Project 2:

Connor McCurley
Deep Learning, Fall 2019
University of Florida
Gainesville, FL, USA 32611
Email: cmccurley@ufl.edu

***Abstract*—**

***Index Terms*—**

I. Introduction

Autonomous image classification is a challenging problem which offers potential for significant advancement in the areas of biometrics, biology, medical diagnosis, security, and more [1], [2]. This paper focuses on the use of dimensionality reduction/ manifold learning in conjunction with multi-layer perceptron artificial neural networks to automatically classify clothing items from the well-known Fashion MNIST dataset [3].

A wide variety of approaches have been taken in attempt to solve detection and classification problems in imagery. [4] used dissimilarity-based classifiers along with metric learning to dually drive samples toward their respective class representatives while also enforcing separation between classes. [5], [6] utilized vector embeddings with linear support vector machines to discriminate between low-dimensional image representations. The work in [7] found sparse weighted combinations of dictionary atoms to accurately reconstruct images where specific bases equated to the various classes. The authors of [8] utilized statistical properties to match samples to generating distributions. The work in [9] employed traditional template matching to locate objects or compositions in imagery. The review in [10] demonstrated that expansive uses of artificial neural networks in image classification. This, of course, is just a small sample of image classification techniques. The reviews in [1], [2] elaborate extensively on the myriad of methods. A commonality among all of the discussed methods is that they suffer from high-dimensionality. Because of this fact, this work explores the use of dimensionality reduction as a preprocessing procedure for classification with fully-connect multi-layer perceptrons.

The remainder of this paper is organized as follows. Section II describes the methodology used to perform dimensionality reduction and classification with multilayer perceptrons. Classification results are presented in Section III. Practical insights to results are given in Section IV. Finally, Section V reveals concluding remarks and discusses future lines of research.

II. Methodology

This section describes the methodology used throughout this work. Analysis of the data is performed, dimensionality reduction techniques are described, various network architectures under analysis are elaborated on and the experimental procedure is outlined.

A. Data Analysis

B. Network Architecture

C. Experiments

Experiments

1) *Baseline:* Model trained on the original data (described in section II).
2) *PCA 100:* Principal vectors estimated from the training set were used to project the data down to 100 dimensions, thus retaining 91.1% of the original variance. The model was the same as the baseline excluding the input layer.
3) *PCA 400:* Principal vectors estimated from the training set were used to project the data down to 400 dimensions, thus retaining 98.5% of the original variance. The model was the same as the baseline excluding the input layer.
4) *UMAP 100:* UMAP was used to project the data down to 100 dimensions. 15 neighbors were used in the local neighborhood and the minimum distance of points in the latent space was constrained to 0.1, which was measured by Euclidean distance. The model was the same as the baseline excluding the input layer.
5) *UMAP 400:* UMAP was used to project the data down to 400 dimensions. 15 neighbors were used in the local neighborhood and the minimum distance of points in the latent space was constrained to 0.1, which was measured by Euclidean distance. The model was the same as the baseline excluding the input layer.
6) *AE 100:* An additional multi-layer perceptron was first trained in the form of an autoencoder. This MLP demonstrated layers as (784-500-400-150-100) in the encoder with the reverse in the decoder. A hyperbolic tangent activation was applied at the output layer. The network was trained with mean-square error loss and updated with Adamax. The original dataset was passed through the encoder to reduce dimensionality. The model was the same as the baseline excluding the input layer.
7) *AE 400:* An additional multi-layer perceptron was first trained in the form of an autoencoder. This MLP demon-

strated layers as (784-500-400) in the encoder with the reverse in the decoder. A hyperbolic tangent activation was applied at the output layer. The network was trained with mean-square error loss and updated with Adamax. The original dataset was passed through the encoder to reduce dimensionality. The model was the same as the baseline excluding the input layer.

III. Results

In this section, classification results for each experimental method tested are presented in the form of confusion matrices. A confusion matrix demonstrates the discrepancies between predicted and true class values for groups of samples. Essentially, it is a way to measure how accurate a classifier is, while providing insight into how the network confuses samples. A diagonal matrix signifies zero mis-classifications among all categories.

A. Confusion Matrices

Figures

TABLE I: Cross Entropy Loss on 10000 Blind Test Samples

Dimensionality Reduction	**Cross-Entropy Loss**
Baseline	0.61
PCA 100	**0.34**
PCA 400	2.23
UMAP 100	0.86
UMAP 400	0.87
AE 100	1.30
AE 400	0.99

IV. Discussion

In this sections, observations are made on results and insight is given to potential influences.

A. Results

B. Potential Improvements

In

V. Conclusions

Three

Future research endeavors toward this topic include,

Honor Statement

* I confirm that this assignment is my own work, it is not copied from any other person's work (published or unpublished), and has not been previously submitted for assessment either at University of Florida or elsewhere.

References

[1] S. Prasad, T. S. Savithri, and I. V. M. Krishna, "Techniques in image classification; a survey," vol. 15, 2015.

[2] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *International Journal of Remote Sensing*, vol. 28, pp. 823 – 870, 03 2007.

[3] H. Xiao, K. Rasul, and R. Vollgraf. (2017) Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms.

[4] T. Mensink, J. Verbeek, F. Perronnin, and G. Csurka, "Distance-based image classification: Generalizing to new classes at near-zero cost," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 11, pp. 2624–2637, Nov 2013.

[5] J. Sanchez and F. Perronnin, "High-dimensional signature compression for large-scale image classification," in *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition*, ser. CVPR '11. Washington, DC, USA: IEEE Computer Society, 2011, pp. 1665–1672.

[6] Y. Lin, F. Lv, S. Zhu, M. Yang, T. Cour, K. Yu, L. Cao, and T. Huang, "Large-scale image classification: Fast feature extraction and svm training," in *CVPR 2011*, June 2011, pp. 1689–1696.

[7] S. Shao, Y.-J. Wang, B.-D. Liu, W. Liu, and R. Xu, "Label embedded dictionary learning for image classification," 2019.

[8] R.Timofte, T. Tuytelaars, and L. van Gool, 2013.

[9] P. R. S. Swaroop and N. Sharma, "An overview of various template matching methodologies in image processing," 2016.

[10] S. B. Driss, M. Soua, R. Kachouri, and M. Akil, "A comparison study between mlp and convolutional neural network models for character recognition," in *Commercial + Scientific Sensing and Imaging*, 2017.