

Comparing Image Recognition Algorithms in Artificial Intelligence

Devanshi Patel^a, Nishith Parmar^{b,*}

^aGandhinagar Institute of Technology, Khatraj-Kalol Road, Moti Bhoyan, Kalol and 382721, India

^bGandhinagar Institute of Technology, Khatraj-Kalol Road, Moti Bhoyan, Kalol and 382721, India

Abstract

Image Recognition is a crucial component of artificial intelligence which finds many applications in fields from the self-driving car to the detection and diagnosis of diseases. Different Algorithms have been developed to give high accuracy and effectiveness in image recognition. Each has its strengths and weaknesses. This paper compares the most widely used algorithms, including convolutional neural networks (CNNs), decision trees, k nearest neighbors (KNN), support vector machines (SVMs) and random forests. The comparison is based on their accuracy, training and testing time and performance on standard image recognition datasets. Our analysis reveals that CNNs are the most favourable algorithm for image recognition due to their high accuracy and their capability to learn complex features from images. Overall, this paper delivers insights into the performance and practicality of different image recognition algorithms and can serve as a valuable resource for researchers and practitioners in the field of artificial intelligence.

Keywords: Image Recognition; Convolutional neural network; Random forests; Support vector machines; K nearest neighbors

1. Introduction

Image recognition is one of the most important and popular applications of artificial intelligence (AI). Image recognition is the process of identifying and detecting objects, patterns, or features in digital images. It involves using computer algorithms and techniques to analyze the visual content of an image and make sense of it. Image Recognition finds its application in many fields like Object recognition, Autonomous vehicles, Security, Entertainment, Agriculture, Medical diagnosis, Face recognition and Environmental monitoring. For example, in healthcare, image recognition can be used to diagnose diseases and detect abnormalities in medical images such as X-rays and MRI scans. In security, image recognition can be used to identify individuals and track their movements in surveillance footage. In manufacturing, image recognition can be used to inspect products for defects and ensure quality control. In agriculture, it can be used to monitor crops and yield prediction. It can help farmers detect diseases and pests, track plant growth, and optimize harvests. In the entertainment industry, it can be for video and image analysis. It can help identify copyright violations, monitor content for inappropriate material, and provide personalized recommendations to users. As technology is evolving, we can expect to see even more innovative and impactful uses emerge in the coming years. Image recognition systems typically work by analyzing the features of an image and comparing them to known patterns or templates. This is done using complex algorithms that are designed to identify specific features or characteristics of an object or scene, such as shape, color, texture, and spatial arrangement. These algorithms can be trained using large datasets of labeled images, allowing them to recognize patterns and make accurate predictions about new images.

Image Recognition is one of the important and old problems in artificial intelligence. It is a challenging task to detect objects from the image [1]. The capability to accurately identify and classify the objects in the image has been a problem for researchers. AI has made this difficult task possible by making devices as intelligent as humans to recognize patterns in different conditions [2]. There are many image classification algorithms where the most commonly used are machine learning algorithms and deep learning algorithms [3]. The process of Deep learning includes self-learning by building a multilayer model and training it with large amounts of data. It helps in improving the accuracy rate and performance of the classification or prediction [4]. The image recognition process based on deep learning mainly gives the image into the neural network. Then it uses the deep learning forward propagation and backpropagation error algorithms so that it can minimize the loss of input function [5]. Here, several machine learning and deep learning algorithms are selected such as convolutional neural networks (CNNs), support vector machines (SVM), k nearest neighbors (KNN), decision trees and random forests.

Artificial Neural Network (ANN) has been studied for many years to solve complex classification problems including image classification [6]. One advantage of artificial neural networks is that the algorithm could be generalized to use in different kinds of situations using the same designs [7]. Convolutional Neural Networks (CNNs) are a class of deep learning neural networks that are generally used for image recognition tasks. CNNs are designed to learn spatial orders of features from input data, such as

* Corresponding author. Tel.: +91 8849694865
E-mail address: devanship824@gmail.com

images, on their own without the need for manual feature extraction [8]. CNN can learn the local and global structures from image data effectively. SVM is a mathematical extension of Neural Network, and it performs classification by converting the original training data into multidimensional space and constructing hyper-plane in higher dimensional [9]. Decision Trees are a class of supervised machine learning algorithms used for classification [10]. The effectiveness of Decision Tree algorithms is that they provide rules of classification which are easily understandable [11]. It has some drawbacks too, one of them is when the tree chooses to split a node, the sorting of all numerical attributes is required. It can become costly if Decision Trees are set on the larger size of data i.e. it has more instances [12]. KNN classification algorithms determine the class of an image by looking for the k images of the training set which are most similar to the image to be classified, and then it performs a class-weighted frequency analysis [13]. A random forest, a combination of decision trees, is constructed by randomly selecting trees from a set of possible trees, where each tree is trained on a random subset of input data and features [14]. Random trees can be created effectively and the collection of large sets of random trees commonly leads to accurate models. There has been comprehensive research on Random trees in current years [12]. These algorithms are compared based on their precision, recall, f1 score, accuracy, and speed.

