Journal of Artificial Intelligence General Science (JAIGS)

ISSN: 3006-4023 (Online), Volume 6 , Issue 1, 2024 DOI: 10.60087

Home page https://ojs.boulibrary.com/index.php/JAIGS

Analysis of Machine Learning Algorithms for Image Classification

Shanika Kithulgodage

Swansea University, UK

ABSTRACT

This research aims to use a subset of the CIFAR100 dataset to develop and present an empirical examination of the performance of powerful Convolutional Neural Networks (CNN) and Support Vector Machines (SVM) for object detection. Image datasets, CIFAR10, CIFAR 100, and MINIST image datasets are all common benchmark datasets for testing performance. This research examines the well-known dataset CIFAR 100. Both classifiers' important attributes are combined in the constructed models. The suggested model's algorithm is trained and tested using images from the CIFAR-100 dataset. The CIFAR 100 is made up of 100 different classes, each with 600 photos. CNN and SVM's receptive fields aid in automatically extracting the most identifiable features from these images. The experimental results confirm the effectiveness of the implementation by reaching a maximum recognition accuracy of 0.4919 percent. Results were obtained as maximum accuracy of 0.4817% For the CNN on Coarse Super Label and maximum accuracy of 0.3696% For the CNN on Fine Super Label, while SVM on Coarse Super Label resulted in an accuracy of 0.0795% and SVM on Fine super Label resulted in 0.0285%. According to the results, CNN was able to exceed the benchmark of 39.43% and 24.49%, but SVM was unable to do so.

Keywords: Convolutional Neural Networks (CNN), Support Vector Machines (SVM), CIFAR-100 Dataset, Dataset Image Classification Techniques, Deep Learning for Object Detection, Feature Extraction in Image Processing, HOG (Histogram of Oriented Gradients) Features

ARTICLE INFO: *Received: 01.10.2024 Accepted: 20.10.2024 Published: 11.12.2024*

1. Introduction

Due to its expanding applicability in intelligent robotics and visual surveillance, object recognition is currently one of the most concentrated research areas. However, the researchers are still having issues with correct object detection in this domain, such as recognising an object's shape and spotting a slight change, among numerous things. For the correct recognition of complex objects, a sustainable system is necessary to retain its performance despite changes in the object's nature (Krizhevsky et al.,2009). The cornerstone of a long-term optical surveillance system is object detection. Object classification is used in various fields, including intelligent robotics, face and action recognition, video watermarking, pedestrian tracking, autonomous cars, semantic scene analysis, content-based picture retrieval, etc. We believe that a truly sustainable object recognition system must overcome several challenges, including a complex background, different shapes and colours for different objects, continuously moving objects, different angles, and others, because conventionally used unsustainable systems have failed to perform well for complex object classification (Bobić et al.,2016).

Many approaches in computer vision and data analysis have been created to tackle the previously mentioned issues linked to complex objects. The majority of them attempted to develop an ideal solution that would work for a variety of problems, but this proved to be a difficult task (Canny, 1986). Although traditional approaches such as Handcrafted Features (HCF) were used in the last few decades, as time passed, the items and their backgrounds became increasingly complicated, limiting their use. Geometric features Scale Invariant Feature Transformation (SIFT), Difference of Gaussian (DoG), Speeded-Up Robust Features (SURF), and texture features (HARLICK) were among the handcrafted features. (Gupta, & Singh, 2019). On the other hand, recent strategies proposed using a hybrid set of features to improve an object's representation. Unfortunately, the methods failed to comprehend the increasing complexity of objects and images (Camargo and Smith, 2009).

Faced with the constraints mentioned above, the notion of deep learning has recently been proposed in this context, which has shown greater performance while requiring less computing effort. A vast number of pre-trained convolutional neural networks (CNN) models have been proposed as a result of this. Acceptable accuracy has been challenging to obtain even with these contributions. This has given rise to the notion of features fusion, which is a method for integrating multiple feature populations into a single feature space used in various applications, including medical imaging and object categorisation (Majid et al., 2020). The concept of feature fusion does succeed in improving classification accuracy, but only at the expense of greater computing cost.

