Introduction

This project aims to find the influence of manipulating datasets for network performance. As studied in class, the manipulation could be data augmentation or pre-processing of images used before training. In paper [1], we learned the method that normalizing images in color space to erase the objective difference caused by machines. Also, rotation and flipping are good ways to increase datasets using limited information. In article [2], though it's for radiology, a method of windowing different organs arose. Images of this project are all cells not organs, however. I was thinking about whether extracting the nucleus feature of cells is feasible (pathologists may identify cell types by recognizing nucleus or finding mitosis). According to that, I split the nucleus and other tissues apart to black and white separately, which is like a kind of ultra-windowing (only keeping the nucleus part in black). In paper [3], translating pictures (shifting) and adding noise are inspired me. In this project, two kinds of methods for enhancing network performance using limited datasets are discussed -- data Augmentation and data preprocessing.

Methods and settings

1. *segmentation of test dataset and verification dataset.*

   The verification dataset was set for 1500 images, each type has 500 pictures (each type of cell accounts for one-third). The number 1500 is decided through several tests (result in appendix), making sure that the network can improve performance. Meanwhile, there's enough data for training. The verification dataset has 1500 images. The same as training data. The verification dataset has no overlap with training data. For every method below, the original training dataset, as well as the verification dataset, did not change. This is for the variable control, to find the influence of methods.

2. *Software for manipulation.*

   All the methods related to image processing are finished by MATLAB. No external package was used.

3. *Evaluation criterion.*

   In this project, the final performance of the deep learning network is evaluated by the Micro-AUC generated by the test dataset. In each method, a curve formed by original network performance vs manipulated network shows the influence (all listed in the appendix) of the method. the highest AUC was selected to evaluate the performance and saved for testing.

4. *Data Augmentation methods.*

4.1 Rotation

As mentioned in class and paper [3], rotation is a good basic method for augmenting a limited dataset. By rotating inputs, the network can recognize the same cell at different angles, which makes the algorithm performs more like a human.

Figure 1 Different rotation Angle Image

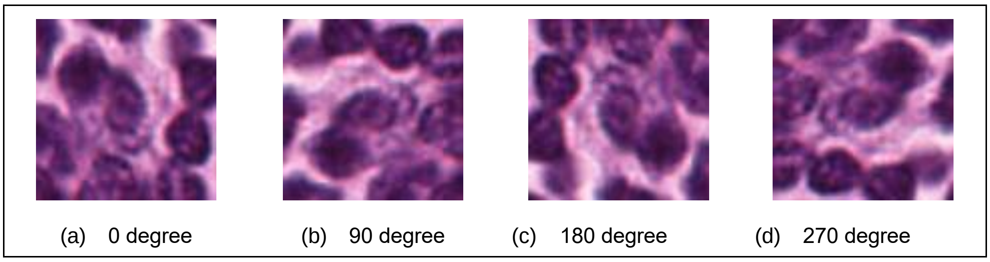


Table 1 Performance of Rotating Image

Original Network Best AUC	Augmented Network Best AUC
0.8371	0.8523

As in **figure 1**, three extra directions rotation of image added to the dataset, the performance of verification was enhanced as **table 1**.

4.2 Adding noise

In paper [4], adding noise can help the network learn features more robustly. In general life, people's decisions are robust for lots of disturbing things. It is reasonable that the network should learn this kind of ability for a more robust view. In this method, an average gaussian white noise matrix was generated, which has the same size as images used as training input. The original picture and added noise version are like **figure 2** below:

Figure 2 Image with AWGN Matrix

(a) Original Image (b) Added Noise Image

Table 2 Performance of Adding Noise

Original Network Best AUC	Augmented Network Best AUC
0.8371	0.8541

As the result in **table 2**, this method enhanced the performance of the network.