2. Methodology

2.1. Model Diagram:

Here, First the data has been taken from dataset then after pre-processing, we separated it into two parts, training set and testing set. The data in training set was trained using different algorithms. Then testing conducted on the testing set and it gave results which helps us to evaluate the performance and accuracy of an algorithm. The following figure shows the flow of process conducted in constructing the model.

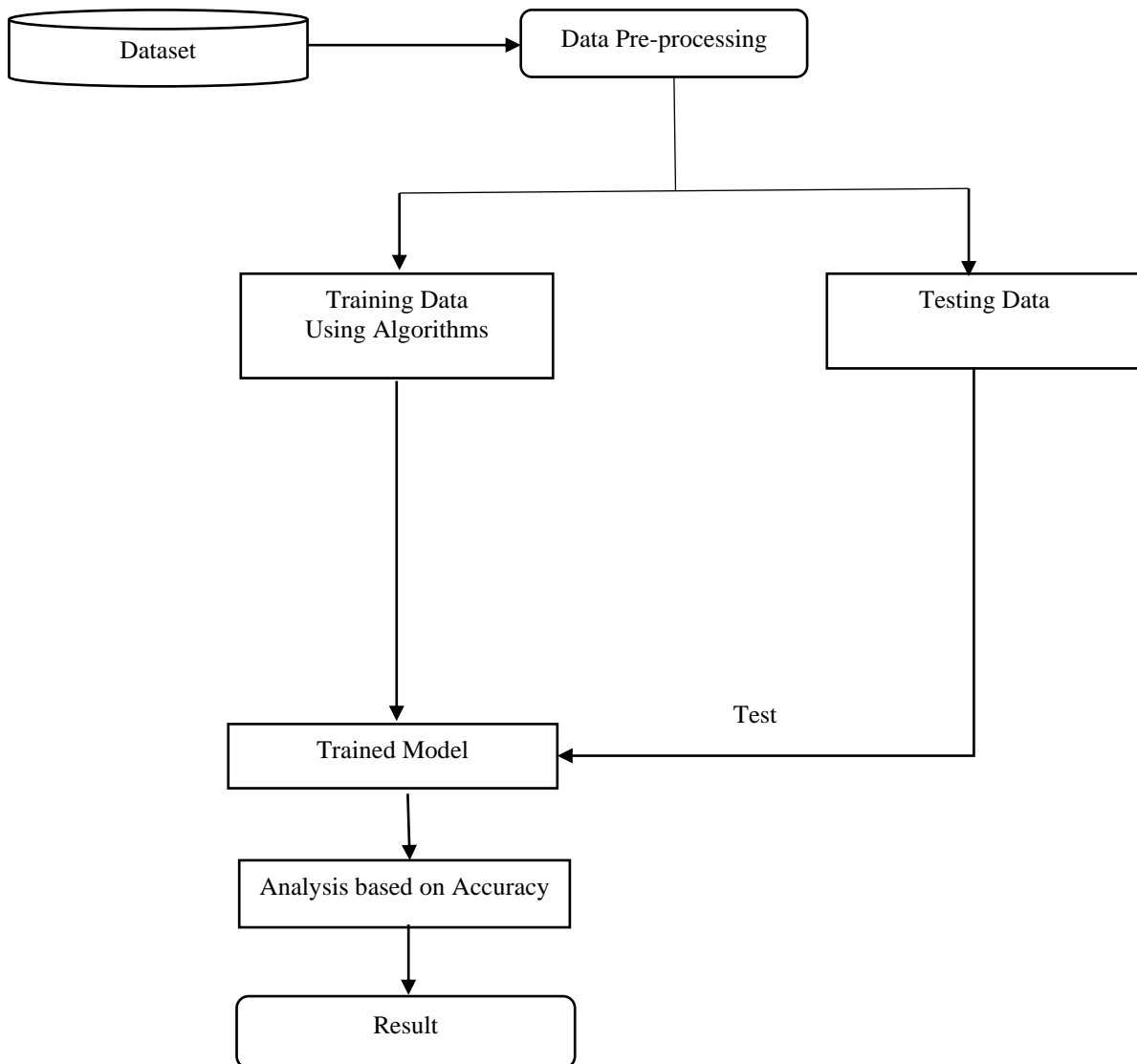


Fig. 1. Illustration of process flow of model

2.2. Dataset Used:

For this work, the MNIST dataset is used for an image classification task. It consists of a collection of 70,000 handwritten digits (0-9) that have been normalized and centered in a fixed-size image of 28x28 pixels. The dataset is commonly used as a standard for training and evaluating machine learning models. The dataset is divided into two sets: a training set of 60,000 images and a test set of 10,000 images. The goal of the MNIST dataset is to classify the digits correctly based on their pixel values. Each image is represented as a 784-dimensional vector, where each element represents the grayscale value of a single pixel. The task is to train a machine learning model on the training set to predict the correct label of each digit in the test set.

2.3 Algorithms Used:

2.3.1. Convolutional Neural Network (CNN):

In recent years, CNN finds its application in many fields such as face recognition, speech recognition [15], natural language analysis [16], brain wave analysis [17], [5]. The convolutional neural network is a type of artificial neural network used as a deep learning algorithm which is designed to process inputs like images and videos. CNN is composed of several layers of processing, including convolutional layers, pooling layers, and fully connected layers [18]. A convolutional model takes input images and convolves them with kernels to extract features of the image [3]. The result of the convolution operation is a single number that shows the degree of similarity between the filter and the input region. The pooling layer reduces the spatial dimensions of the feature maps and makes them more manageable for the following layers. The fully connected layer performs a classification or regression task based on the features extracted by the earlier layers.

2.3.2 Support Vector Machines (SVM):

Support Vector Machine is a supervised machine learning algorithm which is used for classification and regression analysis. The SVM algorithm works by finding the optimal hyperplane that splits the data into different types. The hyperplane is selected in such a way that it maximizes the distance between the closest data points of different classes, which are called support vectors [19]. This hyperplane is called the Optimal Separating Hyperplane whose task is to minimize the risk of misclassifying the examples in the training set as well as the unseen examples of the test set [20].

2.3.3 Decision tree:

