Enhancing convolution network with non-linear scaling for iris recognition

Keyur Shah, Dr. S.M. Shah, Dr. Dipika Pathak

Abstract:
Convolution neural network (CNN) based iris recognition models have been proven superiority over traditional models in terms of accuracy. In this paper, we introduce preprocessing method to enhance the performance of CNN-based iris model. By employing non-linear scaling transformation technique, we examine the influence of varying hyperparameter value in CNN performance. An experiment is carried out to evaluate the performance of the proposed model and the results are discussed.

Introduction:

Convolution neural network (CNN) based iris recognition models have received significant attention from research community as they overcome the limitation of conventional image processing schemes and make the system more robust, see [1 – 5].

Considering iris recognition, images are represented by matrices of pixel values. Pixel values range from 0 (black) to 255 (white). In order to improve the convergence speed of the CNNs based model, pixel values are scaled or normalized prior to feed into the neural network model. A typical and simple approach could be to scale by 255 to have pixel values ranging in [0, 1]. Besides, scaling is applied to change the visual appearance of an image, to alter the quantity of information stored in a scene representation, or as a low-level preprocessor in multi-stage image processing chain which operates on features of a scale. Hence, instead of pixel scaling in linear manner, we propose non-linear form for pixel scaling technique whose shape is flexibly determined by changing the shape parameter $\gamma > 0$ (See, equation (1))

$$f(x) = 1 - \left(\frac{x - x_{\min}}{x_{\max} - x_{\min}} \right)^{\gamma} \tag{1}$$

This non-linear elevation function over pixel value (x) is a generalization of many shapes depicted in Figure 1: For $0 < \gamma < 1$, convex increase; for $\gamma = 1$, linear increase; for $\gamma > 1$, concave increase. By taking different values for γ, the changes in the visual appearance of an image are shown in Figure 2. Smaller value of γ (say $\gamma = 0.2$) darkens the image, as many values of the pixel are between 0 and 0.5. On the other hand, as γ increases, the image has exposed more and washed-out the appearance.

