

Personalized Photo Enhancement Using Artificial Neural Network

Jennifer C. Dela Cruz¹, Ramon G. Garcia¹, Glenn V. Magwili¹, Jerico C. Cordon¹, Jesser Paul P Suplico¹ and Zaliman Sauli²

¹*School of Electrical, Electronics and Computer Engineering, Mapua University, Intramuros 658 Muralla St., Intramuros, Manila 1002, Philippines*

²*School of Microelectronic Engineering, Universiti Malaysia Perlis, Pauh Putra Campus, 02600 Arau, Perlis, Malaysia. jcdelacruz@mapua.edu.ph*

Abstract—Artificial Neural Network (ANN) is applied to create a photo enhancement program that automatically adjusts image parameters on the face based on the preference of its own user. Viola Jones algorithm was used for face detection, and a Graphical User Interface (GUI) is created to enable users to edit the photos easily. Input data sets are essential in the learning progress of ANN. Variety of users inputted their respective image data into the program for training the neural network. Regression plot developed will be used to determine the performance of the system. The authors would relate the consistency of the users in editing their photos to the produced regression plot. On the other hand, actual tests were conducted to determine the time spent editing the photos manually and the amount of time the system automatically adjusted the photo. There is a difference in editing time between average manual adjustment and automatic adjustment by the ANN.

Index Terms—ANN; Photo Enhancement; Automated; GUI; Viola Jones.

I. INTRODUCTION

Photo enhancement is used to improve the quality/definition of a picture using a computer-based program. This includes reducing grey scale values to black at lowest and improving it as white as possible at highest. These pictures may either be from cellphone cameras, DSLRs, webcams, scanners, and even microscopes. The role of images is widespread ranging from medical images up to personal photography images. From the medical point of view, image clarity regarding contrast is highly required to diagnose the medical condition from the images. The main tenacity of image enhancement is to augment the details and visibility of an image. An enhanced image will be able to provide a well-defined view, and it will assist computer vision system for feature extraction dispensation. There is various image enhancement method [1], and one should consider the factors that influence the enhancement such as computational constraint, user intervention, enhancement competence and application appropriateness in deciding the suitable method for image enhancement.

An Artificial Neural Network (ANN) is a brain-like model which undergoes complex and logical computer-based computations which are based on the values being fed to it. It is a network based on previous information that its memory identifies. It is highly reliable in the system's input and output for its learning capabilities because the network is affected by the changes in its neural network.

Existing photo-enhancement software demands manual adjustment to every photo which indeed takes time especially

in a photo studio and algorithms are commonly provided by the default source of the software which eliminates the freedom of people to adjust it further with their preference. In this study, the authors aim to automatically adjust an image's parameters specifically the contrast, brightness, saturation, hue, warming/cooling and sharpness based on the preference of the users by using the unique characteristics of an artificial neural network which helps users achieved their preferred photo settings parameters.

A study regarding image enhancement with the use of Artificial Neural Network and Fuzzy Logic wherein an input image will be used, and the neural network would detect the noises present in it. These noises are the unwanted part of the image which will be filtered out using the fuzzy logic. Then, image enhancement would take place [1].

Another study includes the use of an integrated recurrent neural network (IRNN) [2]. The IRNN comprises a set of the feed-forward neural network which functions the aptitude of having a response to info from its output to its input cooperatively. It has both of the capabilities of learning and searching optimal solutions in the solution space for optimisation problems. Combining the notion of adaptive filters with neural networks for analysing the content of an image in a detailed manner is also being highly investigated [3].

Conventional image improvement guidelines are principally determined empirically. Software such as Microsoft Office Picture Manager, Adobe Photoshop and Google Auto Awesome are tools utilised to employ automatic colour correction and tone alteration for a given image. Research is also being carried out for automatic colour and tone modification [4].

Automatic methods naturally function on the whole image in a universal method without considering the image content. Thus, to provide an option for this issue, an automatic method was proposed where it initially distinguishes the semantic content of the image and performed saturation, exposure and contrast adjustments based on a determined sequential empirical step [5]. Nevertheless, the output of this method is hardcoded in the algorithm and is difficult to modify for desired output style which is a main limit of this approach [6]. Thus, the work proposed in this study will be based on a data-driven approach that can effortlessly be educated to generate a variety of styles. This work intends to give the users the freedom to personalise and enhance the image instantly and of course in his desired preference.

