Educational Administration: Theory and Practice
2024,30(2), 638-643
ISSN:2148-2403
https://kuey.net/

Research Article

Handwritten Digit Recognition Accuracy Comparison Using Knn,Cnn And Svm.

Anukriti Rajput[1*], Anish Kumar Singh[2]

[1*]Computer Science and Engineering Galgotias University Gr. Noida, India Email:- anu25.kriti@gmail.com
[2]*Computer Science and Engineering Galgotias* UniversityGr. Noida, India Email:- Anish.21scse1010790@galgotiasuniversity.edu.in

Citation: Anukriti Rajput (2024, Handwritten Digit Recognition Accuracy Comparison Using Knn, Cnn And Svm. *Educational Administration: Theory And Practice*, *30*(2), 638-643
Doi:10.53555/kuey.v30i2.1676

ARTICLE INFO	ABSTRACT

This project Handwritten digit recognition is a vital area in the field of computer vision. This paper compares the performance of convolutional neural networks (CNNs), k-nearest neighbors (KNN), and support vector machines (SVMs) for digit classification using the MNIST dataset. The dataset consists of 60,000 labeled images of handwritten digits 0-9 for training and 10,000 images for testing. Experimental results demonstrate that CNNs achieve the highest test accuracy of 98.6%, outperforming KNN (96.8%) and SVM (97.1%). The multi-layer feature extraction capability and depth of CNN models are better at capturing the unique visual patterns of different digit classes in the images. The study provides quantitative results that illustrate that CNN architectures yield state- of-the-art performance on MNIST digit classification compared to other common approaches.

Keywords: Handwritten digit recognition, Computer vision, Convolutional neural networks (CNNs), k-nearest neighbors (KNN), Support vector machines (SVMs), MNIST dataset, Classification accuracy, Feature extraction, Visual patterns

I. INTRODUCTION

Machine learning and deep learning techniques have transformed diverse fields such as computer vision, natural language processing, and pattern recognition. These algorithms enable systems to automatically learn features from raw data and make predictions without being explicitly programmed. Image classification is an important application area that has benefited immensely from advances in deep neural networks. Handwritten digit recognition, in particular, is a fundamental task in image classification that serves as a benchmark for evaluating machine learning algorithms. The MNIST dataset of handwritten digits has been pivotal for development of new machine learning models over the past several years. The dataset contains 60,000 labeled images of handwritten digits 0-9 for training machine learning models. It also includes 10,000 test images to assess model performance. There have been intense efforts towards achieving human-comparable accuracy levels by using MNIST.
This research aims to evaluate different machine learning approaches for recognizing handwritten digits. The models used include k-Nearest Neighbors (KNN), Support Vector Machines (SVM) and Convolutional Neural Networks (CNN). KNN is based on finding similar training images to a test image by nearest neighbor search in feature space. SVMs [3] identify optimal decision boundaries between classes by maximizing margin. CNNs automatically learn hierarchical visual features through convolutional layers. The models are trained on MNIST digits, and their accuracy is recorded on the test set. The key research question is - which model provides the best test accuracy on MNIST handwritten digits? The performance analysis will determine if sophisticated deep learning approaches can outperform traditional machine learning on this dataset. The findings will contribute towards advancing handwritten digit recognition using machine learning. Recognizing handwritten digits is a long-standing task in the field of machine learning and computer vision. This research aims to compare three popular algorithms - k-Nearest Neighbors (KNN), Support Vector Machines (SVM) and Convolutional Neural Networks (CNN) for classifying the MNIST dataset of 70,000 handwritten digits images into 10 classes (0-9).
.KNN is a [9] simple instance-based learning technique that makes predictions by searching through the entire labeled training dataset to find the k closest matches to the input test image. The test image is then

