NAME: Meith Navlakha
SAP: 60004190068
AI Practical

EXPERIMENT 6

AIM: Case study on Artificial Intelligence applications.

TOPIC: Application of Convolutional Neural Network in Handwritten Chinese Character Recognition

ANALYSIS:

The paper was based on the recognition and classification of the HWDB dataset which was established by the National Laboratory for Pattern Recognition (NLPR) of the Institute of Automation, Chinese Academy of Sciences. Results obtained by the earlier models on HWDB handwritten Chinese character recognition, based on convolutional neural networks differ noticeably at different learning rates; if the learning rate was too low, the CNN's convergence speed slowed, resulting in low training efficiency. This research mainly focused on improving the accuracy by pre-processing of the dataset, selecting an apt activation function, construction and parameterizing the CNN and finally optimizing the model. The research successfully achieved an accuracy of 93% in the recognition and classification of handwritten Chinese characters.

In the HWDB dataset, the handwriting samples were made on paper by 1,020 writers using Anoto pens. The size and the label of the pictures were not uniformly processed which were needed to be pre-processed, including unify the size of the pictures and label the corresponding pictures. The font shape of Chinese characters were square, and the relative size of the original image were distributed between 40–100 pixels, thus unified to a pixel size of $50*50$. Four activation functions: sigmoid, Tanh, ReLU and Mish were selected for data analysis using the CNN model to compare the influence of each activation functions on accuracy. For the Sigmoid activation function, the accuracy rate was very low, around 10%, clearly indicating that it wasn't suitable in that scenario. When compared to Mish and ReLU activation functions, Tanh had a greater overall learning efficiency, could converge to a higher accuracy rate faster, and the final recognition accuracy could be stabilised at around 85%. Although, in the final recognition rate, Mish can be stabilized at about 90% with the fluctuation range is about 2%. Tanh, ReLU, and Mish had produced convincing results but among these three activation functions the recognition rate and stability of the Mish activation function produced the best results.

After selecting an apt activating function, an optimising algorithm was formulated in order compensate with the large fluctuations in the value of loss function around the minimum value, which hinder to reach to the optimal value. Three Loss functions: SmoothL1Loss(), BCEWithLogitsLoss() and MSELoss() were selected to measure the inconsistency between the predicted value and the real value of the model. No significant influence was observed on the final recognition accuracy, but a certain influence was seen on the initial convergence speed.

Finally, through the learning rate tuning, the weights of the CNN model were adjusted. As the learning rate reduced, the accuracy rate had a significant decrease after training 30,000 times. The appropriate range of learning rate determined was between 0.01 ~0.08. With the learning rate being 0.08, the highest accuracy is obtained, which could be stabilized to about 93%.

CONCLUSION:

In this experiment, I performed case study on ‘Application of Convolutional Neural Network in Handwritten Chinese Character Recognition’ research paper published on IEEEE. The paper was based on the recognition and classification of the HWDB dataset which was established by the National Laboratory for Pattern Recognition (NLPR) of the Institute of Automation, Chinese Academy of Sciences. The research mainly focused on improving the accuracy by pre-processing of the dataset by unifying the images to 50*50 pixels, selecting an apt activation function, construction and parameterizing CNN layers. and finally optimizing the model. The research successfully achieved an accuracy of 93% in the recognition and classification of handwritten Chinese characters using learning rate 0.08 along with Mish activation function.