4.3 Translation

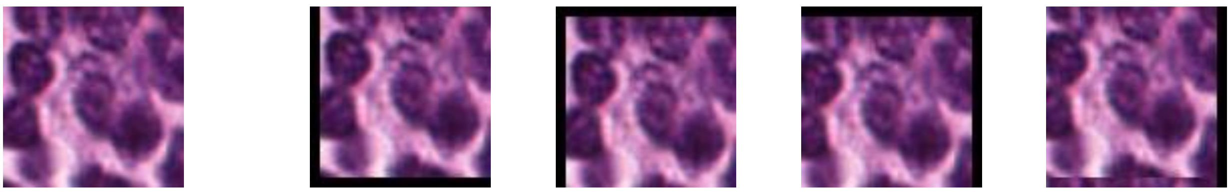
Shifting images to left, right, up, or down can be a very useful transformation to avoid positional bias of the dataset [3]. At first, I tried to shift images like **figure 3**:

Figure 3 Translated Image

(a) Original Image (b) Translated Image

The AUC was not been better. The aim of translation (shifting) is similar to adding noise – enhancing the robustness of the system. So, I shifted the image to different directions, which can eliminate the errors from pixel sites. **Figure 4** shows how pictures shifted. After shifting, the accuracy becomes higher as table 3

Figure 4 Translated Image (four directions)



Original Image (a) Top-Right (b) Down-Right (c) down-Left (d) Top-Left

Table 3 Performance of Translation

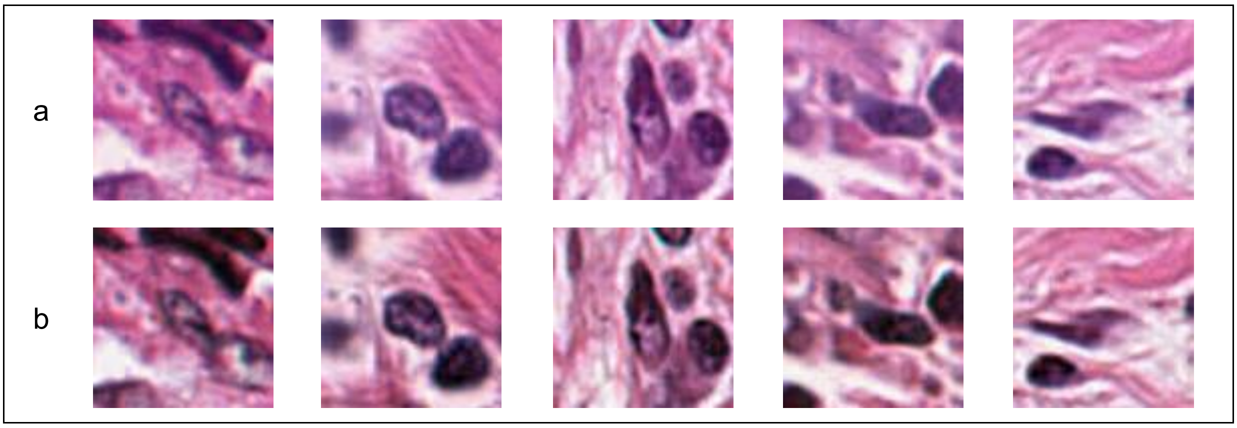
Original Network Best AUC	Augmented Network Best AUC
0.8371	0.849

5. *Image pre-processing methods.*

5.1 Normalizing RBG channel

A good and normal way to manipulate data is to normalize the image [1]. In this method, images were preprocessed by uniform three-color channels' range. Every image's RBG data range was extended to [0,255], making them seem similar and erasing the difference caused by the machine. Images after processing shows below in **figure 5**:

Figure 5 Uniformed Images. a. Original images. b. Normalized images.



By observing the images above. I guess this method will have a good performance influence. As in the picture, the nucleus' edge is sharpened. The network should learn the nucleus' behavior easier. Because cells classification is related to nucleus splitting, system accuracy should increase. The verification result is like below in **table 4**. The accuracy was indeed increased.

