MNIST Handwritten Digit Dataset: A Review of Literature

Shivam. S. Kadam
Rajarambapu Institute of Technology, India

Abstract— This paper deals with MNIST dataset(60,000 images of training and 10,000 images of testing) for classifying handwritten digits by reviewing few methods used in literatures and concluding which method is more accurate in classifying images.

Keywords: MNIST, Handwritten Digit Dataset

I. Data

MNIST is dataset of handwritten digits which has 60,000 images of training images and 10,000 testing images. It is subset of larger dataset NIST. The original images from NIST dataset are black and white of 20 x 20 pixels whereas MNIST dataset is centered in 28 x 28 pixels. MNIST dataset can be obtained through http://yann.lecun.com/exdb/mnist/.

II. Literature

"Extraction Method of Handwritten Digit Recognition Tested on the MNIST Database" [1] facilitates the use of neural network for optical character recognition (OCR) system of handwritten digits. Emphasizing on learning as an important step in process of recognition descent of the gradient algorithm is used along with the modification of the synaptic weight of the connections between the neurons. MLP multilayer perceptron's are used in the same task, which contain 3 layers of neutrons. The first layer functions similar to retina in matching the input image, the second layer performs extraction, while the third gives the output. This task is done with steps as Acquisition of Image followed by pretreatment, followed by Extraction and then finally classification. In the process of extraction dilatation of image is being done with 2 methods processed in four directions, after which characteristic zones of image are being detected. After extracting characteristic data of an image, it is classified using neural networks that is multi-layer perceptron's. Proposed method is tested on MNIST database with 60000 training and 10000 testing images which records approximately 80% of success rate.

"Comparison and Combination of State-of-the-art Techniques for Handwritten Character Recognition: Topping the MNIST Benchmark" [2] is been aimed at proposing different approaches in field of the recognition of isolated handwritten digits to prove that the yielded results vary enough to allow improvements. This paper focuses on comparing the errors made on the standard MNIST benchmark data which facilitate analysis of four of the state-of-the-art methods for the recognition of handwritten digits [3, 9, 15, 26]. Bootstrapping techniques are used for statistical analysis which records no of errors as well as specification of errors which displayed substantial differences leading to use of classifier combinations. Which were proven to be better than single best classifier. Four systems are described to compare and combine for handwritten digit recognition which include Shape context matching, Invariant support vector machine, Pixel-to-pixel image matching with local contexts, Convolutional neural net and virtual data. Using the described hypothetical combination, the resulting error rate is 0.35%. These combinations are trained and tested on MNIST database to achieve low error rates.

"In CUDA implementation of deformable pattern recognition and its application to MNIST handwritten digit database" [3] the challenging topic of Deformable approaches in the field of computer vision and pattern recognition is been discussed. Earlier in regularization-based deformable recognition approach which is used for computing sub-pixel correspondence between input and prototype images which yields very simple iterative equations given from calculus of variations. But in this case the whole computation cost increases in case with large sized images. Due to complexity of implementation on GPU (graphics processing units) the compute unified device architecture (CUDA) succeeded in providing a simple and powerful platform. In this paper deformable pattern recognition method is been proposed based on regularization framework. Reduced computational time is achieved by CUDA implementation with prototype parallel displacement computation (PPDC) and gradual prototype elimination (GPE). Eight directional derivatives and the computed displacement are used to find dissimilarity between the input and prototype images. The proposed method was evaluated based on the MNIST database, which includes 10,000 input and 60,000 prototype images which gave the lowest error rate of 0.57%.

"A snapshot of image pre-processing for convolutional neural networks: case study of MNIST" [4], first of all intensifies the quality of good classifier to generalize on new data, i.e., correctly classifying new instances that do not belong to the training set. This paper is presented with a short description of image preprocessing for deep learning algorithms in general and CNNs in particular, which was used on the competitive MNIST handwritten digits classification problem. To analyze and compare the performance of proposed method, 3 different types of CNN are used as LeNet, Network3 and Drop Connect with their ensembles. The task was to analyze the transformations which were observed to be centering, elastic deformation, translation, rotation and different combinations of them as well. Results of which demonstrate that the combination of elastic and rotation improves the accuracy of the three analyzed networks up to 0.71% and together with ensembles have a high potential to further improve the state-of-the-art accuracy in MNIST classification.

In "TOWARDS THE FIRST ADVERSARIALLY ROBUST NEURAL NETWORK MODEL ON MNIST" [5], the vulnerability of deep neural network to tiny input perturbations and even for MNIST is been mentioned which further exhibit lack of robustness. Hence a robust classification model is been proposed in this study which will perform analysis by synthesis using learned class-conditional data distributions which is inspired by unrecognizable images or distal adversarial which are images that do not resemble images from the training set but which typically look like noise. While providing with the robust classification model, it is mainly facilitated with the many defenses against adversarial attacks which can be divided into four categories as Adversarial training, Manifold projections, Stochasticity and Preprocessing. For which Analysis by Synthesis model (ABS) is been proposed which includes Class-conditional distributions, Optimization-based inference, Classification and confidence, Binarization (Binary ABS only), Discriminative finetuning (Binary ABS only). Having acknowledged the structure that it is not easy to reliably evaluate a model's adversarial robustness , the ABS model prevents the computation of gradients which might give the model an unfair advantage which is evaluated using a large collection of powerful attacks, including one specifically designed to be particularly effective against the ABS model (the Latent Descent attack) which showed that the proposed approach has great potential to reduce the vulnerability against adversarial attacks.

