International Conference on Recent Trends in Data Science and its Applications
DOI: rp-9788770040723.007

MNIST Digit Classification using Machine Learning Algorithm

MandaNihari
Department of Data Science and Business Systems, School of Computing, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Chennai, 603203, India
mn1944@srmist.edu.in

Priyadarsini K
*Department of Data Science and Business Systems, School of Computing, College of Engineering and Technology, SRM Institute of Science and Technology,*Kattankulathur, Chennai, 603203, India
priyadak@srmist.edu.in

Badhri S
*Department of Data Science and Business Systems, School of Computing, College of Engineering and Technology, SRM Institute of Science and Technology,*Kattankulathur, Chennai, 603203, India,
be6061@srmist.edu.in

***Abstract*—During these modern days everything revolves around technology, thus everything is getting digitized. From teenagers to working age groups use technology, so it is important for computers to classify digits based on handwriting. Many of them have different handwriting which sometimes cannot be identified by humans themselves, so what we have to do is by using machine learning we could make the computer read with an accuracy of 95%. This will surely benefit the process of storing old documents without retyping everything as it will consume a lot of time, instead they can take a picture of the documents and upload into the system which will in turn go through it and show it in text form. On this topic we have done our share of research, and also we have implemented it practically. All types of handwritten documents can be read with a high percentage of accuracy, which makes many people's lives easier. The project is still in its development and early stages, but once it reaches its destination it could be very useful.**

Keywords*— Artificial Intelligence, MNIST Dataset, convolutional Neural Networks, digital reading, Machine Learning, Deep learning

I. INTRODUCTION

MNIST digit classification using a machine learning algorithm for handwritten digit recognition has a great importance in recognition of the optical character , not only that this digit classification can also be used in theories as a test case and not only that they can also be used for machine learning algorithms. The digits that are handwritten are first preprocessed, which includes both normalization and segmentation So that it is possible to compare recognition results on some common basis and reduce their work for the researchers.Not only machines but even humans have a problem reading few types of handwriting as they are unique and different but nowadays everyone are opting for the digital technology.Retyping all the handwritten documents manually into the computer is a very difficult and tiring job .Artificial intelligence computer vision techniques had made it very easy because of which everything can be digitized and not only that, it will also make it easy for the teachers as students answer script can be corrected directly with the help of machines, just by uploading the scripts and marks. Also can be easily updated in the computer with ease .That is why conventional neural networks was developed so as to read whatever type of
\
handwriting more than 95% accuracy.This type we can see in google lens that is currently used by many people .This

also can be implemented in account and finance department for both private and public sectors in which they can upload a number of querie. This will be of tremendous benefit to us as it reduces the chance of human error and also can reduce time consumption and make our lives easier .So it reduces human error and also saves time. One more thing is that it as the main method in schools are paper and pen mode there will be many documents from which they are supposed to be digitized every year and manually doing that will take them forever so the best method would be to upload it and that should be digitized by the help of AI which will make their work easier . Our final goal is to make as much as accuracy while predicting handwritten dataset with the help of Convolutional Neural Network and machine learning. One of the most essential data sets for evaluating the effectiveness of the convolutional neural networks and learning algorithms is indeed the MNIST handwritten character recognition classification data set. Learning algorithms such as k-nearest neighbors (KNN), random forests, svm (SVM), and simple neural network models may easily achieve 97%-98% accuracy on a test set of 10,000 photographs when using 1 million photos as the training set. Convolutional neural networks (CNN) improve this accuracy to over 99% with fewer than 100 misclassified photographs in the test set. The last 100 photographs are becoming more difficult to correctly identify. More complex models, careful tuning of hyperparameters such as learning rate and batch size, regularization methods like batch normalization anddropout, and so on. augmenting of training data are required to enhance accuracy after 99%. We obtain a model capable of achieving extremely high precision on the MNIST test set without the requirement for sophisticated structural elements or learning approaches. One of the most frequent model designs is a collection of convolution layers followed by a completely linked layer at the end. We employ fundamental data augmentation strategies such as translation and rotation. We train three different models with comparable architectures and use majority vote to choose the best model.The final forecast The designs of the three models are identical, but the kernel sizes with in convolution layers change.Experiments demonstrate that merging images with various kernel sizes

DOI: rp-9788770040723.007

improves precision more than mixing models with a similar kernel size.