The general objective of this study is to create a photo enhancement program that automatically adjusts image

parameters on the face based on the preference of its own user using ANN. Image parameters specifically the contrast, brightness, saturation, hue, warming/cooling, and blurriness/sharpness would be involved. Specifically, this study aims; (1) to isolate the facial feature for enhancement using viola jones algorithm, (2) to enhance an image based on individual preference with the use of ANN, (3) to create a graphical user interface for the program, and (4) to compare manual photo editing time versus the proposed system.

II. METHODOLOGY

The framework for this study comprises of 5 parts and information on the neural network modelling. It is explained in the following subsections.

A. Part 1: Collections of Photo Training Sets

The collection of training sets involves several pictures being modified into the system as it would be used as the initial input variables to be fed into the Artificial Neural Network. This is also the part of the study, in which the users could personalise the images they would be modifying as it will result in the output image that would be produced by the Neural Network. It involves loading the image to be part of the training set. Next, this will be followed by Viola-Jones face detection for the loaded images [8]. Next, the detected face will be cropped followed by quantification step, manual adjustment step, and attainment of quantified values input and adjusted output values and finally storage to training sets as input or output variables.

B. Part 2: Quantification Process

In quantification, the image is given a numerical value that can be understood by the ANN. The quantification process done by the authors are done in two ways, by getting the average hue, average saturation, the average value (HSV) and Histogram of the grayscale in the loaded image. The Grayscale histogram was divided into five regions, the white portion, semi-white portion, grey portion, semi-black portion and the black portion of the loaded image. Viola Jones algorithm is responsible for detecting faces [8]. The yellow box detecting the face signifies the Viola-Jones Algorithm. A feature is added wherein when the image is composed of several faces; the user will just click the "Next Face" button to set it to the right image. As the face image is cropped using the viola jones algorithm [8], the cropped face will now go under the process of quantification and will start by converting the RGB value of the cropped face to HSV. Take note that the viola jones detected portion is the only parameter in the image to be processed in the quantification [8].

C. Part 3: Image Processing

In Image Processing, the image is adjusted manually by the user according to their preferred quantification values. The Contrast, Brightness, Hue, Saturation, Lightness, Warming - Cooling, and Blur - Sharpen of the image are provided one slider each for adjustment purposes. If the slider is moved, the adjust image function is executed. The adjust image function gets all of the current positions (values) of all slider and performs image processing based on all of the slider values. Normalization is used to convert the slider values from 0 – 1. All of the sliders are placed in default with the value of 0.5 which means that no change has occurred for that particular function. Note that the adjust image function works in order

and gets the slider values as follows: Brightness, Contrast, Hue, Saturation, Lightness, Warming - Cooling, and Blur – Sharpen.

D. Part 4: Training Process

When the network weights (inputs and outputs) are gathered, the network is ready for training. The training process requires a set of examples of proper network behaviour—network inputs and target outputs. The user would just click the ANN Adjust button on the GUI, and the Artificial Neural Network will train all the training sets. The collection of training sets involves several pictures being modified into the system as it would be used as the initial input variables to be fed in the ANN. This is also the part of the study wherein the users could personalise the images they would be modifying as it will result in the output image that would be produced by the ANN.

E. Part 5: Artificial neural network auto adjustment process

As soon as the collection of training sets are done, the user can test the learned adjustments of the Artificial Neural Network. When ready, load the Neural Network, and it will use the quantified values of the image, and the ANN will return its learned adjustments.

F. Neural Network Modelling

MATLAB Neural Network Toolbox is used in creating the network. Figure 1 is the neural network model that will be in use for this system. The Neural Network consists of Input Layer, Hidden Layer and Output Layer. The input layer contains the HSV Region and Segmented Histogram Regions. There are seven input elements in this model. They correspond the hue, saturation, hue, black, semi-black, grey, semi-white and white.