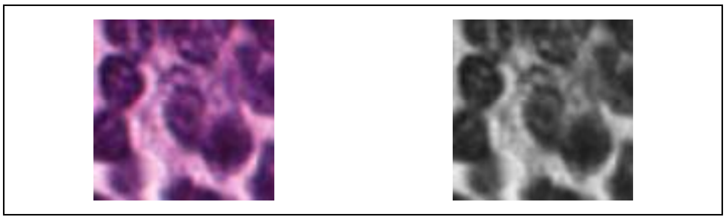
Table 4 Performance of Normalization

Original Network Best AUC	Augmented Network Best AUC
0.8371	0.8434

5.2 Turning picture to gray image

When learning the radiology part lecture, I was curious about how will the network behave when inputs are gray pictures. In radiology, everything scanned without color. In the classification of cells, I want to find whether colors will offer plenty of information. The processing of the image is like below in **figure 6**. AUC of the network trained by gray images is listed in **table 5.**

Figure 6 Gray Image Processing



(a) Original Image (b) Gray Image

Table 5 Performance of Gray Image

Original Network Best AUC	Augmented Network Best AUC
0.8371	0.7936

The performance is decreased compared with the original image. The color domain offers lots of information, which cannot be ignored.

5.3 Windowing

When experimenting with normalizing images' RBG channels, I start to think about the windowing studied in the lecture. As mentioned in class, a good signal from tumors is that the cell mitosis happening disorderedly. If the network can learn the nucleus' performance, it might behave more like real doctors. After that though, windowing from paper [2] existed in my mind. For radiology, doctors limited the range of each kind of organ to observe them. If only want network figure out the change of nucleus, a direct way is to delete any other information. So, I tried to window the nucleus and turn the part to pure black. Left other tissues part as totally white. Images after processing like below in **figure 7**:

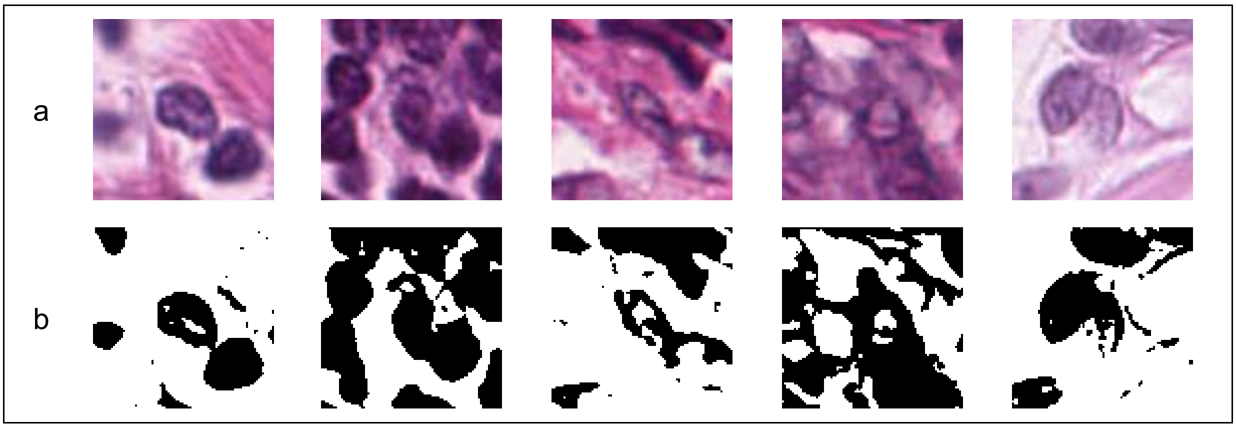


Figure 7 Uniformed Images. a. Original images. b. Windowed Images.

The fourth picture above is a tumor cell, as in the windowed version, the nucleus part is chaotic as I supposed. However, this method does not have good accuracy. The performance is listed in **table 6**.

