A Novel Solution to the Curse of Dimensionality in Using KNNs for Image Classification

Aditya D Bhat
Department of
Computer Science and Engineering
PES University
Bangalore, India
adityadb24@gmail.com

Harshith R Acharya
Department of
Computer Science and Engineering
PES University
Bangalore, India
harshithracharya@gmail.com

Srikanth H.R.
Department of
Computer Science and Engineering
PES University
Bangalore, India
srikanthhr@pes.edu

***Abstract*— The k-Nearest Neighbors (KNN) is one of the simplest and widely used algorithms in Machine Learning applications such as Image Classification. Being based on the Euclidean distance the algorithm is quite simple and effective in most cases. However, it suffers from the problem of "The Curse of Dimensionality" as the Euclidean distance becomes meaningless when the dimension of data becomes significantly high. In this paper we present a novel solution to this problem by making use of the Convolutional Neural Network (CNN) which can extract the most important features automatically from the images. These features extracted by the CNN are of reduced dimensions and can effectively be used by the KNN to recognize the images. The results and comparisons show that the proposed method is also seen to reduce the time taken for testing while retaining high accuracy. The proposed technique achieved an accuracy of 96.92% on MNIST, 85.09% on Fashion MNIST and 95.17% on the A-Z Alphabets databases respectively.**

***Index Terms*—Curse of Dimensionality, k-Nearest Neighbor (KNN), Convolutional Neural Network (CNN), Image Classification**

I. INTRODUCTION

Image Classification and Character Recognition have been widely researched areas in Computer Vision having numerous applications in various fields. Present day research [1] is focused towards developing accurate, efficient and robust classification systems using Machine Learning techniques such as K-Nearest Neighbors (KNN) [2], Support Vector Machines (SVM) [3], etc.

The developments in Deep Learning has achieved groundbreaking results in Image Classification with the emergence of Convolutional Neural Networks (CNN). CNNs are capable of extracting automatically the important features from images and have proven to be quite successful in Image Classification tasks. CNNs can be effectively used as feature extractors to improve the accuracy metrics of the stand-alone classifiers [4].

Problems in classification could arise when the features are having high dimensions. The problems arising due to high dimensions in Euclidean space was termed as "The Curse of Dimensionality" by Bellman [5]. This phenomenon has been studied in great detail as traditional techniques based on distances fail to be meaningful in these high dimensional spaces [6].

The KNN [7] [8] is a simple and lazy instance based algorithm which is based on the Euclidean distance, and is widely used for classification. Lazy meaning that the training involves only storing of the features and class labels from the training images. Classification is simply done taking a majority vote from the k nearest neighbours of an instance.

Although the KNN has been widely and effectively used in various applications with good results, it suffers from the problem of Curse of Dimensionality and irrelevant features as the dimension of the feature list increases to a high number. This arrives due to fact that the algorithm makes use of the Euclidean distance which becomes irrelevant when the dimension of features becomes substantially high [9] [10].

In this paper we propose a novel solution to this problem in using KNNs for Image Classification by using a CNN, which extracts the most important features automatically from the image. These features extracted by the CNN are of reduced dimensions and can effectively be used by the KNN to recognize the images. By doing so, we have overcome the curse that was cast by the dimensionality of the data. The hybrid model obtained from using the CNN in conjunction with the KNN is also seen to be highly accurate while also being more efficient than the stand-alone model making it more reliable as depicted from the results obtained. To test the performance of the proposed technique and compare it with that of the stand-alone models, three benchmark databases namely: MNIST, Fashion MNIST and the A-Z handwritten alphabets databases were used.

The rest of this paper is organized as follows: Section II describes the working of the k-Nearest Neighbour algorithm, Convolutional Neural Networks (CNNs) and the proposed CNN-KNN hybrid model. Section III highlights the databases used in this paper to test the models. Section IV shows the experimental results obtained on the databases and the conclusions drawn from these results are shown in Section V.