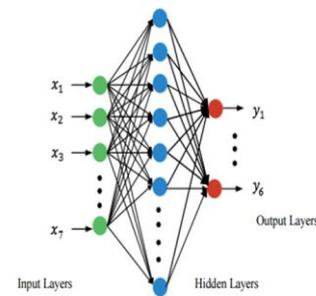


Figure 1: Artificial Neural Network Model

The output layer refers to the layer of manually adjusted by the user or the image sliders. The model is designed with six output elements. The Contrast, Brightness, Hue, Saturation, Lightness, Warming - Cooling, and Blur – Sharpen. Each output element is the decision of the users on how they adjust the image sliders of each of the image parameters a neural network with ten hidden layers has been developed after series of the training process. The number of input nerve cells and the number of output nerve cells together will determine the number of nerve cells in the Hidden Layer. It is generally calculated $L=Nm/(n|m)$ or $L= n0.5m0.5in$ which L denotes the number of nerve cells in the Hidden Layers, and n and m respectively correspond to the number of nerve cells in the Input Layer and Output Layer, while N refers to sample capacity[7].

G. Multi-Layer Perceptrons (MLP)

Multilayer perceptrons are biologically enthused neural models comprising of a multifaceted network of interconnections among fundamental computational units, called neurons which functions for patterns recognition and regression of nonlinear functions. A single neuron processes multiple inputs pertaining an activation function on a linear combination of the input.

$$y_i = \varphi_i \left(\sum_{j=1}^N w_{ij}x_j + b_j \right) \quad (1)$$

In Equation (1), the {x_j} is the set of inputs, w_{ij} is the synaptic weight connecting the jth input to the ith neuron, b_i is a bias, φ_i(·) is the activation function, and y_i is the output of the ith neuron considered. Figure 2 shows a neuron. The activation function is usually nonlinear, with a sigmoid shape.

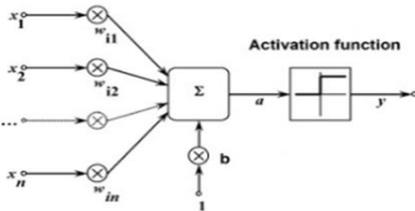


Figure 2: Representation of an artificial Neurons

A modest network devising the general approximation element is the feed-forward MLP with a single hidden layer. The Levenberg-Marquardt is an iterative algorithm to estimate the vector of synaptic weights (a single output neuron is considered) of the model, minimising the sum of the squares of the deviation between the predicted and the target values as shown in Equation (2).

$$E(\vec{w}) = \sum_{i=1}^N (d_i - y(\vec{x}_i\vec{w}))^2 \quad (2)$$

These model equations were simplified through MATLAB and can be easily used in the program in the simple coded language. The researcher was able to gather 50 data sets for each user. Among the samples, 70% were used for training; 10% for validation and 15% for testing. Figure 3 shows the neural network model diagram created in MATLAB tool.

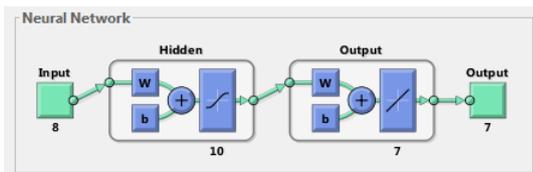


Figure 3: Artificial Neural Network Diagram

H. Analyze Artificial Neural Network after training

When the collection of training sets is complete for a user, the ANN performance can be checked by creating a regression plot which shows the relationship between the outputs of the network and targets. Figure 4 shows a regression sample produced by the Neural Network. The dashed line in each plot represents the perfect result – outputs = targets. The solid line represents the best fit linear

regression line between outputs and targets. The R-value is an indication of the relationship between the outputs and targets. If R = 1, this indicates that there is an exact linear relationship between outputs and targets. If R is close to zero, then there is no linear relationship between outputs and targets. For this example, the training data indicates a good fit. The validation and test results also show R values that greater than 0.9. The scatter plot is helpful in showing that certain data points have poor fits. For example, there is a data point in the test set whose network output is close to 35, while the corresponding target value is about 12.