II. Literature Review

[20]MNIST (Modified National Institute of Standards and Technology) is a classic dataset used in the field of machine learning for the task of image classification. It consists of 70,000 images of handwritten digits from 0 to 9, with each image being a 28x28 grayscale image. The MNIST dataset has been widely used as a benchmark dataset for evaluating the performance of various machine learning algorithms, including deep learning models, for image classification tasks.

In this literature survey, we will review some of the recent studies on MNIST digit classification using machine learning algorithms.[19]Authors compared the performance of various deep learning models, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Autoencoders, for MNIST digit classification. They found that CNNs outperformed the other models, achieving an accuracy of 99.35%. The authors also showed that data augmentation techniques, such as rotation and translation, could improve the performance of the models.[2]Few other authors have also investigated the effectiveness of transfer learning for MNIST digit classification. They compared the performance of various transfer learning approaches, including fine-tuning and feature extraction, using different pre-trained models. The authors found that transfer learning could significantly improve the performance of the models, achieving an accuracy of up to 99.7%. Usage of some deep learning and neural networks for us to get a image representation is high on demand in this modern era.Image processing has a huge number of publications already.In our paper we are going to show how to classify images based on using CNN. The fashion domain also uses classification of digits and it has a lot of benefits .It also has a lot of work that was published on it. All humans have a different handwriting which is very unique so the computer finds it very difficult to understand each type of handwriting.For us to improve the accuracy to over 99% then we have to view complex methods like some careful tuning For some like learning rate and batch size which have few regulations such as training data augmentation and normalization of batch is need for hyperparameters .MNIST test accuracy is approximately 98% accurate on a few papers .In this paper we will produce a model that can be achieved which has a higher accuracy on MNIST test which will also be very simple.The model also uses convolution layers which have fully connected layers at the end. ,these are one of the commonly used model architectures .Basic data augmentation schemes,translation and rotation are used .Three models which have similar architectures and also use majority voting between the models to obtain some final prediction .They have architectures which are similar but have different kernel sizes .we can give better accuracy by combining different kernel sizes more than combining different models.We present a model that achieves very good precision here on MNIST testing sample without the use of sophisticated structural elements or learning approaches. One of the most frequent model designs is a collection of convolution layers followed by a completely linked layer at the end. We employ fundamental data augmentation strategies such as translation and rotation. We train three models with comparable architectures and then use majority vote to get the final prediction. The designs of the three models are identical, but the kernel sizes in the convolution layers change.Experiments demonstrate that combining models with different kernel sizes improves accuracy more than combining models with the same kernel size.

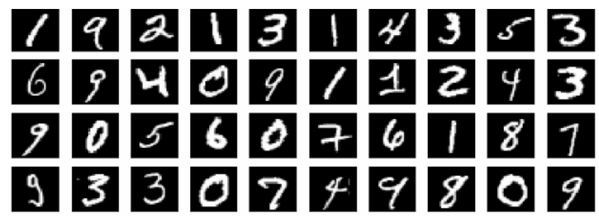

Fig. 1. Image used for testing the digit classification

III. Previous Advancement

This section discusses the most modern strategies for handwritten numeric digit categorization, as well as the benefits and drawbacks of each methodology.

Shulan et al. introduced a cascade CNN which has two stage with outstanding performance. Here a cascade CNN approach is based on complimentary differentiation of few objectives between Stage I and the consecutive next stage. By feeding directly poorly recognised training samples, learning is little discriminative which is added to train Stage II. Experiments were carried out on the highly competitive MNIST handwritten digit database. Their cascade model outperformed all others in terms of performance. The drawback of this approach is that it takes more time than other procedures.

Hubert Cecotti studied the influence of several parameters and pre-processing methods of a distance based on picture distortion models on accuracy. One major problem is to lower the processing time of closest neighbor classification by taking rejection criteria and adaptive distances into account. The author evaluated the efficacy of this single character identification algorithm on three datasets of Indian handwritten numbers, one for each common Indian plot: Bangla, Devanagari, and Oriya. The authorsdemonstrated that extracting information relating to four orientations improves accuracy significantly. This method makes use of GPUs and high-performance clusters to deliver cutting-edge results.

