Satellite Image Classification using Spectral Signature and Deep Learning

Zainab Hussain Jarrallah [*], Maisa'a Abd Ali Khodher
Department of Computer Science, University of Technology, Baghdad, Iraq

Abstract

When images are customized to identify changes that have occurred using techniques such as spectral signature, which can be used to extract features, they can be of great value. In this paper, it was proposed to use the spectral signature to extract information from satellite images and then classify them into four categories. Here it is based on a set of data from the Kaggle satellite imagery website that represents different categories such as clouds, deserts, water, and green areas. After preprocessing these images, the data is transformed into a spectral signature using the Fast Fourier Transform (FFT) algorithm. Then the data of each image is reduced by selecting the top 20 features and transforming them from a two-dimensional matrix to a one-dimensional vector matrix using the Vector Quantization (VQ) algorithm. The data is divided into training and testing. Then it is fed into 23 layers of deep neural networks (DNN) that classify satellite images. The result is 2,145,020 parameters, and the evaluation of performance measures was accuracy = 100%, loopback = 100%, and the result F1 = 100 %.

Keywords: Image preprocessing, Fast Fourier Transform, Spectral Signature, Deep Learning, Data set, Deep Neural Network, Artificial Neural Network.

تصنيف صور الأقمار الصناعية باستعمال التوقيع الطيفي والتعلم العميق

زينب حسين جارالله [*] ، ميساء عبدعلي خضر
قسم علوم الحاسوب ، الجامعة التكنولوجية ، بغداد ، العراق

الخلاصة

عندما يتم تخصيص الصور لتحديد التغيرات التي حدثت باستعمال تقنيات مثل التوقيع الطيفي ،والتي يمكن استعمالها لاستخراج الميزات ،يمكن ان تكون ذات قيمه كبيرة .في هذا البحث تم اقتراح استعمال التوقيع الطيفي لاستخراج المعلومات من صور الاقمار الاصطناعية المسحوبة من موقع Kaggle ومن ثم تصنيفها الى اربع فئات. هذه الفئات هي: السحب ، الصحارى، المياه ، والمساحات الخضراء. بعد المعالجة المسبقة لهذه الصور ، يتم تحويل البيانات إلى توقيع طيفي باستعمال خوارزمية تحويل فورير السريع FFT ثم يتم تقليل بيانات كل صورة عن طريق تحديد أهم 20 ميزة وتحويلها من مصفوفة ثنائية الأبعاد إلى مصفوفة احادية الأبعاد باستعمال خوارزمية تكميم المتجهات. بعد تقسيم البيانات إلى تدريب واختبار ، ومن ثم تغذيتها في 23 طبقة من الشبكة

---

*Email: cs.20.26@grad.uotechnology.edu.iq

العصبية العميقة التي تصنف صور الأقمار الصناعية ، تكون النتيجة 020،145،2 معلمة. حصلنا على النتائج

التالية : الدقة= 100%، الاسترجاع =100% و 100%= F1.

1. Introduction

Satellite images cannot be obtained continuously, thus most satellite image time sequences are unequal (sampled irregularly over time with perhaps large holes) [1]. Researchers have been trying to extract information from images since the dawn of remote sensing. While significant progress has been made in many applications, more work remains to identify high-level elements such as buildings and roads [2].

Convolution Neural Networks (CNN) have paid a lot of attention to machine vision in recent years. CNN's can be trained to extract robust features from raw pixel values while also learning classes for object recognition tasks [3]. This paper used deep learning to classify satellite images into four categories after preprocessing operations to obtain clearer images without noise. Then comes the feature extraction stage using the Fast Fourier Transform technology and obtaining the best features using the Vector Quantization algorithm. Using a deep neural network as a classifier after preprocessing satellite images and extracting features using the images' spectral signature can obtain the best possible classification, high accuracy, recall, and F1 score. Therefore, we used a (DNN) classifier, and the satellite image classification accuracy was very high with low loss.

