❒   750

Pre-convoluted neural networks for fashion classification

Mustafa Amer Obaid, Wesam M. Jasim
College of Computer Science and Information Technology, University of Anbar, Iraq

Article Info

ABSTRACT

In this work, concept of the fashion-MNIST images classification constructed on convolutional neural networks is discussed. Whereas, 28×28 grayscale images of 70,000 fashion products from 10 classes, with 7,000 images per category, are in the fashion-MNIST dataset. There are 60,000 images in the training set and 10,000 images in the evaluation set. The data has been initially pre-processed for resizing and reducing the noise. Then, this data is normalized for ensuring that all the data are on the same scale and this usually improves the performance. After normalizing the data, it is augmented where one image will be in three forms of output. The first output image is obtained by rotating the actual one; the second output image is obtained as acute angle image; and the third is obtained as tilt image. The new data set is of 180,000 images for training phase and 30,000 images for the testing phase. Finally, data is sent to training process as input for training model of the pre-convolution network. The pre-convolution neural network with the five layered convoluted deep neural network and do the training with the augmented data, The performance of the proposed system shows 94% accuracy where it was 93% in VGG16 and 92% in AlexNetnetworks.

Corresponding Author:

Mustafa Amer Obaid
College of Computer Science and Information Technology
University of Anbar
Anbar governorate-Ramadi-Republic of Iraq
Email: must85afa@uoanbar.edu.iq

1. INTRODUCTION

Nowadays, statistical and machine learning techniques undergoes to some of the reasonable techniques to withstand certain constraints in the estimation of high-dimensional data. The previous setbacks can be solved by reducing the number of input variables before applying the data mining techniques [1]. Therefore, in machine learning the dimensionality reduction techniques analyzed in two different ways; basically, the required variables was extracted from the primary dataset, this techniques was formulated as feature selection. Otherwise; by exploring the redundant input data and eliminating the data to form a new set of smaller input variables dataset. Each and every column is the combination of the input variables, and the information is same as the input variables. These techniques were depicted as dimensionality reduction. The above process is exploration techniques in statistical modeling [2].

Han Xiao *et al.* [3] were determines the fashion-MNIST images to be the grayscale images of 28×28 with 70,000 fashion products commencing in 10 classes, through 7,000 images for each class. The first 60,000 images for the training, along with the final 10,000 images for the test phase. This original MNIST dataset remains highly remarkable for machine learning techniques that could be served by the alteration of direct drop-in. the similar image size has been shared along with their design for splitting the training and testing sets. Several researches were introduced for the fashion-MNIST images data set classification. ArildNøkland *et al.* [4] were explains the training of the layer wise, which could arrive at up to date ranges in

image datasets. Hirata *et al.* [5] were proposed a model comprised of one base convolution neural network (CNN) along with multiple fully connected sub networks (FCSNs) known to be EnsNet. This model creates the group of feature-maps through the final layer of convolution through the CNN base that has been classified and the channels comprised within the disjoint subsets. In FCSNs, the subsets have been allotted with the disjoint subsets. Each subset of the FCSNs is prepared autonomously with the goal that it can anticipate the class name from the subset of the element maps relegated to it. The yield of the general model was controlled by the dominant part vote of the base CNN and the FCSNs. Trial results utilizing the MNIST, fashion-MNIST and CIFAR-10 datasets show that the anticipated technique additionally enhances the demonstration of CNNs. Specifically, an EnsNet accomplishes the best in class mistake pace of 0.16% on MNIST. The effect on the fashion-MNIST dataset of various hyper-parameter optimization (HPO) strategies and regularization procedures with deep neural organizations was studied in [6]. As profound learning requires a heap of information, the inadequacy of picture tests can be grow through different information expansion techniques like rotation, cropping, shifting and flipping. Therefore outcomes of the validated results show great outcomes on this new benchmarking dataset. A cutting-edge model for grouping of fashion article pictures was proposed in [7]. The authors prepared three convolutional neural organizations to characterize the pictures in the fashion-MNIST dataset. The model shows amazing outcomes on the benchmark dataset. LeNet-5 designed network was applied on the fashion-MNIST datasetto direct an extensive correlation between various CNN structures, (for example, VGG16) on different garments datasets, (for example, Image Net). For instance, a custom VGG 16 CNN type with stacked convolution layers obtained 93.07 percent accuracy [8]. Various models of CNN were introduced to distinguish which of them presents better exactness in characterization and ID. LeNet-5, AlexNet, VGG-16 and ResNet were the deep learning architectures. The tensor stream was used to construct the separation work planning period.