Decision trees are commonly used supervised machine learning algorithms that are used for both classification and regression. Decision trees consist of nodes. At each node, the algorithm picks the feature or attribute that provides the most information gain to divide the data into the different classes. The decision tree is built by selecting the best attribute to separate the data at each node until all the data points are correctly classified.

2.3.4 K nearest neighbors (KNN):

The k-nearest neighbor algorithm is the simplest algorithm out of all machine learning algorithms. The algorithm works by finding the k closest data points in the training set. The algorithm works by finding the distance between the input data point and every other data point in the training set, using a distance metric such as Euclidean distance or Manhattan distance. The k closest data points, also known as the nearest neighbors, are then used to determine the predicted label of the input data point [21]. KNN is an easy-to-understand algorithm that can be used for both classification and regression tasks.

2.3.5 Random Forests:

A random forest is a popular decision tree ensemble learning algorithm used in machine learning for classification and regression tasks. It means that it combines multiple weaker models to create a more powerful and accurate model. In the random forest algorithm, each tree grows with some type of randomization [22]. This algorithm creates a forest of decision trees. Each tree will be trained on a random subset of training data and features. At the time of prediction, each tree in the forest is used to make a prediction and the final prediction is determined by aggregating the predictions of all the trees.

3. Evaluation

The evaluation is done based on various factors such as confusion matrix, precision, f1 score and recall. A confusion matrix is used to show the number of correct and incorrect predictions made by the model. It consists of two dimensions: predicted class and actual class. It has four possible outcomes: True Positive (TP) which means the prediction is correct, and False Positive (FP) which means the model predicted a positive class but the prediction is incorrect. Then we have True Negative (TN) which means the model predicted a negative class with a correct prediction, and False Negative (FN) where the model predicted a negative class with an incorrect prediction. Precision, recall and F1 score are the parts of the classification report which help to evaluate the performance of the model. Precision is a measure which calculates the percent of correctly identified positive cases out of all positive cases identified by a model. Precision can be given as:

$$Precision = \frac{TP}{(TP + FP)} \quad (1)$$

Recall is a measure that calculates percent of correctly out of all the actual positive cases that exist in a population. Recall can be expressed as follows:

$$Recall = \frac{TP}{(TP + FN)} \quad (2)$$

F1 score is a measure that combines precision and recall into a single metric. The F1 score is a harmonic mean of precision and recall and is expressed as below:

$$F1 = \frac{2 * Precision * Recall}{(Precision + Recall)} \quad (3)$$

4. Result

The classification report for different algorithms is shown as from table 1 to table 5. The classification report contains the Precision, recall and f1 score of the algorithm.

