Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 5
AIDML 2024

A Research on Image Recognition and Classification Based on Traditional Machine Learning and Deep Learning

Zelong Huang
Faculty of Engineer University of New South Wales Sydney, Australia
z5489331@ad.unsw.edu.au

Abstract. With the rapid development of smart devices, a huge amount of images are appearing in people's lives, so the processing of images, especially recognition and classification processing, becomes more and more important. This paper aims to provide researchers with an overview of current traditional machine learning (ML) and deep learning (DL) in solving image recognition and classification problems, specifically analyzing architectures such as KNN, SVM, Random forest algorithms, and CNN Architecture, and finding available data training results for comparative analysis. In summary, it is found that deep learning algorithms perform well in handling large-scale complex datasets and automatic feature extraction, and are applied to most image recognition and classification problems, but the traditional machines have obvious advantages of accessibility and speed in handling small-scale data, and deep learning algorithms are still not an effective choice in some special cases or resource-limited environments. At the end of the article, the full text is summarized and outlook.

Keywords: image classification; image recognition; DL, ML.

1. Introduction

In this era characterized by an explosion of information, pictures are becoming a necessary part of everyday life. The proliferation of smartphones, digital cameras, and surveillance cameras has resulted in a growing number of photographs being taken each year. The advent of image classification and recognition technology plays a pivotal role, such as in precision agriculture based on image recognition and remote diagnosis for Medical IoT based on image classification and recognition [1, 2]. The application of image classification and recognition spans diverse fields, seemingly liberating individuals from arduous, mechanical tasks and significantly enhancing work efficiency, thereby affording people more time to allocate to other pursuits [3].

For the classification and recognition of images, it is crucial to accurately analyze and extract pertinent information from the image within a constrained timeframe, with the efficacy of various algorithms being pivotal in addressing this class of problems. Common techniques used to address picture classification and recognition problems include deep learning and classical machine learning. The classical machine learning algorithm framework has matured significantly, with most models being adaptable to image classification and recognition predicaments, thus facilitating the utilization of traditional machine learning methods to address prevailing challenges and computational capabilities [3]. These methods typically rely on manually extracted features and intricate pre-processing steps to attain commendable performance. While these techniques excel in specific tasks, they are often constrained by the efficacy of feature extraction and generalization capabilities. Deep learning has been a well-known trend in recent years, demonstrating a clear advantage over alternative algorithms in fields including natural language processing, computer vision, and image processing . Deep learning finds broader application in the sphere of image recognition and classification, particularly through convolutional neural networks (CNNs), where deep learning algorithms enhance the completeness of image information by autonomously learning layered feature representations of the data, thereby rendering them more precise when addressing image recognition challenges. Nevertheless, deep learning is not devoid of challenges, as its robust learning capabilities in certain training instances can give rise to overfitting issues, thereby leading to a deterioration in practical results [3]. Hence, by scrutinizing and contrasting the performance, advantages, disadvantages, and application scenarios of traditional ML and DL methods in addressing image recognition and classification challenges, researchers and developers can discern the most suitable approach based on the specific problem at hand. This holds significant implications for bolstering the accuracy and efficiency of image recognition and classification endeavors.

This paper will delineate the process of image recognition and classification based on traditional machine learning in Section II, elucidating the functioning of various classifier algorithms, while Section III will delve into the utilization of deep learning for image recognition and classification, with a focus on the application of CNNs to tackle such conundrums. Section IV will furnish a comparative analysis and comprehensive discussion through existing data and case studies exemplifying the application of these two methodologies in addressing real-world challenges. The final section will encapsulate the key findings and expound upon the future trajectories and trends pertaining to classical machine learning and deep learning.

2. Image Recognition and Classification Based on Traditional Machine Learning

The process of developing classical machine learning for image identification and classification has involved several stages, from the earliest mainly based on simple pattern matching as well as geometric feature extraction to the emergence of feature engineering, and then later a series of statistical methods, which has made the field of image processing from the early gaps to a large extent. The process of traditional machine learning for image recognition and classification at this stage is roughly divided into the following steps, as shown in Figure 1, the obtained images are first preprocessed, followed by feature extraction, dimensionality reduction operations, classification of the images using classification algorithms, and finally applied to practical tasks.