II. THE CNN-KNN HYBRID MODEL

A. *k-Nearest Neighbors (KNN)*

The k-nearest neighbor (KNN) algorithm [7] [8] is one of the simplest Machine Learning techniques for Image Classification. Being lazy and instance based, the classification

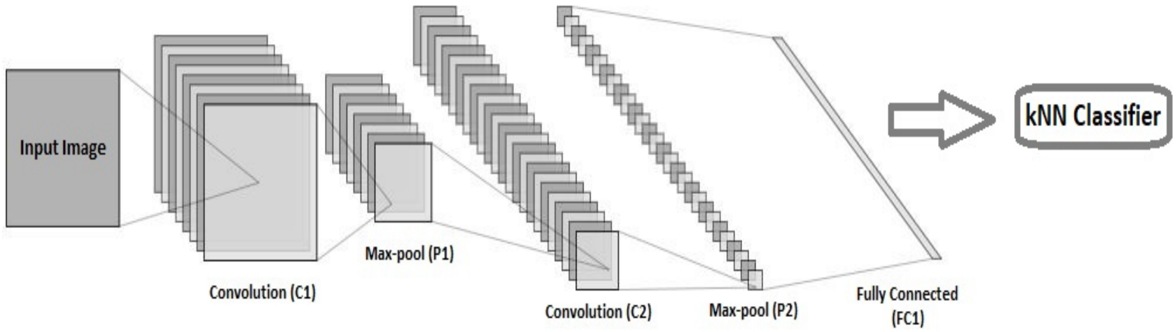


Fig. 1. The proposed CNN-KNN Architecture.

algorithm determines the label of an instance based on the majority vote of its k closest neighbors, which is determined based on the Euclidean distance. We can write the K-nearest neighbor algorithm precisely as shown in Algorithm 1, where X is the training data set, Z is the test set, there are C possible classes, and r is the distance metric (typically the Euclidean distance).

Algorithm 1 k-Nearest Neighbors (KNN) Algorithm

1: 1. For each test example $z \in Z$:
2: i) Compute the distance $r(z, x)$ between z and each training example $(x, c) \in X$
3: ii) Select $X(k, z) \subseteq X$, the set of the k nearest training examples to z
4: 2. The label c of the test example z is then given by: $c(z) = \text{argmax}_{v \in \{1,\ldots,C\}} \sum_{(x,c) \in X(k,z)} \delta(v, c)$

B. Convolutional Neural Network (CNN)

Yann LeCun [11] developed the Convolutional Neural Network, a hierarchical network which is trained through the backpropagation algorithm to effectively extract the features in a robust manner automatically, having a powerful capacity to learn well the good features at every level of the visual hierarchy. The features extracted by the CNN has invariance against shift and distortion in the input images. This had resulted in high success of using CNNs for various Computer Vision tasks. The network has been experimented on a various computer vision applications like object classification [12].

C. The CNN-KNN hybrid Model

The hybrid CNN-KNN Model makes use of the CNN to automatically extract the most important features from the image. These features extracted by the CNN are of reduced dimensions and can effectively be used by the KNN to recognize / predict the label of the images. This is done by replacing the final fully connected layer of the CNN with an KNN Classifier. Fig. 1 shows the architecture of the CNN-KNN model used in this paper. Table I shows the set of parameters used in the hybrid CNN-KNN model in this paper.

TABLE I
PARAMETERS USED IN THE PROPOSED CNN-KNN MODEL

Model	Layer	Parameter	Value
CNN	Convolution (C1)	Number of Filters	64
		Kernel Size	5
	Max-pool (P1)	Pool Size	2
		Strides	2
	Dropout (D1)	Dropout Rate	0.25
	Convolution (C2)	Number of Filters	32
		Kernel Size	3
	Max-pool (P2)	Pool Size	2
		Strides	2
	(FC1)	Number of Neurons	256
KNN		k Value	4
		Metric	Euclidean

III. DATABASES

This section highlights the databases used to conduct experiments in this paper. Three databases namely, The MNIST handwritten digits database [13], The A-Z handwritten alphabets database [14] and the Fashion-MNIST [15] which are widely used for image classification tasks were used.

A. MNIST

The MNIST handwritten digits database consists of 60,000 training and 10,000 testing grayscale images of the handwritten digits 0-9 each centered and normalized to 28 X 28 pixels. Fig. 2 shows a sample of images from the MNIST database.

B. The A-Z handwritten alphabets database

The A-Z handwritten alphabets database consists of over 3700000 grayscale images of the handwritten alphabets A-Z each centered and normalized to 28 X 28 pixels. Fig. 3 shows a sample of images from the A-Z handwritten alphabets database.