2. Related Works

In 2018, Ovidiu Csillik et al. [1] were able to distinguish citrus trees from other trees by building a simple convolutional neural network, and the work was going smoothly. They achieved high accuracy (general accuracy = 96.24%, precision (positive prophetic value) = 94.59%, recall = 97.94%). This was the first time CNN was utilized with drone images focused on citrus trees.

In 2019, Omid Ghorbanzadeh et al. [2] used RapidEye satellite data, and topographical factors were analyzed using machine learning capabilities such as SVM, RF, ANN, and various convolutional neural networks for landslide detection. The CNN16,5 approach had the highest precision of 83.31%, followed by random forest RF 5 and random forest RF 8, at 81.95% and 80.9%, respectively. CNN22,5 had the lowest false negative (FN) score and the highest recall metric value of 92.85% . However, this method's accuracy and F1 values were lower as a result. The CNN16,5 model has an excellent F1 value of 87.8%.

In 2020, Ibrahim Ghadirpour and Tijana Vujadinovic proposed providing a reliable approach to detect time-series change that involves identifying and describing only jumps in the spectrum and trends. The objectives of this study were to detect abrupt changes in the trend component of an unevenly spaced time series, improve the estimation of trend, seasonal components, and jumping sites by considering uncertainty in time series values, and estimate jump direction and magnitude. To characterize gradual and rapid changes in the environment, JUST has been used to successfully replicate time series of plants with different jump sites and sizes in time series of different species across a range of time series (fake and realistic plants of Southeast Australia) [3].

In 2021, Saurabh Kumar and Shwetank [4] proposed development of a spectral signature and feature extraction of land use classes. Using the multi-temporal and multispectral (MTMS) Landsat image dataset, they suggested this work in 2021 to establish a spectral signature and

feature extraction of land use classifications. From 2003 to 2017, the imagery dataset collected three photos from the Landsat satellite system's various sensors. The preprocessing of imagery is critical for extracting geographical data and analyzing land-use features. For different years (2017, 2010, and 2003), the classification accuracy utilizing the ANNs approach was 90.10%, 75.75%, and 78.37%, respectively.

3. Prepossessing

After gathering satellite images, all of these images were processed to improve details and prepared for the training phase. It involves several procedures, including[5]:

1- HSV*:* The HSV color model defines color based on the three primary characteristics of color: hue, saturation, and brightness. The simplest property of color is hue(H), which is just the color name, such as red or yellow. It goes from 0 to 360, depending on the position of the standard color wheel. Color purity is measured by saturation (S). The higher the value, the better. The purer the hue, the better. It has a scale from 0% to 100%. Luminance (V) is also known as brightness, and it goes from 0 to 100 [6].The following are the reasons for utilizing HSV color space in picture segmentation:

- The HSV color space is designed to approach human vision, i.e., to describe colors in a way that is comparable to how the human eye perceives color.
- The color is represented in RGB color space as a mixture of three primary colors: Red, Green, and Blue. HSV is a color description system that uses three components: hue (color), saturation (vibrancy), and value (brightness).
- The intensity information may be separated from the color information in HSV color space. This comes in handy in various situations, such as ensuring stability in the face of lighting fluctuations or eradicating shadows.
- The HSV color space can provide more information [7]. As explained in the definition earlier.

2- Color to Gray Conversion: The majority of the benefits of converting a color image to a grayscale domain include having less data because the grayscale domain has one channel rather than three in the RGB domain, allowing for faster processing in other stages (feature extraction and training phase) with minimal brightness influence [8].

3- Histogram Equalization (HE): A step in image preprocessing called equalization of the histogram modifies disparity based on the histogram of the image [9]. It is a typical technique for increasing contrast in digital photographs. Equalization of histograms (HE) has been shown to be a simple and effective picture contrast technique. The method for improving the traditional histogram, on the other hand, equalization approaches, in most cases, result in overwhelming contrast. Augmentation results in an artificial appearance and visual artifacts in the image after it has been treated. This method is effective. Because of its versatility, it is frequently used for picture enhancement. Simplicity and superior performance on practically all fronts for pictures of all kinds. The HE procedure is carried out by remapping the image's gray levels depending on the gray level. A probability distribution is just a probability distribution of the gray levels that were entered [10].