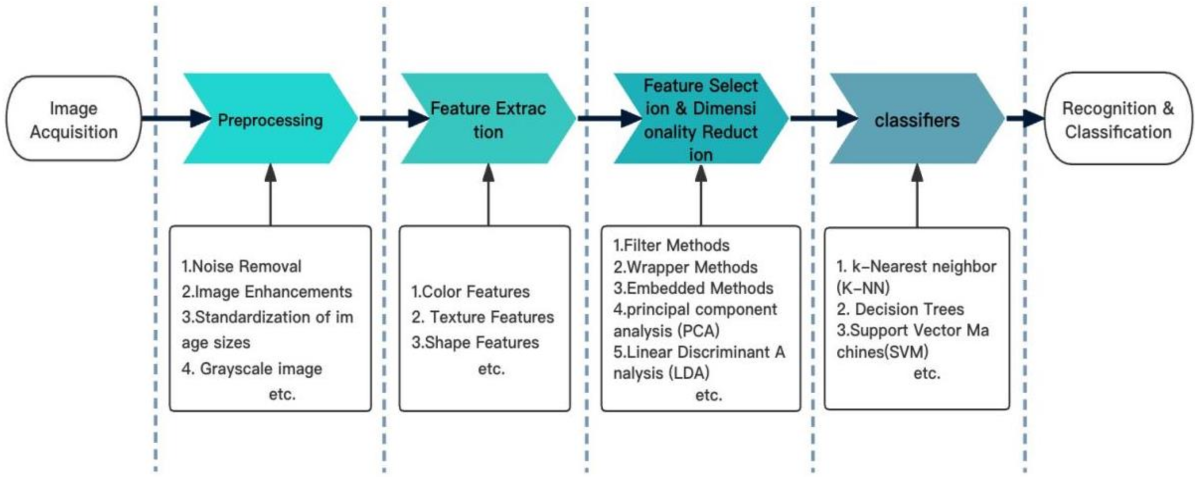

Figure 1. Classification Process for Image Recognition with Traditional Machine Learning(Photo/Picture credit :Original)

2.1. Image Preprocessing

In image recognition and classification based on traditional machine learning, image preprocessing is very critical. Picture preprocessing, which is comparable to a dataset's mathematical normalization, can significantly impact the feature extraction's accuracy and the outcomes of picture analysis. The core of image preprocessing is the use of several techniques to raise the original image's quality and enhance the representation of key information while removing extraneous noises and interferences to ensure that the data being fed into the model is clean and normalized, a typical image preprocessing flow, as illustrated in Figure 2, starts with image reading. The next step involves converting the image to a grayscale map to eliminate color information and preserve only the brightness. Next, denoising is performed to remove random noise and then the image is normalized to adjust the brightness and contrast. Finally, the image contrast is further optimized by histogram equalization to make the brightness distribution more uniform. Grayscaling is the process of converting captured color images into grayscale images, which leads to a substantial decrease in both data volume and processing complexity, and grayscaling helps to highlight the structural and shape information of the image for subsequent operations; Common image denoising methods are Mean Filter, Median Filter, Gaussian Filter, Wavelet Transform; Normalization is a technique used to increase the stability and effectiveness of model training that modifies the range of picture pixel values; Histogram equalization enhances the contrast of the entire image by adjusting the histogram distribution of the image to make it more uniform [4].

Figure 2. Image Preprocessing Process (Photo/Picture credit :Original)

2.2. Feature Extraction

Feature extraction is a crucial stage in picture identification and classification in classical machine learning, which refers to the extraction of representative features from an image, which is transformed into numerical patterns for subsequent training of classification models. There are various ways of feature extraction for different types of images.