assigned the class label of the majority of the k- nearest neighbors identified. KNN compares images using distance metrics in high dimensional feature space. A key aspect is finding the optimal value of k neighbors by hyperparameter tuning. SVM is a discriminative classifier that constructs optimal linear decision boundaries between classes in higher dimensional space. It maximizes the margin between different classes for reliable separability. The digit images, which cannot be linearly separated in input space, can become separable by SVM's kernel functions like radial basis function. This provides nonlinear decision boundaries. Unlike the previous two, CNN is able to automatically learn deep hierarchical visual features from digit images through consecutive convolutional and pooling layers. This alleviates the need for manual feature engineering. Each neuron connects to a region in input and scans for specific patterns. The relevant activations produced form distinctive features for each digit class. This deep learned representation combined with classifiers then identifies digit classes. The key difference between the selected models lies in how they learn to transform raw input images for reliable discrimination between digit classes to maximize test accuracy on this problem. The findings will provide insight into design of optimal handwritten digit classifiers.

.

II. LITERATURE SURVEY

Handwritten text recognition is an important application of deep learning and machine learning which is helpful in detecting forgeries and Numerous research endeavors have undertaken an exhaustive exploration, incorporating in-depth examinations and practical applications of diverse well- known algorithms. Pioneering works by researchers such as S M Shamim [3], Anuj Dutt [4], Norhidayu binti [5], and Hongkai Wang [8] have been pivotal in systematically comparing various CNN models against fundamental machine learning algorithms. These comparative assessments span multiple facets, including performance rates, execution times, complexities, and other pertinent criteria, aiming to provide explicit evaluations of each algorithm's efficacy. [3] concluded that the Multilayer Perceptron classifier gave the most accurate results with minimum error rate followed by Support Vector Machine, Random Forest Algorithm, Bayes Net, Naïve Bayes, j48, and Random Tree respectively while
[4] presented a comparison between SVM, CNN, KNN, RFC and were able to achieve the highest accuracy of 98.72% using CNN (which took maximum execution time) and lowest accuracy using RFC. [5] did the detailed study-comparison on SVM, KNN and MLP which we are not considering at this point in our study models to classify the handwritten text and digits. In the study by Hongkai Wang [8], emphasis is placed on the comparative analysis between deep learning and machine learning methodologies.

The primary objective is to discern the suitability of these approaches for classifying mediastinal lymph node metastasis in non-small cell lung cancer using 18 F-FDG PET/CT images. Furthermore, the research delves into assessing the discriminative capabilities of PET/CT texture features, a contemporary trend, against conventional diagnosticfeatures, contributing valuable insights to the ongoing discourse in the field.
The study's findings suggest that the effectiveness of CNN is comparable to the top classical approaches and human medical professionals in the classification of mediastinal lymph node metastasis of NSCLC based on PET/CT images. Nevertheless, CNN neglects the utilization of crucial diagnostic features, which have demonstrated greater discriminative potential than texture features in the classification of small-sized lymph nodes. Consequently, an encouraging avenue for future research involves integrating these diagnostic features into the CNN framework. The key to achieving our objectives lies in amassing extensive data and information, enabling the training of substantial neural networks capable of fulfilling specific tasks. In essence, a convolution can be construed as an analytical process, scrutinizing the surroundings to formulate precise prognoses about outcomes [6].

Additionally, in the work by [7], a convolutional neural network was applied to handwritten digit recognition using MNIST datasets. Employing a 7-layered CNN model with 5 learning to document recognition, providing a foundational reference for subsequent investigations in the field. In parallel, Ceresin et al. (2010) explored the efficacy of deep neural networks for handwritten digit recognition, shedding light on the potential advantages of leveraging complex architectures. The comparative analysis extends to traditional methods, with Zhang and LeCun (1998) delving into efficient feature extraction techniques for handwritten digit recognition using SVMs. Cortes and Vatnik's pivotal work (1995) introduced Support Vector Networks, while Burges' tutorial (1998) comprehensively outlined the principles of support vector machines, providing essential insights into their role in digit recognition.