C. Fashion-MNIST

The Fashion-MNIST database consists of 60,000 training and 10,000 testing grayscale images belonging to 10 classes. Each image has been centered and normalized to 28 X 28

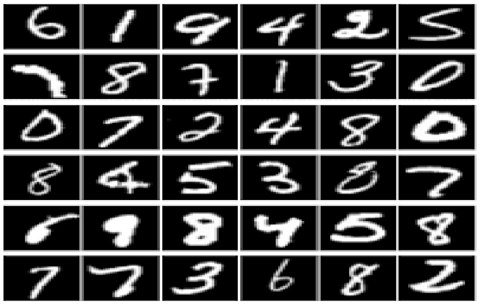


Fig. 2. Sample of images from the MNIST database.

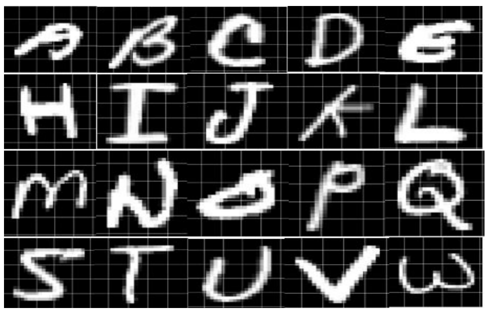
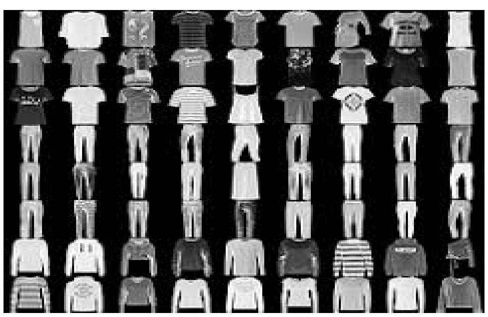


Fig. 3. Sample of images from the Alphabets database.

pixels. Fig. 4(a) shows a sample of images from the Fashion-MNIST database and Fig. 4(b) shows the different classes in the database.



(a) Fashion MNIST

Description	Label	Description	Label
T-shirt/top	0	Sandal	5
Trouser	1	Shirt	6
Pullover	2	Sneaker	7
Dress	3	Bag	8
Coat	4	Ankle boot	9

(b) Classes in Fashion MNIST

Fig. 4. A sample of images taken from the Fashion-MNIST database and the classes present in the database.

IV. Experimental Results and Inferences

To evaluate the hypothesis, we compared the results of the experiment conducted on the proposed system along-side the standard KNN method. In this section, we highlight the results and the techniques from the experiments conducted on the considered datasets.

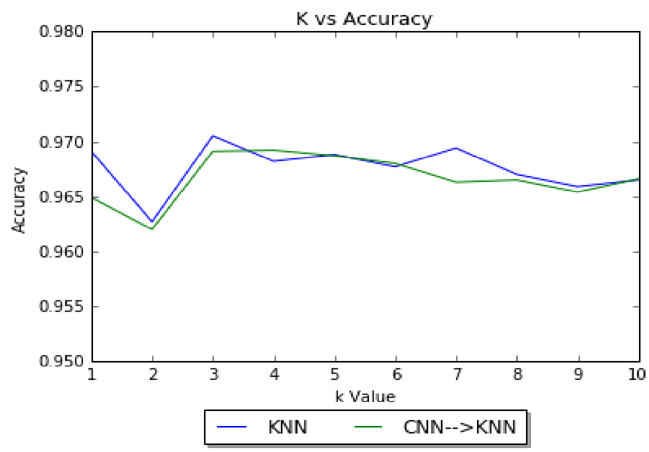
The hybrid CNN-KNN architecture that was proposed in

TABLE II
TEST TIME OBTAINED ON BENCHMARK DATASETS.

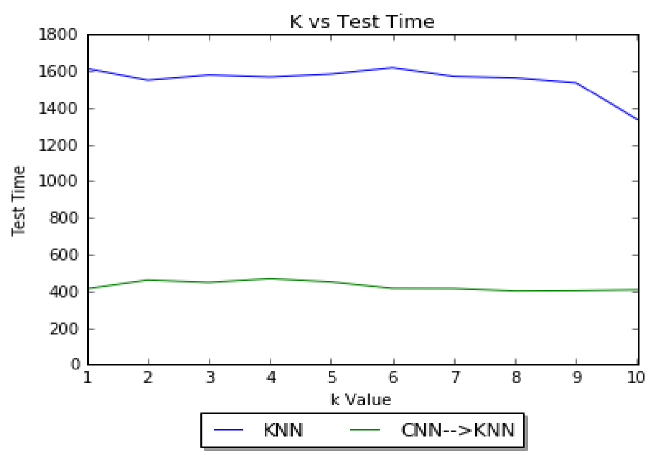
Method	Test Time (seconds) MNIST	Fashion MNIST	A-Z alphabets
KNN	1567.94	1240.28	5844.15
CNN-KNN	**467.73**	**255.76**	**1794.78**

TABLE III
ACCURACY OBTAINED ON BENCHMARK DATASETS.