Image feature extraction is mainly based on the following features, Color Features: defined by a specific color space or model, common methods include histogram, CM, CCV, etc.; Texture features: a set of measurements used to characterize the internal surface structure and appearance properties of an image, common methods based on texture feature extraction are the Gabor Filter, etc.; Shape features: Based on the description of the shape of the object in the picture to extract relative features, two important methods: contour-based and region-based[5]. In addition to this, there are methods , which extracts the main features of an image by dimensionality reduction, and Independent Component Analysis (ICA), a method that decomposes an image into separate parts and extracts features independently.

2.3. Traditional Machine Learning Image Classification Algorithms

Different algorithms play an important role in the processing of image classification. Among the primary algorithms are AdaBoost, KNN, Bayesian, decision trees, random forests, logistic regression, and so on. Each algorithm has specific application scenarios and advantages and disadvantages, and several classification algorithms are described below.

KNN:K-nearest neighbor classification is performed by assigning unlabeled observations to the class of the same labeled instances[6]. So based on this algorithm, it is very critical to define the distance, common ways to define distance are the Minkowski distance, Euclidean distance, Manhattan distance, etc. The outcomes of this algorithm can be greatly impacted by the selection of the k-value; smaller values of k may be more susceptible to noisy data and therefore the model may be too complex, leading to overfitting, while larger values of k may lead to a model that is too easy and does not adequately capture the complexity of the data, which can trigger underfitting. For this reason, some writers advise taking the square root of the training dataset's observation count and setting it to k[6].

SVM: Creating higher-dimensional hyperplanes and converting the initial training data into a multidimensional space allows SVM to categorize pictures[7]. In order to discover boundaries that maximize the borders between various classes of data points, the SVM algorithm will search for a hyperplane, or many hyperplanes, that can best discriminate between distinct classes of data points, SVM uses the kernel function to deal with nonlinearly divisible data, the kernel function is a reliable mathematical method for nonlinear mapping of high dimensional data [7], which is especially useful when dealing with image data, which tends to be highly nonlinear, and this is one of the important reasons why SVM algorithms have a good performance.

Random Forest: Random forest is an integrated learning method based on decision trees, by constructing a collection of decision trees and voting for the most likely to be in the category, this method usually generates random vectors to control the growth of each decision tree, which also has a significant improvement in the performance of the final result. Because only a random subset of the characteristics is taken into consideration each time during training and each tree is randomly picked from the original training set when finding the best segmentation features, this improves the randomness and at the same time makes the model better able to handle high dimensional data like images.

Each of the above three image categorization methods has its own characteristics and Table I shows the advantages and disadvantages of each of these three methods.

Table 1. Comparison of the advantages and disadvantages of three common algorithms

	Advantage	**Disadvantage**
KNN algorithm	KNN does not require training and is easy to implement in smaller datasets.	KNN performs poorly on larger datasets or high-dimensional data, and is also more sensitive to noisy data
SVM algorithm	SVM, by searching for the optimal segmentation hyperplane and using the kernel technique to deal with nonlinear problems, demon-strates a better generalization ability and applicability to relat-ively large-scale datasets	The performance of SVM relies heavily on the selection of the correct kernel function. Choosing an inappropriate kernel function may lead to overfitting or underfitting.
Random Forest algorithm	Random forest as an integrated learning approach may be able to handle large-scale and high-dimensional data well.	Random Forest is robust to missing data and unbalanced data, but its modeling results may not be easy to interpret, and it is prone to overfitting in certain noisy situations.

3. Image Recognition and Classification Based on Deep Learning

The main feature of deep learning is the ability to automatically learn complex representations and features from large amounts of data that are difficult to define manually or through traditional programming methods. Deep learning has been rapidly advancing in the field of image classification and recognition in recent years. This involves building a multilayer network to extract higher-dimensional and abstract information, from which the computer automatically learns and acquires the implicit relationships within the data, resulting in more expressive learned features. Among different network structures, CNNs are extensively utilized in image recognition and classification tasks. The upcoming paper will focus on demonstrating how CNNs are applied in the process of image recognition and classification [8].

3.1. Convolution Neural Network

CNN is a key advancement in deep learning techniques suitable for dealing with the problem of image recognition and classification. CNN architectures autonomously extract features from images from simple to complex by constructing multiple layers such as Convolutional Layer, Pooling Layer, and Fully-Connected Layer, and avoids a lot of feature extraction work in image preprocessing.