In the context of KNN, Wu et al. (2002) augmented the understanding of KNN classification in the spatial domain, emphasizing enhancements to this classical algorithm. On the deep learning frontier, Krizhevsky et al. (2012), Simonyan and Zisserman (2014), and Szegedy et al. (2015) pioneered the application of deep convolutional neural networks to image classification, showcasing the transformative potential of CNNs. Further contributions from Goodfellow et al. (2016) encapsulate a comprehensive exploration of deep learning principles, establishing a theoretical foundation for the subsequent advancements in neural network architectures. Additionally, Sermanet et al. (2013) applied CNNs to house numbers digit classification, further underlining the versatility and efficacy of these models. This literature survey encapsulates a rich tapestry of research, highlighting the nuanced distinctions in accuracy among KNN, SVM, and CNN models for handwritten digit recognition. These seminal studies collectively propel the field forward, offering a comprehensive understanding of the strengths and limitations inherent in each approach.

III. METHODOLOGY

The comparison of the algorithms (Support vector machines, KNN (K-Nearest neighbors) and Convolutional neural network) is based on the characteristic chart of each algorithm on common grounds like dataset, the number of epochs, complexity of the algorithm, accuracy of each algorithm, specification of the device (Windows 20.04 LTS, i5 10th gen processor) used to execute the program and runtime of the algorithm, under ideal condition.

A. DATASET

hidden layers, coupled with gradient descent and The dataset that I have downloaded from the MNIST database backpropagation, [6] systematically investigated and contains 60,000 images of handwritten digits, from zero to nine, all compared accuracy across different epochs, contributing grouped in one file. Each of the images is of size 28 by 28 pixels valuable insights to the field and getting maximum accuracy and represents a digit. I have noticed that there is no pattern or order of 99.2% while in [7], they have briefly discussed different to the way the images were organized in the file. The images are

components of CNN, its advancement from LeNet-5 to Senet represented as matrices, of which the elements represent the pixels.

and comparisons between different model like Alex Net, Also, each image has a label that indicates the digit represented.

DenseNet and ResNet LeCun et al.'s seminal work (1998Moreover, the dataset did not exhibit any noise or significant issues

laid the groundwork for the application of gradient-based that required preprocessing; hence it was utilized in its raw form. Special Database 1 and Special Database 3, which contain digits written by high school students and employees of the United States Census Bureau, respectively, were included in the analysis. MNIST comprises a total of 70,000 images of handwritten digits (60,000 in the training set and 10,000 in the test set), each presented in a 28x28 pixel bounding box with anti-aliasing.

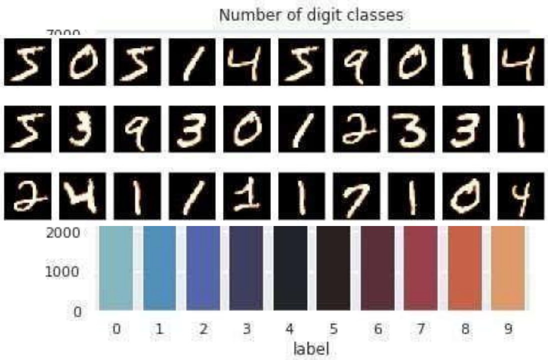

Figure 1. Bar graph illustrating the MNIST handwritten digit training dataset (Label vs Total number of training samples).

Figure 2. Plotting of some random MNIST Handwritten digits

B. KNN

The methodology for employing K-Nearest Neighbors (KNN) in handwritten digit recognition begins with dataset preparation. Grayscale values are normalized, scaling them between 0 and 1 to expedite model training and mitigate local optima challenges. Data arrays are reshaped from 3 dimensions to 2 dimensions, aligning with KNN's 2D array input expectation. The MNIST dataset [10] is then divided into training and testing sets using the train_test_split function from sklearn's model_selection module. A test_size of 0.2 reserves 20% of the data for testing, ensuring label distribution consistency in both sets through the stratify=y parameter. Reproducibility is ensured by utilizing the random_state parameter. KNN's performance hinges on hyperparameter selection. Three crucial hyperparameters are considered: n_neighbors (number of neighbors), weights (weight function - 'uniform' or 'distance'), and p (power parameter for the Minkowski metric). A range of values for n_neighbors is explored (3 to 11), and two weight functions are considered.

