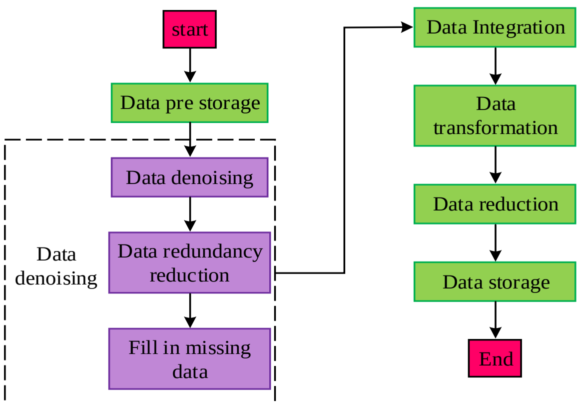


Fig. 6. Image data preprocessing steps.

Standardizing image data can adjust images of different sizes and resolutions to the same range, facilitating subsequent processing and analysis. Data standardization processing can be divided into normalization processing and z-score processing. The difference in comparison results of different data sets is due to the diversity of the characteristics of data sets and the number of samples. The performance of the algorithm on different data sets depends on the characteristics of the data, such as the resolution of the image, the complexity of the background, and the clarity of the target style. Normalizing the data can balance the weights of various dimensional features and avoid the interference of features with large or small numerical scales on the model. Among them, normalization processing can be divided into minimax normalization and mean normalization. The calculation of minimax normalization is Eq. (9).

$$\begin{cases} X = \dfrac{X_{old} - \min\left(X_{old}\right)}{\max\left(X_{old}\right) - \min\left(X_{old}\right)}, [0,1] \\ X = \dfrac{2\left(X_{old} - \min\left(X_{old}\right)\right)}{\max\left(X_{old}\right) - \min\left(X_{old}\right)} - 1, [-1,1] \end{cases} \tag{9}$$

In Eq. (9), X_{old} is the original dataset. $\max\left(X_{old}\right)$ and $\min\left(X_{old}\right)$ take extreme values for each column feature of the original dataset. Mean normalization maps the values of features to the range of [0,1], eliminating the influence of dimensionality on the eventual result, making different features comparable, and allowing features with potentially large distribution differences to have the same weight impact. The expression of mean normalization is Eq. (10).

$$X = \frac{X_{old} - mean\left(X_{old}\right)}{\max\left(X_{old}\right) - \min\left(X_{old}\right)} \tag{10}$$

In Eq. (10), $mean\left(X_{old}\right)$ defines taking the mean of each column feature of the original dataset. Standardizing data using z-score involves scaling the data to a distribution centered around 0 with a standard deviation of 1. This method preserves the original data information without changing the distribution type of the original data, making the different features of the original data comparable. The formula for z-score is shown in Eq. (11).

$$\mu \ X = \frac{X_{old} - \mu}{\sigma} \tag{11}$$

In Eq. (11), μ is the vector of the mean of each column feature of the original dataset. σ is the vector of the standard deviation of each column feature in the original dataset. To better evaluate the detection model's performance, this study utilizes the $F-Score$ method for evaluation, and the formula for $F-Score$ is Eq. (12).

$$F-Score = \left(1+\beta^2\right)\frac{Precision \cdot Recall}{\beta^2 \cdot Precision + Recall} \tag{12}$$

In Eq. (12), $Recall$ represents the recall rate of abnormal data in the results. $Presion$ is the accuracy of abnormal data in the result. β is the degree of importance to the outcome. When β =1, it indicates that both have the same impact on the result. At this point, $F-Score$ is expressed as Eq. (13).

