CONCEPTUAL RESEARCH PAPER

COMPUTING TRENDZ Vol. XII; No. 1,2, January - December 2022 Print ISSN : 2230-9152, Online ISSN 2456-138X

Machine Learning Algorithms for Image Classification: A Comparative Review

Basu Dev Shivahare[1], Radha Raman Chandan[2]

[1]Galgotias University, Greater Noida, India, [2]SMS, Varanasi

ABSTRACT

In the realm of computer vision, one of the most important challenges is image categorization. Its goal is to assign semantic labels to photographs based on a predetermined set of categories, and it does this via a process called semantic labeling. In attempt to find a solution to this issue, several distinct machine learning algorithms have been developed throughout the course of time; each strategy has its own unique mix of benefits and drawbacks. This article gives an in-depth review and comparison of a wide variety of well-known machine learning approaches for the classification of pictures. This review covers a wide range of algorithms, including more traditional approaches such as Support Vector Machines (SVMs), Convolutional Neural Networks (CNNs), and Transfer Learning using pretrained models such as VGG, Reset, and Inception.

Keywords: Image Classification, Machine Learning, Transfer Learning, Convolutional Neural Networks

Computing Trendz (2022). DOI: https://doi.org/10.21844/cttjetit.v12i1-2.1.14001

Introduction:

Computer vision has only very recently been included into the work of the industrial revolution. The fields of automation, robotics, medicine, and surveillance are just few of the businesses that make substantial use of deep learning. As a result of its success in a variety of applications, including language processing, object identification, and picture classification, deep learning has emerged as the most talked about technology. The projection for the market indicates that extraordinary growth will occur in the next years. The availability of powerful graphics processing units (GPUs) and a sizable number of datasets are often identified as the key reasons for this phenomenon. Image classification and the identification of objects inside images are the two most important components of object detection. There are many different datasets that may be accessed. In the field of image categorization, Microsoft COCO is an industry leader. This dataset is used as a standard for object detection. It offers a big dataset for the identification and categorization of visual content. A comparison of SSD, Faster-RCNN, and YOLO is going to be the focus of this essay. In the current investigation, the first method for comparison is SSD. SSD is a technique that adds layers of numerous features to the final network, which makes it easier to identify

Corresponding Author: Basu Dev Shivahare
E-mail: basuiimt@gmail.com
How to cite this article: Shivahare; B.D., Chandan; R.R. (2022). Machine Learning Algorithms for Image Classification: A Comparative Review. Computing Trendz 12(1&2): 1-8
Source of support: Nil
Conflict of interest: None

individuals [1]. The R-CNN is a unified approach for detecting objects that makes use of a convolutional neural network. It is also quicker and more accurate than other methods. Joseph Redmon is the brains behind the creation of YOLO, which is an end-to-end network. This study compares the relative performance of the three algorithms with the different architectures stated above. The Microsoft COCO dataset is used as a common element in the analysis, and the same metrics are measured across all of the listed implementations. The results of comparing the effectiveness of several algorithms on the same dataset may offer insight into the unique properties of each algorithm, how they vary from one another, and which technique of object identification is more successful in any particular circumstance. This information can be gained by analyzing the outcomes of the comparison.

Objective

The research aimed to fulfill the following objectives:

- Image Classification
- Deep learning architectures for image classification
- Result and discussion

Methodology

We describe the basic ideas, architecture, and essential components of each method, focusing on how they relate to the task of picture categorization. In addition to that, we investigate their performance indicators, the difficulty of their training, and their scalability. Furthermore, we investigate whether or not they are suitable for a variety of picture classification tasks by considering the size of the dataset, the computing resources available, and the level of accuracy that is sought. When it comes to dealing with a wide variety of picture classification issues, such as differences in lighting conditions, object occlusions, and image resolutions, each method has its own set of advantages and disadvantages that we discover via an in-depth comparison examination. In addition, we investigate the effect that hyperparameter tweaking and regularization approaches have on the overall performance of these algorithms and compare the results. The findings of this comparative evaluation provide academics and practitioners with a thorough guidance for selecting the machine learning algorithm that is best suited to accomplish the image classification tasks that are unique to their fields of study. In conclusion, we will discuss the current developments and future prospects in the area of image classification, as well as the possibilities for hybrid techniques. These hybrid approaches combine classical algorithms with deep learning models in order to further improve accuracy and efficiency. Image classification, machine learning techniques, comparative review, Convolutional Neural Networks (CNNs), Support Vector Machines (SVMs), k-Nearest Neighbors (k-NN), Transfer Learning, deep learning, computer vision are some of the keywords that may be found in this article.