This paper proposes a complete long-term framework based on a deep learning architecture. I used the CNN (Convolutional Neural Network) model and SVM (Support Vector Machine) for data classification and object detection.

2. Implementation

First CNN (Convolutional Neural Network) model have implemented for the fine sub and coarse super label classification. CNN is a deep learning system used for image recognition. This algorithm groups photos based on their similarity and recognizes objects within scenes. CNN employs picture features such as a cat's tail and ears, an aeroplane's wing and engine, and so on to identify objects in the image. In fact, this mechanism is remarkably similar to how our brain recognizes objects.

We use CNN because traditional neural networks aren't suited for image processing. CNN, on the other hand, is not dissimilar to ANN. Because while the CNN technique uses Artificial Neural Networks in the end, it first gathers information and determines specific features from the image using layers.

We train the model two times for the fine sub and coarse super kinds of labels. We made the CNN model using basic building blocks (ConV, Max pooling and Flatten Layer.) The loss function is set to categorical cross-entropy, and the optimizer is adam. For the training, the batch size is 32.

Support Vector Machine (SVM) aims to represent a multi-dimensional dataset in a space where a hyperplane separates data elements belonging to different classes. The SVM classifier has the ability to minimize the generalization error on unseen data. The separating hyperplane is also called an optimal hyperplane. SVM is found to be good for binary classification and is considered poor for noisy data. The shallow architecture of SVM presents some difficulties in learning deep features.

We also applied SVM (Support Vector Machine) for our two kinds of levels in the present work. We have used an rbf kernel where gamma is set to auto. Before feeding to SVM, we have extracted features by HOG.

3. Methodology.

The model is evaluated for CIFAR 100, and the experimental setup involves the following steps.

At first, the Import Libraries and Read Data were done for CNN training. We processed the target variable and encoded it categorically. Then we transpose the train feature (X_test, X_train) arrays. In these circumstances, transposing is better than reshaping. The arrays are fed to CNN sequential models. At the last Layer, we applied softmax because it is a multiclass classification. Then separately train the fine sub and coarse super dataset. Later we Trained the label dataset by compiling the model and fitting data to the model. We also tested our model with different test sets and then calculated the accuracy. The epoch value is set to 15. **For better accuracy, we can increase the number of epochs, change parameters on layers, or add additional Layers to the model. But for, the fitting process takes a lot of time.** CNN is trained after running several epochs and until the training process converges.

For SVM training, we extracted hog features using normalization; orientations value 9, pixels per cell 8 x 8 and cell per block 2 x 2. We trained both label datasets using the SVM rbf kernel where gamma is set to auto.

The main advantage of utilising the CNN model is that it can better understand spatial relations (relations between neighbouring pixels of an image) between pixels of images; hence it will outperform MLP for complicated images. MLP models, on the other hand, never take into account extensive topology information from the input and are therefore unsuitable for complex situations.

4. Experimental Setup

4.1. Data Set used

The CIFAR-100 is a well-known benchmark dataset utilized to train and test the suggested classification method in this study. The CIFAR-100 picture collection contains a number of super classes of general object images and a number of subclass categories for each superclass. CIFAR-100 is divided into 100 picture classes, each including 600 photos. These 600 photographs are divided into 500 training images and 100 testing images for each class, resulting in 60,000 unique images. (Sharma et.al, 2018). A total of 20 super classes have been created from these 100 classes. This dataset has 100 object classes, such as bus, chair, table, train, and bed, with each class containing 100 samples, making it more difficult to work with. There are 50,000 photos available for training and 10,000 images available for testing in this dataset. This dataset is used to evaluate the proposed technique in this research.

5. Results

The results are given in Table 1. The training results are provided, which show the maximum accuracy of 0.4817% For the CNN on Coarse Super Label and maximum accuracy of 0.3696% For the CNN on Fine Super Label, while SVM on Coarse Super Label resulted in an accuracy of 0.0795% and SVM on Fine super Label resulted in 0.0285%.