Table 1: Classification report of CNN

	precision	recall	f1 score
0	0.98	0.99	0.98
1	0.99	1.00	0.99
2	0.98	0.98	0.98
3	0.98	0.99	0.98
4	0.98	0.99	0.98
5	0.99	0.97	0.98
6	0.98	0.99	0.98
7	0.98	0.98	0.98
8	0.99	0.95	0.97
9	0.98	0.98	0.98

Table 2: Classification report of SVM

	precision	recall	f1 score
0	0.95	0.98	0.96
1	0.97	0.99	0.98
2	0.92	0.94	0.93
3	0.91	0.93	0.92
4	0.93	0.96	0.94
5	0.91	0.89	0.90
6	0.96	0.95	0.95
7	0.95	0.93	0.94
8	0.93	0.90	0.91
9	0.95	0.92	0.93

Table 3: Classification report of Decision trees

	precision	recall	f1 score
0	0.91	0.94	0.93
1	0.96	0.96	0.96
2	0.87	0.86	0.86
3	0.83	0.85	0.84
4	0.87	0.88	0.87
5	0.84	0.83	0.84
6	0.90	0.89	0.89
7	0.92	0.90	0.91
8	0.82	0.81	0.82
9	0.85	0.85	0.85

Table 4: Classification report of KNN

	precision	recall	f1 score
0	0.98	0.99	0.99
1	0.96	1.00	0.98
2	0.98	0.96	0.97
3	0.96	0.96	0.96
4	0.97	0.96	0.97
5	0.95	0.96	0.95
6	0.98	0.99	0.98
7	0.96	0.97	0.96
8	0.98	0.94	0.96
9	0.95	0.95	0.95

Table 5: Classification report of Random forests

	precision	recall	f1 score
0	0.97	0.99	0.98
1	0.99	0.99	0.99
2	0.97	0.97	0.97
3	0.96	0.96	0.96
4	0.97	0.98	0.98
5	0.98	0.96	0.97
6	0.97	0.98	0.98
7	0.97	0.96	0.97
8	0.96	0.96	0.96
9	0.96	0.96	0.96

From the above data, it is clear that convolutional neural network algorithm is showing maximum precision for all the classes. Also, it is the algorithm with maximum recall value for all the classes. CNN's f1 score is highest and consistent among all the classes as compared to other algorithms. Accuracy of different algorithms are shown in the table 6. Accuracy of an algorithm can be calculated as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (4)$$

From table 6, it is analysed that Convolutional neural networks showing the maximum accuracy. So, we can say that CNN can recognize and classify the images with more accuracy than any other algorithms.

Table 6. Performance of Algorithms

Algorithms	Accuracy %
Convolutional neural networks (CNN)	98.00
Support Vector Machines (SVM)	94.00
Decision Tress	88.00
K nearest neighbour (KNN)	97.00
Random Forest	97.00

5. Conclusion

In conclusion, the comparison of different algorithms for image recognition is an important area of research that can provide valuable insights into the strengths and weaknesses of different approaches. We have reviewed and compared several popular algorithms for image recognition, including convolutional neural networks (CNNs), support vector machines (SVMs), and k-nearest neighbors (KNNs), decision trees and random forests. The finding of this paper indicates that Convolutional neural networks are an effective deep learning algorithm for image recognition. This paper has compared different image recognition algorithms to evaluate their performance based on accuracy, precision, recall and f1 score. Our analysis has shown that CNN is currently the most effective algorithm for image recognition as it has achieved high accuracy and speed on a standard dataset. However, SVMs and KNNs can also be effective for certain types of image recognition tasks, such as object recognition and image segmentation. Other factors, such as the size and quality of the training dataset, the complexity of the image features, and the computational resources available, can also impact the performance of different algorithms. It is important to note that the performance of each algorithm depends on the quality and complexity of the image. The study underscores the potential of image recognition technologies to drive innovation and create new opportunities in various industries.