Image Classification

Image Classification is a work in Computer Vision Image classification is computer vision work that includes giving a label or category to an input picture based on the image's visual content. Image classification is also known as visual content analysis [3]. The goal is to educate computers to recognize and differentiate between unique items, situations, or patterns that are contained inside images. This will be accomplished via the use of image recognition software. It is a kind of supervised learning, in which algorithms learn from training data that has been labeled in order to make predictions on new photos that they have not seen before. In other words, it is learning from data that has been labeled.

Importance of Image Classification and Its Applications

Image classification plays an important part in a wide variety of real-world applications, boosting automation and decision-making processes in a wide range of industries. Some significant uses include: Assisting in the identification of illnesses by categorizing medical pictures, such as X-rays, MRIs, and CT scans, for the purpose of detecting anomalies and medical disorders is what is referred to as medical imaging.

Autonomous Vehicles:

The capability of self-driving automobiles to recognize and categorize many elements of the surrounding environment, including people, other vehicles, traffic signs, and barriers. Surveillance and security include determining whether or not there are possible dangers or suspicious goings-on in recorded film, which helps with the prevention of crimes and maintains public safety.

Agriculture:

evaluating the condition of crops and identifying any illnesses, pests, or deficits in nutrients using photos captured by satellites or drones.

Enhancing the shopping experience for customers in brick-and-mortar stores and online stores via the use of visual product search, in which customers may identify products that are comparable to a picture they have uploaded.

Artificial Intelligence Assistants:

Increasing the capacity of virtual assistants to comprehend user inquiries and answer appropriately depending on the context of images seen by the user.

Image inspection is used for both quality control and the detection of flaws in the products that are being manufactured. This helps to ensure that products are of a high standard.

Environmental Monitoring is the process of analyzing satellite photos to detect changes in land use, deforestation, and the monitoring of animal populations.

Image Classification Tasks Present a Number of Difficulties and Complexities:

Despite its practicality, picture categorization presents a number of difficulties and complexity, including the following:

- *Variability in the Data:*

  Since images might differ greatly owing to changes in lighting, perspectives, occlusions, and object deformations, it is difficult to generalize across all of the potential variations because of the variability in the data. If deep learning models that include a large number of parameters are not adequately regularized, there is a possibility that the models would overfit the training data, which will result in poor generalization on data that has never been seen before. This risk can be mitigated by properly regularizing the models.

- *Labeling the Data:*

  Producing accurate and comprehensively labeled datasets may be a costly and time-consuming endeavor, particularly for some specialized fields of study.

- *Incorrect Class Distributions:*

  Incorrect class distributions in the training data may lead to biased models, which prefer the class that has the majority of members and perform badly on classes that have less members.

- *Computing Resources:*

  Training deep learning models for image classification often needs a large amount of computing power. This may be a barrier for people or organizations that have limited resources, such as computer hardware or software.

It may be difficult to interpret decisions made by deep learning models since these models, particularly convolutional neural networks, are sometimes referred to as "black boxes." This makes it difficult to comprehend why these models make the predictions that they do. Image classifiers may be subject to minor perturbations that are invisible to humans but can mislead the model's predictions if an adversarial attack is launched against them.

Improving the reliability and performance of picture classification models is the topic of current research that is being conducted in order to solve these difficulties. This research is primarily focused on building more robust algorithms, data augmentation approaches, transfer learning, and interpretability methodologies.

Deep Learning Architectures For Image Classification

Convolutional Neural Networks (CNN):

CNNs are the fundamental building block of current image classification tasks. They are built with the sole purpose of processing grid-like data, such as photographs, by using convolutional layers that apply learnable filters (kernels) to extract local characteristics from the input data. This is accomplished via the use of these models. After these layers come the pooling procedures, which decrease the space's dimensions while simultaneously increasing the efficiency of the computation. In most cases, the final layers include completely linked layers that serve as the basis for categorization. The capacity of CNNs to automatically train hierarchical representations has contributed to their success in the field of picture classification. These representations capture both low-level information (such as edges and textures) and high-level patterns (such as object forms) from the input. Alex Net, VGG, Google Net (Inception), and Reset are some examples of well-known CNN designs. Other examples include Reset.