The KNN model is initialized using KNeighborsClassifier(), followed by a grid search to determine the optimal hyperparameter combination. GridSearchCV from sklearn's model_selection module facilitates this process, training KNN models for each hyperparameter combination and employing cross-validation for performance evaluation.

C. SUPPORT VECTOR MACHINE

SVMs are known as high performance pattern classifiers. While Neural Networks aim at minimizing the training error, SVMs have a goal to minimize the "upper bound of the generalization error". SVMs are trained on extracted feature vectors from the input digit images to create a model that can classify new examples. Common features include pixel densities, contours, histograms of gradient orientations, etc. The SVM model tries to find an optimal hyperplane that maximizes the margin between different digit classes in the high-dimensional feature space. Kernel functions like radial basis functions are often used to transform the feature space to make it more linearly separable for the SVM. Since [1] SVMs are binary classifiers, multi-class SVMs are created to classify the 10-digit categories (0-9). This is done by training multiple binary SVMs and combining their outputs. SVMs are effective for digit recognition because the model is robust to variations in input patterns. The maximal margin improved generalization. SVMs are relatively memory efficient since only the support vectors are needed during prediction. This is useful for handheld devices. SVMs provide probabilistic class outputs, which can be used to improve prediction accuracy through further post processing. SVMs avoid overfitting on limited training data compared to neural networks. This is important as labeled digit data is scarce.

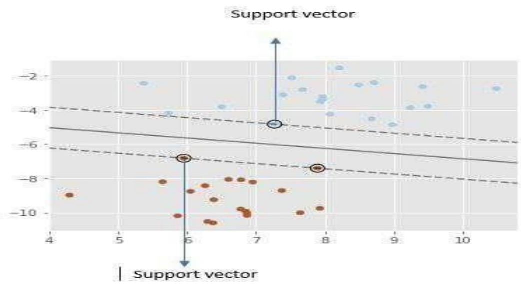

Figure 3. This image describes the working mechanism of SVM Classification with supporting vectors and hyperplanes.

The methodology for employing Support Vector Machines (SVM) in handwritten digit recognition involves utilizing supervised learning for classification tasks. SVM, adept at handling both classification and regression problems, operates by establishing optimal hyperplanes to distinguish distinct categories. The [12] process begins with plotting data points in a two-dimensional space, followed by the identification of a suitable hyperplane or linear/nonlinear plane to effectively separate the two classes. In the context of binary classification, SVM strives to pinpoint the correct hyperplane while maximizing the margin—measured as the maximum distance between the hyperplane and the nearest data point from either class—on both sides. This emphasis on margin maximization enhances the robustness of the classifier and diminishes the likelihood of misclassification. SVM's resilience to outliers is acknowledged, and the algorithm actively seeks a hyperplane with a maximum margin to improve generalization across diverse data distributions, contributing to its efficacy in handwritten digit recognition tasks.

D. CONVOLUTIONAL NEURAL NETWORK

CNN, a deep learning algorithm widely employed for image recognition and classification, belongs to the category of deep neural networks that necessitate minimal pre- processing. Instead of processing a single pixel at a time, CNN takes in the image as small chunks [13], enhancing its ability to efficiently detect uncertain patterns, such as edges. Comprising three layers—input, output, and multiple hidden layers, including Convolutional layers, Pooling layers (Max and Average pooling), Fully connected layers (FC), and normalization layers—CNN utilizes a filter (kernel), an array of weights, to extract features from the input image. Each layer in CNN employs different activation functions, introducing non-linearity to the process. Progressing through the CNN, there is an observable decrease in both height and width while the number of channels increases. Ultimately, the generated column matrix is employed to predict the output. The methodology for utilizing Convolutional Neural Networks (CNN) in addressing handwritten digit recognition involves a well-structured neural network architecture with input, output, and multiple hidden layers.

CNN's architecture begins with the Convolutional Neural Layer (CNL), the first layer that memorizes input image features using vertical and horizontal sliding filters. The rectified linear unit (ReLU) activation function is applied to the output of this layer. Subsequently, the Pooling Layer (PL) performs max pooling or subsampling, reducing data volume for more efficient network computation. The fully connected layer follows, with a neuron for each pixel corresponding to the predicted class number.

