J23 The Application of the Data Augmentation Technique to the Data Generation in Handwriting Classification

Yoko Seki, MA, National Research Institute of Police Science, Kashiwa-shi, Chiba 277-0882, JAPAN*

Learning Overview: After attending this presentation, attendees will understand the effectiveness of the comparison between handwritten letters that were different in the letter type but similar in shape in writer classification using a neural network.

Impact on the Forensic Science Community: This presentation will impact the forensic science community by showing the validity and the problems of the application of a neural network to handwriting identification.

Machine image recognition techniques along with deep learning have been advancing rapidly and their application to forensic document examination has also been expected. Writer classification experiments using a neural network, conducted as one of the feasibility studies, were reported in this presentation.

Handwriting classification experiments using LeNet were done to investigate the effectiveness of using handwritten samples that were different in the letter type but similar in shape as the known handwriting. Samples used for the experiments were lowercase letters “h” and “n” written by ten Japanese writers repeated five times. Image data of samples were obtained as 8-bit grayscale images under conditions where the image size was 32 pixels x 32 pixels with a resolution of 85 dots per inch (dpi).

Experiments were done as follows: 50 samples of either “h” or “n” for the test (five samples per writer x ten writers) and 1,000 samples of the other “n” or “h” generated by Data Augmentation (DA) for the training (100 samples per writer x ten writers), ten samples of either “h” or “n” for the test (one sample per writer x ten writers) and 50 samples of the other “n” or “h” for the training (five samples per writer x ten writers), ten samples of either “h” or “n” for the test and 1,000 samples of the same letter generated by DA and ten samples of either “h” or “n” for the test and 40 samples of the same letter for the training (four samples per writer x ten writers). Average F-measure was 0.4516 (max: 0.7333, min: 0.2666). Classification using “h” only showed higher F-measure value than others. Two writers showed higher recall value than others in all experiments. Their handwriting showed the unique writing manner observed both in “h” and “n.” One writer’s samples were small in size and written at the center of the writing box, which were not observed in other writers’ samples. Moreover, the letter shape was similar between “h” and “n” in the upward stroke and the stroke terminal. The other writers’ samples were thin in ink color, though all writers used the same kind of pen and ink. On the other hand, there were “h” and “n” samples that were written by the same writer and were similar in the letter shape but showed low recall value.

These results may show that the feature extraction was applied not only to the shape but also to the placement, the size, and the line quality. Thinking of the low recall value of the writers whose “h” and “n” were written similarly in shape, the local similarity in the letter shape was less emphasized than the placement, the size, and the line quality in these experiments.

Writer classification using the local similarities in the shape between two different letters was not so effective in the experiments done. More experiments such as classification using samples with higher resolution or using more complicated algorithm will be necessary for the future study.

Handwriting Classification, Neural Network, Data Augmentation

*Presenting Author