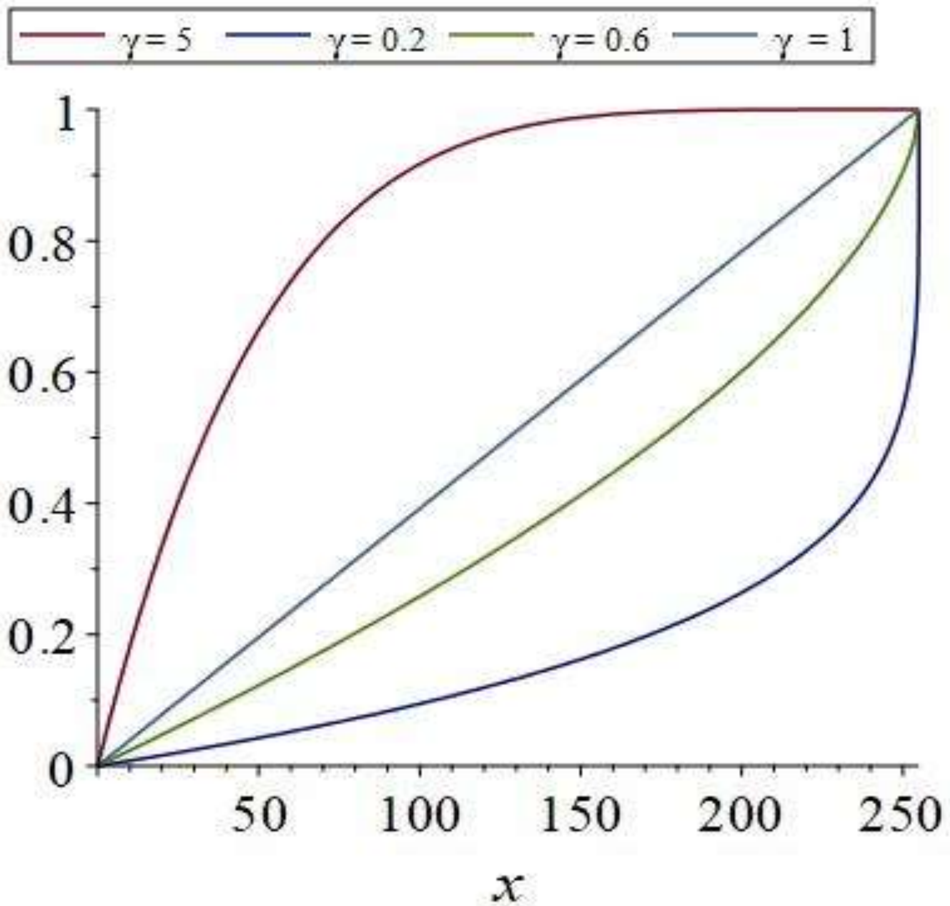

Figure 1: Scaling under different value of γ

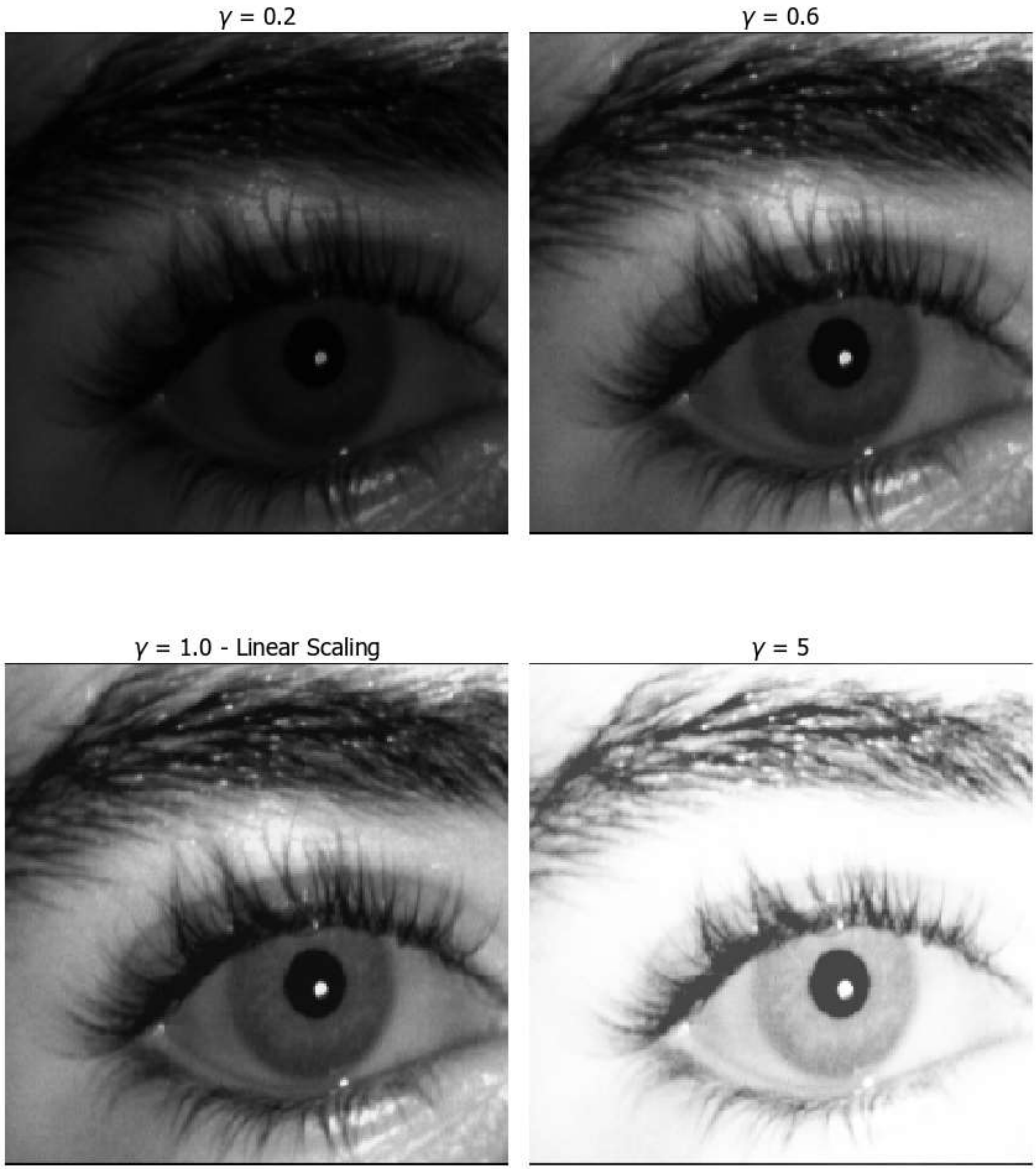

Figure 2: Image under different scaling

Proposed method:

In order to examine the effect of pixel scaling, we decided to use smaller model whose architecture is shown in Table 1.