The initial segment of the work was centered on a similar investigation of various CNNs for a similar informational index to classify product images. The traditional Alex Net CNN style with stacked convolution layers obtained 92.34 percent accuracy [9]. Deep learning architectures on basis of neural network were trained to characterize pictures on standard fashion-MNIST and CIFAR-10 dataset. The different CNN-based arrangement design and RNN-based characterization engineering were prepared just as tried on those standard datasets. In the case of group classification, either training from scratch or fine tuning is choices for researchers. Classification based on pre-training obtained a performance value of around 88 percent [10]. An assessment of preparing size effect on approval exactness for an upgraded CNN was presented in [11]. The authors were utilized Amazon's AI environment to prepare and test 648 models to locate the ideal hyper parameters to apply a CNN towards the fashion-MNIST dataset to get 91.08 percent accuracy [12].

Different CNN activation functions were tested for fashion-MNIST image classification in [13]. This data set was also tested based on Long Short-Term memory network which obtained accuracy of 88.26% in [14]. Moreover, it was examined based on histogram of oriented gradient and multiclass support vector machine in [15]. It was showed 86.53% accuracy. The fashion-MNIST data set was modified in [16] and a part of it was used in [17] to find the optimal CNN network. It was found that the use of 40% of the data obtained 90% validation accuracy. A. Ferreira *et al* [18], the CNN was used to classify the granite tiles in several conditions. It was also used in [19] for person objects wearing detection such as, hat, shoes and bags. Changing the hyperparameters and filter size of the network leads to a higher performance in [20]. E. Dufourq *et al* [21], the authors proposed an evolutionary deep networks algorithm that includes deep network and genetic algorithm strengths for image classification. The obtained results were good even in using one GPU.

It is clear that from the above literature, in the fashion domain databases were not of enough size and the obtained accurccy still below the requirments in several applications. Therefor; the main contribution of this work is to make a good benchmark dataset with all of MNIST's accessibility, namely its straightforward encoding, permissive license and its small size. This is attained by increasing the number of images with three image types. The first one is by rotating the image, the second one is obtained as acute angle image, and the third one is the shift image. This new augmented data was used as input in training model of pre-convolution neural network. Moreover, our CNN is enhanced by adding third layer to the convolution layers and the max pooling is introduced for all of them. The remaining sections are systematized as follows. Section 2 presents a brief history of neural networks with its components and the used metrics in performance analyzing. Section 3 illustrates the proposed architecture and the standard data set with its augmentation. In section 4 the proposed CNN with pre-convolution is discussed. In section 5 the used metrics in performance analysis is illustrated with the obtained simulation results. In section 6 the relevant conclusions is discussed.

2.   CONVOLUTIONAL NEURAL NETWORK

Convolutional neural network is a deep learning model based on neural network. It extensively well demonstrated techniques to utilize the image classification and the object detection. It incorporates [22] with three main layers; convolutional (Conv) layer, pooling, and fully-connected (FC) layers. The convolutional layer includes filter, stride, and padding [23]. The CNN main structure is illustrated in Figure 1.

−   Convolution layer-the input image is resizing to the CNN model of standard size, i.e., 3×224×224. The convolution is a stag with a series of mathematical operation; it executes by convoluting sliding kernel matrix over the input matrix to extract the features and maps it to the consecutive layers. The outcomes are cumulated to get the feature matrix.
−   Pooling layer-pooling layer is the consecutive layer of the convolution and it is used to reduce the spatial domain of the representation and the computation in the network [24, 25]. The max pooling usually the size of the pooling kernel which is 2×2 and the stride is 2.
−   Fully connected layers: These layers are simulated in CNN with the help of convolution. The format of its size is n1× n2, where n1 is a triplet (7×7×512) output tensor and n2 is an integer output tensor.
−   Dropout: It generally used to remove the over fitting in the input. It is a technique to enhance convolution of deep learning algorithm. It assigns weights to the linked nodes in the network.
−   Softmax: The deep learning model followed by a stack of layer where the convolution layer is subsequently followed by a ReLU layer in CNN. The nonlinearity in CNN model is determined by a ReLU layers.
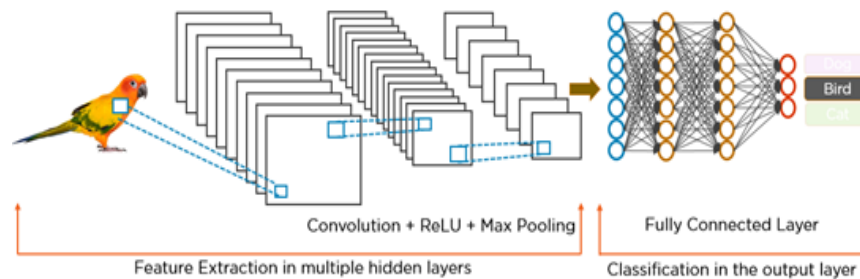