$$F-Score = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \tag{13}$$

IV. Analysis of APS Effect based on FST Algorithm Optimization of DL Technology Construction

A. *Experimental Parameter Setting and Performance Verification*

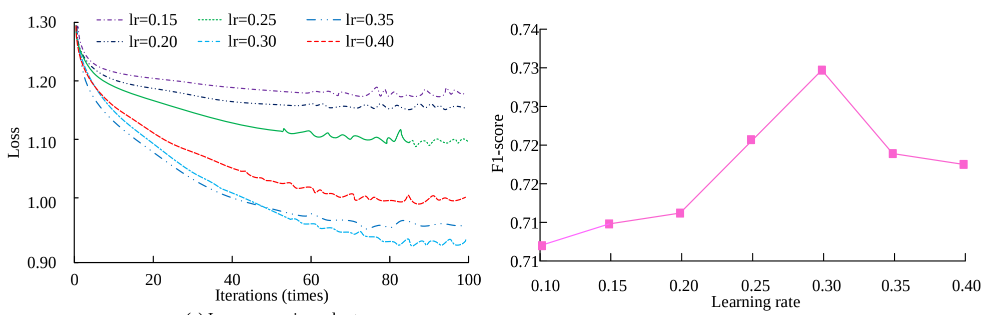
The software selected for the experiment is Python v3.5.2, and the learning framework is TensorFlow. The operating system is Windows 10, the memory is 128GB, the CPU is Intel Core (TM) i7-6700CPU @3.40GHz, and the GPU is NVIDIA GeForce GTX 750 Ti. Table I provides information on the backend functions of the system.

The size of the LR directly affects the convergence speed and state of the CNN model. Excessive LR may cause the model to oscillate near the optima and fail to converge. If the LR is too low, it may lead to slow convergence speed of the model, and even get stuck in local optimal solutions. Therefore, to obtain the best LR, this study set the LR to different values and evaluated it through the loss value and F1 Score. The statistical results are displayed in Fig. 7. In Fig. 7(a), when iterating up to 70 times, the loss curves for different LRs tend to stabilize. When the LR = 0.3, the loss value of the model is minimized. When the LR < 0.3, the loss value decreases as the LR increases. When the LR > 0.3, the loss value grows with the growth of the LR. In Fig. 7(b), when the LR is 0.3, the F1 Score value is optimal, and when the LR is < 0.3, the F1 Score value rises with the sire of the LR. When the LR is > 0.3, the F1 Score value decreases as the LR increases.

The maximum tree depth parameter (max_depth) controls the complexity of the decision tree. However, if this value is set too small, the CNN model may underfit, resulting in poor performance. When the max_depth is set too high, it may cause overfitting issues. Therefore, to obtain the optimal x_depth, this study sets the x_depth to different values and evaluates it based on the loss value and F1 Score. The statistical results of the loss value and F1 Score are shown in Fig. 8. In Fig. 8(a), when the iteration reaches 48 times, the loss curves of different max_depth tend to be stable. When the max_depth=3, the loss value of the model is the smallest; When the max_depth<3, the loss value descends with the rise of the max_depth; When the max_depth>0.3, the loss value grows with the increase of LR. In 8 (b), when the max_depth=3, the F1 Score value is optimal; When the max_depth<3, the F1 Score value grows with the increase of the max_depth; When the max_depth>3, the F1 Score value decreases with the rise of the max_depth.

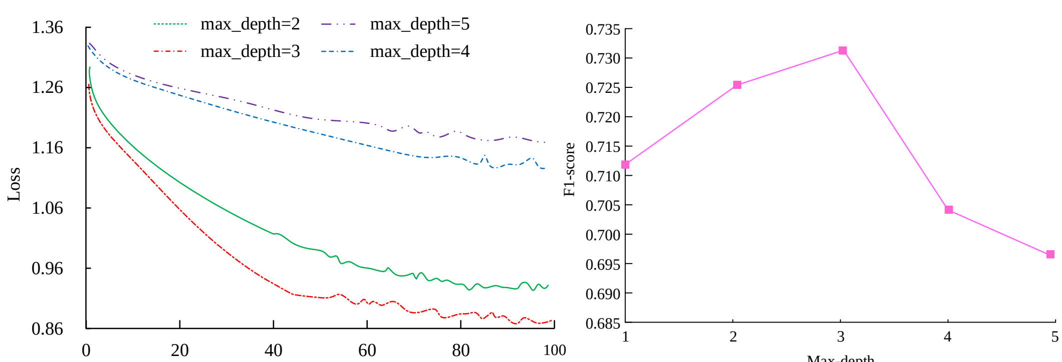
TABLE I. SYSTEM BACKEND FUNCTION SETTINGS