4- Blur the Image: The image is convoluted using a low pass filter kernel to achieve blurring. It is useful for noise reduction and smoothing images to remove minor details of texture or noise. It's frequently useful when using image processing algorithms that look at the image's finer details [11].

5- Resize: the resize algorithm was used to make the features in the image clear and reduce the size to make it more accurate by removing the extra features, so the processing time becomes less.

4. Fast Fourier Transform (FFT) for Feature Extraction

Many proposed methods are used to extract features. They include simple and complex frequency features such as Fourier and wavelet domains [12].

The following processing step is to use the Fast Fourier Transform (FFT), which translates each frame from the time domain into the frequency domain. It is necessary to obtain the volume frequency response [10]. This study describes how to use the FFT, a computer method that calculates the discrete Fourier transform faster than current algorithms. The time savings can be significant. For example, an $N = 2^{10}$ point transformation can be calculated 100 times faster using the FFT than using the straight method. The direct technique wants a time proportional to N to transform a point N, while the FFT needs a time proportional to N log2 N. The estimated FFT to direct calculation time ratio is as follows:

$$\frac{N log_2 N}{N^2} = \frac{N log_2 N}{N} = \frac{K}{N} \tag{1}$$

where $N = 2^k$, e.g., if $N = 2^{10}$, the FFT requires less than 1/100 of the usual calculation time. Convolution and correlation, both beneficial mathematical techniques in time-series analysis, are usually computed digitally by forming the wadded product [13].

5. Vector Quantization (VQ)

It is a process for taking a large group of feature vectors and producing a small group of feature vectors that reflect the spreading centroids, or points spaced so that the average distance between each of the other locations is minimized. Because accumulating each of the feature vectors is inefficiently generated from training utterances, VQ is used[14] . While the VQ technique is time-consuming to compute, it saves time during the testing phase. The Euclidean distance will be given as:

$$d(a,b) = \sqrt{\sum_{j=1}^{a}(aj - bi)^2} \qquad\qquad i=1,2 \ldots\ldots k \tag{2}$$

where aj denotes the jth component of the input vector, and bi denotes element of the code-word [15], i denotes the code-word, k denotes the number of clusters.

6. Evaluating Performance Measures

Precision, recall, and the F1-score are only a few of the statistical measurements utilized to boost performance power. Precision is measured by dividing the total number of optimistic forecasts by the number of actual positive forecasts. Recall is one of the most critical metrics in models with unbalanced datasets. The true positive rate is calculated in the model. The F1-score may be viewed as the average of recall and accuracy [16] .

$$Precision = \frac{true\ positive}{true\ positve + fales\ positve} \tag{3}$$

$$Recall = \frac{true\ positive}{true\ positve + fales\ negative} \tag{4}$$

$$F1 = 2 * \frac{precision - recall}{precision + recall} \tag{5} \quad [17]$$

7. Deep Learning Classifier

Deep learning allows computational models with numerous processing layers to learn and represent input at various levels of abstraction, simulating how the brain receives and analyzes modal information and automatically capturing sophisticated data structures. Neural networks, hierarchical probabilistic models, and a range of unsupervised and supervised feature learning algorithms are part of the deep learning family of approaches[18]. A deep neural network is made up of numerous layers of neural networks that are linked together by a deep learning model architecture. The output from neurons in one layer is supplied as an input to the next layer in a neural network. Because the data cannot be read directly, the layers between the input and output layers are concealed layers. The weights associated with the input, hidden, and output layers are adjusted to train the deep neural network using input documents and output

properties[19] **.** Before presenting the proposed system, it is necessary to refer to some of the important terms used in it and to include a simplified definition of them.

• Convolution layers (Conv1D): the activations are mapped from one layer to the next using a filter. A 3-dimensional weighted filter with the same depth as the current layer but a lower spatial extent is used in a convolution process [20]. A 1D convolution layer generates a tensor of outputs by applying a convolution kernel(k) to a single spatial (or temporal) dimension.