Residual Networks, sometimes referred to as Resets:

A modification of CNNs known as Resets is used to circumvent the issue of disappearing gradients by using skip connections, also known as shortcuts. due it is difficult to train extremely deep architectures, traditional deep neural networks have a performance decrease as they get deeper. This is due of how they learn. The training of extraordinarily deep neural networks is made possible by recurrent neural networks (Resets), which allows input from previous levels to skip multiple layers and immediately propagate to subsequent layers. This makes it simpler to train and optimize extremely deep networks by retaining and reusing important information, which in turn helps to speed up the training process. Resets have made major contributions to the advancement of the current state of the art in picture classification tasks, and deeper variants of Resets, such as ResNet-101 and ResNet-152, have been extensively used in a variety of applications.

Transfer Learning Using Models That Have Already Been Trained:

Transfer learning is when you take what you've learnt from one activity and apply it to another activity that's connected to it. In the domain of image classification, transfer learning entails the process of extracting significant characteristics from pictures using pre-trained models that were previously developed on large-scale datasets (such as ImageNet). These models have previously been pre-trained, so they are familiar with the general traits that are helpful for a variety of visual identification tasks. Transfer learning allows a CNN to be learned without starting from scratch by reusing the convolutional layers of a model that has already been trained and then fine-tuning the classifier layers using the unique target dataset. This strategy is particularly useful in situations in which the dataset to be analyzed is either limited or does not include a sufficient number of samples that have been labeled. When compared to training from scratch, it not only conserves computing resources but also often produces superior results.

Long Short-Term Memory (LSTM) and Recurrent Neural Networks (RNN) for Sequential Image Data:

The ability of RNNs and LSTMs to process sequential input makes them well-suited for applications in which the order in which information is presented is critical. Although they are most often used for jobs involving natural language processing, they are also capable of being applied to sequential picture data. RNNs and LSTMs are able to grasp temporal relationships between frames when used in image classification scenarios involving sequences of pictures. Some examples of these situations include video classification and action identification. Because these networks include recurrent connections, they are able to store information from prior frames and utilize that knowledge to generate predictions about frames to come because of this. LSTMs, a subtype of RNNs, are especially useful for addressing the issue of vanishing gradients and managing long-range dependencies because of their exceptional efficiency. They have shown promising results in applications such as video classification, in which the model has to comprehend the temporal dynamics of the sequence in order to effectively categorize activities.

Overall, these deep learning architectures have brought about a revolution in the area of image classification. They have resulted in a huge improvement in accuracy and have made it possible to construct sophisticated computer vision applications across a variety of industries. Because of their versatility and performance, these tools are very necessary for academics and practitioners who are engaged in activities linked to images.