Serial number	Function Name	Module to which it belongs	Function Introduction
1	path	URL request forwarding module	Filter and match the links sent by the front-end
2	open/write	System logic function module	Provide read and write functionality for images
3	json	System logic function module	Responsible for processing JSON related formatted data
4	base64	System logic function module	Base64 encoding and decoding of images
5	get_unique_key	System logic function module	Generate a unique key for each user image
6	apply_async	System logic function module	Asynchronous execution of rendering process
7	settings	System logic function module	Access and modify some system settings options
8	check_img	System logic function module	Check image format
9	pool_exec	System logic function module	Provide scheduling and execution of process pools
10	style_rendering	Style conversion implementation module	Render the received image into the specified style



(a) Loss comparison chart
(b) F1-Score Comparison Chart

Fig. 7. Comparison of loss values and F1 Score at different LRs.



(a) Loss comparison chart
(b) F1-Score Comparison Chart

Fig. 8. Comparison between loss values and F1 Score under different max_depth.

B. APS Quality Analysis based on CNN

This study compares the recall rate and F1 Score value of image features pooled using a Maximum-Median-Mixed Pooling (MMMP) strategy with other pooling approaches. The statistical results are shown in Table II. Single Pooling Strategies (SPS) are lower than Mixed Pooling Strategies (MPS). In the SPS, the maximum recall rate and F1 Score value of the SPS are the highest, at 79.58% and 0.7350, respectively. In the MPS, the two values of the MMMP are the highest, at 81.3% and 0.7408. The two value of the three MPSs are lower than those of the MMMP strategy and the average-median-MPS, which are 80.25% and 0.7362.

TABLE II. COMPARISON OF DIFFERENT POOLING STRATEGIES

Pooling strategy	Recall (%)	F1 Score
Maximum-single	0.7958	0.7350
Average-single	0.7846	0.7328
Middle value single	0.7909	0.7335
Maximum-average-mixed	0.7966	0.7357
Average-median-mixed	0.8058	0.7370
Maximum-intermediate-mixed	0.8103	0.7408
Three mixed	0.8025	0.7362

The results of comparing the accuracy of extracting image style information based on CNN with other algorithms under different iteration times are shown in Fig. 9. The accuracy of extracting image style information using different algorithms increases with the increasing number of iterations. The accuracy of CNN in extracting image style information is higher than other algorithms, with an accuracy of 80%. The highest accuracy of the four layer Deep Neural Network algorithm (DNN) is 68.9%, because the algorithm requires numerous parameters and computational resources, which lifts the difficulty of training and tuning. The highest accuracy rates of User-based Collaborative Filtering Algorithm (UserCF), Content-based Algorithm (CBF), Probability-based Matrix Factorization (PFM), and Item-based Collaborative Filtering Algorithm (ItemCF) are 58.6%, 68.8%, 68.8%, and 67.9%, respectively.