• Leaky ReLU layers (Lrelu): An additional parameter (0, 1) is used to define the leaky ReLU. Although it is a hyperparameter that the user chooses, it can also be learned. As a result, the leaky ReLU layers are used to reduce the high negative values that affect the results and to bring them closer to zero. The leaky ReLU also aids in the normalization of each neuron's output to a range from 1 to 0 or -1 to 1[20].

• Max-Pooling layers: Although the former happens considerably less frequently in deep architectures, the max-pooling layers are interspersed with the convolutional/ReLU layers. This is because pooling greatly decreases the spatial size of the feature map, requiring just a few pooling operations to reduce the spatial map to a tiny constant size[20].

• Flatten layer: The data moves on to the flattening layer after multiple iterations of the convolution layer, non-linear layer, and pooling layer [21]. The flattening layer takes the output from the previous layers and "flattens" it into a single vector that may be used as an input for the next level.

• Dense layer: Because the completely connected layers are tightly linked, the fully connected layers include the vast majority of parameters. It is the regular layer of a densely connected neural network. It is the most popular and widely utilized layer. In this system, the dense layer is used as SoftMax functions.

• SoftMax finds a maximum of a set of values using the logistic function [22].

• Stride(s): A neural network's filter parameter that controls the amount of movement across an image or video. The filter will move one pixel (or unit) at a time if the stride is set to 1.

8. Data Set

The Satellite Image Classification Dataset-RSI-CB256 is used [23]. This dataset has 4 different classes mixed from sensors and Google map snapshots. The whole dataset has 5631 images in jpg format, and each class has about 1500 images. The training dataset was used to calibrate the chosen model, whereas the testing dataset was used to evaluate the models' performance using a dataset that was not used to train them [24].

Table 1: Classes of Dataset

CLASS	NO. OF PICS
Cloudy	1500 cloud pics at different times
Green area	1500 green area pics at different time
Desert	1131 desert pics at different time
Water	1500 water lake pics at different time

9. Proposed System

Classification based on Deep Neural Networks Classifier (DNN) is one of the most common classifiers used to recognize images. It takes an image as input, processes it, and assigns it to one of several categories. In Figure (1) the proposed system, each input image will be processed through a set of convolution layers that include filters (Kernels), Maxpooling, leaky relu, activation function, and a flattening layer.
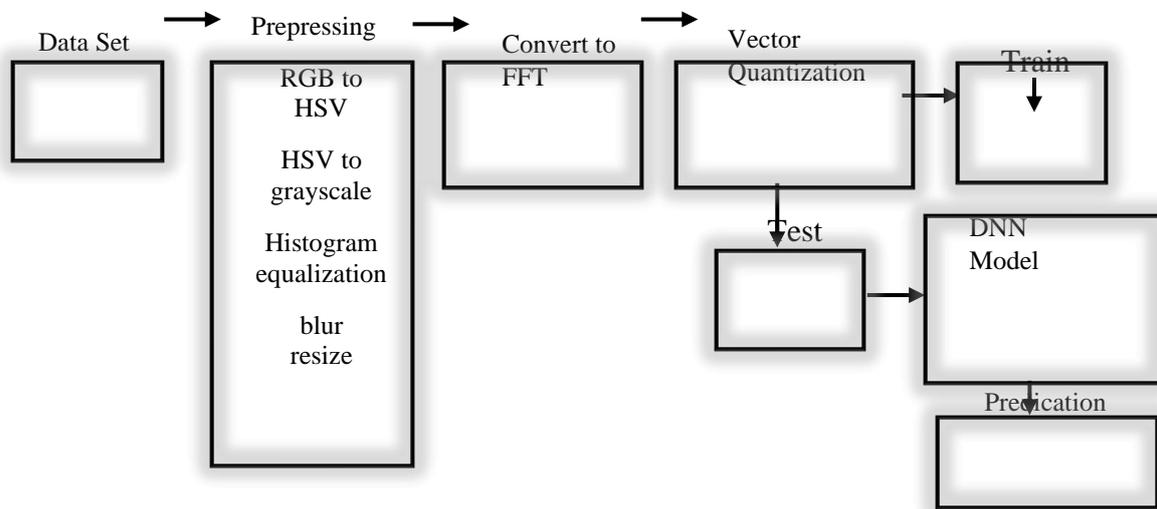

Figure 1: The Block Diagram of The Propose System

The proposed system's algorithm is as follows.