Figure-1 Architecture of convolutional neural network

3.   PROPOSED ARCHITECTURE AND DATA SET

In this paper, the dataset was taken from fashion-MNIST data as input for both training and testing phases. First the input data has been preprocessed for resizing the image and to filter out the noise. This filtered data was augmented, where the image has been rotated, shifted and tilted to obtain 3 various sets of input. The images were sent to augmented data generator to obtain the training set of 60,000 images *3 and the testing set of 10,000 images*3. Then, the generated data set was pre-convoluted using CNN with its softmax. Its output is used as the trained data, whereas the tested output is compared with the trained data and then the prediction process takes place. The output was the 10-classes output. Finally, the accuracy was measured from this classified output. Figure 2 shows our proposed architecture.
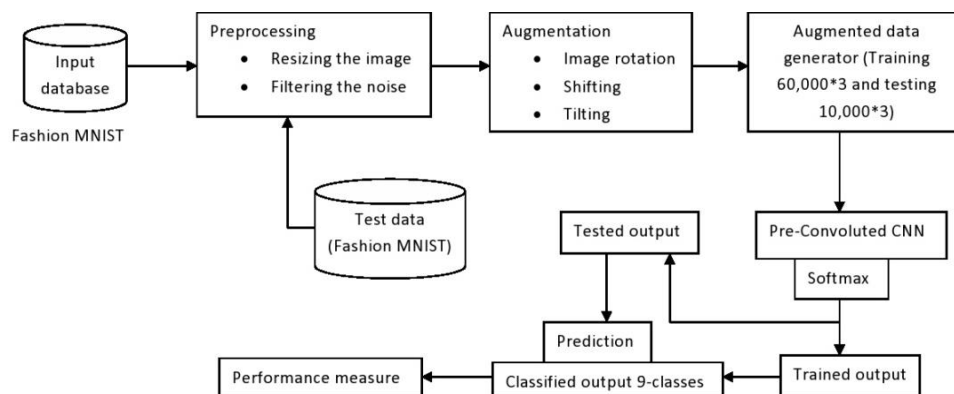

Figure 2. The proposed architecture

4. PROPOSED CNN WITH PRE-CONVOLUTION

The initial strategy in neural network is focused in accumulating the process of the data. The input data dimension looks equivalent to the 2D image to be used in CNN pre-convolution. Pre-convolution is used becuase deep convolutional neural network models make days or even weeks to train on very large datasets, a way to short-cut this process is to re-use the convluted trained model weights that were designed for standard convolution. The CNN has been widely applied for image recognition along with pattern detection and image classification. The image is measured by pixels of the matrix based on it is shape. The distant training in pre-convolution CNN is utilized in image detecting and classification. CNN is a convolution network where its operation of convolution is generally carried out through the input I and the kernel K generates output which utilized in getting the knowledge of shapes modification. The sum of weight has been convoluted as sliding window through the entire image. The overall convolution process is to convert the image through a weighted matrix to become another image of the similar size and convention dependent. Moreover, the operation of convolution is to extract the feature map. In extraction of image characteristics, the first layer used is the convolution layer. The mathematical function needs multiple inputs; two input namely image matrix and filter or kernel as shown in (1) [9].

$$S\,(i,j) = (I * K)(i * j) = \sum_{m}\sum_{n} I(m,n)\,K\,(i - m, j - n) \tag{1}$$

For 3*3 kernel, (1) becomes,

$$S\,(i,j) = (I * K)(i * j) = \sum_{m=1}^{3}\sum_{n=1}^{3} I\,(i - m, j - n)K\,(m,n) \tag{2}$$