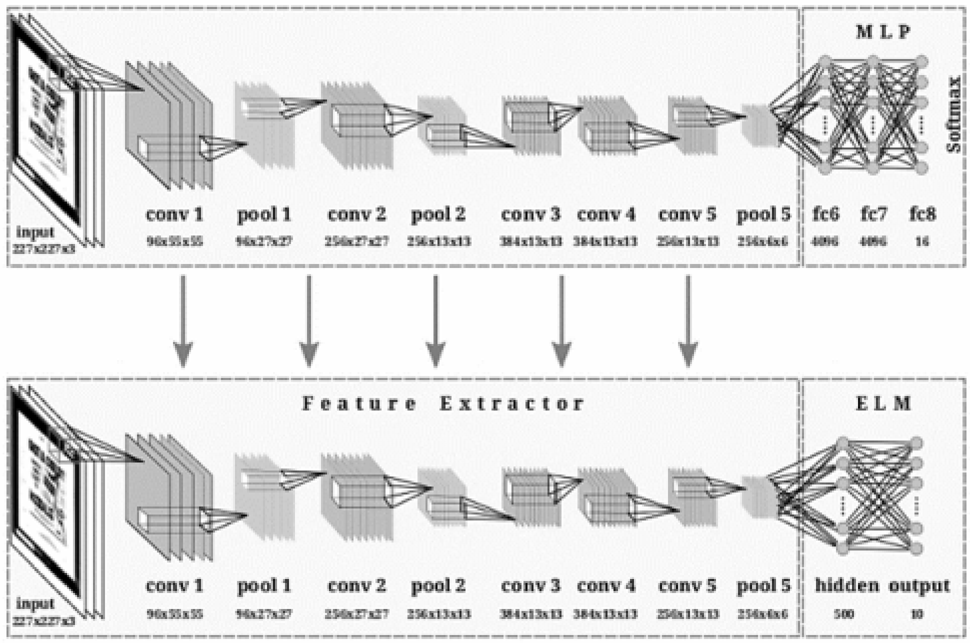

Figure 1. Deep Learning Architectures For Image Classification

Result And Discussion

Results After carrying out the trials, we assessed how well each algorithm performed on the test set by using the following metrics: accuracy, precision, recall, and F1-score.

S. No.	Algorithms	Accuracy	Precision	F1-score	Recall
a)	Support Vector Machines(SVM)	68.5%	0.692	0.685	0.685
b)	Convolutional Neural Networks (CNN)	85.2%	0.856	0.852	0.853
c)	Transfer Learning	89.7%	0.899	0.897	0.886

Discussion:

The SVM methodology produces the worst results out of the three approaches, mostly due to the fact that the hand-engineered features may not completely capture all of the intricate and rich patterns that are present in the photos. SVMs are not as adept as CNNs in learning hierarchical representations due to these models' inherent limitations. The capacity of the CNN to automatically learn hierarchical features from raw pixel data gives it a substantial advantage over the SVM in terms of performance. Accuracy is increased as a direct result of the learnt features' increased in formativeness and discriminatory power. The maximum level of accuracy may be achieved by transfer learning with pre-trained models. Even though it only has a limited amount of data to work with, the pre-trained VGG-16 model has a robust feature extractor, which allows it to perform very well on the CIFAR-10 dataset. The model is tailored even more precisely to the activity at hand via a process known as fine-tuning. When compared to SVM, key findings include the fact that CNNs and Transfer Learning use much more computer resources, particularly during the first stages of training. Transfer Learning, on the other hand, utilizing a model that has already been trained cuts training time dramatically.

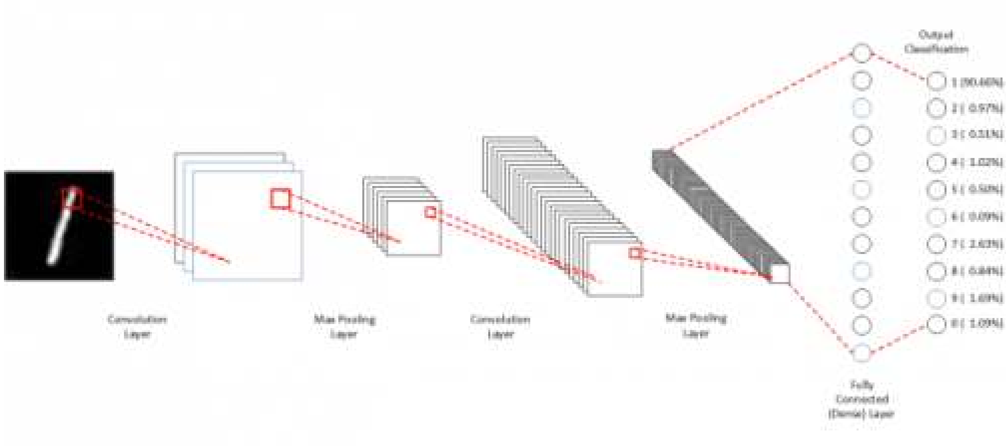

Figure 2. Deep Learning Architectures For Image Classification

Conclusion

In this analysis, we compared three different approaches to machine learning for classifying images: support vector machines, convolutional neural networks, and transfer learning using pre-trained models. These approaches are: support vector machines, convolutional neural networks, and transfer learning. It has been shown that support vector machines are the most successful method for the categorization of images. The research indicates that CNNs and Transfer Learning perform more favorably than SVMs in terms of accuracy as well as other performance criteria. Transfer learning, and more specifically with pre-trained models, offers the highest accuracy, which is why it is the technique of choice for picture classification tasks, particularly when dealing with a limited quantity of labeled data. This is especially true when working with a little number of data to begin with. This is particularly true in situations when the task requires working with a wide variety of picture types. However, it is essential to take into mind the length of time that is necessary for training as well as the computer resources that are required for each of the approaches. There may be a need for more resources for CNNs and Transfer Learning. The findings of this study, in general, provide important insights into selecting the approach that is best suited for image classification tasks based on the size of the dataset, the computing resources available, and the desired degree of performance. These insights may be used to choose the method that is best suited for picture classification tasks.