Table 1: Architecture of CNN

Name	Type	Kernel/Pool Size	Output feature map size (height × width × number of channel)	Parameters
Conv1	Convolution	7	32 × 32 × 64	3136
BN1	Batch Normalization		32 × 32 × 64	256
Pool1	Max-pooling	2	16 × 16 × 64	---
Conv2	Convolution	3	16 × 16 × 128	73728
Conv3	Convolution	3	16 × 16 × 128	147456
BN2	Batch Normalization		16 × 16 × 128	512
Pool2	Max-pooling	2	8 × 8 × 128	---
Conv4	Convolution	3	8 × 8 × 256	294912
Conv5	Convolution	3	8 × 8 × 256	589824
BN3	Batch Normalization		8 × 8 × 256	1024
Pool3	Max-pooling	2	4 × 4 × 128	---
FC1	Fully Connected	256	---	1048832
Drop1	Dropout – 50%			
FC2	Fully Connected	128	---	32896
Drop1	Dropout – 50%			
FC3	Fully Connected	Number of classes		

Each convolution layer is constructed with Rectified Linear Unit (ReLU) activation function, same padding and no bias. The input of our net is a gray iris image of resolution size 32 × 32 which is preprocessed by fixing the value of γ in equation (1). The net is trained using different scaling without applying data augmentation. Training of CNN model is done by using Nadam optimizer. Nadam is Adam optimization plus the Nesterov trick, so it will often converge slightly faster than Adam. After each training epoch, we compute the error on validation dataset and update the model that provides highest validation accuracy.

Experiment and results:

We use MMU.2 Database to carry out the experiments. MMU.2 iris database consists of 995 iris images which are stored in BMP format with resolution 320×240. We perform our experiments using TensorFlow 2.x framework with GPUs on Google Colaboratory. We split the dataset into training, testing and validation using stratified sampling. In our experiment, we use 795 images for training, 100 images for validation and 100 images for testing. Loss and accuracy on the dataset are summarized in Table 2 and depicted in figures 3 – 5.

Table 2: Loss and Accuracy for different values of γ

γ	Loss (train/valid/test)	Accuracy (train/valid/test)
0.8	0.3128/0.7865/0.9967	0.9161/0.8400/0.8100
1.0	0.3034/0.8401/1.0744	0.9088/0.8100/0.8000
1.1	0.3854/0.9457/0.9880	0.8632/0.8100/0.8100

We can observe from the figures 3 – 5 that the accuracy steadily increases whereas the loss decreases for the training dataset. On the other hand, oscillatory trajectory is observed for accuracy and loss in the validation data. However, the linear scaling shows higher oscillation for accuracy than that of non-linear scaling in the validation data. Moreover, accuracy is higher for non-linear transformation than that of linear transformation.