A non-linear activation function is functional to the input neurons and it is multiplied with dot product of (2) in the next phase. The result will be x=max (0,x) and the operation is,

$$A(i,j) = \begin{cases} 0, & if\ (i,j) < 0 \\ S(i,j) & if\ (i,j) \geq 0 \end{cases} \tag{3}$$

In convolution, the computation part has been given with the same signals also by using this process the image has been identified and classified. The CNN has multiple convolution layers which this means many altered convolutions have been generated. For calculation the matrix weight has been used and finally the tensor forms of 5x5xn, where n is indicated as convolution numbers of CNN is obtained. The proposed CNN is a pre-convoluted neural network, and therefore the function of distance is trained for evaluating identities among the fashion images. Semantic noise problem has been managed with optimization using CNN. The fashion images calculating are evaluated using CNN in classification obtains the optimal accuracy. The pre-convolution and fine tuning has been done in preparation of distance function that has been assisted with the performance.

The proposed CNN model has five network models that process on datasets of three convolutions discussed above is shown in Figure 3. Initially the network model 1 will have single 128 filter layers. Its kernel size is 3*3. Then ReLU layer activated and process the convolution. Next for maxpooling 2*2 kernel size is applied. Their dropout will be 20% that is utilized to prevent overfitting. At the end, computation of loss function is done through implemented the softmax layer.

The network model 2 consists of 32 and 64 filters with 2 convolutional layers correspondingly. In both layers they use 3*3 size for kernel. The activation of ReLU has been utilized similar to the initial network model for every layer. For the network model 2 the kernel size for maxpoolingused of 2*2 in every layer. The dropout is 20% is used for the first layer and 25% of dropout has been utilized in layer 2 in prevention of over fit. The loss function has been evaluated by the softmax which is deployed at the end. Network model 3 consist of 32,64, and 128 filters for three convolution layers correspondingly. Layer1 uses 5*5 as kernel size. Layer 2 and layer 3 uses 3*3 kernel size. All three layers is implemented in ReLU Activation. After layer 1 max pooling is functioned as Kernel size 3*3 and after layer 2 and 3 the maxpooling is implemented as 2*2 kernel size. After implementation of layer 1 and 2 there is 25% of dropout, the dropout of 25% is taken from leyars 1and 2. The dropout of 20% is taken from leyar 3 to reduce the overfitting. Finally the loss functionis calculated best on fully connected fully network.

The network model 4 and 5 are on the basis of the pre-convoluted CNN. In general CNN are unstable sometimes, eventually during gradient propagation by the extended window that might generate gradient desertion and over burst. Hence the design developed uses pre-convolution network model. A sequential model is used by the network model 4 which remains stack at linear layer. On 150 batch size the