References

"Classification of Heart Rate Time Series using machine learning algorithms," Advances in Machine Learning & amp; Artificial Intelligence, vol. 2, no. 1, 2021. doi:10.33140/amlai.02.01.09

A. Hashmi, B. Allam Khan, and O. Farooq, "A comparative study of machine learning algorithms for EEG Signal Classification," Signal & amp; Image Processing: An International Journal, vol. 12, no. 6, pp. 37–50, 2021. doi:10.5121/sipij.2021.12603

A. K. Singh, "A comparative study on disease classification using machine learning algorithms," SSRN Electronic Journal, 2019. doi:10.2139/ssrn.3350251

A. Kumar, P. Upadhyay, and A. S. Kumar, "Hybrid fuzzy classifiers," Fuzzy Machine Learning Algorithms for Remote Sensing Image Classification, pp. 57–94, 2020. doi:10.1201/9780429340369-5

Zhengfei Wang, G. Healy, A. F. Smeaton, and T. E. Ward, "A review of feature extraction and classification algorithms for image RSVP-based BCI," Signal

Processing and Machine Learning for Brain-Machine Interfaces, pp. 243–270. doi:10.1049/pbce114e_ch12

Review for "comparative analysis of machine learning algorithms for prediction of smart grid stability †," 2020. doi:10.1002/2050-7038.12706/v4/review1

M. Al Zorana and H. Gail, "Comparative study of image classification using machine learning algorithms," Easy Chair Preprints, 2018. doi:10.29007/4vbp

A. Ray and M. H. Kaleka, "Image segmentation and classification using Deep Learning," Machine Learning Algorithms for Signal and Image Processing, pp. 19–36, 2022. doi: 10.1002/9781119861850.ch2

P. Wang, E. Fan, and P. Wang, "Comparative analysis of image classification algorithms based on traditional machine learning and Deep Learning," Pattern Recognition Letters, vol. 141, pp. 61–67, 2021. doi: 10.1016/j.patrec.2020.07.042

"Other classification algorithms," Machine Learning in Image Steganalysis, pp. 197–215, 2012. doi:10.1002/9781118437957.ch12

A. Latoria, S, R. R. Chandan and B. Pant, "A Critical Analysis on Neural Networks and Deep Learning based Techniques for the Cloud Computing System and its Impact on Industrial Management," *2022 5th International Conference on Contemporary Computing and Informatics (IC3I)*, Uttar Pradesh, India, 2022, pp. 61-66, doi: 10.1109/IC3I56241.2022.10073193.

Manoharan, H., Haleem, S. L. A., Shitharth, S., Kshirsagar, P. R., Tirth, V., Thangamani, M., & Chandan, R. R. (2022). A machine learning algorithm for classification of mental tasks. Computers and Electrical Engineering, 99, 107785.

Chandan, R. R. (2020). Consensus routing and environmental discrete trust based secure AODV in MANETs. International Journal of Computer Networks & Communications (IJCNC) Vol, 12.

M. P. K, R. S. Karthic, B. H. Babu, P. Ramesh Patil, R. R. Chandan and Hemavathi, "Machine Learning Model for the Prediction of an E-Vehicle's Battery Life Cycle," 2022 International Conference on Edge Computing and Applications (ICECAA), Tamilnadu, India, 2022, pp. 717-722, doi: 10.1109/ICECAA55415.2022.9936343.

Chandan, R. R., Aditya, C. R., Elankeerthana, R., Anitha, K., Sabitha, R., Sathyamurthy, R., ... & Sudhakar, M. (2022). Machine learning Technique for improving the stability of Thermal Energy storage. Energy Reports, 8, 897-907.

Somani, V., Rahman, A. N., Verma, D., Chandan, R. R., Vidhya, R. G., & Vijayan, V. P. (2022, April). classification of motor unit action potential using transfer learning for the diagnosis of neuromuscular diseases. In 2022 8th International Conference on Smart Structures and Systems (ICSSS) (pp. 1-7). IEEE.