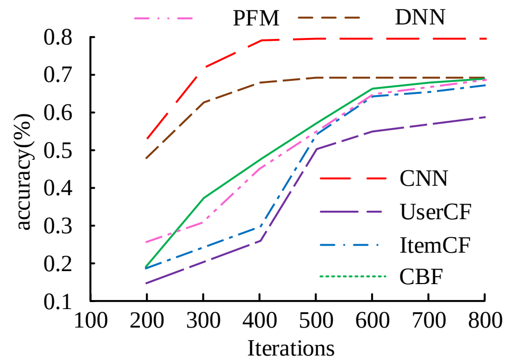
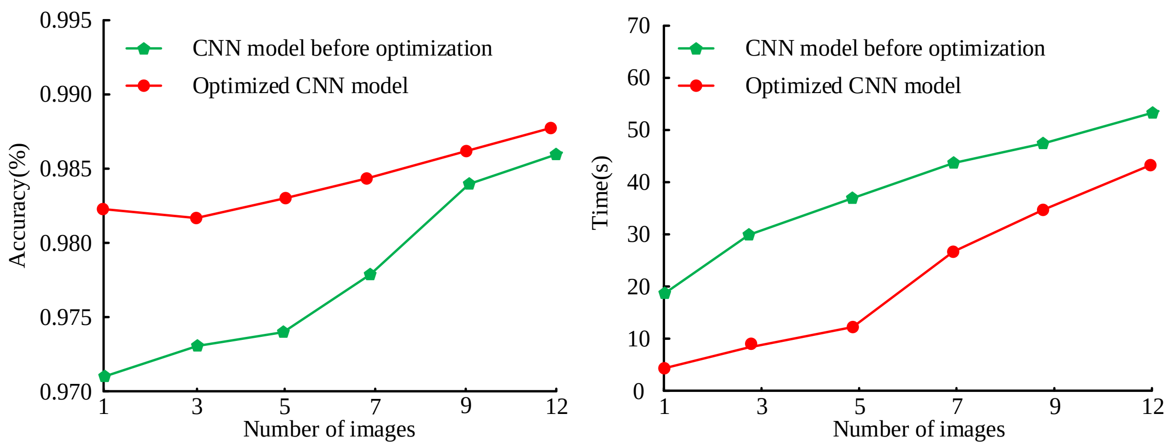


Fig. 9. The accuracy of extracting image features using different models.

In the case of different numbers of images, this study compares the accuracy and time of extracting image features based on the optimized CNN model. Fig. 10 shows the comparison results. In Fig. 10 (a), the accuracy of the optimized CNN in extracting image features increases with the increase of the number of images. When there are 12 images, the highest accuracy is 98.3%, which is 1.1% higher than before optimization. In Fig. 10(b), the running time of the pre-optimized model for extracting image features increases with the rise of the amount of images, while the optimized time decreases. When the image is 1, the optimized CNN takes a minimum of 4.7 seconds, which is 15.1 seconds shorter than before optimization. The proposed algorithm is more suitable for complex and detailed image style data sets, because the algorithm can effectively extract and process the style details in the image through the hybrid pooling strategy and the optimization of the FST layer, and improve the accuracy and efficiency of the model.



(a) The accuracy before and after optimizing the CNN model

(b) The running time before and after optimizing the CNN model

Fig. 10. Comparison of the accuracy and time of extracting image features.

In this paper, we propose an auxiliary spray system (APS) based on finite state sensor (FST) algorithm to optimize deep learning (DL) technology, and realize automatic image style extraction through convolutional neural network (CNN) model. Compared with the optimized model, it shows higher accuracy and speed in processing multiple images. When the number of images is 12, the accuracy of feature extraction reaches 98.3%, and the running time is reduced to 4.7 seconds. In addition, the maximum-median-mixed pooling (MMMP) method was used in the pooling strategy, and the recall rate and F1 score reached 81.3% and 0.7408, respectively, which was superior to the single pooling strategy. The experiment also verified the performance of CNN under different learning rate (LR) and max_depth Settings, with the best parameter configurations being LR=0.3 and max_depth=3.