Algorithm (1): The Proposed System
Input: Dataset
Output: Accuracy
Begin
A: No. of images from the data set
B: HSV of image
C: Grayscale of the image
D: Histogram equalization
E: A blur of the image
F: Resize of image
G: Feature extraction FFT
H: Vector quantization of features
Step 1: Read satellite images
Step 2: Preprocessing phase
a. Convert Color to HSV Call B.
b. Convert Color to Grayscale Call C.
c. Histogram Equalization Call D.
d. Apply Median Filter Call E.
c. Image Resize Call F.
Step 3: Feature Extraction phase using FFT, Call G.
And Vector Quantization, Call H.
Step 4: Data Splitting, the data set divided to 70% of the dataset is for training and 30% for testing Call A.
Step 5: Initialize a random value for all parameters learning-rate = 0.003
Max_No._epoch = E (batch's size = 64, K=3 for each epoch) S = 1, padding="same"
Step 6: Call 1D CNN: Apply network layers

a. Conv 1, with 16 filter, act = Leaky ReLU, Max_pool, with size (2,2),Leaky ReLU
b. Conv 2, with 32 filter, Max_pool, with size (2,2)
c. Conv 3, with 64 filter, act = Leaky ReLU, Max_pool, with size (2,2).
d. Conv 4, with 128 filter, act = Leaky ReLU, Max_pool, with size (2,2).
e. Conv 5, with 256 filter, act = Leaky ReLU, Max_pool, with size (2,2).
f. Conv 6, with 512 filter, act = Leaky ReLU, Max_pool, with size (2,2).
g. Conv 7, with 1024 filter, act = Leaky ReLU
h. Conv 8, with filter= (NO. Classes+5) *5, K=1, padding="valid", act= Linear
i. Flatten output of conv 8
j. Fully connected layer with Dense layer, act = SoftMax
k. Classify the outcome for a give dataset
l. update weight using ADAM optimizer
Step 7: Return to Step 6
Step 8: CNN performance measures by confusion matrix
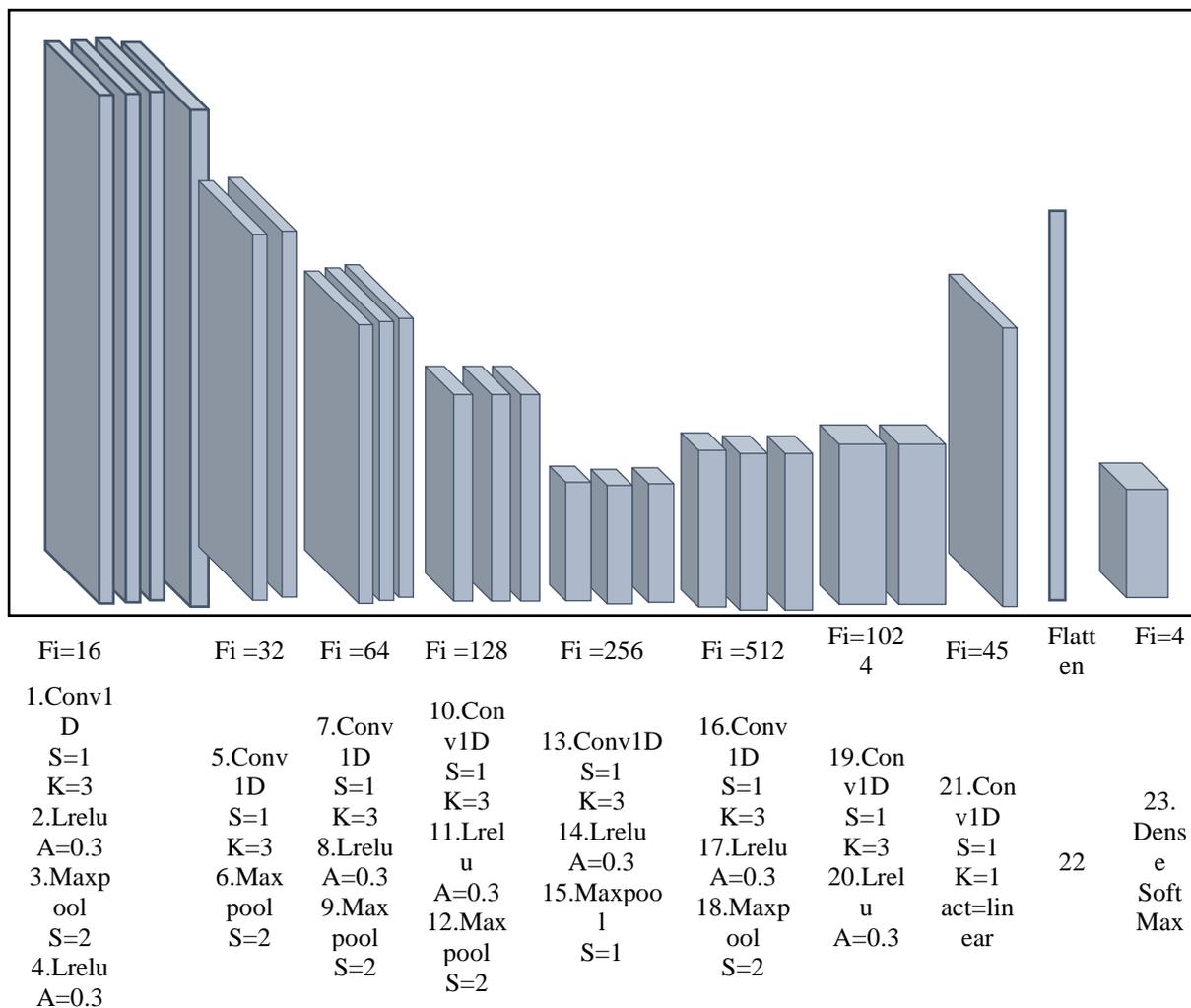Step 9: Classifier Results = Best (Accuracy. Results)
END

