e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 6, Issue 5, May 2023

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.54

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54|

| Volume 6, Issue 5, May 2023 |

A Study on Image Categorization Techniques

Shailja Sharma[1], Princy[1], Kirti Bhatia[1], Rohini Sharma[2]

P.G. Student, Department of CSE, Sat Kabir Institute of Technology and Management, Haryana, India[1]

Assistant Professor, Department of CSE, Sat Kabir Institute of Technology and Management, Haryana, India[1]

Assistant Professor, Department of CSE, Sat Kabir Institute of Technology and Management, Haryana, India[1]

Assistant Professor, Department of CS, GPGCW, Rohtak, Haryana, India[2]

ABSTRACT:This research does a literature review on the various methods for object-based picture classification. The most important and difficult task in the realm of computer vision is classification. The description, texture, or likeness of objects or entities serve as the basis for classification. The labelling of photographs into one of several specified categories is known as image classification. An image is represented by units called pixels. Pixels in an image are categorised into various classes. Image acquisition, image pre-processing, and picture segmentation are all included in image categorization. For the classification of images, numerous classification approaches have been created. Many fields, including medical imaging, object recognition in satellite images, traffic control systems, brake light detection, machine vision, etc., use classification. Supervised image classification techniques and unsupervised image classification techniques make up the two primary types of image classification techniques.

KEYWORDS:Feature Extraction, Classification of Images, Neural network

.

I. INTRODUCTION

Fields including machine learning (ML), and its subdivisions, deep learning and neural networks (NN), have greatly advanced in recent years as a result of the growing turbulence, requirement, and applicability of artificial intelligence (AI). Software and tools like classification algorithms, which feed enormous amounts of data, analyse them, and retrieve valuable characteristics, are required for training. All of the pixels in a digital image are intended to be categorised into one of multiple groups using the classification process. The classification process is often carried out using multi-spectral data, and each pixel's spectral pattern is employed as the numerical basis for categorization. The goal of image classification is to recognise and describe the features that appear in an image with reference to of the real object that these characteristics actually indicate on the ground as a distinct grey level (or colour). The most crucial aspect of digital image analysis is probably picture classification. Since distinguishing between different objects is a difficult operation, image classification has been a crucial task in the area of computer vision. The tagging of images into one of a number of predetermined classifications is referred to as image categorization [1]. A given image can be categorised into potentially n different classes. It could be time-consuming to manually verify and categorise large numbers of photographs, thus it would be really helpful if we could automate the entire procedure utilising computer vision. Another excellent example of how picture categorization is used in practise is the development of autonomous vehicles. We need classification systems to achieve the highest level of accuracy because of the applications, that involve computerised image organisation, stock images and videos on websites, visual search for better product discoverability, massive pictorial databases, image and face differentiationon social media etc. Fig. 1 illustrates categorization of images using different popular procedures.

