An Efficient Hardware Implementation of Competitive Learning Classifier for Gender Classification

James Mnatzaganian, *Student Member, IEEE,* and Qutaiba Saleh, *Student Member, IEEE,*

Abstract—TEXT

Index Terms—?, ?, ?.

I NTRODUCTION TEXT

A. Dataset

The Labeled Faces in the Wild (LFW) [1] dataset was modified for use with gender classification. This dataset is comprised of one or more faces per person, with no labels to denote gender. To obtain the genders, the genderize.io [2] API was utilized. This API can be used to obtain the gender, with a confidence interval, for a given first name. Only results with over 90% accuracy were utilized.

The faces in the dataset are not all frontalized nor are they cropped to remove background noise. To remove noise, the frontalized images from Hassner et al. [3] were used as the dataset. This dataset frontalized and cropped all images from the LFW dataset, creating a cleaner dataset.

To ensure that the dataset would not be biased towards a particular individual, a single instance of each person was randomly chosen. The remaining images were then randomized and split by gender, to ensure an equal distribution of males and females. 400 of each gender were used for training and 200 of each gender were used for testing. All code for generating the dataset as well as the software and hardware models are available at [4].

B. Design Approach

The primary design constraints were area and power. With this in mind, the system still had to perform its primary goal of gender classification. It was desired to obtain the highest possible accuracy; however, if a minor loss in accuracy resulted in major hardware improvements, favor was given to the hardware.

1) Algorithm: An algorithm similar to competitive learning was designed. This algorithm is coined Competitive Learning Classifier (CLC). The algorithm is comprised of n competitive learning networks, where n is the number of labels. Each network has i inputs and o outputs, where each output represents a cluster. The implementation of this algorithm used for this work is shown in Fig. 1. It consists of two competitive learning networks (one for males and one for females), 49 inputs per

J. Mnatzaganian, and Q. Saleh are with the NanoComputing Research Laboratory, Rochester Institute of Technology, Rochester, NY, 14623.
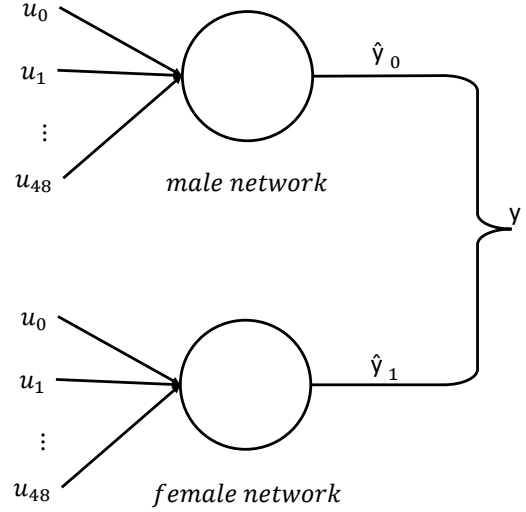
Fig. 1. A CLC consisting of two networks (male and female), 49 inputs per network, and a single output for each network.

network (one for each feature), and a single cluster for the output of each network.

In CLC, there is a randomly initialized weight for each input in each network. The weights are updated using the online update equation in (1), where $w_{i,n}$ is the weight of input i for network n, α is the learning rate, $\hat{y}_{o,n}^{(p)}$ is output o of network n for pattern p, and $u_i^{(p)}$ is input i for pattern p. Each output is computed by finding the squared Euclidean distance between the weight and the input. The corresponding cost function is shown in (2), where P is the total number of patterns and O is the total number of outputs.

If $O > 1$ dead neurons may occur. To combat this, boosting is used, where frequently winning neurons are negatively boosted and infrequently winning neurons are positively boosted. Boosting is only updated during the learning phase, so after training the boost values are fixed.

Training occurs by selecting the network representing the current label and updating its weights. The other networks are left untouched. During testing, both networks receive the same input and compute their respective outputs. The neuron having the lowest value is the winning neuron and the network that it belongs to becomes the output of the network.

$$w_{i,n} \mathrel{+}= \alpha\hat{y}_{o,n}^{(p)}(u_i^{(p)} - w_{i,n}) \qquad (1)$$

$$J(\boldsymbol{w}_n) = \frac{1}{2} \sum_{p=1}^{P} \sum_{o=1}^{O} \hat{y}_{o,n}^{(p)} |\boldsymbol{w}_n - \boldsymbol{u}^{(p)}|^2 \qquad (2)$$

2) Software: TEXT
3) Hardware: TEXT

C. Results

TEXT

I. CONCLUSION

TEXT

REFERENCES

[1] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," University of Massachusetts, Amherst, Tech. Rep. 07-49, October 2007.

[2] "Genderize.io," Available at https://genderize.io, 2012, accessed on 2015-05-04.

[3] T. Hassner, S. Harel, E. Paz, and R. Enbar, "Effective face frontalization in unconstrained images," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[4] J. Mnatzaganian and Q. Saleh, "lfw_gender," Available at https://github.com/jwmqms/lfw_gender, 2015.