model has been tested. Layer 1 is 128-unit pre-convolution layer and reverts back as series. This has to ensure that the series has been received at next convolution layer and not indiscriminately discretion of data. Therefore, the final layer of softmax was enabled by a fully connected layer and equivalent neurons were it explores the different classes. The network model 5 is same as network model 4 since the network layer 5 also consist of stack of linear layers. Layer 1 is a layer of pre convolution that has 300 units of memory in spot of previous network models that have 128 units and revert back as series. In layer 2 there are 300 units. The dropout layer is implemented to prevent the overfitting of the model after pre-convolution layer 2. At the end, the last layer has softmax activation with fully connected layer.
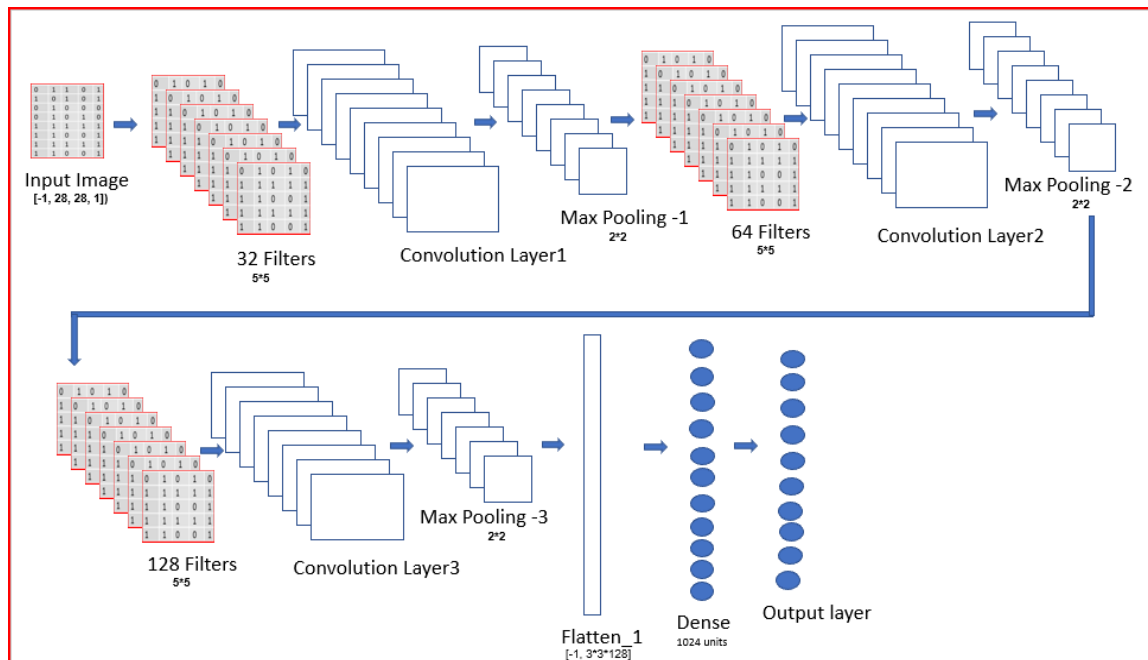

Figure 3. Implementation architecture for proposed pre-convolution CNN

5.   SIMULATION RESULTS
The performance analysis of the proposed method is illustrated below. The parameter to be considered for evaluation is accuracy, precision, recall, and F1 score. Confusion matrix is used to calculate the various performance metrics.
− Accuracy: It shows correctly the classified instances percentage in course of classification. It is evaluated as in (4).

$$Accuracy\ rate = \frac{True\ Positive + Negative}{Total\ instances} * 100 \tag{4}$$

− Precision: It gives the proportion of data that transmit to the network, actually had intrusion. The predicted positives (Network predicted as intrusion is TP and FP) and the network actually having an intrusion are TP. This is used to measure the quality and exactness of the classifier as shown in (5).

$$Precision = \frac{True\ positive}{True\ positive + False\ Positive} \tag{5}$$

− Recall: Recall is the ratio of real positives which are correct the predicted positive and it is defined as in (6).

$$Recall = \frac{True\ Positive}{True\ positive + False\ Negative} \tag{6}$$

– F1 score: F1 score is determined from the precision and recall of the test values as in (7). The F1 score is an indicator of the test's accuracy to assess the binary classification. Where the precision is the number of true positive outcomes, divided by the number of all the positive outcomes, the recall is the number of true positive outcomes, divided by the number of all the positive samples that should have been detected.

$$F1 - Measure = \frac{2x\ Precision\ x\ Recall}{Precision + Recall} \qquad (7)$$

Table 1 illustrates the observation of classification analysis of ten classes of fashion-MNIST dataset. The performance measures of various techniques of VGG16 26M parameters, 3 Conv+BN+pooling, 2 Conv+pooling is compared with the proposed techniques Proposed_PreConv. It shows the comparison of the performance in precision, recall and F1-score. It has been analysed from the actual and predicted value from the objective of ten classes in confusion matrix. The obtained precision, recall, F1-score, and accuracy are compared in Figures 4-7 respectively. In Figure 4, the Proposed_PreConv achieves precision with supreme percentage in all the ten classes than the existing techniques. Figure 5 and 6 depicts very clearly comparison for various techniques in terms of recall and F1-score.