The clustering algorithm used gives great performance but takes longer to cluster the input and construct the feature.Khedidja and Hayet studied the application of various differentiators or as we say classifiers which help in the identification of printed digits by describing the digits in a novel manner utilizing hybrid feature extraction. The study comprised three separate features estimated from extraction,cavities, zonal extraction, and retinal representations, and also nine distinct predictors, K-Nearest Neighbor - KNN - with varied distance measurements, SvmClassification - decision tree, and linear svm

DOI: rp-9788770040723.007

classification - LDA.-. Majority Voting takes into account classifier combinations. Experiments were done out on a dataset of printed numbers with multiple fonts and sizes. This approach has only been assessed for printed databases; it has yet to be analyzed and reviewed for handwritten digit databases.

Few people suggested a novel hybrid classification strategy for identifying printed numerals in their paper .To extract features, object region perimeter assessment, Fourier Descriptors, and a Chained code-based approach were utilised. A unique curve tracking Chain code based approach (CTCC) was introduced to recover curve knowledge from digit photos. Few multi layer and Dynamic programming employing a back propagation approach were used to achieve recognition (MLP-BP). The accuracy was increased to 99%. Although confined to printed digits, the proposed methodologies were simple, with higher identification accuracy and decreased time complexity.

Gattal et al. studied the use of several statistical and structural factors in the detection of solitary handwritten digits, a classic pattern recognition challenge. By merging multiple representations of non-normalized handwritten numbers, the authors attempted to enhance identification rates.

Some of these characteristics include statistical information, moments, feature and projection-based characteristics, and features derived from the contour and skeleton of the digits Some of these characteristics are taken from the whole picture of the digit, while others are retrieved from various sections of the image after the image has been subjected to uniform grid sampling. One-against-all SVM is used for classification. Experiments on the CVL Single Digit Database yielded high recognition rates similar to state-of-the-art approaches in this field.

Alkhateeb and Alseid introduced a Handwritten Arabic digits multi-class classification approach utilizing Dynamic Bayesian Network (DBN), with technological details provided in 3 parts: pre-processing, feature extraction, and categorization. The digits are first pre-processed and their sizes are standardized. Then, using the discrete cosine transform (DCT) coefficients technique, features are reconstructed from each normalized digit, and even a set of additional handwritten characteristics are added. digits are provided. Finally, theseTo develop a deep neural network for classification, attributes are employed. The proposed method was successfully evaluated on a Handwritten arabic digit database (AD Base), which had 70k digits produced by 700 different authors, as well as the results were encouraging and promising.

El et al. compared the efficiency To capture discriminative characteristics of handwritten digits, four feature extraction techniques based here on Discrete Cosine Transform (DCT) were developed. Upper left corner (ULC) DCT coefficients, zigzag DCT coefficients, and block-based DCT ULC coefficients and block-based DCT zigzag coefficients are the techniques. To assess the performance. The parameters of the each DCT variation are used as data input for the Svm Classifier. Their objective was to identify the optimum feature extraction strategy for enhancing classification accuracy while speeding up learning algorithms. The data demonstrated that perhaps the block-based DCT zigzag extraction of features performed well in terms of accuracy in classification and reduction rate. outperforms its competitors.

Babu et al. introduced a novel solution to off-line handwritten digit detection based on structural characteristics that eliminates the need for thinning and size normalization techniques. For digit identification, they employ four distinct kinds of structural features: number of openings, water tanks in four directions, optimum profile distance in 4 directions, and fill-hole density. To identify minimal distances, a Euclidean minimum distance criteria is utilized, and a k nearest neighbor classifier is used to categorize the digits. The MNIST database is used to train and test the system. A total of 5000 numerical pictures are evaluated to validate the suggested approach, with a high identification rate.

Singh and Lai [9] proposed a method for digit recognition based on a single layer neural network classifier and Principal Component Analysis (PCA). The created model minimizes the characteristics in order to decrease computing needs while accurately classifying the digit into ten groups (0 to 9). The developed system is a backward propagation (BP) neural network that has been trained and tested on the MNIST dataset of handwritten digits. On the MNIST 10K test dataset, the suggested approach achieved good accuracy. They took into account not just accuracy, but also training time, recognition time, and memory needs for the whole process.

They have also discovered the digits that the system misclassified.