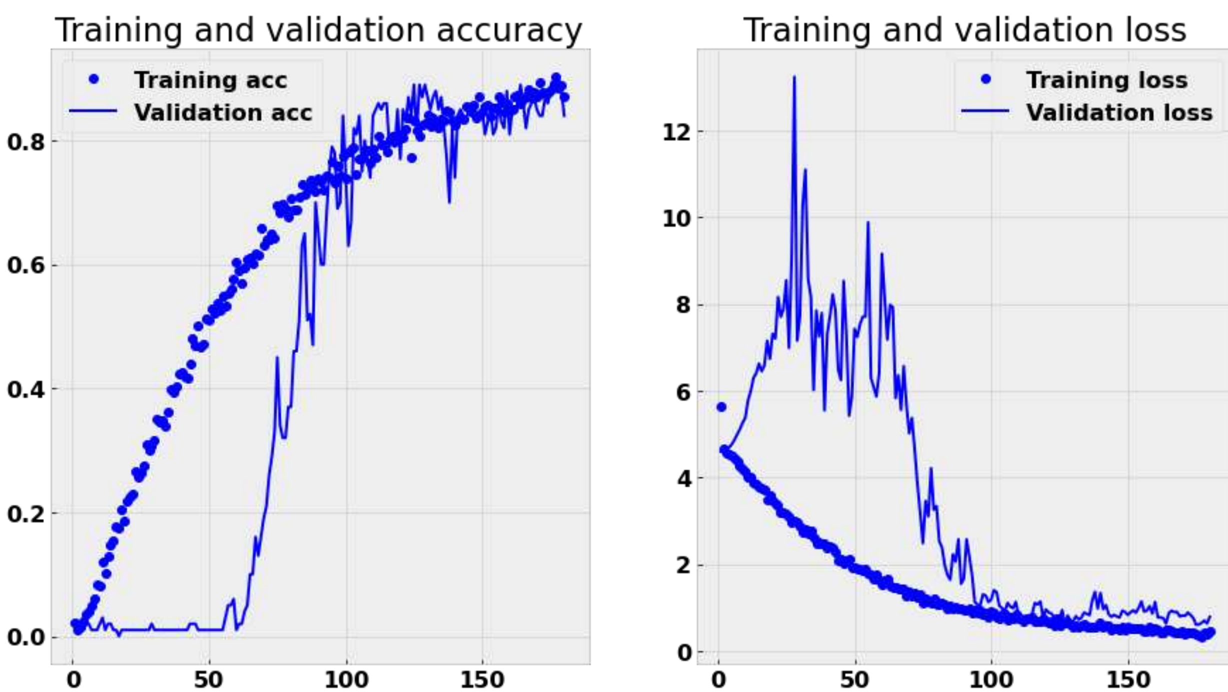

Figure 3: Loss and Accuracy for $\gamma = 0.8$

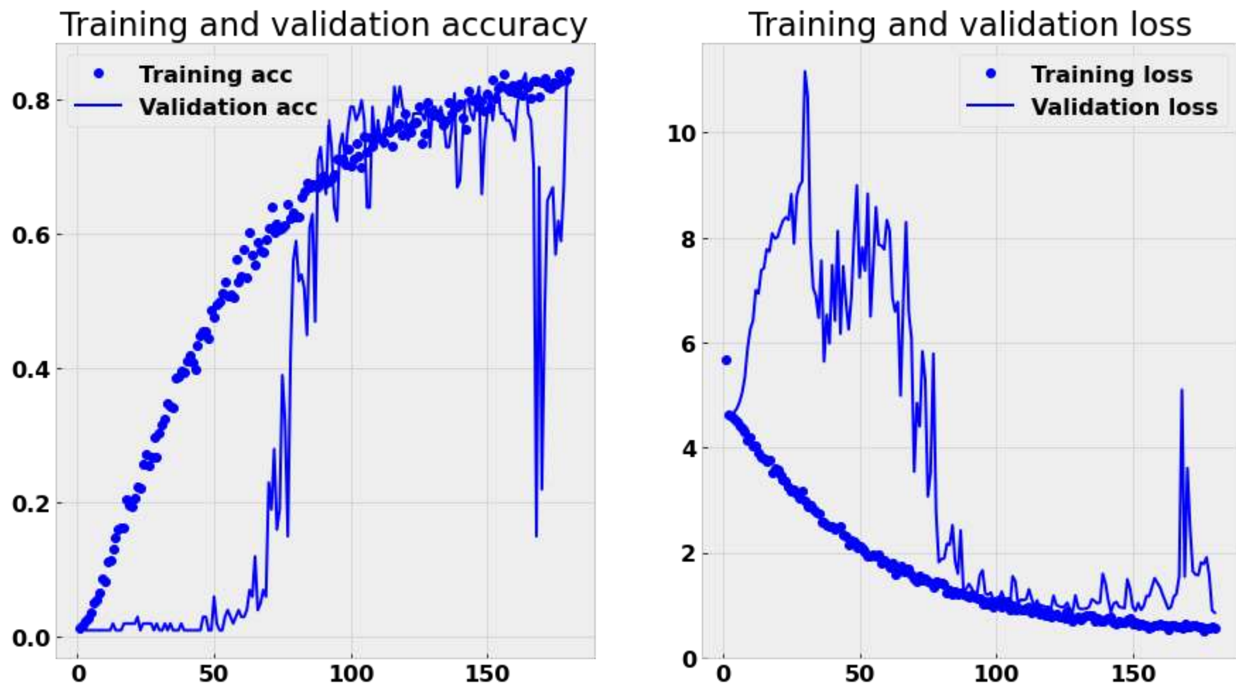

Figure 4: Loss and Accuracy for $\gamma = 1.0$

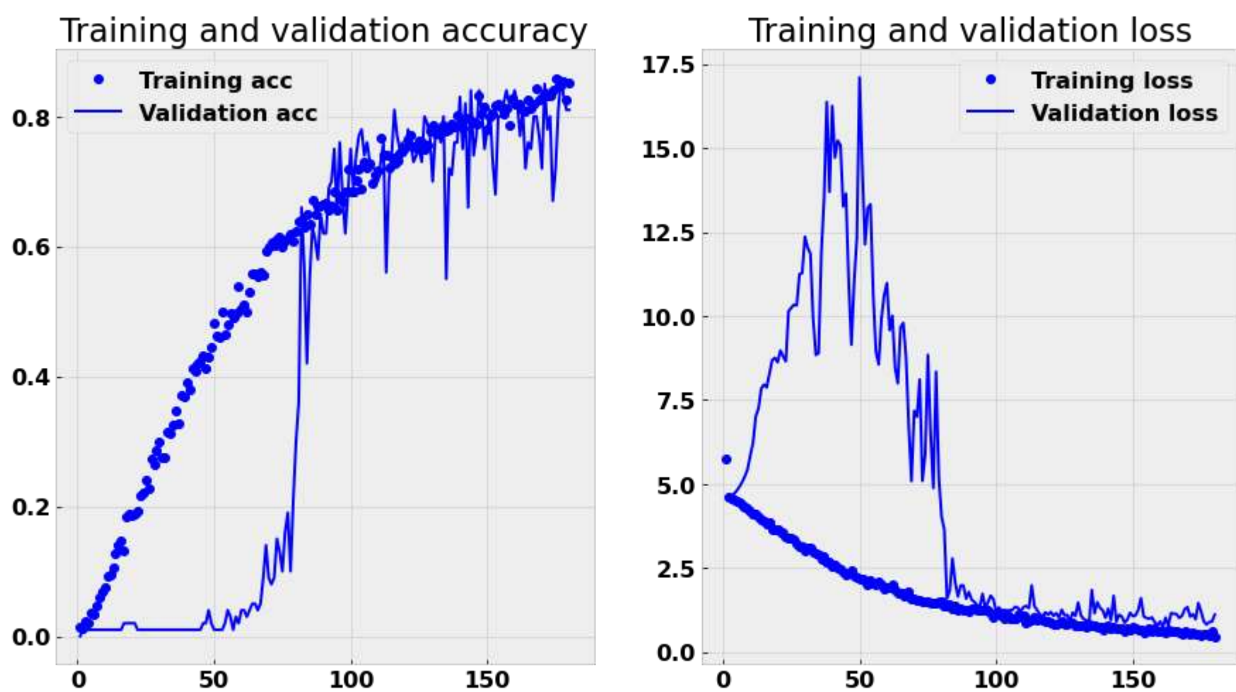

Figure 4: Loss and Accuracy for $\gamma = 1.1$

Conclusion:

In deep learning, the images are generally preprocessed by applying the scaling transformation such that the resultant matrix of the image is mapped into [0, 1]. In this paper, we propose non-linear scaling transformation technique and analyze the performance of CNN for iris recognition. The results of the experiments exhibit the improvement in the performance of the CNN model when non-linear scaling is employed. This research can be extended in many ways. One immediate probable extension will be to apply the same model on another public iris dataset. One could apply the proposed scaling technique on several well-known networks with different resolutions.

References:

[1] Arsalan, M., Kim, D., Lee, M., Owais, M., FRED-Net: Fully residual encoder – decoder network for accurate iris segmentation, Expert Systems With Applications, 2019, 217 – 241.
[2] Li, Y. H., Huang, P. J., Juan, Y., An Efficient and Robust Iris Segmentation Algorithm Using Deep Learning, Mobile Information Systems, 2019, 1 – 14.
[3] Long, J., Shelhamer, E., Darrel, T., Fully Convolutional Networks for Semantic Segmentation, EEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 640 – 651.
[4] Tobji, R., Di, W., Ayoub, N., FMnet: Iris segmentation and recognition by using fully and multi-scale CNN for biometric security, Applied Sciences (Switzerland), 2019, 1 – 17.
[5] Trokielewicz, M., Czajka, A., Maciejewicz, P., Post-mortem iris recognition with deep-learning-based image segmentation, Image and Vision Computing, 2020, 103866.
[6] Tustison, N. J., Avants, B. B., Gee, J. C., Learning image-based spatial transformations via convolutional neural networks: A review, Magnetic Resonance Imaging, 2019, 142 – 153.

AUTHORS PROFILE:

Keyur Shah
System Analyst – GTU.
MCA, M.Phil, Ph.D. – Pursuing.

Dr. S.M. Shah
Professor – LDCE, Ahmedabad

Dr. Dipika Pathak
Pro. VC - APJAKU