Table 1. Comparison of performance of Proposed_PreConv system and existing algorithm

CLASS 0-T-shirt/TOP, CLASS 1-Trouser, CLASS 2-Pullover, CLASS 3-Dress, CLASS 4-Coat, CLASS 5-Sandal, CLASS 6-shirt, CLASS 7-Sneaker, CLASS 8-Bag, CLASS 9-Ankle Boot

| NO	Classes	VGG16 26M parameters			3 Conv+BN+pooling			2 Conv+pooling			Proposed_PreConv		
		Precision (%)	Recall (%)	F1-score (%)	Precision (%)	Recall (%)	F1-score (%)	Precision (%)	Recall (%)	F1-score (%)	Precision (%)	Recall (%)	F1-score (%)
1.	0	87	86	85	85	84	82	83	82	81	90	89	89
2.	1	95	93	91	93	91	89	91	89	87	100	99	99
3.	2	89	91	85	87	89	83	85	87	81	91	92	91
4.	3	91	93	91	89	91	89	87	89	97	93	95	94
5.	4	89	87	85	87	85	83	85	83	81	91	91	91
6.	5	95	93	93	93	91	91	91	87	88	99	98	99
7.	6	81	80	78	79	78	75	77	75	73	83	82	82
8.	7	93	95	94	91	93	92	87	91	90	95	98	97
9.	8	97	93	91	95	91	89	93	89	87	99	99	99
10	9	97	93	95	95	91	91	93	89	87	99	96	97
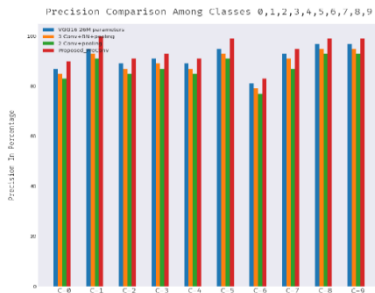
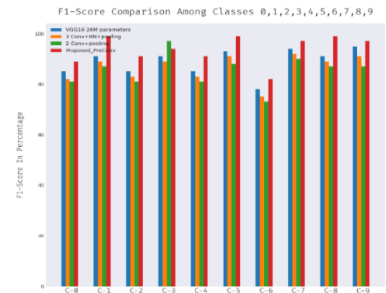Figure 4. Precision comparison
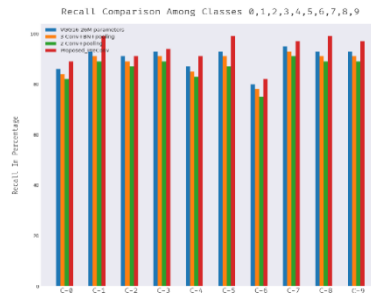
Figure 5. Recall comparison

Figure 6. F1-score comparison

In Figure 7, the proposed Proposed_PreConv utilizes maximum accuracy than existing techniques. The VGG16 26M parameters approach has resulted to worst performance by furnishing a minimum accuracy of about 93.4%. Simultaneously, 2 Conv+pooling approach has resulted to worst performance by furnishing a minimum of accuracy value of about 92.4%. The 3 Conv+BN+pooling model acquires lessa accuracy score compared to the previous one of about 91.8%. It is clear that the Proposed_PreConv technique operates more efficiently when compared with other models by acquiring the maximum recall, accuracy, precision, and F1 score. Figure 8 depicts the training and validation accuracy for the Proposed_PreConvtechnique. The accuracy increases with the epochs exponentially and attain the maximum accuracy at 125 Epochs. The training and evaluation accuracy is predicted at the same epochs.
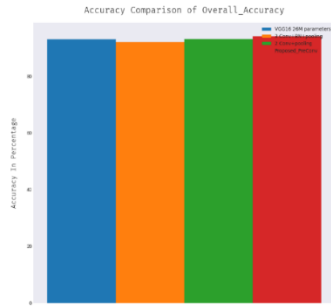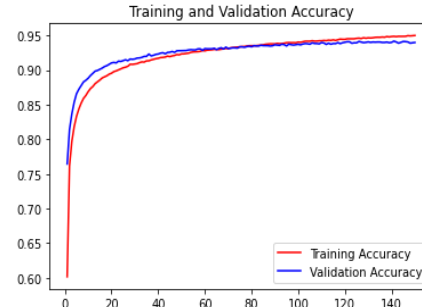
Figure 7. Accuracy comparison

Figure 8. Training and validation accuracy for the possible epoch steps

5.1. Confusion matrix

The table frequently applied to determine a classification model's output on a collection of test data for which the true values are known as the confusion matrix. Therefore, the estimated performance of our classification model on the data using confusion matrix is shown in Figure 9.

Figure 9. Confusion matrix for Proposed_PreConvtechnique

5.2. Correctly predicted classes

Figure 10 determines the outcomes pictorially that the projected model predicts the predicted classes and the correct classes.