Fi=16	Fi =32	Fi =64	Fi =128	Fi =256	Fi =512	Fi=1024	Fi=45	Flatten	Fi=4
1.Conv1D S=1 K=3 2.Lrelu A=0.3 3.Maxpool S=2 4.Lrelu A=0.3	5.Conv1D S=1 K=3 6.Maxpool S=2	7.Conv1D S=1 K=3 8.Lrelu A=0.3 9.Maxpool S=2	10.Conv1D S=1 K=3 11.Lrelu A=0.3 12.Maxpool S=2	13.Conv1D S=1 K=3 14.Lrelu A=0.3 15.Maxpool S=1	16.Conv1D S=1 K=3 17.Lrelu A=0.3 18.Maxpool S=2	19.Conv1D S=1 K=3 20.Lrelu A=0.3	21.Conv1D S=1 K=1 act=linear	22	23. Dense Soft Max

Figure 2: The Proposed System DNN Layers

Where :

Fi refers to a filter. Conv 1D refers to one diminution convolution layer. K refers to kernel size. S refers to strite. Lrelu refers to the LeakyReLU layer. Act refers to the activation function.

9.1 Preprocessing

First Step: Convert satellite images from RGB to HSV, then convert these images from HSV to grayscale, as shown in Figure (3).
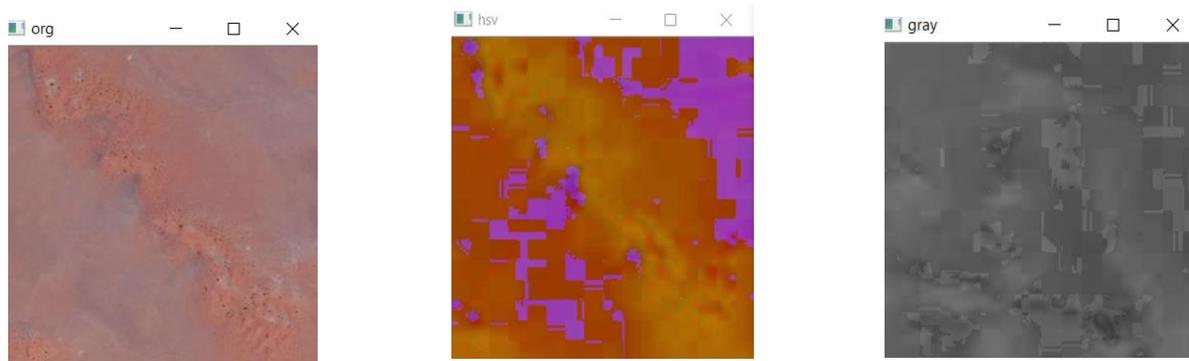