IV. Algorithm

The CNN is a learning technique which is deep and where it automatically classifies input . The past few years CNN has been good at classifying images and also it is being used in many domains like healthcare and academic domain . One of the most reliable algorithm was indicated that it was CNN which is good at automated prediction from start to end. CNN also extracts very important features from the given input that makes it easy for us.Data entry professionals devote hundreds of hours to typing handwritten data into computers. It is a very time-consuming activity that also necessitates a high level of precision and fast typing due to the large number of entries that must be input.Various organizations spend tons of money transferring their documents from one format to another.Deep learning and machine learning are critical components of computer technology and artificial intelligence. Deep learning can minimize human effort in identifying, learning, predicting, and many other areas. As a result, we developed a Convolutional Neural Network System capable of automatically converting handwritten pictures to digital format with a 93% accuracy.

DOI: rp-9788770040723.007

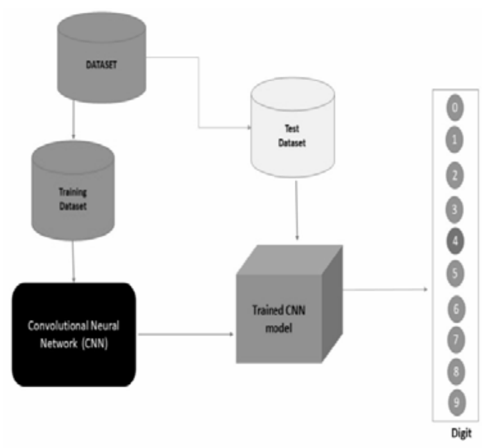

Fig. 2. CNN model

[1] It has very high features from few input data which are passed onto the other layer which is a feature map. This is the conventional layer.

[2] The data dimensions are reduced by feature map pooling,where dimensions are reduced by generating new feature maps.PL takes average or it might take maximum,this is done at Pooling Layer.

[3] FC layer completes the task of classification.Softmax function is where for each class the Probability scores are calculated and labeled a popular activation function.

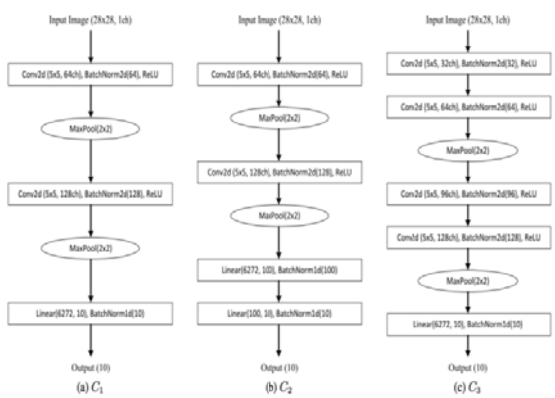

Fig. 3. Network model used for classification

V. DATASET

MNIST digit classification dataset is being used .The keras is in Python and MNIST as an API which is a dataset that is provided by the API. 60k -10k training images are available in it. This is a great way to individuals who need to go through pattern recognition just in a minimal amount of time.

The Keras API have four values returned namely- *x_train, y_train, x_test, and y_test*.

VI. LOADING DATASET

The language used here is python. We used googlecolab for doing this python code. We can also use jupyter notebook or vs code

The steps that we followed :-

1. Add and create a googlecollab notebook.
2. Then we load the necessary libraries, we must load the MNIST dataset

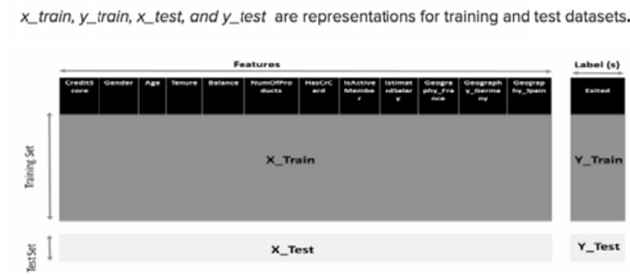

Fig. 4.dataset into training and test set

VII. PROCESSING THE DATASET

To improve the quality of the data it should be rectified,cleaned and processed .The best form of dataset has no null values and all numeric values are scaled.So in order for the data to be good for CNN we will be training the dataset .If we write X_train[0] we get images between 0-255. The output is a 2-dimensional matrix. writey_train[0] we will get 5 as output. The 0th image of this training dataset represents the number 5. Training and test datasets .we convert the 2-d matrix to a 1-d array after scaling.