In "Echo State Networks-based Reservoir Computing for MNIST Handwritten Digits Recognition" [6] writer states that Reservoir computing is part of Recurrent Neural Network architecture which has existing neuromorphic implementation with ease of training. This article uses Echo State Networks to classify digits of MNIST dataset. Though Echo State Networks are unable to outstand CNN they are better comparatively due to image preprocessing techniques. Here best performance is obtained with 4000 Neurons. Error rate of 1.42% was observed for reservoir of 1200 neurons and 0.93% when reservoir of 4000 neurons were used. Image presented to reservoir had great impact on error rate. Deformation and rotation of image reduced the error rate.

"Handwritten Digit Classification using the MNIST Data Set" [7] mentions that due to scalability of applications of Handwritten Digit Classification with convenience of input, processing and output, it has become the area of great interest for many researchers. The major challenge identified during the study was in class variance as the same digit is not written in same way for every single time, hence shape invariance is induced to increase the discrimination ability. The study facilitates 3 kinds of extraction as direction features, local structure features or curvature features, which are proven to increase accuracy as well as efficiency of the classifiers. Multivariate Gaussian, Mixture Gaussian are the 2 kinds of generative models used along with Support Vector Machine (SVM), Neural Network(NN) as discriminative models. Various combinations of these classifiers are then tried and tested on MNIST database where the lowest error rate 1.19% is achieved using 3-NN with potential to further improvements.

"Resource Efficient Arithmetic Effects on RBM Neural Network Solution Quality Using MNIST"[8] is focused on new neural network algorithm (mainly concerned with ANN) Restricted Boltzmann Machine (RBM) which has proven to be strong alternative for most popular Multi-Layer Perceptron with Back Propagation learning (MLP-BP) but with a drawback of long running hours. A connectivity reduction is made to RBM to enable a type of learning used similarly to how back propagation (BP) is used in MLP with neuron nodes in two layers, the hidden and visible. GPUs and FPGAs can both be used to implement ANNs with potentially significant speed up hence RBM is implemented on FPGA, in which limited numeric range resulting into generalization typically results into parameter saturation. After analysing the results produced by RBM, it can be concluded that if the parallelism of computation is made maximum FPGA implementations of RBM yields greater performance for a particular chip family, which have further potential applications in the areas of low power mobile or personal use environments with limited precision arithmetic on FPGAs.

III. Conclusion

Many methods like State-of-art technique [2], Echo State Networks [6] are been used to classify digits of MNIST handwritten digit dataset but [3] proved to be more accurate in classifying digits as the error rate observed is just 0.57% though [6] is close enough with just 0.93% of error rate. More of this accuracy is been obtained by deformation and rotation of image.

References

[1] El Kessab, B., Daoui, C., Bouikhalene, B., Fakir, M., & Moro, K. (2013). Extraction method of handwritten digit recognition tested on the mnist database. International Journal of Advanced Science & Technology, 50, 99-110.
[2] Keysers, D. (2007). Comparison and combination of state-of-the-art techniques for handwritten character recognition: topping the mnist benchmark. arXiv preprint arXiv:0710.2231.
[3] Mizukami, Y., Tadamura, K., Warrell, J., Li, P., & Prince, S. (2010, August). CUDA implementation of deformable pattern recognition and its application to MNIST handwritten digit database. In 2010 20th International Conference on Pattern Recognition (pp. 2001-2004). IEEE.
[4] Tabik, S., Peralta, D., Herrera-Poyatos, A., & Herrera, F. (2017). A snapshot of image pre-processing for convolutional neural networks: case study of MNIST. International Journal of Computational Intelligence Systems, 10(1), 555-568.
[5] Schott, L., Rauber, J., Bethge, M., & Brendel, W. (2018). Towards the first adversarially robust neural network model on MNIST. arXiv preprint arXiv:1805.09190.
[6] Schaetti, N., Salomon, M., & Couturier, R. (2016, August). Echo state networks-based reservoir computing for mnist handwritten digits recognition. In 2016 IEEE Intl Conference on Computational Science and Engineering (CSE) and IEEE Intl Conference on Embedded and Ubiquitous Computing (EUC) and 15th Intl Symposium on Distributed Computing and Applications for Business Engineering (DCABES) (pp. 484-491). IEEE.
[7] Wu, M., & Zhang, Z. (2010). Handwritten digit classification using the mnist data set. Course project CSE802: Pattern Classification & Analysis.
[8] Savich, A. W., & Moussa, M. (2011, November). Resource efficient arithmetic effects on rbm neural network solution quality using mnist. In 2011 International Conference on Reconfigurable Computing and FPGAs (pp. 35-40). IEEE.