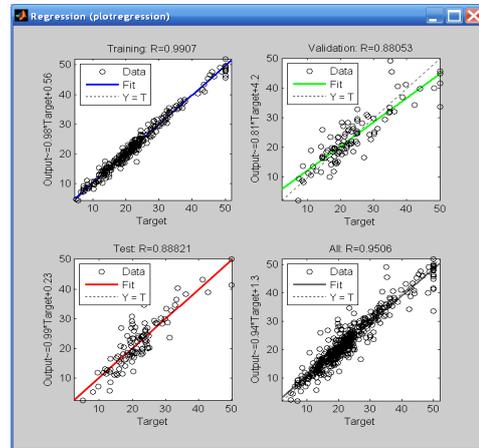


Figure 4: Regression Sample produced by the Neural Network

I. Graphic user interface (GUI) system

Figure 5 illustrates the main GUI of the system. The Main GUI consists of the loaded image, detected face, input quantifications, sliders, output quantifications, and ANN related buttons. The allotted space for the loaded image shows the image to be quantified. The yellow box signifies the Viola-Jones Algorithm, and the face detected [8]. The face quantification or the input quantifications are the original quantification values of the loaded image. This includes the histogram values which is represented by a bar graph.

After logging in, the user now is enabled to load an image that wanted to be edited by clicking the Load Image to Edit button. If multiple faces are detected, the user could click the Next face button and select the face to be edited. Afterwards, the user could use the image parameter sliders located at the left side of the GUI to start adjusting the image parameters of the image such as the Brightness, Contrast, Hue, Saturation, Lightness, Cooling-Warming and Blur-Sharpen. After editing the image based on the user’s preference, the image is ready to be added to the Training set by clicking the Add to Train Set button. The time a user edits an individual photo for training is recorded and averaged to get average manual adjusting time.

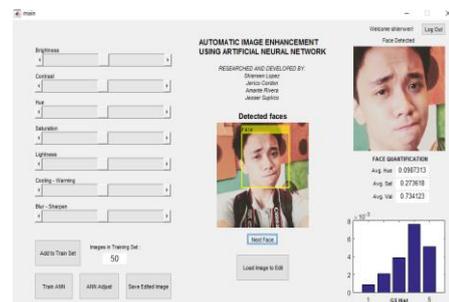


Figure 5: Main Graphical User Interface

III. RESULT AND DISCUSSION

The training set for each user profile consists of 50 images each. Different user profiles were made to create groupings of whether the training sets are composed of a fixed background and random background. These groupings are made to analyse the performance of the artificial neural network on different inputs being made. Each row represents the quantified value of one image, while the column shows what particular image variable is represented. The regression values must ideally be close to 0.8 since it is not too close to the value of 1 which leads to overfitting and it is not too low which would mean that more training sets should be added to fit the user's preference. An additional adjustment has also been added for each of the users instead of only having an image adjustment where the Viola-Jones detected face [8] would be 100% adjusted with its background having 50% adjustment concerning that of the detected face. In this adjustment, the whole image would be 100% adjusted since not all of the adjustments show great outputs.

Figure 6 shows the Regression value between the targeted data and the output data of the adjusted images of the user "Ryan". A regression value of 0.74449 is good enough. This means that the user has at least been very consistent during his adjustments of the slider values since the values on the output are not too far apart to each other which made the regression value not too low and not too high. This means that the neural network would need a few more training sets to analyse the user's preferred adjustment correctly.

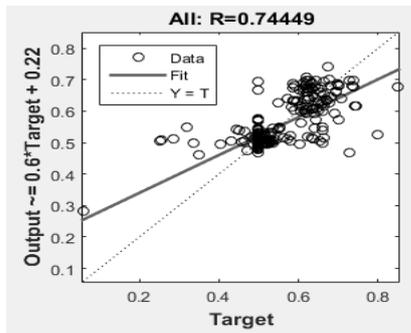


Figure 6: Regression for user: Ryan (White Backgrounds)

Figure 7 shows the sample images with white backgrounds of the user "Ryan", starting from the original images, the images with only 50% background adjustment, and the images with 100% background adjustment.



Figure 7: Image Comparison (White Background 50% vs 100% Background Adjustment – user: Ryan)

Figure 8 shows the Regression value between the targeted data and the output data of the adjusted images of the user "Justine" for a different background. A regression value of 0.689 is a bit low. This means that the user has adjusted the slider values of some of the image to be a bit far from that of the other images or with a bit of inconsistency, which caused the regression to be a bit less than that of the ideal value. This means that the neural network needs some more training set to get further analyse the user's preferred adjustment.