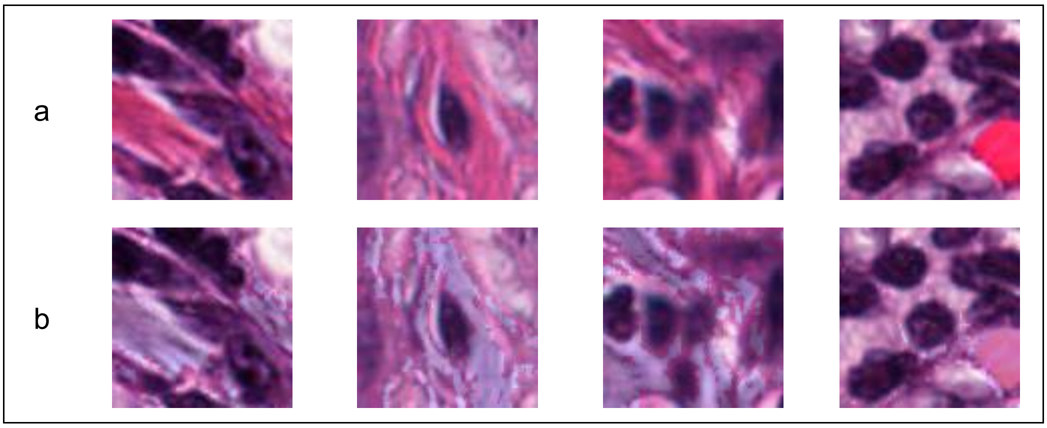
Table 6 Performance of Windowing

Original Network Best AUC	Augmented Network Best AUC
0.8371	0.7557

5.4 Erasing red patches

When observing images, the blood will clover part of tissues sometimes. To find out whether this kind of cover affects network decisions, I tried to erase some images' red patches as in **figure 8**.

Figure 8 Erasing Red Patches Processing



The test result is not better than the original one. It is reasonable because though red patches are gone, other color patches exist, which formed more noises.

Table 7 Performance of Erasing Blood Patch

Original Network Best AUC	Augmented Network Best AUC
0.8371	0.7687

Results & Discussion

Result:

According to the experiment result on the verification dataset, the performance of test data was generated from the experiment. All AUC changes result listed in **table 8** as below:

Table 8 Performance of Each Method

	Best AUROC			
Method	Original	Manipulated	Test Data in Original	Test Data in Manipulated
Rotation	0.8371	0.8523	0.7655	0.8367
Translation	0.8371	0.849	0.7655	0.8367
Normalization	0.8371	0.8434	0.7655	0.7707
Windowing	0.8371	0.7557	0.7655	0.7242

Gray Image	0.8371	0.7936	0.7655	0.8095
Adding Noise	0.8371	0.8541	0.7655	0.8358

Discussion:

From **table 8** above, before test data was released, I thought augmentation always perform better than preprocessing. As lots of experiments I did, preprocessing always does not have better accuracy than the original network, in which the training data is handled by nothing (except normalizing). But after testing on the test dataset, I find the accuracy changed with data (obviously). Simply turning images to gray is working on the test dataset. In several methods, rotation, translation, and adding noise get the best performance on both verification dataset and test dataset. Though these three methods are very simple, they got good AUC enhancement. This is amazingly effective for domains like medical DL in which data are hard to get. Before I start to experiment, I thought maybe windowing is a good and simple method for classification. After all, the aim of windowing in my project is to extract the nucleus feature. As in my mind, this is interpretable as human behavior. But the result is not as good as I imaging. Maybe I was making mistakes on how real doctors are doing. Though it's still worth having a try.

In three simple methods. I found that translation and noise adding do have better robustness. In figure 9 and figure 11, the AUC has a small ripple changing with epochs. But rotation has a valley on the curve. It might be because I rotate the image every 90 degrees. In the reference paper, the author said small degrees will lead to better performance. But rotating a small degree is a little hard for me right now. These could be later steps.

For normalization, I believe there're multiple better ways to pre-process images. I only realized the simple method in my project. If there's any chance. I can figure out how other ways work and have a try.

Also, about noise adding, I only add one random noise metric. Maybe adding several versions of metrics will enhance the performance better. Like in translation and rotation. At first, I use original data and a kind of augmentation to training the network. The AUC does not change at all. After reading the papers, I find eliminating particularity is an important thought. By adding other directions (angles and shifting way), the performance changed a lot. According to that, the performance might be much better when all good methods are applied together. I'll try to mix them later.