CNN architectures usually contain a variety of different layers, each with a specific role, that together form the power of a CNN to process image data, the figure 3 shows the network layer structure of a CNN as an example of recognizing handwritten numbers.

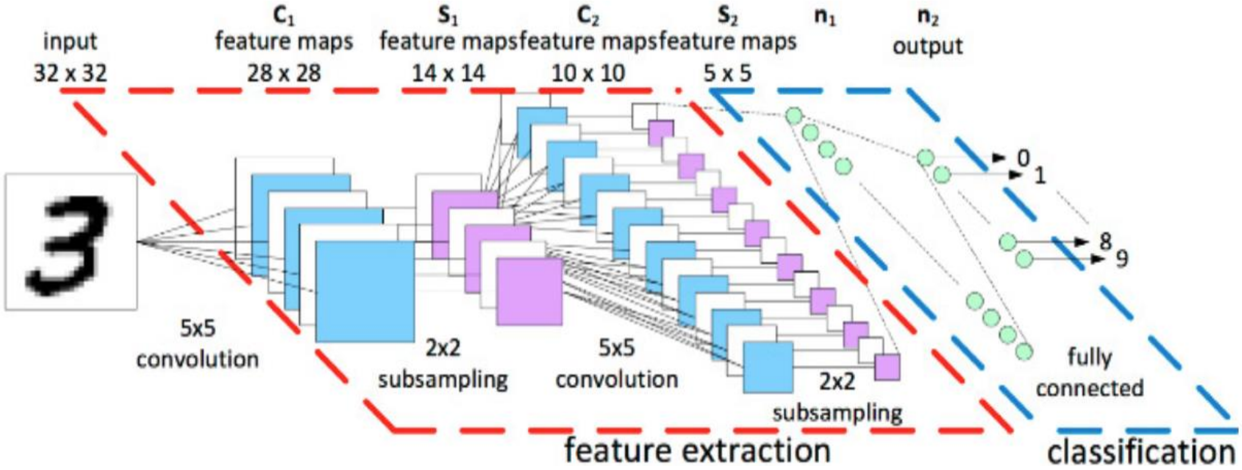

Figure 3 Internal Layers of CNNs [9]

Convolutional Layer: The convolutional layer acquires features in the image by using filters, which are gradually combined into more advanced features for subsequent operations as the layers of the network deepen.

Pooling Layer (subsampling): Pooling layers are used to minimize an image's features while keeping the layer's feature invariance and lowering the number of parameters. Maximum pooling and average pooling are two common pooling procedures [9].

Fully-Connected Layer: The fully connected layer, which can synthesize all the characteristics acquired from the preceding layers for image classification tasks, is often found at the very end of the network structure.

3.2. CNN Network Architecture

There are a variety of CNN architectures, each of these architectures has its own characteristics and performs well when used to solve certain problems. Common architectures include LeNet-5, an early CNN architecture, mainly used for handwritten digit recognition problems; AlexNet convolutional neural network architecture, which incorporates five convolutional layers, three fully-connected layers, and techniques such as the application of Dropout to reduce model overfitting, significantly reduces the error rate for image classification tasks; VGG convolutional neural network architecture, which is characterized by the use of a uniform convolutional layer design and maximum pooling between convolutional layers to reduce the size of the feature maps, allows VGG to use a relatively simple structure while also obtaining a large number of features. With the continuous development of deep learning in recent years, some new CNN architectures have emerged that can be used to solve image classification problems, EfficientNet proposes a composite scaling strategy to simultaneously scale width, depth, and image resolution, an approach that demonstrates high efficiency and accuracy as well as having excellent performance; RegNet is able to efficiently find the optimal network configuration given resource constraints by predicting the number of covariates and performance of the model, proving its effectiveness in image classification tasks on many datasets.

These architectures show how convolutional neural networks have advanced in handling image classification tasks and highlight the significance of broad advancements through novel model structures, optimization of training methods, and increased model efficiency and accuracy. As deep learning continues to evolve, we can expect more innovative network architectures to emerge in the future.