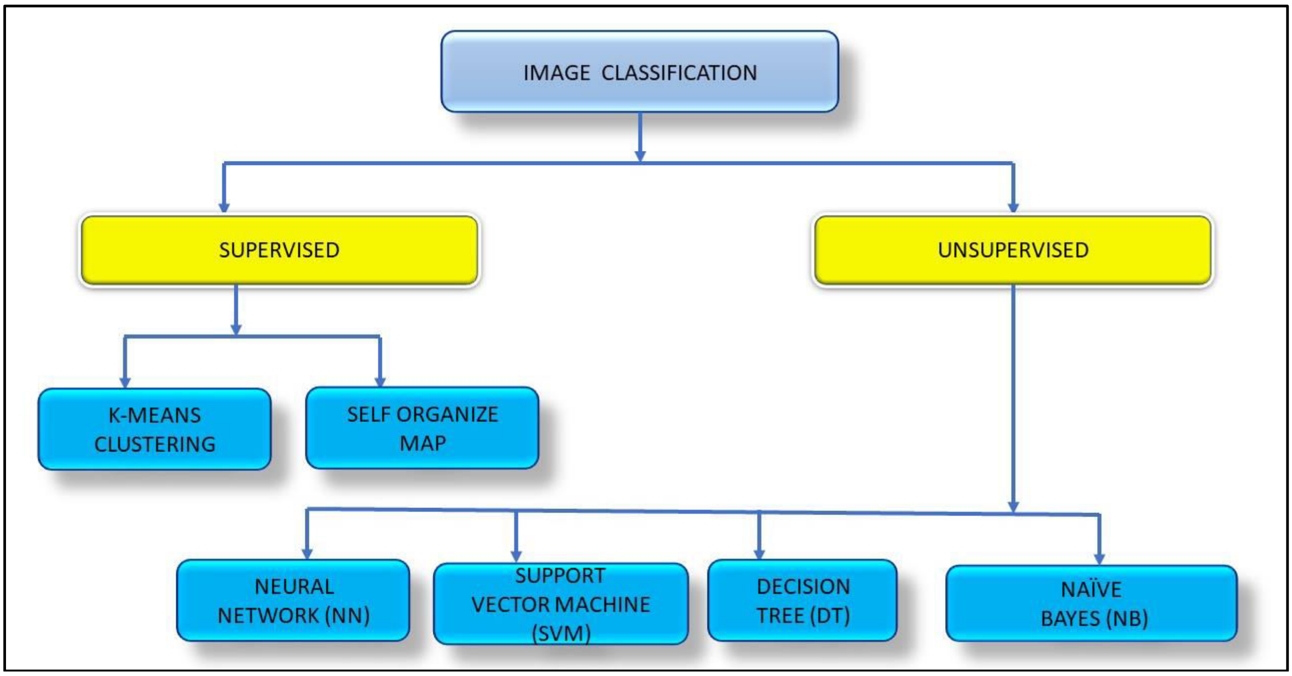

Fig1. Image categorization

II. RELATED WORK

Image feature extraction, the cornerstone of image categorization, has always been the subject of research. The conventional methods for extracting image features place more emphasis on manually defining individual picture attributes. Both the generalizability and portability of this approach are lacking. Therefore, researchers envision giving computers the ability to interpret images similarly to biological vision. A vast number of interconnected neurons make up an abstract BNN, a mathematical operation model. It mimics the comprehension of brain impulses by neural networks roughly. At first, MP neuron model was proposed by McCulloch and Pitts based on their analysis of biological brain networks and internal logic as a mathematical representation of neuron activity [2]. There are lots of applications in the field of image categorization [3-4]. The multilayer perceptron has been used by researchers to try and learn features, and the backpropagation (BP) algorithm was used to train the model [5]. The idea of building a computer neural network resembling a biological visual system was sparked by this discovery, and CNN was subsequently created. The CNN model's first batch, LeNet-5, was introduced by Lecun et al. [6]. Yet the LeNet-5's recognition results on complicated images were not perfect because to a lack of enough training data, as well as limitations imposed by the theoretical underpinnings and computational power [7].

The application of CNNs in remote sensing analysis of images finally took off after they were successfully used to large-scale visual classification tasks [8]. By utilising various methodologies for leveraging CNNs, a range of CNN-based scene categorization methods have been developed [9]. Ideally, there are three different types of CNN-based remote sensing picture scene categorization methods: (i) A feature generator is employed with pre-trained CNNs [10]. (ii) Use the dataset to refine the pre-trained CNNs [11]. (iii) Initialise CNN training weights universally [12].

III. SUPERVISED CATEGORIZATION

The concept behind supervised classification is that a user can choose an appropriate set of pixels from an image that best represents a given class, and then instruct the image processing software to utilise these training sites as references when classifying all other pixels in the picture. The user's knowledge is taken into consideration while choosing training sites, sometimes referred to as testing sets or input classes. The threshold for how similar other pixels must be to be grouped together is likewise chosen by the user. These limits are frequently established using the training area's spectral properties.A person can additionally choose how many classes an image will be divided into. After each information class has been statistically characterised, the image is next categorised by determining which of the signatures each pixel's reflectance most closely resembles. To create predictive theories, supervised categorization employs methods of regression and classification systems[13].

IV. UNSUPERVISED CATEGORIZATION

Despite the assistance of the user giving sample categories, software analyses an image and produces the results (groupings of pixels with similar features). The computer employs algorithms to identify related pixels and classify them. The user can choose the software's algorithm and the needed number of output sets, but this does not otherwise help with categorization.When the computer-generated clusters of pixels with similar properties have to be matched to actual features on the ground, the user must be familiar with the region being classed. Cluster assessment, identifying anomalies in neural networks, and methods for learning latent variable prototypes are some of the most frequently used strategies in unsupervised learning. Fig. 2 shows comparison between supervised and unsupervised process of image categorization