Figures (curve of AUC vs Epochs for Verification dataset)

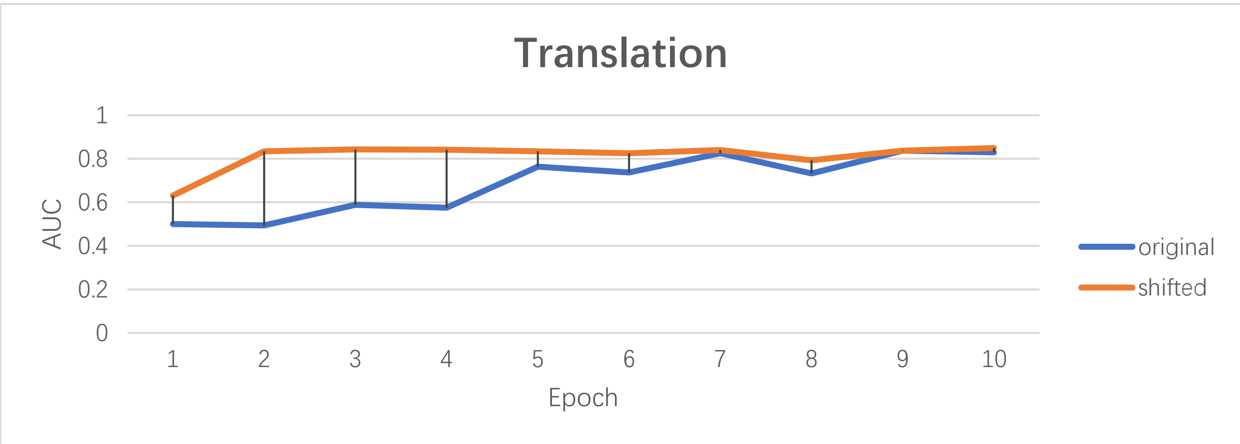


Figure 9 Translation Processing

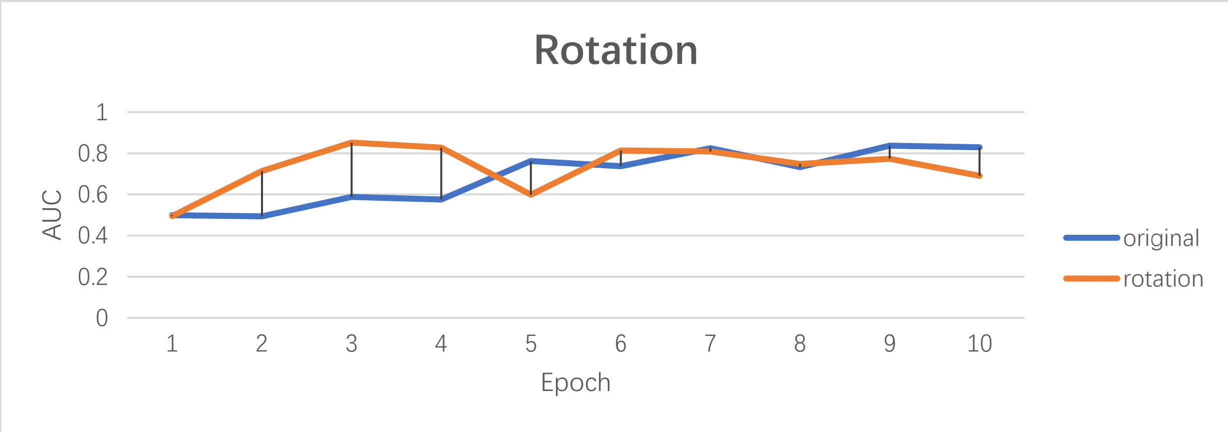


Figure 10 Rotation Processing

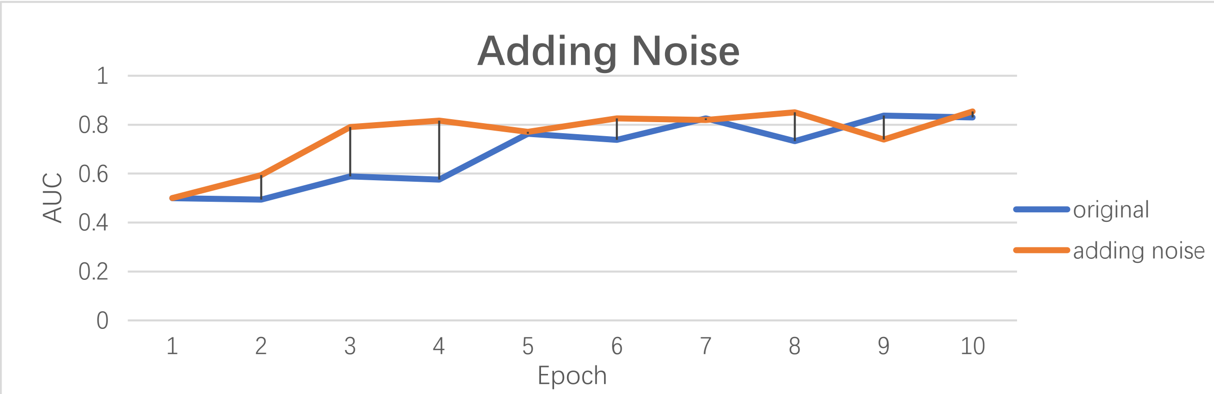


Figure 11 Noise Adding Processing

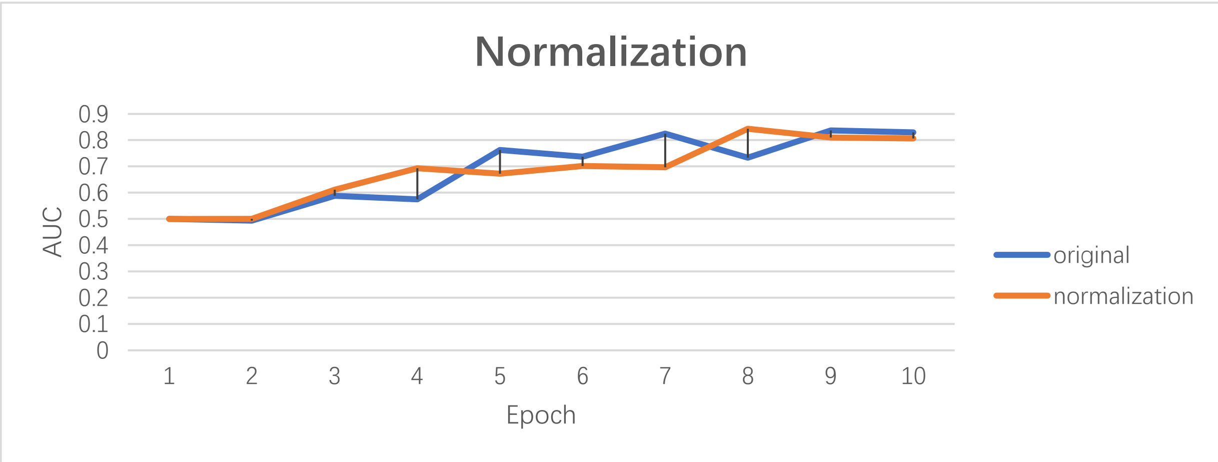


Figure 12 Normalization Processing

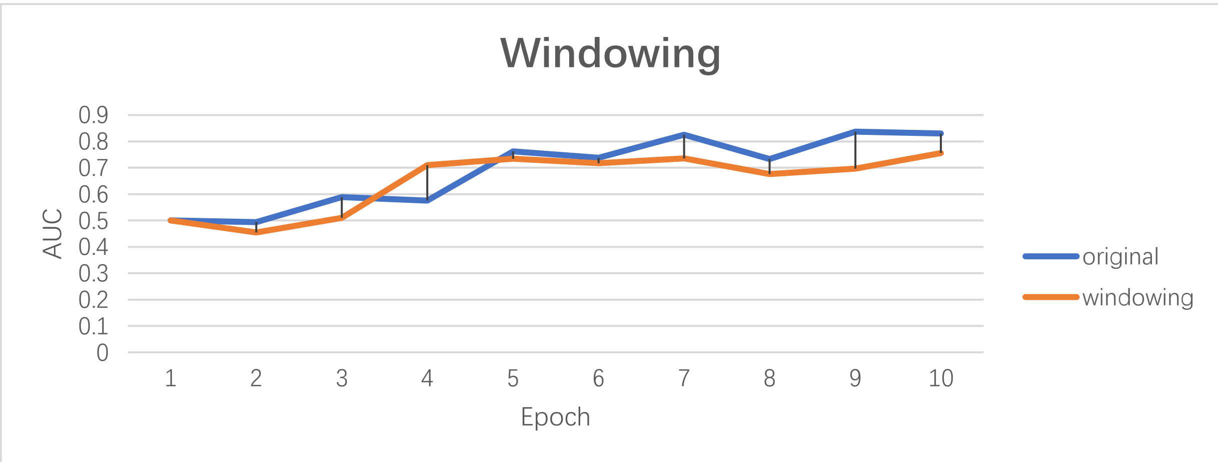


Figure 13 Windowing Processing

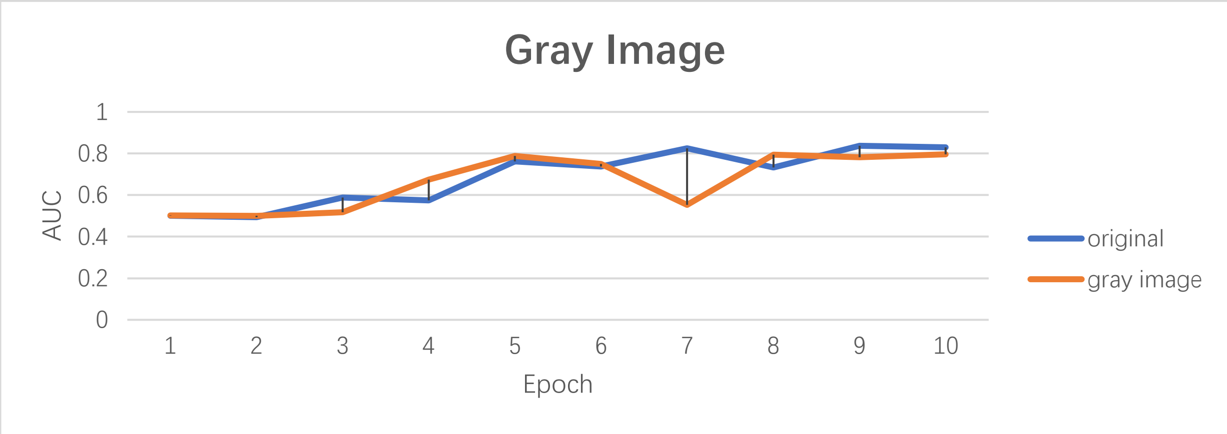


Figure 14 Gray Image Training

Bibliography

[1] P. Salehi and A. Chalechale, "Pix2Pix-based Stain-to-Stain Translation: A Solution for Robust Stain Normalization in Histopathology Images Analysis," 2020 International Conference on Machine Vision and Image Processing (MVIP), 2020, pp. 1-7, doi: 10.1109/MVIP49855.2020.9116895.

[2] Samira Masoudi, Stephanie A. A. Harmon, Sherif Mehralivand, Stephanie M. Walker, Harish Raviprakash, Ulas Bagci, Peter L. Choyke, Baris Turkbey, "Quick guide on radiology image pre-processing for deep learning applications in prostate cancer research," J. Med. Imag. 8(1) 010901 (6 January 2021) https://doi.org/10.1117/1.JMI.8.1.010901

[3] Shorten, C., Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. J Big Data 6, 60 (2019). https://doi.org/10.1186/s40537-019-0197-0

[4] Dua D, Karra TE. UCI machine learning repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science; 2017