Figure 3: Convert from RGB to HSV and HSV to Grayscale.

Second Step: Apply histogram equalization and media blurs to obtain the best images without noise, as shown in Figure (4). Then resize images to a lower dimension 20×20 for best feature extraction.
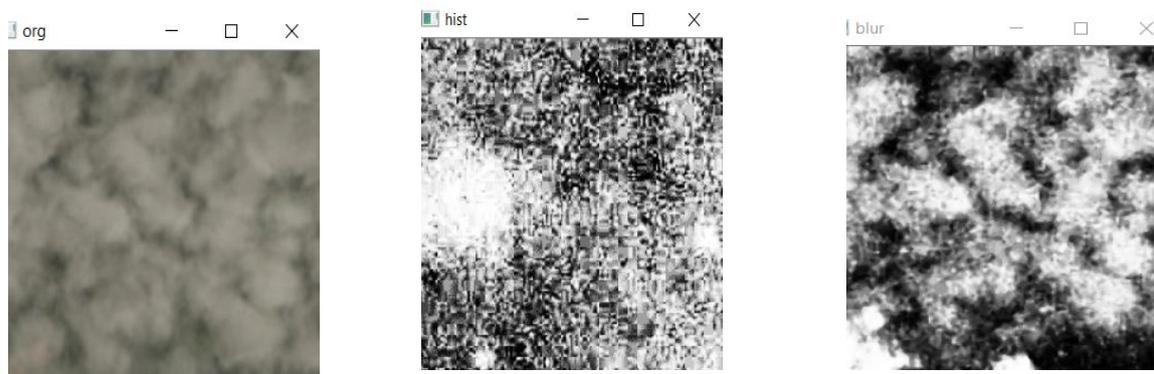

Figure 4: Apply Histogram Equalization and Blur.

9.2 Feature Extraction

First Step: This paper used Fast Fourier Transform for feature extraction using Eq.(1). That explains the work of converting image data after resizing to its spectral signature, as shown in Figure (5).

Second Step: Apply vector quantization as in Eq. (2). It is the process of turning a continuous range of values into a finite range of discrete values to reduce a group of features and obtain a one-dimensional vector.
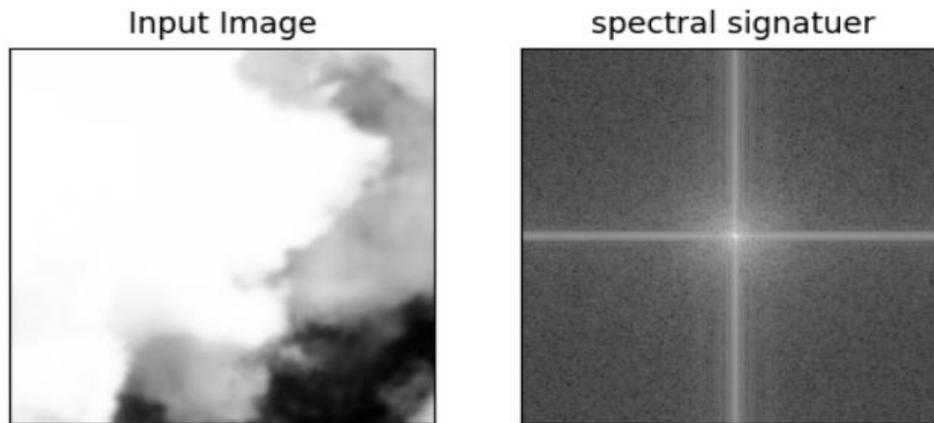
Figure 5: The Spectral Signature of an Image.