In this context, the number of neurons aligns with the digit classes (0–9). CNN's key strength in optical character recognition (OCR) lies in its integrated feature extraction and classification capabilities, eliminating the need for separate algorithms. The use of the Deeplearning4j (DL4J) framework with the CNN architecture is deemed optimal for character recognition from handwritten digit images. The [11] experiments employ the normalized standard MNIST dataset to validate the system's performance, leveraging CNN's ability to map datasets accurately with high accuracy by adjusting trainable parameters and hidden layer configurations.

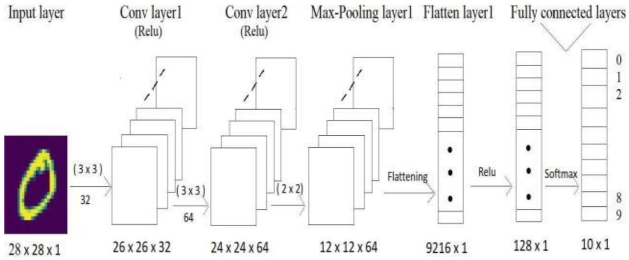

Figure 4. Detailed architecture of Convolutional Neural Network with apt specifications of each layer

IV. RESULT

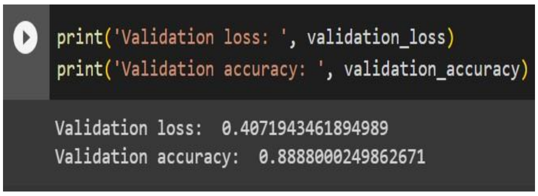

Figure 5. shows accuracy achieved by CNN.

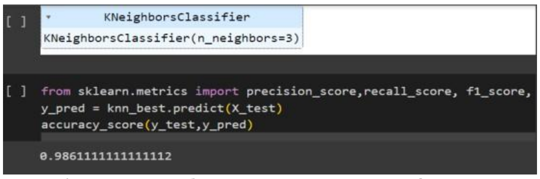

Figure 6. shows accuracy of KNN

The SVM model required the least execution time as it involves simpler mathematical operations relative to deep neural networks. In contrast, the CNN model incurred the highest execution time due to greater computational requirements for propagating data through multiple layers. However, model inference time is more crucial than training time for real-world usage.

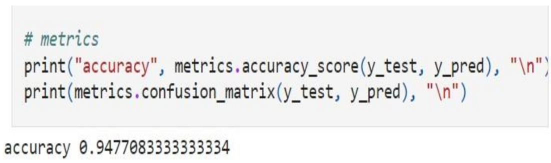

Figure 7. shows accuracy of SVM model

In summary, CNNs achieve state-of-the-art accuracy on image datasets but require greater training resources. SVMs offer faster training but lag in accuracy. SVMs fall between CNNs and KNNs on both metrics. These findings will guide appropriate model selection based on accuracy, speed, and resource constraints for a given handwritten digit recognition application. We show different model's accuracy in Table.1. Further research can explore model ensembles and architecture optimizations to improve efficiency.

MODEL	ACCURACY
CNN	0.89
SVM	0.947
KNN	0.98

Table 1. Shows the accuracy of different models for handwritten digit recognition.

This research implemented and evaluated three machine learning models for handwritten digit recognition on the MNIST dataset - Support Vector Machine (SVM) as shown in Fig.7, Convolutional Neural Network (CNN) as shown in Fig.5 and KNN as shown in Fig.6. The models were trained on 60,000 images and tested on 10,000 images to compare their performance across key metrics. The KNN model attained highest training accuracy, indicating strong memorization of the training data. However, the CNN model achieved superior testing accuracy by a slight margin, demonstrating better generalization on new data. This highlights the generalization capability of deep neural networks like CNNs. Across models, testing accuracy trailed training accuracy due to overfitting, an expected occurrence in machine learning pipelines. Execution time serves as a proxy for computational complexity and training efficiency.

