2020

# Fusion-Net: Integration of Dimension Reduction and Deep Learning Neural Network for Image Classification

Mohammad Masum
mmasum@students.kennesaw.edu

Philippe Laval
plaval@kennesaw.edu

# Fusion-Net: Integration of Dimension Reduction and Deep Learning Neural Network for Image Classification

Mohammad Masum
*Analytics and Data Science Inst.*
*Kennesaw State University*
Kennesaw, USA
mmasum@students.kennesaw.e

Philippe B. Laval
*Department of Mathematics*
*Kennesaw State University*
Marietta, USA
plaval@kennesaw.edu

***Abstract*— Building a deep network using original digital images requires learning many parameters which may reduce the accuracy rates. The images can be compressed by using dimension reduction methods and extracted reduced features can be feeding into a deep network for classification. Hence, in the training phase of the network, the number of parameters will be decreased. Principal Component Analysis is a well-known dimension reduction technique that leverage orthogonal linear transformation of the original data. In this paper, we propose a neural network-based framework, named Fusion-Net, which implements PCA on an image dataset (CIFAR-10) and then a neural network applies on the extract principal components. We also implemented logistic regression on the reduced dataset. Finally, we compare between results of using original features and reduced features. The experimental results show that Fusion-Net outperformed other methods.**

***Keywords—neural network, principal component analysis, image classification***

## I. Introduction & Related Works

Building a deep network using original digital images requires learning many parameters which may reduce the accuracy rates. The images can be compressed by using dimension reduction methods and extracted reduced features can be feeding into a deep network for classification. Hence, in the training phase of the network, the number of parameters will be decreased. Principal Component Analysis is a well-known dimension reduction technique that leverage orthogonal linear transformation of the original data. Machine learning algorithms have been showing promising results in classifying images. Machine learning approaches like Support Vector Machine (SVM), Logistic Regression (LR), and Decision Tree (DT) were previously proposed for image classification.

Many dimension reduction techniques have applying for compressing image without compromising the quality of the image. SVD-WDR and SVD-WDR with PCA are applied to image compression where the image is initially compressed by WDR technique and then wavelet transform is applied on it [1]. NMF (Non-negative Matrix Factorization) and PCA are used to transformed images into low dimensional feature space and then a graph-based ranking algorithm is implemented to classify images of 20 classes [2]. The combined approach achieved maximum accuracy comparing with other individual classifier (NMF, PCA, graph-ranking). An iterated large-margin discriminant technique is applied on medical image for dimension reduction of the images [3].

Neural networks are currently widely used for many applications due to the capability of highly non-linear systems and flexibility in architecture design. It has been showing promising results in classifying images in recent years. An artificial neural network was applied to classify specific features of a tooth on 3d scan data [4]. A compress learning method is applied to MNIST and CIFAR dataset for image classification [5]. In the compress learning method consists of a sensing stage is followed by an inference stage where an end-to-end convolutional neural network is applied.

In this paper, we propose a deep neural network framework named Fusion-Net for image classification. Our contributions include: (1) We implement dimension reduction technique (PCA) to the original dataset to reduce the number of parameters that learned in the training phase of a neural network. and (2) Fusion-Net provides high accuracy score based on deep neural network architecture which suggests that classifying using deep learning technique is promising.

The rest of the paper is organized as follows: Section II describes the methodology of our proposed method Fusion-Net along with the logistic regression classifier. The experimental setting and results are explained in Section III. Finally, Section IV concludes the paper.

## II. Methodology

### A. Principal Componnet Analysis

Principal Component Analysis (PCA) is a dimension reduction technique to suppress excess information of original data and transforming data into fewer dimension while preserving the trends and patterns of underlying information. It is an unsupervised machine learning method that discovers patterns of the data without utilizing the information of label of the observations [6].

PCA extensively leverages orthogonal linear transformation method to convert the original data to a new coordinate system so that the first coordinate (first principal component) in the new space expresses the maximum variance by some scalar projection of the data [7]. The second coordinate (second principal component) expresses second largest variance of the data and consequently, remaining components express variance of the data in a descending order.