Model	Accuracy of the test set
CNN on Coarse Super Label	0.4817
CNN on Fine Super Label	0.3696
SVM on Coarse Super Label	0.0795
SVM on Fine super Label	0.0285

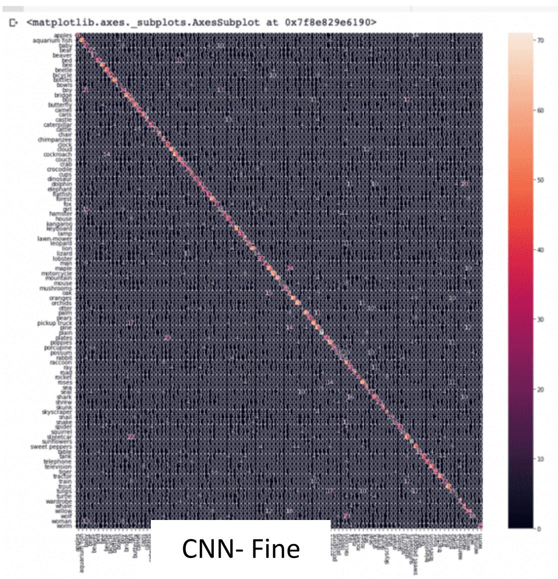
Table 1.

So based on the above result, the CNN achieves the maximum accuracy, although we implemented Hog for feature extraction in SVM. The training accuracy of SVM can be improved by further training, but it cannot cross the CNN performance in these circumstances. The accuracy of CNN models exceeded the benchmark of 39.43%, averaged across all 20 super categories, and 24.49% for the finer granularity

ISSN: 3006-4023 (Online) DOI: 10.60087 Page: 507

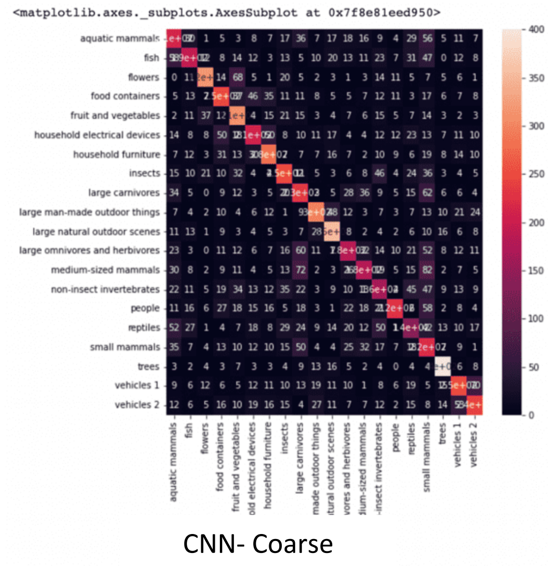
categories, while SVM models unable to reach that level of benchmark. The accuracy of these datasets can be further verified through confusion matrixes shown in figure1. Figure 2., figure 3., and figure 4.

Confusion matrixes for CNN and SVM



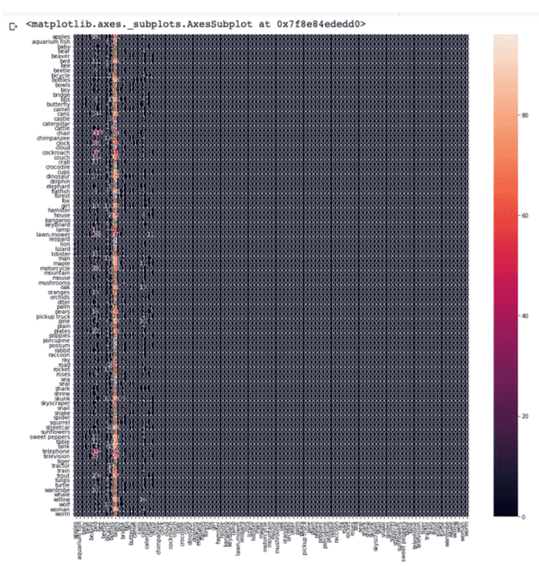
CNN- Fine

Figure1.