Method	Accuracy (%) MNIST	Fashion MNIST	A-Z alphabets
KNN	96.82	85.77	94.26
CNN-KNN	**96.92**	**85.09**	**95.17**

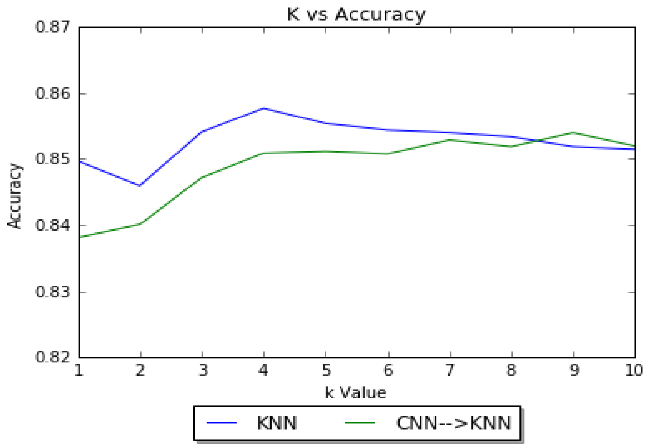


(a) k vs Accuracy

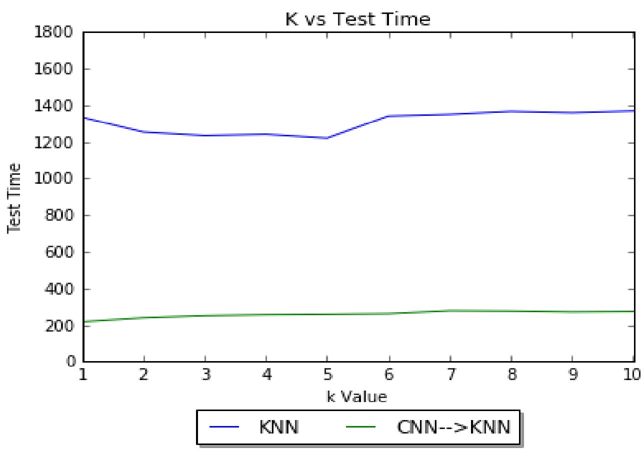


(b) k vs Test Time

Fig. 5. Outcomes from the experiments on the proposed CNN-KNN model on MNIST Dataset

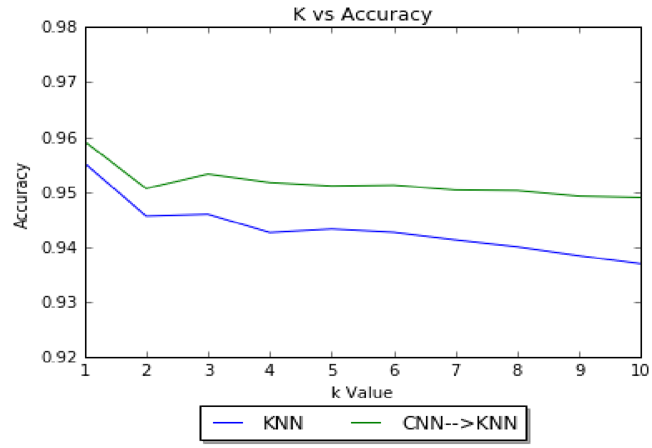


(a) k vs Accuracy

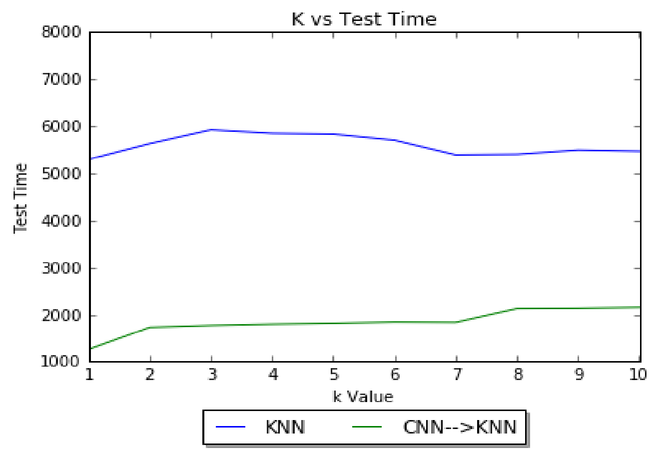


(b) k vs Test Time

Fig. 6. Outcomes from the experiments on the proposed CNN-KNN model on Fashion MNIST Dataset