4. Result

ML and DL are the two primary technological approaches for resolving picture classification problems. and recognition problems and these two types of methods have obvious differences in terms of accuracy application scenarios, etc. The following is their comparative data on solving several different image classification and recognition problems.

1) Contrasting deep learning with traditional machine learning methods for classifying human activities: The implementation recognizes six basic human activities and evaluates the accuracy of fourteen classifier models, as shown in Figure 4.

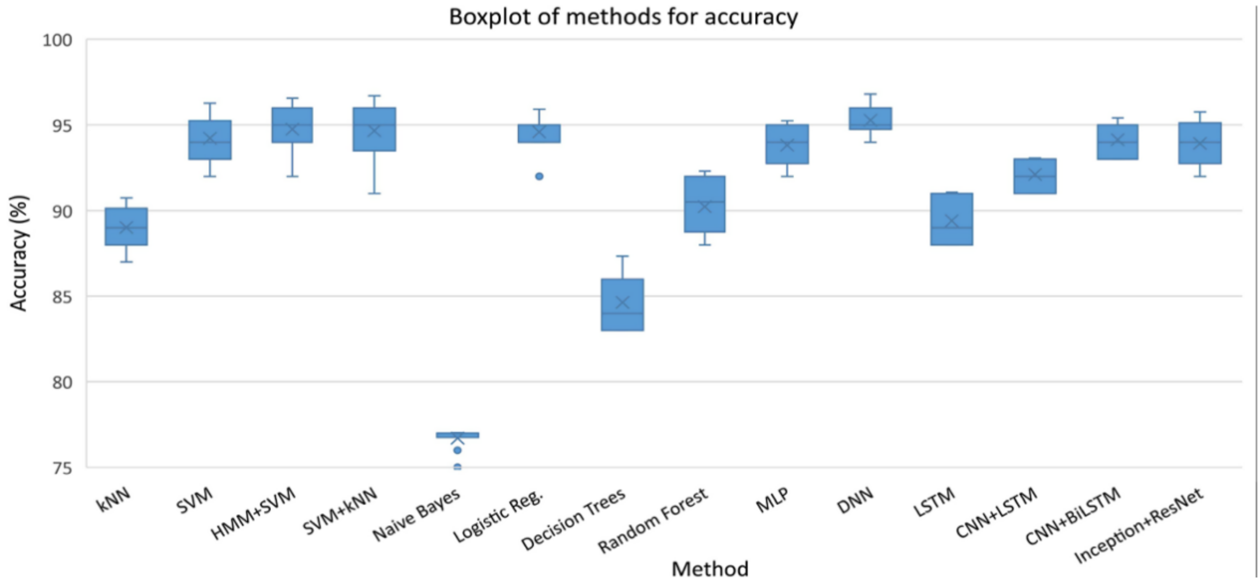

Figure 4. Boxplot of methods for accuracy

2) Comparing DL and ML Methods for Handwritten Number Recognition: The normal MNIST database. We have picked part of the studied data from ML and DL algorithms in Table II, the test set contains samples from an entirely different author, and the handwritten data samples are from about 250 distinct writers

Table 2. Comparing the classifiers different methods [11]

Classifier	**Sensitivity(%)**	**Precision(%)**	**Specificity(%)**
KNN	97.18	97.92	98.58
SVM	98.16	98.64	98.96
Decision tree	78.24	79.36	80.42
Random forest	81.72	82.88	83.64
Naïve Bayes	85.74	86.36	86.80
CNN	98.48	98.86	99.14

3) Comparison of NL and DL based automated detection of agricultural diseases: The dataset used and the performance of several classifiers on the potato disease classification problem are mainly shown in Table III.