V. Discussion

In this paper, a deep learning technology based on FST algorithm optimization is proposed. The accuracy of the model is increased by 1.1% when processing 12 images, and the processing time of a single image is shortened by 15.1 seconds, which is better than the existing methods and meets the requirements of rapid response in the field of industrial spraying. Compared with the scheme used in the literature [8] for the auxiliary diagnosis of breast cancer, this method has achieved a breakthrough in image processing speed. Although the literature [9] and [10] perform well in the identification of human activity (HAR), there are shortcomings in dealing with complex industrial spraying tasks; The APS model designed in this study is more suitable for the changeable construction environment. In addition, compared with the real-time HAR method in literature [11], MMMP pooling strategy is superior in terms of recall rate and F1 score, which highlights its high efficiency in image feature extraction. In terms of image style information extraction, the accuracy of this model reached 80%, which was higher than the plant disease detection model proposed in literatures [12] and [13]. At the same time, although literature [14] combines weighted FST and BERT architecture to perform well in terms of recall rate and F1 score, this study reduces the dependence on large-scale labeled data and simplifies the data preprocessing process through CNN optimization. In general, the introduction of FST optimization deep learning technology not only improves the accuracy and processing speed of image feature extraction in APS, but also shows broad application prospects in the field of industrial spraying.

VI. Conclusion

To optimize the accuracy and efficiency of the spraying system to improve product quality and market competitiveness, reduce human error and adapt to the changing construction environment, so as to achieve intelligent and efficient spraying process, reduce operating costs and reduce material waste. This study proposed APS based on FST-optimized DL technology. After optimizing DL technology based on FST algorithm, this study constructed APS based on the optimization model. The experiments indicated that when the max_depth and LR of the model were set to 0.3 and 3, the loss value and F1 Score value of the CNN were optimal. Compared with the pre-optimized CNN, the accuracy and time of extracting image features by the optimized CNN were 98.3% and 4.7s, under different numbers of images. Compared with other pooling strategies, the MMMP strategy used in this study had the best pooling effect on image features, with recall and F1 Score values of 81.3% and 0.7408. The research model has been effectively applied in assisting painting, and compared to existing models, this model has higher efficiency in extracting features. However, CNN has a large number of layers and parameters, requiring a significant amount of computation time and storage space, thus requiring high computational resources. Future research will focus on the computational efficiency of the system to reduce the computing resource requirements and further improve the performance of the assisted painting system (APS) in image style conversion. Specifically, image style recognition is enhanced by improving the feature extraction accuracy of convolutional neural network (CNN). Adjust the learning rate (LR) and maximum tree depth (max_depth) to get the best model parameters. The hybrid pooling strategy is used to optimize image feature processing to improve the recall rate and accuracy of the system.

References

[1] Ilyaevna P U, Bahodirovich T I, Davlatovich A X, Mahmudovich T M. DERMATOSCOPIC PAINTING SOME COMMON DERMATOSES. World scientific research journal, 2023, 16(2): 107-120.

[2] Cimini B A, Chandrasekaran S N, Kost-Alimova M, Miller L, Goodale A, Fritchman B, Carpenter A E. Optimizing the Cell Painting assay for image-based profiling. Nature protocols, 2023, 18(7): 1981-2013.

[3] Bakhtiyorovich K I, Shodikulovich T A. The Role of Painting in the Folk Art of Surkhandarya. International Journal of Human Computing Studies, 2021, 3(2): 26-28.

[4] Rakhmatullaevna U L. Aesthetic role of folk traditions in the development of miniature painting of the east[C]//Archive of Conferences. 2020, 8(1): 34-35.

[5] Wahyuni R, Erdiyanti E. Meningkatkan Kemampuan Motorik Halus Anak Melalui Finger Painting Menggunakan Tepung Singkong. Murhum: Jurnal Pendidikan Anak Usia Dini, 2020, 1(1): 28-40.

[6] Shlezinger N, Whang J, Eldar Y C, Dimakis A G. Model-based deep learning. Proceedings of the IEEE, 2023, 111(5): 465-499.

[7] Janiesch C, Zschech P, Heinrich K. Machine learning and deep learning. Electronic Markets, 2021, 31(3): 685-695.