(a) k vs Accuracy



(b) k vs Test Time

Fig. 7. Outcomes from the experiments on the proposed CNN-KNN model on Hand-written Alphabet Dataset

this paper was tested on the three datasets mentioned in Section III. The results obtained have been summarized in Table II and Table III. Table II shows the time taken by the proposed model and the standard KNN method to converge and Table III shows the accuracy obtained. It can be clearly observed that the proposed CNN-KNN architecture performs much better in terms of convergence time, retaining the accuracy to the standard. Another observation that is note-worthy is that on the Hand-written alphabet dataset, the CNN-KNN architecture attains accuracy scores higher than that of the standard KNN method while also converging at much a faster rate.

Fig. 5(a), Fig. 6(a) and Fig. 7(a) show the variation of Accuracy with the value of k for both the proposed CNN-KNN and the stand-alone KNN models when tested on the MNIST, Fashion MNIST and the Hand-written Alphabets databases respectively.

Fig. 5(b), Fig. 6(b) and Fig. 7(b) show the variation of time taken to converge with the value of k for both the proposed CNN-KNN and the stand-alone KNN models when tested on the MNIST, Fashion MNIST and the Hand-written Alphabets databases respectively.

From the above mentioned figures, it can clearly be seen that the proposed CNN-KNN technique takes far less test-time while being able to retain / improve on the accuracy of the stand-alone KNN model.

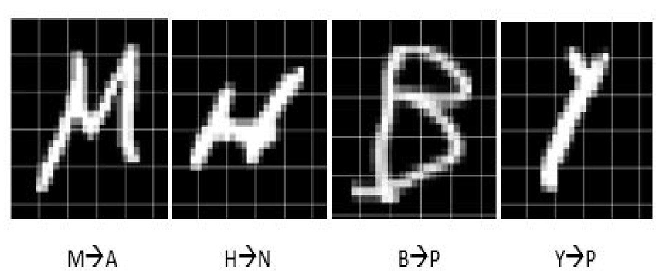


Fig. 8. Sample of images where the CNN-KNN model classified correctly, but the stand-alone model classified wrong. The labels under each image correspond to ”Prediction by the hybrid CNN-KNN model” -> ”Prediction by stand-alone KNN model”.

Fig. 8 shows a sample of images that were predicted right

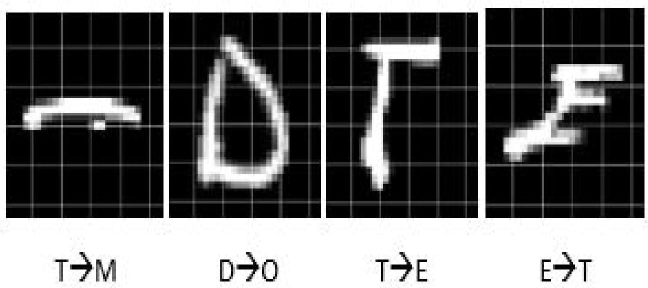


Fig. 9. Sample of images that were wrongly classified by the CNN-KNN model. The labels under each image correspond to "truth" -> "predicted"

by the proposed hybrid model, but were incorrectly predicted by the KNN model. Fig. 9 shows a sample of images that were wrongly classified by the hybrid model. These images seem to contain noise that could have been introduced due to poor handwriting, normalization or segmentation, etc. making it very ambiguous to give a correct prediction of the class of the image.

The outcomes of the experiment from Table II and Table III establishes a significant improvement in effectiveness and the accuracy to the existing state-of-the-art techniques. The best performing KNN algorithm converged in 5845 seconds with an accuracy of 94.26%, whereas the proposed CNN-KNN model converged in a lower time period of 1795 seconds with an accuracy as high as 95.17%. It was obviously noted that the proposed model clearly outperformed others in terms of the testing time involved. This proves that the "Curse of Dimensionality" is hence overcome by considering the proposed architecture. We come to a conclusion that the CNN model clearly extracts only the few significant features, thus eliminating the rest. This without doubt proves that the proposed method has a statistically significant improvement over the prevailing state-of-the-art methods in the field.