CONCLUSION

In this research endeavor focused on handwritten digit recognition using MNIST datasets, we systematically implemented and compared three distinct models based on deep and machine learning algorithms. Evaluating their characteristics, we aimed to identify the most accurate model for this task. Among the classifiers, Support Vector Machines (SVM) emerged as a fundamental and efficient choice, demonstrating a commendable training accuracy rate. However, its simplicity renders it less adept at classifying complex and ambiguous images when compared to more robust algorithms like KNN and more intricate Convolutional Neural Networks (CNN).

Upon comprehensive assessment, CNN notably stood out by providing the most accurate results for handwritten digit recognition. This outcome underscores the effectiveness of CNN in handling image data inputs and positions it as the optimal choice for prediction problems of varied complexity. Furthermore, our comparison of execution times revealed a crucial insight – increasing the number of epochs without modifying the algorithm's configuration proved futile. This observation is attributed to the inherent limitations of certain models, leading to overfitting beyond a specific epoch count and yielding biased predictions. Consequently, our research not only highlights the supremacy of CNN in handwritten digit recognition but also emphasizes the importance of thoughtful epoch selection to mitigate overfitting challenges in machine learning applications.

V. FUTURE ENHANCEMENT

The future development of the applications based on algorithms of deep and machine learning is practically boundless. In the future, we can work on a denser or hybrid algorithm than the current set of algorithms with more manifold data to achieve the solutions to many problems.

In future, the application of these algorithms lies from the public to high-level authorities, as from the differentiation of the algorithms above and with future development we can attain high-level functioning applications which can be used in the classified or government agencies as well as for the common people, we can use these algorithms in hospitals application for detailed medical diagnosis, treatment and monitoring the patients, we can use it in surveillances system to keep tracks of the suspicious activity under the system, in fingerprint and retinal scanners, database filtering applications, Equipment checking for national forces and many more problems of both major and minor category. Advancement[14] in this field can help us create an environment of safety, awareness and comfort by using these algorithms in day to day application and high-level application (i.e. Corporate level or Government level). Application-based artificial intelligence and deep learning is the future of the technological world because of their absolute accuracy and advantages over many major problems.

VI. References

[1] [Online]. Available: https : //en.wikipedia.org/wiki/Handwritingrecognition.

[2] [Online]. Available: https : //www.quora.com/What can a digit recognizer be used for.

[3] M. B. A. M. A. S. M. R. &. A. A. J. S M Shamim, ""Handwritten Digit Recognition using Machine Learning Algorithms"".

[4] A. D. a. A. Dutt., ""Handwritten Digit Recognition Using Deep Learning",".

[5] N. N. B. A. S. Norhidayu binti Abdul Hamid, ""Handwritten recognition using SVM, KNN, and Neural networks"".

[6] S. S. M. A. B. S. Fathma Siddique, ""Recognition

of Handwritten Digit using Convolutional Neural Network in Python with Tensorflow and Comparison of Performance for Various Hidden Layers",".
[7] A. S. &. P. D. Farhana Sultana, """Advancements in Image Classification using Convolutional Neural Network""".
[8] Z. Z. Y. L. C. P. L. W. W. W. L. a. L. Y. Hongkai Wang, """Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from 18 F- FDG PET/CT images".
[9] [Online]. Available: https : //www.javatpoint.com/machine learning support vector machine algorithm.
[10] [Online]. Available: https://machinelearningmastery.com/neural- networks-crash-course/.
[11] [Online]. Available: researchgate.net/publication/285164623 An Introduction to Convolutional Neural Networks.
[12] Ü. Ç. Kübra Gülgün Demirkaya, "Handwritten Digit Recognition With Machine Learning Algorithms".
[13] S. V. S. P. Bharadwaj Yellapragada, "Effective Handwritten Digit Recognition using Deep Convolution Neural Network".
[14] A. S. I. S. K. Yevhen Chychkarova, "Handwritten Digits Recognition Using SVM, KNN, RF and Deep learning neural network".
[15] H. B. A. B. A. S. D. A. W. I. Ayushi Sharma, "A Machine Learning and Deep Learning Approach for Recognizing Handwritten Digits".