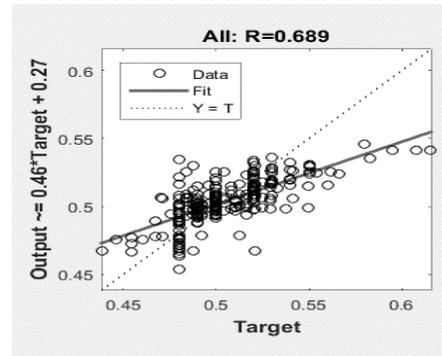


Figure 8: Regression for user: Justine (Different Backgrounds)

Figure 9 shows the sample images with white backgrounds of the user "Justine", starting from the original images, the images with only 50% background adjustment, and the images with 100% background adjustment.



Figure 9: Regression for user: Justine (Different Backgrounds)

Table 1 shows ANN takes the time comparison by the user for manual editing vs the time for automatic editing. Users are asked to edit an image for the training of the ANN. For every image that is edited for the training, a time is recorded. Each user's manual adjusting time is recorded to keep track how fast a user modifies a single image. Based on the results in Table 1, the adjustment time produced by the ANN is way less than the manually adjusted by the users.

Table 1
Time Comparison of Average Manually Adjusted and Automatically Adjusted

User	Average Manual Adjusted (sec)	Automatically Adjusted (sec)
Ryan	33.2	0.69
Justine	38.52	0.71

IV. CONCLUSION

In this paper, we have demonstrated the effectiveness of ANN in automatic photo adjustment based on the preference of the user. Application of automation to photo enhancement procedures significantly decrease time spent on editing photos. The GUI provides easy access and uses for the whole program. A trained neural network only needs 0.70 seconds to enhance an image. Using an automated algorithm is an appealing alternative to manual work, but such an algorithm faces many hurdles. Performance of neural network was tested by comparing the regression fit of ANN trained using fixed background and random backgrounds. Both show good regression fit. The study reveals that a user with consistent adjustment throughout the collection of training sets produce good fit and performance to the ANN meanwhile a user with inconsistent adjustments throughout the collection of training sets produce bad fit. Additionally, it shows that time is a factor when it comes to manually adjusting the image. Thus, the output of this study confirmed that there is a reduction in time and effort when using the auto enhancement system compared to manual editing.

REFERENCES

- [1] L. Feng, L. Xiaoyu, and C. Yi, "An efficient detection method for rare colored capsule based on RGB and HSV color space," in *Granular Computing (GrC), 2014 IEEE International Conference on*, 2014, pp. 175–178.
- [2] Q. Li, U. Niaz, and B. Merialdo, "An improved algorithm on Viola-Jones object detector," in *Content-Based Multimedia Indexing (CBMI), 2012 10th International Workshop on*, 2012, pp. 1–6.
- [3] G. Dong and V. Taslimitehrani, "Pattern-aided regression modeling and prediction model analysis," *IEEE Trans. Knowl. Data Eng.*, vol. 27, no. 9, pp. 2452–2465, 2015.
- [4] T. W. Cenggoro, A. H. Kridalaksana, E. Arriyanti, and M. I. Ukkas, "Recognition of a human behavior pattern in paper rock scissor game using backpropagation artificial neural network method," in *Information and Communication Technology (ICoICT), 2014 2nd International Conference on*, 2014, pp. 238–243.
- [5] C. Zhang, J. Gao, O. Wang, P. Georgel, R. Yang, J. Davis, J.-M. Frahm, and M. Pollefeys, "Personal photograph enhancement using internet photo collections," *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 2, pp. 262–275, 2014.
- [6] S. Ghosh, S. Hazra, S. P. Maity, and H. Rahaman, "A new algorithm for grayscale image histogram computation," in *India Conference (INDICON), 2015 Annual IEEE*, 2015, pp. 1–6.
- [7] M. Yikui and D. Jun, "Fundamentals of intelligent public transportation dispatching systems planning," in *Computing, Communication, Control, and Management, 2009. CCCM 2009. ISECS International Colloquium on*, 2009, vol. 4, pp. 41–44.
- [8] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, 2004.