Let, $X$ be a matrix with column-wise zero mean of a given data of $n$ observations and $p$ features and $x_i$ is row vector of $X$. Then, a set of $p$ dimensional vectors of weights $w_k = (w_1, w_2, \dots, w_p)_k$ can be determined to transform the data by mapping each of the row vectors $x_i$ to a new coordinates or principal components $t_i = (t_1, t_2, \dots, t_l)_i$. Mathematically, we can write:

$$t_{k(i)} = x_i w_k \;\; for\ i = 1, \dots., n;\ k = 1, \dots., l$$

Where each weight vector is a unit vector.

Let $w_1$ is the first weight vector, then to obtain first principal component of the data $w_1$ must satisfy the following optimization criteria:

$$w_1 = \arg \max_{\{||w=1||\}} \{\sum_i (t_1)_i{}^2\} = \arg \max_{\{||w=1||\}} \{\sum_i (x_i \boldsymbol{w})^2\}$$

## B. *Logistic Regression*

Logistic regression is a classical classifier of supervised learning. It utilizes the sigmoid function to squeeze the output of a linear equation between 0 and 1. Thus, the output of logistic regression can be used to predict the probability of a class [8]. Fig. 1 shows an example of a sigmoid function.

## C. *Neural Network*

At present, neural networks are widely used for many applications due to the capability of highly non-linear systems and flexibility in architecture design. The neural network's basic architecture contains input layers, one or more hidden layers, and output layers where each of the layers includes a certain number of neurons. Weighted linear combination of neurons of a layer is computed and then used as input to another neuron in the succeeding layer. To capture the non-linearity of the data, a non-linear function, called activation function, can be applied to the weighted sums of neurons. All the weights of a neural network are set to random values at the initial stage of training. Data is fed into the input layer of the network, then it travels through the hidden layers, and finally output is produced in the output layer. The network continually updates the weights applying backpropagation based on the output and desired target of the neural network. The network consequently reduces the error between the output and target in each iteration. In the process, a loss function is used to calculate the error of the network and the error is minimized by applying optimization function during backpropagation.

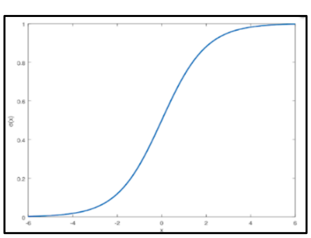

*Fig. 1: An example of sigmoid function*

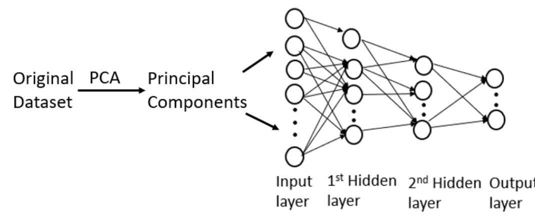


*Fig. 2: Architecture of Fusion-Net*

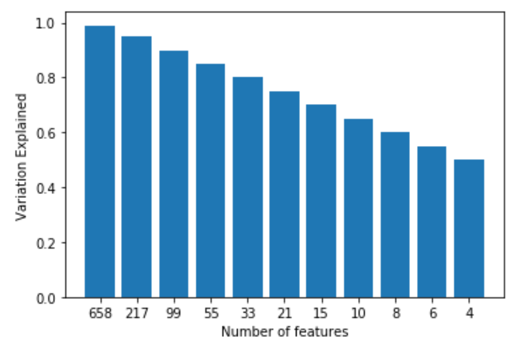


*Fig. 3: Variation explained by Principal Components*

## D. *Fusion Approach*

In this paper, we experimented both logistic regression and neural network methods with a fusion approach. In the fusion approach, we at first reduced dimension of the original dataset leveraging PCA and later applied logistic regression and NN methods. We also experimented the model performance by varying the number of principal components that were extracted by PCA.

In this paper, we propose a Fusion-Net framework to classify the images. The Fusion-Net consists of two major portions: dimension reduction and neural network. Fig. 2 shows the architecture of Fusion-Net where the neural network contains four layers: input layer, two hidden layers, and output layer. The input layer contains neurons which is the number of features extracted by PCA. We present three different versions of the fusion approach where the reduced features preserve 99%, 95%, and 90% of variation of the data for first, second and third version respectively.

We transformed each of the color images of three dimension into one dimensional space so that we can apply PCA. After the conversion, each of the observations consists of $32 \times 32 \times 3 = 3072$ features and therefore PCA is implemented on the converted dataset. Fig. 3 shows the variation explained by the number of features. 99%, 95%, and 90% of variation of the original data can be explained by first 658, 217, and 99 number of principal components respectively. 80% of the variation of the data can be extracted by using only the first 21 number of components.