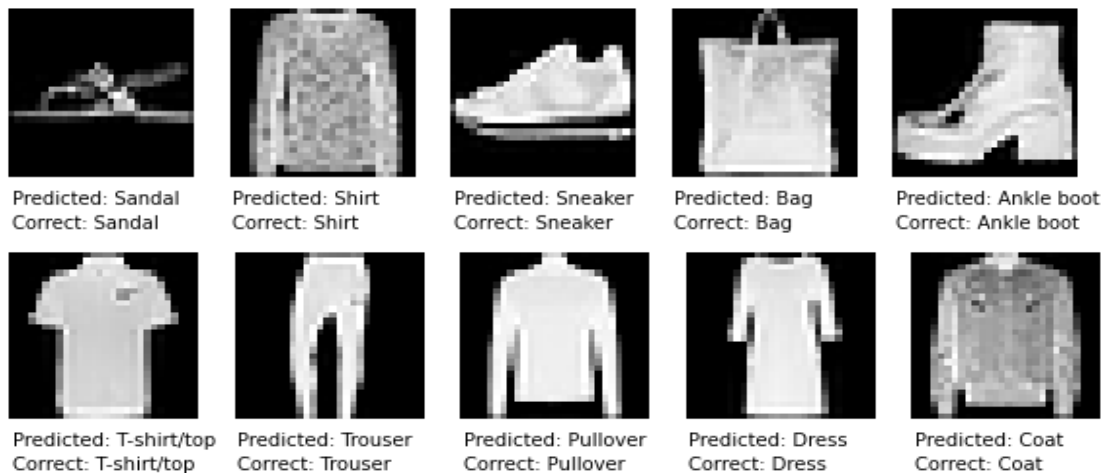Figure 10. Correctly predicted classes

6. CONCLUSION

One of the main objectives of this research was to develop pre_convolution neural networks to evaluate the difficulty of modern CNN architectures that have achieved state-of-the-art success in terms of classification error. A small CNN model based on two key ideas of original fashion-MNIST dataset was augmented by three different types of images to be three ways its original one. The projected pre_convolution architecture had a few parameter numbers and low computational cost with a high accuracy than the existing strategy. The origin images were rotated, taken with acute angle and tilted to obtain the new dataset. Regarding the proposed pre-conv, the convolution layer of the standard CNN is improved by increasing the number of layers to be three layers with max pooling. Classifying the obtained new dataset based on the proposed pre-conv network resulting an improvement in the classification performance of 0.6% better classification accuracy than VGG16 26M. It has been shown that the proposed model has a much increase in accuracy of 2.2% than the 2 Conv+pooling and 1.6% than the 3 Conv+BN+pooling. Wherfore, the limited architecture model and acceptable accuracy at very low parameter numbers, the proposed method is more suitable for all the high definition classification. Our future work in this direction is to apply the new dataset to determine the classification methods and different convolution structures.

REFERENCES

[1] N. Singh, N. Garg, and V. Mittal, "Big Data–insights, motivation and challenges," *International Journal of Scientific & Engineering Research*, vol. 4, no. 12, 2013.

[2] D. Singh and C. K. Reddy, "A survey on platforms for big data analytics," *Journal of big data*, vol. 2, no. 8, pp. 1-20, 2014.

[3] A. Schindler, T. Lidy, S. Karner, and M. Hecker, "Fashion and apparel classification using convolutional neural networks," *arXiv preprint arXiv:1811.04374*, 2018.

[4] T. Zebin, P. J. Scully, N. Peek, A. J. Casson, and K. B. Ozanyan, "Design and implementation of a convolutional neural network on an edge computing smartphone for human activity recognition," in *IEEE Access*, vol. 7, pp. 133509-133520, 2019.

[5] H. Xiao, K. Rasul, and R. Vollgraf, "Fashion-mnist: A novel image dataset for benchmarking machine learning algorithms," *arXiv preprint arXiv:1708.07747*, 2017.

[6] A. Nøkland and L. H. Eidnes, "Training neural networks with local error signals," *arXiv preprint arXiv:1901.06656*, 2019.

[7] D. Hirata and N. Takahashi, "Ensemble learning in CNN augmented with fully connected subnetworks," *arXiv preprint arXiv:2003.08562*, 2020.