VLII. CONVOLUTIONAL NEURAL NETWORK BEING TRAINED AND CREATED

For the above code the input layer follows two hidden layers and a output layer.The activation function make key decisions if they should move forward or not.In CNN it has many neurons which are based on activation functions,the neurons fire up and the network moves forward. ‘relu’ is a activation function.The model is compiles and fit after it been created.The relations are viewed in the dataset during fitting. They will learn throughout the process as many times as has been defined. For every mistake it makes it learns and gets better accuracy each time.At the end of the epoch it gets the highest form of accuracy .A image has three dimensions: width, height, and channel. The MNIST dataset is a monochromatic image with the dimensions 28x28. In the shape parameter, we set the sample size to -1 so that it adopts the form of both the attributes ["x"]. The advantage is that the batch size may be tuned to hyperparameters. The first convolutional layer seems to have 18 filters also with kernel size of 7x7 with equivalent padding. The same padding has the same height and width for both the output and input tensors. TensorFlow then adds zeros to the columns and rows to guarantee that they are of the same size.Following the convolutional is the pooling process. This pooling computation will shorten the data's extension. With a dimension of 3x3 and a stride of 2, we can employ the max pooling2d module. We use the previous layer as input. The output size of the second CNN is precisely 32 filters. The convolution operation has the same size as before, and output shape is [batch_size, 14, 14, and18].to define the completely linked layer Before combining with the thick layer, the feature map must be compressed. With a size of 7*7*36, we may utilize the module reshape.

DOI: rp-9788770040723.007

The thick layer will link neurons. We can and will implement an activation feature for Relu. A dropout regularization term will be added which has 0.3 as its rate, implying that 30percent of the weights are 0. Dropouts occur only during the training time. The mode input to the classification algorithm fn() function specifies if the model should really be trained or evaluated.

TABLE 1: OTHER ALGORITHM

Rank	Model	Percentage Error	Accuracy	Paper	Year
1	Heterogenous ensemble with simple CNN	0.09	99.91	An Ensemble of Simple CNN models for MNIST DigitRecognition	2020
2	Branching/Merging CNN+Homogenous Vector Capsules	0.13	99.87	No Routing Needed Between Capsules	2020
3	EnsNet (Ensemble learning in CNN augmented with fully connected subnetworks	0.16	99.84	Ensemble Learning in CNN Augmented with fully connectedsubnetworks	2020
4	Efficient CapsNet	0.16	99.84	Efficient CapsNet: Capsule Network with Self-AttentionRouting	2021
5	SOP CNN	0.17	99.83	Stochastic Optimization of Plain CNN with Simple Methods	2020

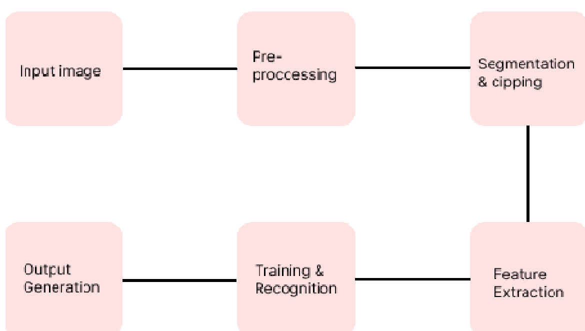

Fig. 5. Architecture of CNN

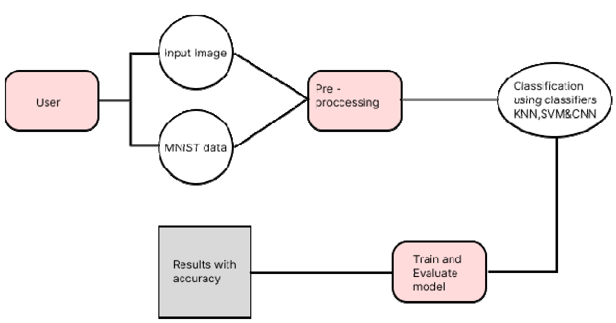

Fig. 6. Data flow Diagram of the system model

IX. MAKING PREDICTIONS

The test dataset has some predictions that are made and then they are stored in the data frame which is called the y_predicted_by_model . Calculation will be done for each dataset to see the probability score. The prediction made by the model is the highest probability score. We must define a tensor with the data's shape. We may accomplish this by using the reshape module. Here we must specify the tensor in order to shape it. The first for this argument is a data characteristic that is defined in a function argument.