The first and second hidden layer contains 128 and 64 neurons respectively, and the output layer includes 10 neurons since the problem is a 10-class classification. We applied categorical cross-entropy as loss function and Adaptive Moment Estimation (Adam) optimizer for calculating eerror and updating the parameters.

We implemented logistic regression and neural network with the same architecture on the original dataset and finally compare the results with the proposed Fusion-Net.

## III. EXPERIMENT & RESULTS

### A. Dataset specification

We performed all experiments on an image dataset, named CIFAR. The CIFAR data set is a well-known image data for multi-class classification of 10 classes of images which are collected by Alex Krizhevsky, Vinod Nair, and Geoffrey Hinton [1]. The images in the dataset have 10 different classes of object: airplane, automobile, bird, cat, deer, dog, frog, horse, ship, and truck. Fig. 4 illustrates some of the sample images with class labels of the images.

In this paper, we used the CIFAR-10 dataset consists of 60,000 images which is a subset of 80 million tiny images dataset (CIFAR). Each of the images in the dataset is color image (containing 3 channels) with $32 \times 32$ dimension. CIFAR-10 dataset is a balanced data where each of the classes contains 6,000 images. We split the dataset into 50,000 training images and 10,000 test images preserving the proportion of classes.

### B. Model evaluation metrics

The dataset is balanced. Therefore, we consider the "accuracy" metric to assess the performance of the models. Accuracy is the most intuitive performance metric which is the proportion of the number of correctly predicted images to the total number of images.

$$Accuracy = \frac{number\ correctly\ predicted\ observation}{total\ number\ of\ observation}$$

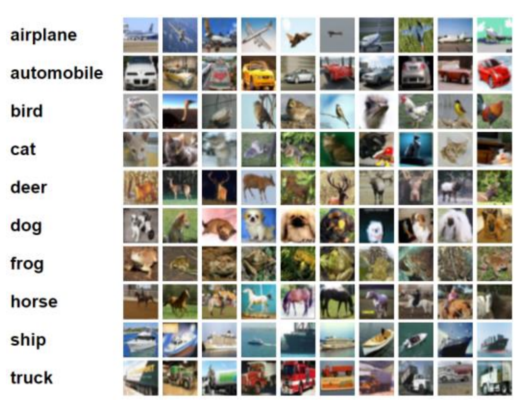


*Fig. 4: Variation explained by Principal Components*

### C. Experimental Design

We evaluated our model performance by comparing performance of LR, and Neural Network methods with original dataset. Trained data was used to train each of the models we experimented with while test data was used for evaluating the performance of the models.

The LR and Fusion-Net classifiers were applied to both datasets (original and reduced) for comparing results with our proposed Fusion-Net. The algorithms were implemented using Python scikit-learn library with available hyperparameter options.

We used 'ReLu' activation function in the hidden layer and 'softmax' function in the output layer. 'Adam' and 'categorical cross-entropy' were used for optimizer and loss function respectively. We set the number of epochs to 100 and implemented an early stopping method to stop training once the model performance stops improving on the test data. We selected validation loss to be monitored for early stopping and set minimum delta to $1e-4$ (checks minimum change in the monitored quantity to qualify as an improvement) and patience to 5 (checks number of epochs that produced the monitored quantity with no improvement after which training will be stopped). Mini-batch gradient descent was considered and a batch size of 64 was chosen to train the model. The initial learning rate was set to 0.0001. The $L^2$ regularization technique was applied to the output of the hidden layer to prevent the network from overfitting and the regularization parameter 'lambda' was set to 0.000001. All the parameters and hyperparameters used in the model were randomly optimized.

We implemented our experiment on Keras framework in Python 3.7 version.

### D. Experimental results

We compared the results of Fusion-Net with the other classifiers with reduced and original datasets. The accuracy score was used to evaluate the models' performance. The same configuration was applied to both datasets for maintaining consistency.

*1) Performance evaluation on CIFAR-10 original dataset*: We trained the neural network for 100 epochs with early stopping and Logistic Regression also applied on the original dataset. Model performance was measured by implementing the trained model on the test data. Table I illustrates the experimental results of the two models on this dataset. Neural Network outperformed LR method and achieved the highest accuracy score of 0.46, whereas the LR achieved 0.3954 accuracy score.