Table 3. Potato disease classification performance [12]

Year	Dataset	Classifier	Performance Accuracy
2019	Three hundred photographs of potato leaves taken from the PlantVillang Dataset	Maximum-Minimum Color Difference and Euclidian Distance	91.67%
2017	892 photos of potato leaves, some from the PlantVillang Dataset and some from the Department of Pathology	SVM	84%
2017	892 photos of potato leaves, some from the PlantVillang Dataset and some from the Department of Pathology	Random forest	79%
2019	4004 photos taken from the PlantVillage Dataset	CNN	90.85%
		CNN based AlexNet Model	98.33%
2019	2250 potato leaf photos taken from the PlantVillage Dataset	DCNN	98.33%

Through a comparative analysis of the ML and DL algorithms' respective performances on various issues, we can find that for dealing with image classification tasks with small datasets and relatively simple problems traditional machine learning is still effective and can achieve a high accuracy rate through the design of the algorithms, and at the same time in traditional machine learning, the SVM algorithm performs better in most problems. As for deep learning, it outperforms traditional machine learning in most problems, especially on large-scale image datasets or when dealing with complex image classification and recognition tasks, the performance of deep learning models usually improves as the amount of data increases, while the performance of traditional machine learning models has a limited improvement.

5. Conclusion

In order to solve image recognition and classification problems, this paper first introduces both traditional machine learning and deep learning. It then lists a few popular machine learning algorithms, like KNN, SVM, and Random Forest, focuses on CNN in deep learning and its common network architectures, and concludes a comparative analysis of the available experimental data. The models of traditional machine learning are relatively simple, easy to understand and implement, and usually perform well on small-scale data. Deep learning, especially CNN, demonstrates excellent capabilities in dealing with image classification problems, especially when dealing with large and complex datasets, but at the same time requires high performance and time. With the enhancement of hardware capabilities, deep learning techniques have become the core of the image recognition field. Nevertheless, traditional machine learning algorithms still show their irreplaceable utility in environments with specific needs or limited computational resources. Future developments in the domain of picture recognition and classification might lead to the creation of straightforward, very accurate, effective, and safe algorithms.

Reference

[1] H. S.Abdullahi, R.Sheriff, & F.Mahieddine, Convolution neural network in precision agriculture for plant image recognition and classification. INTECH, Vol. 10, pp:256-272, 2017.

[2] Z. Guo, Z., Y. Shen, Wan, W. Shang, K. Yu, Hybrid Intelligence-Driven Medical Image Recognition for Remote Patient Diagnosis in Internet of Medical Things, in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 12, pp: 5817-5828, 2022.

[3] Y. Lai. A Comparison of Traditional Machine Learning and Deep Learning in Image Recognition, 3rd International Conference on Electrical, Mechanical and Computer Engineering, vol. 1314, 2019.

[4] J. Stark Adaptive image contrast enhancement using generalizations of h4stogram equalization. IEEE TIP, 9(5), pp:889-896, 2000.

[5] P. Tian, D. A review on image feature extraction and representation techniques. International Journal of Multimedia and Ubiquitous Engineering, 8(4), pp:385-396, 2013.

[6] Z. Zhang, Introduction to machine learning: k-nearest neighbors. Annals of translational medicine, 4(11), pp: 218, 2016.

[7] M.A. Chandra, S. Bedi, Survey on SVM and their application in image classification. International Journal of Information Technology, 13(5), pp:1-11, 2021.

[8] Y. Dong, G. Liang, Research and discussion on image recognition and classification algorithm based on deep learning. In 2019 MLBDBI, pp. 274-278.

[9] N. Sharma, V. Jain, A. Mishra, An analysis of convolutional neural networks for image classification. Procedia computer science, Vol.132, pp:377-384, 2018.

[10] F. Bozkurt, A comparative study on classifying human activities using classical machine and deep learning methods. Arabian Journal for Science and Engineering, 47(2), pp:1507-1521,2020.

[11] A. Sharma, H. Bhardwaj, A. Bhardwaj, D. Sakalle, D. Acharya, A machine learning and deep learning approach for recognizing handwritten digits. Computational Intelligence and Neuroscience, 2022.

[12] J. Wani, S. Sharma, M. Muzamil, Machine learning and deep learning based computational techniques in automatic agricultural diseases detection: Methodologies, applications, and challenges. Archives of Computational methods in Engineering, 29(1), pp:641-677, 2022.