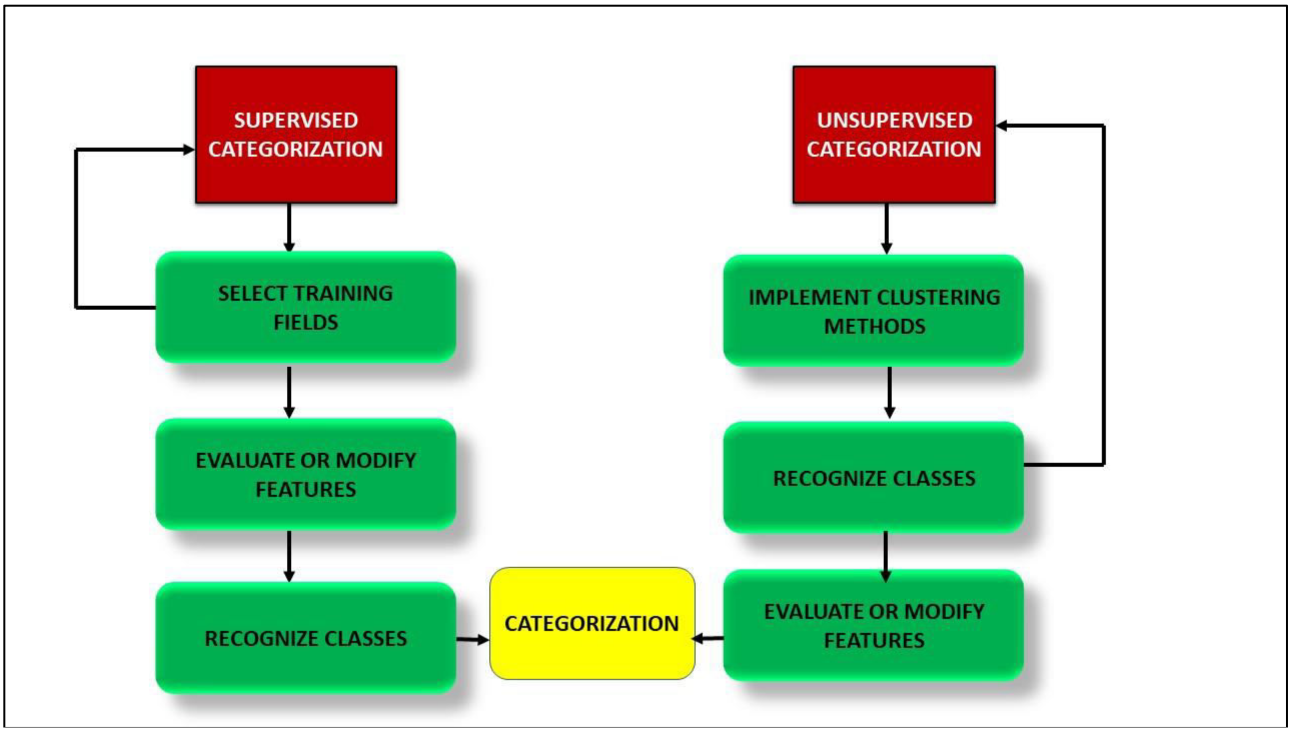

Fig 2.: Comparison of Supervised and Unsupervised Categorization

V. CONVOLUTIONAL NEURAL NETWORK (CNN)

A unique class of multi-layer neural network called a convolutional neural network (CNN, or ConvNet) is made to recognise visual patterns derived from pixel pictures with little to no pre-processing. It is a unique artificial neural network architecture. The CNN have produced advanced results in computer vision problems by utilising some of the visual cortex's properties. Convolutional layers and pooling layers are two incredibly fundamental foundation pieces that go into CNN. Despite their clarity, there are nearly infinite ways to combine these layers for a particular computer vision problem. Convolutional and pooling layers, among other components of a convolutional neural network, are easy to comprehend. The CNNs are difficult to employ in practise because it is difficult to create model structures that make the most use of these basic components. The CNNs are extremely well-liked due of their architecture, and the greatest part is that feature mining is not obligatory. The fundamental idea behind how the system learns to perform feature extraction is that its convolutions images and filters to yield invariant features that are passed on to the next layer. The characteristics in the following layer are twisted with various filters to yield more invariant and abstract attributes, and the procedure is repeated until it obtains the ultimate, consistent characteristic or outcome.