[8] K. Greeshma and K. Sreekumar, "Hyperparameter optimization and regularization on Fashion-MNIST classification," *International Journal of Recent Technology and Engineering (IJRTE)* , vol. 8, no. 2, pp. 3713-3719, 2019.

[9] S. Bhatnagar, D. Ghosal, and M. H. Kolekar, "Classification of fashion article images using convolutional neural networks," *2017 Fourth International Conference on Image Information Processing (ICIIP)*, Shimla, pp. 1-6, 2017.

[10] M. Kayed, A. Anter, and H. Mohamed, "Classification of garments from Fashion MNIST dataset using CNN LeNet-5 architecture," *2020 International Conference on Innovative Trends in Communication and Computer Engineering (ITCE)*, Aswan, Egypt, pp. 238-243, 2020.

[11] Y. Zhang, "Evaluation of CNN Models with Fashion MNIST Data," *Iowa State University*, 2019.

[12] Y. Zhang, E. Zhang, and W. Chen, "Deep neural network for halftone image classification based on sparse auto-encoder," *Engineering Applications of Artificial Intelligence*, vol. 50, pp. 245-255, 2016.

[13] K. Meshkini, J. Platos, and H. Ghassemain, "An analysis of convolutional neural network for fashion images classification (Fashion-MNIST)," in *International Conference on Intelligent Information Technologies for Industry*, Springer, pp. 85-95, 2019.

[14] S. Shen, "Image classification of Fashion-MNIST dataset using long short-term memory networks," *Research School of Computer Science*, Australian National University, Canberra.

[15] K. Greeshma and K. Sreekumar, "Fashion-MNIST classification based on HOG feature descriptor using SVM," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 5, pp. 960-962, 2019.

[16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84-90, 2017.

[17] J. Barry-Straume, A. Tschannen, D. W. Engels, and E. Fine, "An evaluation of training size impact on validation accuracy for optimized convolutional neural networks," *SMU Data Science Review*, vol. 1, no. 4, pp. 1-17, 2018.

[18] A. Ferreira and G. Giraldi, "Convolutional Neural Network approaches to granite tiles classification," *Expert Systems with Applications*, vol. 84, pp. 1-11, 2017.

[19] K. Hara, V. Jagadeesh, and K. Hara, V. Jagadeesh and R. Piramuthu, "Fashion apparel detection: The role of deep convolutional neural network and pose-dependent priors," *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Lake Placid, NY, pp. 1-9, 2016.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778, 2016.

[21] E. Dufourq and B. A. Bassett, "EDEN: Evolutionary deep networks for efficient machine learning," *2017 Pattern Recognition Association of South Africa and Robotics and Mechatronics (PRASA-RobMech)*, Bloemfontein, pp. 110-115, 2017.

[22] S. Fairuz, M. H. Habaebi, and E. M. A. Elsheikh, "Pre-trained based CNN model to identify finger vein," *Bulletin of Electrical Engineering and Informatics*, vol. 8, no. 3, pp. 855-862, 2019.

[23] W. Setiawan, M. I. Utoyo, and R. Rulaningtyas, "Reconfiguration layers of convolutional neural network for fundus patches classification," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 10, no. 1, pp. 383-389, 2020.

[24] I. A. M. Zin, Z. Ibrahim, D. Isa, S. Aliman, N. Sabri, and N. N. A. Mangshor, "Herbal plant recognition using deep convolutional neural network," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 9, no. 5, pp. 2198-2205, 2020.

[25] Z. Hu, H. Yan, and X. Lin, "Clothing segmentation using foreground and background estimation based on the constrained Delaunay triangulation," *Pattern Recognition*, vol. 41, no. 5, pp. 1581-1592, 2008.

BIOGRAPHIES OF AUTHORS

Mustafa Amer Obaid holds a Bachelor of Computer Science. And now a master's student. He is currently working as an employee at the College of Medicine, Anbar University. His current research interests include artificial intelligence and deep learning.

Wesam M. Jasim received the B.Sc. and M.Sc. degrees in control automation engineering from University of Technology, Baghdad, Iraq, and Ph.D. degree in computing and Electronics from University of Essex, Essex, UK. Currently, he is an Assistent Professor with the College of Computer Science and Information Technology, University of Anbar. His current research interests include robotics, multiagent systems, cooperative control, Robust control, linear and nonlinear control, Deep learning. He has published research papers at national and international journals, conference proceedings as well as chapters of book.