*2) Performance evaluation on CIFAR-10 reduced datasest:* We impleneted PCA on the original dataset and reduced the dataset and extracted three different datasets containing 658, 217, and 99 features which explain 99%, 95%, and 90% of variation of the original data respectively. We applied the neural network with the same architecture on the reduced datasets. Logistic Regression also applied on the reduced datasets. Table II illustrates the experimental results of the models on the datasets. Fusion-Net outperformed (PCA+LR) method and achieved the highest accuracy score for three different cases. With 99 principal components Fusion-Net achieved maximum accuracy of 53.41%. Almost similar accuracy of 0.5307 was obtained with 217 features. Fusion-Net achieved 0.5194 accuracy using 658 principal components. Logistic regression achieved almost similar accuracy for the three reduced datasets.

### *E. Discussion*

Neural network produced better results with reduced dataset than the original dataset. With reduced features, the network learns less number of weights which might be the reason of producing better results since we trained model with 100 epochs. A fine tuned neural network model with many epochs may outperform Fusion-Net.

*Table I: Experimental results of different classifiers on CIFAR-10 Original Dataset*

| ***Classifiers*** | ***Accuracy*** |
|---|---|
| LR | 0.3954 |
| NN | 0.4640 |

*Table II: Experimental results of different classifiers on CIFAR-10 Reduced Datasets*

| ***Classifiers*** | ***Accuracy with 658 features*** | ***Accuracy with 217 features*** | ***Accuracy with 99 features*** |
|---|---|---|---|
| PCA+LR | 0.4177 | 0.4120 | 0.4046 |
| **Fusion-Net** | **0.5194** | **0.5307** | **0.5341** |

## IV. Conclusion

Dimension reduction techniques can be applied to an image dataset to reduce the number of features while preserving patterns and trends of the original dataset. Deep learning network can be trained on the reduced dataset instead of the original data so that the number of training parameters. Therefore, the performance of the model might be increased. In this paper, we proposed a neural network-based framework, Fusion-Net, for image classification. We applied PCA for dimension reduction and then trained Fusion-Net with $L^2$ regularization technique, early stopping criteria and mini-batch gradient descent method. We performed all the experiments on datasets (original and reduced) evaluate Fusion-Net. We evaluated Fusion-Nets' performance by comparing it with the performance of fusion LR (PCA+LR) approach. The experimental results show that Fusion-Net outperformed other methods and achieved the highest accuracy score.

## References


[1] Agarwal, Karishma, Arpit Bansal, and Mukesh Rawat. "Image Compression Techniques Comparative Analysis using SVD-WDR and SVD-WDR with Principal Component Analysis." *International Journal on Recent and Innovation Trends in Computing and Communication* 6.2 (2018): 111-116..

[2] Nan, Yao, Qian Feng, and Sun Zuolei. "Image Classification by Feature Dimension Reduction and Graph based Ranking." *arXiv preprint arXiv:1304.2683* (2013).

[3] Jingyan Wang, Yongping Li, E. Marchiori, Chao Wang: Iterated Large-Margin Discriminant Analysis for feature Dimensionality Reduction in medical image retrieval, in: International Journal of Biomedical Engineering and Technology (2011), Volume: 7, Issue: 2, Pages: 116-34, DOI: 10.1504/IJBET.2011.043174.

[4] Raith, Stefan, et al. "Artificial Neural Networks as a powerful numerical tool to classify specific features of a tooth based on 3D scan data." *Computers in biology and medicine* 80 (2017): 65-76.

[5] Zisselman, E., A. Adler, and M. Elad. "Compressed Learning for Image Classification: A Deep Neural Network Approach." *Processing, Analyzing and Learning of Images, Shapes, and Forms* 19 (2018): 1.

[6] Lever, J., Krzywinski, M. & Altman, N. Principal component analysis. *Nat Methods* **14,** 641–642 (2017) doi:10.1038/nmeth.4346

[7] Jolliffe I.T. *Principal Component Analysis*, Series: Springer Series in Statistics, 2nd ed., Springer, NY, 2002, XXIX, 487 p. 28 illus. ISBN 978-0-387-95442-4

[8] Afrin, R., Haddad, H., & Shahriar, H. (2019, July). Supervised and Unsupervised-Based Analytics of Intensive Care Unit Data. In *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)* (Vol. 2, pp. 417-422). IEEE.