[8] Saxe A, Nelli S, Summerfield C. If deep learning is the answer, what is the question?. Nature Reviews Neuroscience, 2021, 22(1): 55-67.

[9] Yu K, Tan L, Lin L, Cheng X, Yi Z, Sato T. Deep-learning-empowered breast cancer auxiliary diagnosis for 5GB remote E-health. IEEE Wireless Communications, 2021, 28(3): 54-61.

[10] Zhou X, Liang W, Kevin I, Wang K, Wang H, Yang L T, Jin Q. Deep-learning-enhanced human activity recognition for Internet of healthcare things. IEEE Internet of Things Journal, 2020, 7(7): 6429-6438.

[11] Wan S, Qi L, Xu X, et al. Deep learning models for real-time human activity recognition with smartphones. mobile networks and applications, 2020, 25(2): 743-755.

[12] Chohan M, Khan A, Chohan R, Katpar S H, Mahar M S. Plant disease detection using deep learning. International Journal of Recent Technology and Engineering, 2020, 9(1): 909-914.

[13] Chowdhury M E H, Rahman T, Khandakar A, Ayari M A, Khan A U, Khan M S, Ali S H M. Automatic and reliable leaf disease detection using deep learning techniques. AgriEngineering, 2021, 3(2): 294-312.

[14] Abro W A, Qi G, Aamir M, Ali Z. Joint intent detection and slot filling using weighted finite state transducer and BERT. Applied Intelligence, 2022, 52(15): 17356-17370.

[15] Mohammadgholiha M, Palermo A, Testoni N, Moll J. De Marchi L. Finite element modeling and experimental characterization of piezoceramic frequency steerable acoustic transducers. IEEE Sensors Journal, 2022, 22(14): 13958-13970.

[16] Dolatian H, Heinz J. Computing and classifying reduplication with 2-way finite-state transducers. Journal of Language Modelling, 2020, 8(1): 179-250.

[17] Martin K, Andreas M, Mereghetti C, Palano B S. Iterated uniform finite-state transducers: descriptional complexity of nondeterminism and two-way motion. Journal of Automata, Languages and Combinatorics, 2023, 28(3): 59-88.

[18] Somerset W E, Feeney A, Kang L, Li Z, Dixon S. Design and dynamics of oil filled flexural ultrasonic transducers for elevated pressures. IEEE Sensors Journal, 2022, 22(13): 12673-12680.

[19] Kumar V, Kalita K, Chatterjee P, Zavadskas E K, Chakraborty S A. SWARA-CoCoSo-based approach for spray painting robot selection. Informatica, 2022, 33(1): 35-54.

[20] Shorten C, Khoshgoftaar T M, Furht B. Deep Learning applications for COVID-19. Journal of big Data, 2021, 8(1): 1-54.

[21] Mohammed A, Kora R. A comprehensive review on ensemble deep learning: Opportunities and challenges. Journal of King Saud University-Computer and Information Sciences, 2023, 35(2): 757-774.

[22] Li Z, Liu F, Yang W, Peng S, Zhou J. A survey of convolutional neural networks: analysis, applications, and prospects. IEEE transactions on neural networks and learning systems, 2021, 33(12): 6999-7019.

[23] P. Preethi and H. R. Mamatha, "Region-Based Convolutional Neural Network for Segmenting Text in Epigraphical Images," Artif Intell Appl, 2023, 1(2): 119-127.

[24] Khan A H, Cao X, Li S, Katsikis V N, Liao L. BAS-ADAM: An ADAM based approach to improve the performance of beetle antennae search optimizer. IEEE/CAA Journal of Automatica Sinica, 2020, 7(2): 461-471.

[25] Parhi R, Nowak R D. Near-minimax optimal estimation with shallow ReLU neural networks. IEEE Transactions on Information Theory, 2022, 69(2): 1125-1140.