VI. ARTIFICIAL NEURAL NETWORK (ANN)

The ANNs are statistical learning algorithms that are employed for an assortment of tasks, from comparatively simple classification tasks to computer vision and speech recognition. They have been influenced by the characteristics of biological neural networks. In ANN, nodes are a set of connected processing components that act similarly to biological neurons. By gradually changing the weights—numerical values that represent the connections between various nodes—the network is finally able to approach the function that is needed. As the data spreads throughout the network, the hidden layers act as independent detectors of features, spotting increasingly intricate patterns in the data. If the network is tasked with recognising faces, for instance, the first hidden layer may serve as a line detector, the second hidden layer may take these lines as input and combine them to form a nose, the third hidden layer may take the nose and align with it with an eye, and so on, as long as the entire face is developed. The network can eventually recognise very complicated things thanks to this structure.

VII. SUPPORT VECTOR MACHINE (SVM)

The SVMs are potent yet adaptable supervised ML algorithms employed for both regression and classification. Compared to other ML methods, SVM are implemented in a different way. They are very well-liked because they can manage numerous continuous and categorical variables. A SVMsysteml is just a hyperplane in multidimensional space that signifies several classes. The SVM will iteratively create the hyperplane in an effort to reduce error. The datasets must be divided into classes in order to identify the maximum marginal hyperplane. The hyperplane that has the greatest distance to the closest training data point of any class separates the two classes well. This is done by building a hyperplane or group of hyperplanes in a high-dimensional space. The kernel function being employed determines the algorithm's true power.

VIII. K-NEAREST NEIGHBOR (K-NN)

A non-parametric methodology for classification and regression is K-NN. The k closest training instances in the feature space make up the input in both scenarios. The algorithm is by far the simplest. It is a non-parametric, lazy method of learning in which any computation is postponed until after the function has been evaluated and only local approximations of the function are made. This technique categorises unidentified data points by identifying the most prevalent class among the k-closest samples, which relies only on the distance across feature vectors.We need to construct a distance metric or similarity function in order to use the k-NN categorization, and two popular options are the Manhattan distance and the Euclidean distance. A class membership is the result. An object is allocated to the class that has the most support from its k closest neighbours (k is an integer with a positive value that is often small) based on a majority vote of those neighbours. The object is merely placed into the class of its one nearest neighbour if k = 1.

IX. NAÏVE BAYES ALGORITHM

Naive Bayes algorithms are a subset of categorization techniques based on the Bayes hypothesis. It is an assortment of processes rather than one process, and they are all based on the idea that each combination of attributes that are categorised has no influence of the other. The NB is a straightforward method for building classifiers. These models assign class labels to problem cases, which are expressed as vectors of attribute values, and the class labels are chosen from a finite set. Given the class variable, all NB algorithms make the assumption that the value of a specific feature is unrelated to the value of any other feature. The NB technique can be utilised for binary and multi-class categorization. It is a quick, highly scalable approach. It depends on performing numerous counts. For text categorization, spam email categorization, etc., it is a well-liked option. On a modest dataset, it is simple to train. It is limited in that it assumes that all features are independent of one another, making it unable to understand how features are related. Naive Bayes can understand the relative relevance of individual features but cannot understand the connections between features.

X. RANDOM FOREST (RF) ALGORITHM

For both categorization and regression, the Random Forest supervised learning method is used. Because a forest is made up of trees, and more trees equal a more resilient forest, the random forest algorithm builds decision trees on samples of data, gets predictions from each one, and then votes to determine which is the best answer. Because it averages the results, the combination method—which is superior to a single decision tree—reduces over-fitting. A system of categorization made up of several decision trees is called the RF. It tries to generate an uncorrelated forest of trees whose estimate by group is more precise by bagging and featuring randomization when developing each individual tree.

XI. MULTI- LAYERED PERCEPTRON (MLP)

The human neurological system, which uses unified neurons to process information in succession, served as the model for multi-layered perceptrons (MLP).Considering that MLP is a FFNN, it means that information spreads from input to output. The principles of each feature are supplied into the inputs, and the outputs provide the class value. The output is a weighted linear blending of the inputs through one layer of neurons. The name of this network is "linear perceptron." A nonlinear mapping between the input and output is possible by adding an additional layer of neurons with nonlinear fundamental functions (the hidden layer). By reducing the mean squared error rate of organisation, the weights pertaining to the neurons are iteratively optimised during the teaching process.The learning rate needs to be chosen as a compromise between making mistakes on the practise set and overtraining. The learning rate regulates the weight modifications during the teaching stage. The quantity of units in the concealed layer is another crucial restriction. There is no doubt that the MLP is susceptible to over fitting and calls for the best selection of regularisation parameters. By creating unlimited decision boundaries, the MLP may create models with any level of difficulty. Additionally, it is resilient to loud characteristics because they will lose weight with time.