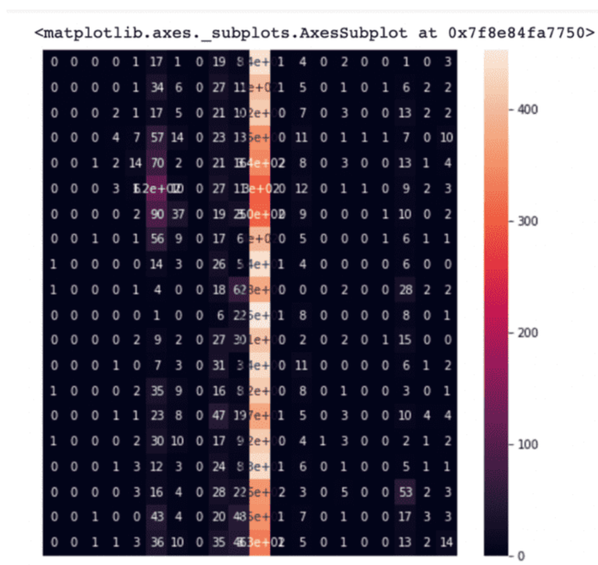
CNN- Coarse

Figure 2.



SVM- Fine

Figure 3.



SVM- Coarse

Figure 4.

6. Conclusion

The study looked at the accuracy of two distinct models, convolutional neural networks (CNN) and support vector machines (SVM), on the most popular training and test dataset, CIFAR100. The major goal of this research was to determine the performance and accuracy of several models on the same datasets, as well as to analyses their results. The results showed that CNN outperformed the benchmark by 39.43 percent and 24.49 percent, respectively, but SVM was unable to do so. It's worth noting that complicated frames can make it difficult for the network to detect and recognize the scene. It was also noticed that, while beds, couches, and chairs are all distinct and easily identifiable objects in the actual world, the trained networks were confused, resulting in differences in accuracy rates.

The results indicated that trained networks with CNN outperformed SVM and had greater accuracy rates. From our experiments, we could easily conclude that the performance of SVM was not much appreciated as it achieved lower accuracy rates and takes more computational time. As a result, the bigger the number of layers in a CNN, the more training it will receive, and thus the higher the rate of prediction accuracy. To summarize, neural networks are the newest and most promising techniques for making a machine intelligent and tackling a variety of real-world item categorization challenges.

References

1. Neha Sharma, Vibhor Jain, Anju Mishra (2018), An Analysis Of Convolutional Neural Networks For Image classification, International Conference on Computational Intelligence and Data Science (ICCIDS 2018).

2. Bobić, V., Tadić, P., &Kvaščev, G. (2016, November) “Hand gesture recognition using neural network-based techniques.” in Neural Networks and Applications (NEUREL), 2016 13th Symposium on (pp. 1-4). IEEE.

3. Camargo, A. and Smith, J.S. (2009) 'Image pattern classification for the identification of disease causing agents in plants,' *Computers and Electronics in Agriculture*, 66(2), pp. 121–125. https://doi.org/10.1016/j.compag.2009.01.003.

4. Canny, J. (1986) 'A Computational Approach to Edge Detection,' *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6), pp. 679–698. https://doi.org/10.1109/tpami.1986.4767851.

5. Krizhevsky, A., & Hinton, G. (2009) “Learning multiple layers of features from tiny images.”
6. Gupta, V.; Singh, J. (2019) “Study and analysis of back-propagation approach in artificial neural network using HOG descriptor for real-time object classification.” In Soft Computing: Theories and Applications; Springer: Berlin, Germany, pp. 45–52.

7. Majid, A.; Khan, M.A.; Yasmin, M.; Rehman, A.; Yousafzai, A.; Tariq, U. Classification of stomach infections: A paradigm of convolutional neural network along with classical features fusion and selection. Microsc. Res. Tech. 2020, 83, 562–576.