V. Conclusion

In this paper, we propose a fast KNN implementation by introducing the novel architecture of integrating the Convolutional Neural Network technique with the KNN method. We show that the implementation of the CNN as the automatic feature extractor enables the deduction of only the effectively distinctive features that differentiate the images, thus making our system more accurate and robust, yet faster. The results of the study warrants an improvement in the convergence time of the conventional KNN method. Experiments particularly in the case of the Hand-written alphabet dataset show that the proposed model outperforms the state-of-the-art KNN model both in terms of accuracy and the convergence time. This without doubt proves that the CNN offers a much smaller, yet reliable feature set, attempting to reduce the Curse of Dimensionality. Thus, the proposed novel architecture has enabled us to achieve higher accuracy rates in lesser time and is demonstrated to outperform the existing state-of-the-art techniques.

The introduction of this novel architecture is an attempt to demonstrate the efficiency of the system at a reduced convergence time to the conventional KNN method on the considered datasets. Overall, we can conclude that the proposed model offers considerably good results at a reduced time frame. In future, the performance can be further enhanced by fine-tuning the parameters of the architecture. This experiment can be extended to other domains other than image classification to reduce the effects of the Curse of Dimensionality prevailing in the conventional KNN models. The enhancement of the architecture by implementing various other optimization techniques is another dimension of development to this work.

References

[1] Kim, Byung-soo and Silvio Savarese. Comparing Image Classification Methods : K-Nearest-Neighbor and Support-Vector-Machines. (2012).

[2] Amato, Giuseppe & Falchi, Fabrizio. (2010). KNN based image classification relying on local feature similarity. Proceedings - 3rd International Conference on SImilarity Search and APplications, SISAP 2010. 101-108. 10.1145/1862344.1862360.

[3] Hua Qian, Chun & Qun Qiang, He & Rong Gong, Sheng. (2015). An Image Classification Algorithm Based on SVM. Applied Mechanics and Materials. 738-739. 542-545. 10.4028/www.scientific.net/AMM.738-739.542.

[4] Abdullah and M. S. Hasan, "An application of pre-trained CNN for image classification," 2017 20th International Conference of Computer and Information Technology (ICCIT), Dhaka, 2017, pp. 1-6.

[5] Bellman R.E. Adaptive Control Processes. Princeton University Press, Princeton, NJ, 1961.

[6] Keogh E., Mueen A. (2017) Curse of Dimensionality. In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning and Data Mining. Springer, Boston, MA

[7] Bremner D, Demaine E, Erickson J, Iacono J, Langerman S, Morin P, Toussaint G, Output-sensitive algorithms for computing nearest-neighbor decision boundaries, Discrete and Computational Geometry, 2005, pp. 593604.

[8] Cover, T., Hart, P. , Nearest-neighbor pattern classification, Information Theory, IEEE Transactions on, Jan. 1967, pp. 21-27.

[9] Aggarwal C.C., Hinneburg A., Keim D.A. (2001) On the Surprising Behavior of Distance Metrics in High Dimensional Space. In: Van den Bussche J., Vianu V. (eds) Database Theory ICDT 2001. ICDT 2001. Lecture Notes in Computer Science, vol 1973. Springer, Berlin, Heidelberg

[10] Verleysen, Michel & Franois, Damien. (2005). The Curse of Dimensionality in Data Mining and Time Series Prediction. Lecture Notes in Computer Science. 3512. 758-770. 10.1007/11494669_93.

[11] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, "Gradient-based learning applied to document recognition", Proc. IEEE, vol. 86, no. 11, pp. 2278-2324, Nov 1998.

[12] Dan C. Ciresan, Ueli Meier, Jonathan Masci, Luca M. Gambardella, Jurgen Schmidhuber (2011). High-performance neural networks for visual object classification. [online] Available at: https://arxiv.org/pdf/1102.0183.pdf

[13] The MNIST Database of Handwritten Digits, http://yann.lecun.com/exdb/mnist, last accessed on 27/10/2018.

[14] The A-Z Alphabets Database, https://www.kaggle.com/sachinpatel21/az-handwritten-alphabets-in-csv-format, last accessed on 27/10/2018.

[15] The Fashion-MNIST Database, https://github.com/zalandoresearch/fashion-mnist, last accessed on 27/10/2018.