XII. KERNEL SUPPORT VECTOR MACHINES

Kernel SVMs use the kernel function with the width of the Gaussian to implicitly translate input feature vectors to a higher dimensional space. A maximal extrication hyper plane is constructed in the altered space while taking a two-class problem into consideration. On either side of the hyper plane that divides the data, two parallel hyper planes are built proportionately. The goal is to utilise the margin—the space between the two exterior hyperplanes.It is stated that the classifier's simplification error will increase the more the margin is increased. While most classifiers strive to minimise the empirical risk, or the error on the training set, SVMs were actually designed with a structural risk minimization attitude, which seeks to minimise an upper clear of the generalisation error. Finding a decision function that minimises the functional is the goal of the SVM algorithm. Using a small training set, the SVMs enable the training of nonlinear classifiers in high-dimensional domains.This is made possible through the choice of a group of vectors (known as the support vectors) that defines the correct limits between the categories of fit.

XIII. CONCLUSION

We covered a variety of image classification techniques in this paper. The most popular methods for classifying images fall into the categories of hard and soft classification, supervised and unsupervised learning, parametric and nonparametric analysis, object-based learning, sub-pixel, per-pixel, and per-field classifiers, spectral classifiers, context-dependent classifiers, and spectral-contextual classifiers. Here, some of the most widely utilised methods are covered. This survey provides theoretical information regarding classification techniques as well as guidance for choosing the best classification techniques. Image classification's main goal is to accurately recognize the features that exist in an image. Utilising both supervised and untrained methods, image categorization. Unsupervised classification is entirely computer-operated, therefore human interaction is not at all necessary.

REFERENCES

[1] https://medium.com/analytics-vidhya/image-classification-techniques-83fd87011cac
[2] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain." Psychol. Rev., vol. 65, pp.386–408, 1958.
[3] S. Fareed, K. Bhatia, S. Bhadola, and R. Sharma,"An Overview of Object Tracing Techniques in Videos and Images", Central Asian Journal of Theoretical and Applied Sciences, vol. 3, no. 7, 2022.
[4] M. Nazir, S. Bhadola, K. Bhatia, and R. Sharma, "A Complete Analysis of Human Action Recognition Procedures", vol. 6, no.5, 2022.
[5] P.J Werbos, "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Science", Ph.D. Thesis, Harvard University, Cambridge, MA, USA, 1974.
[6] Y.Lecun, L. Bottou, Y. Bengio, P. Haffner, "Gradient-Based Learning Applied to Document Recognition", in *Proc. IEEE* , *86*, pp. 2278–2324, 1998.
[7] J. Zhou, Y. Zhao, "Application of convolution neural network in image classification and object detection", Comput. Eng. Appl., vol. *53*, pp. 34–41.
[8] L. Zhang and B. Du, "Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art", IEEE Geosci. Remote Sens. Mag*,* vol. *4*, pp. 22–40, 2016.

[9] G. Cheng, C. Yang, X. Yao, L. Guo, and J. Han, "When deep learning meets metric learning: Remote sensing image scene classification via learning discriminative CNNs", IEEE Trans. Geosci. Remote Sens. Vol. 56, pp. 2811–2821, 2018.

[10] F. Hu, G.S. Xia, J. Hu, and L. Zhang, "Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery", Remote Sens. Vol. 7, pp. 14680–14707, 2015.

[11] H. Sun, Li. S. Zheng, X. Lu, "Remote sensing scene classification by gated bidirectional network", IEEE Trans. Geosci. Remote Sens vol. 58, pp. 82–96, 2019.

[12] N. He, L. Fang, and S. Li, "Skip-connected covariance network for remote sensing scene classification", IEEE Trans. Neural Netw. Learn. Syst. vol. 31, pp.1461–1474, 2019.

[13] D. Yarowsky, "Unsupervised word sense disambiguation rivaling supervised methods",in Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics, Cambridge, MA, USA, June 1995, pp. 189–196.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor
7.54

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com