X. CONCLUSION

In this paper we see how a dataset is divided and trained.MNIST was taken to make predictions using CNN algorithm and TensorFlow .The predictions were made to train the model and make an accuracy to upto 99% for digits 0-9.Using an aggregation of heterogeneous and homogeneous network models might improve performance up to 99.91percent of overall test accuracy, which is among the best available.

Studies using diverse configurations reveal that great performance is attained by a combination of approaches such as batch normalization, data augmentation, and ensemble methods, rather than by a single methodology or model design.

REFERENCE

[1] “Handwriting,Recognition”:https://en.wikipedia.org/wiki/Handwritingrecognition.

[2] S. Ali et al “A Hybrid Machine Learning Model for MNIST Digit Classification”, 2020.

[3] M. Ranzato, F. J. Huang, Y. L. Boureau, and Y. LeCun, “Unsupervised learning of invariant feature hierarchies with applications to object recognition,” in Proc. of the Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1-8, 2007.

[4] E. D. Cubuk, B. Zoph, D. Mane, V. Vasudevan, and Q. V. Le, “AutoAugment: Learning augmentation policies from data,” in Proc. of the Conference on Computer Vision and Pattern Recognition (CVPR), pp. 113-123, 2019.

[5] E. D. Cubuk, B. Zoph, J. Shlens, and Q. V. Le, “Randaugment: Practical automated data augmentation with a reduced search space,” arXiv preprint arXiv:1909.13719, 2019.

[6] S. M/Shamim, Mohammad BadrulAlamMiah, AngonaSarker, MasudRanaand Abdullah Al Jobair. “Handwritten Digit Recognition using Machine Learning Algorithms,” 2018.

[7] AnujDutt and AashiDutt, “Handwritten Digit Recognition Using Deep Learning,” 2022.

[8] NorhidayuBinti Abdul Hamid, and NilamNurBinti Amir Sharif,“Handwritten recognition using SVM, KNN, and Neural networks,” 2020.

[9] Pazhani. A, A. J., Gunasekaran, P., Shanmuganathan, V., Lim, S., Madasamy, K., Manoharan, R., &Verma, A. (2022).Peer–Peer Communication Using Novel Slice Handover Algorithm for 5G Wireless Networks.Journal of Sensor and Actuator Networks, 11(4), 82.

[10] Farhana Sultana, Abu SufianandParamarthaDutta, “Advancements in Image Classification using Convolutional Neural Network,” 2018.

[11] P. Izmilov, D. Podoprikhin, T. Garipov, D. Vetrov, and A. G. Wilson, “Averaging Weights Leads to Wider Optima and Better Generalization,” in Proc. Conference on Uncertainty in Artificial Intelligence (UAI), 2018.

[12] L. Breiman, “Bagging Predictors,” in Machine Learning, vol. 24, no. 2, pp. 123-140, 1996.

[13] Y. Freund and R. E. Shapire, “Discussion of additive logistic regression: A statistical view of boosting,” in Annals of Statistics, vol. 28, pp. 337-374, 2000.

[14] Rajesh, M., &Sitharthan, R. (2022). Image fusion and enhancement based on energy of the pixel using Deep Convolutional Neural Network. Multimedia Tools and Applications, 81(1), 873-885.

[15] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T. Liu, “LightGBM: A highly efficient gradient boosting decision tree,” in Advances in Neural Information Processing Systems, pp. 3149-3157, 2017.

DOI: rp-9788770040723.007
[16] Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang, “Random erasing data augmentation,” arXiv preprint arXiv:1708.04896, 2017.

[17] S. Lim, I. Kim, T. Kim, C. Kim, and S. Kim, “Fast auto augment,” arXiv preprint arXiv:1905.00397, 2019.

[18] P. Bachman, R. Hjelm, and W. Buchwalter, “Learning representations by maximizing mutual information across views,” across views. arXiv:1906.00910, 2019, 9 A PREPRINT - OCTOBER 6, 2020.

[19] S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” arXiv preprint arXiv:1502.03167, 2015.

[20] J. Bjorck, C. Gomes, B. Selman, K. Q. Weinberger, “Understanding Batch Normalization," in Advances in Neural Information Processing Systems, 2018.

[21] Y. Li et al., “An Empirical Study of Transfer Learning for MNIST Digit Classification,” 2020.

[22] J. Liu et al. “A Comparative Study of Deep Learning Approaches for MNIST Digit Classification,” 2020.