References

1. Chauhan, R., Ghanshala, K. K., & Joshi, R. . (2018). Convolutional Neural Network (CNN) for Image Detection and Recognition. 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC). doi:10.1109/iccccc.2018.8703316
2. Sharma, Priyanka, and Manavjeet Kaur. "Classification in pattern recognition: A review." *International Journal of Advanced Research in Computer Science and Software Engineering* 3.4 (2013).
3. Wang, P., Fan, E., & Wang, P. (2020). Comparative Analysis of Image Classification Algorithms Based on Traditional Machine Learning and Deep Learning. *Pattern Recognition Letters*. doi:10.1016/j.patrec.2020.07.042ss
4. K. Yu, L. Jia, Y. Q. Chen, W. Xu, "Deep learning: yesterday, today, and tomorrow," *Journal of Computer Research and Development*, vol. 50, pp. 1799-1804, 2013.
5. Y. Tian, "Artificial Intelligence Image Recognition Method Based on Convolutional Neural Network Algorithm," in *IEEE Access*, vol. 8, pp. 125731-125744, 2020, doi: 10.1109/ACCESS.2020.3006097.
6. Le Cun, Y., Jackel, L., Boser, B., Denker, J., Graf, H., Guyon, I., Henderson, D., Howard, R., Hubbard, W.: Handwritten digit recognition: Applications of neural network chips and automatic learning. *IEEE Communications Magazine* 27(11), 41–46 (1989)
7. Q. Li, W. Cai, X. Wang, Y. Zhou, D. D. Feng and M. Chen, "Medical image classification with convolutional neural network," 2014 13th International Conference on Control Automation Robotics & Vision (ICARCV), Singapore, 2014, pp. 844-848, doi: 10.1109/ICARCV.2014.7064414.
8. T. Guo, J. Dong, H. Li and Y. Gao, "Simple convolutional neural network on image classification," 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA), Beijing, China, 2017, pp. 721-724, doi: 10.1109/ICBDA.2017.8078730.
9. Chandra, M.A., Bedi, S.S. Survey on SVM and their application in image classification. *Int. j. inf. tecnol.* **13**, 1–11 (2021). <https://doi.org/10.1007/s41870-017-0080-1>
10. "Comparison of Decision Tree methods for finding active objects" Yongheng Zhao and Yanxia Zhang, National Astronomical Observatories, CAS, 20A Datun Road, Chaoyang District, Beijing 100012 China
11. Yael Ben-Haim, "A Streaming Parallel Decision Tree Algorithm", Elad Tom-Tov , 2010
12. Ali, Jehad, et al. "Random forests and decision trees." *International Journal of Computer Science Issues (IJCSI)* 9.5 (2012): 272.

13. Amato, G., & Falchi, F. (2010). kNN based image classification relying on local feature similarity. Proceedings of the Third International Conference on Similarity Search and Applications - SISAP '10. doi:10.1145/1862344.1862360
14. Bosh, A., Zisserman, A., Munoz, X.: Image classification using Random Forests and ferns. In: IEEE ICCV (2007)
15. W. Jing, T. Jiang, X. Zhang, and L. Zhu, "The optimisation of speech recognition based on convolutional neural network," Int. J. High Perform. Comput. Netw., vol. 13, no. 2, pp. 222–231, 2019.
16. S. Bacchi, L. Oakden-Rayner, T. Zerner, T. Kleinig, S. Patel, and J. Jannes, "Deep learning natural language processing successfully predicts the cerebrovascular cause of transient ischemic attack-like presentations," Stroke, vol. 50, no. 3, pp. 758–760, Mar. 2019
17. Y. Zhang, X. Zhang, H. Sun, Z. Fan, and X. Zhong, "Portable braincomputer interface based on novel convolutional neural network," Comput. Biol. Med., vol. 107, pp. 248–256, Apr. 2019.
18. Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.
19. Guodong Guo, S. Z. Li and Kapluk Chan, "Face recognition by support vector machines," Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580), Grenoble, France, 2000, pp. 196-201, doi: 10.1109/AFGR.2000.840634.
20. V. N. Vapnik. Statistical learning theory. John Wiley & Sons, New York, 1998.
21. J. P. Jose, P. Poornima and K. M. Kumar, "A novel method for color face recognition using KNN classifier," 2012 International Conference on Computing, Communication and Applications, Dindigul, India, 2012, pp. 1-3, doi: 10.1109/ICCCA.2012.6179151.
22. Ko, B. C., Kim, S. H., & Nam, J.-Y. (2011). X-ray Image Classification Using Random Forests with Local Wavelet-Based CS-Local Binary Patterns. Journal of Digital Imaging, 24(6), 1141–1151. doi:10.1007/s10278-011-9380-3