9.3 Deep Learning

When it comes to dealing with large datasets, the deep neural network (DNN) algorithm, which is quickly becoming one of the most popular deep learning classification algorithms, is thought to be great in classification because it produces high accuracy. Every classifier that uses deep learning techniques is composed of a series of layers. At this stage, 23 layers of the DNN are applied. These layers are as follows:

Eight conv containing filter = 16, 32, 64, 128, 256, 512, 1024, 45 and kernel = 3 except for the last layer, the kernel = 1. It has a linear activation function to withdraw strong and close features and all these layers have Stride(s)= 1. **Seven** layers leaky relu(Lrelu) with alpha=0.3 to reduce high negatives. **Six** layers Maxpooling to select the best features. **One** flatten layer to arrange the parameters according to the input, i.e. a set of vectors, each one consisting of 20 features. **One** dense layer with SoftMax activation function normalizes the weights. It limits them between 0 and 1 for the prediction to be faster, and it distributes the weights on the classes, so the last layer is placed (with learning rate =0.0003 to make learning challenging and slow but firm).

These layers are applied to the training data set and tested by testing the data set to obtain 2,145,092 parameters. The results are explained in section 10 of precision, recall, and F1-score by using Eq. (3), (4), and (5).

10. Results

The program was written using Python, a language that allows reliable writing systems for deep learning layers. The use of the proposed system's layers after preprocessing and conversion to a spectral signature, which was used as a feature extractor, as the FFT algorithm that converts image data into energy and reduces these features to 20 features using vector quantization, gave a high percentage of performance measures evaluation.

When these features enter the DNN network, the features are extracted from the intrinsic features where spaced weights are built before the classification. This results in the accuracy being 100%, precision being 100%, recall being 100%, and the F1 score being 100% with a loss=0.000024656. Combining the DNN with the FFT feature extraction technique is shown in Table (2).

Table 2: The Experimental Results of Implementing the DNN

CLASS No.	Precision	Recall	F1-score	Accuracy
0	1.00	1.00	1.00	1.00
1	1.00	1.00	1.00	1.00
2	1.00	1.00	1.00	1.00
3	1.00	1.00	1.00	1.00

Network training is a method of getting kernels in convolutional layers and weights in fully connected layers that reduces the gaps between performance predictions and actual base truth labels. We used 70% of the data for training, and the framework was trained using adaptive torque estimation (Adam) as an optimization technique with an initial learning rate of 0.003. to me. The loss function was used for categorical loss via entropy. The model was generated over 200 epochs with a batch size of 64. The generated model was validated using the test dataset to determine its validation. For each period, model performance was calculated using training accuracy, training loss, and validation loss. After previous operations, satellite images were discovered and categorized into four distinct types of categories, including deserts, bodies of water, green areas, and clouds**.**

11. Conclusion

It is not easy to identify the types of satellite images because they are taken in different conditions affected by the weather, the atmosphere, the noise they contain, and the difference in lighting. After all, they may have been taken at night. New methods for classifying satellite images are presented in this paper, which combine color saturation and then grayscale conversion with the histogram equation for a uniform distribution of pixel intensity, as well as the application of a Gaussian filter to exercise the Gaussian function in order to remove image noise and reduce dataset size for processing. Then the features extracted through the FFT algorithm finally enter the deep convolutional neural network (DNN) classifier, which consists of 23 layers. Accuracy is 100% with minimal loss. In a related study, Kumar and Shwetank [4] used the same technology and applied deep learning and reached 90.10%. We beat it, but the proposed system was more complex. The results illustrate the efficiency of the proposed method. In the future, deep learning algorithms will be used to extract features and classifiers from another dataset that will yield the same results.

References
[1] O. Csillik, J. Cherbini, R. Johnson, A. Lyons, and M. Kelly, "Identification of citrus trees from unmanned aerial vehicle imagery using convolutional neural networks," *Drones*, vol. 2, no. 4, pp. 1–16, 2018, doi: 10.3390/drones2040039.
[2] O. Ghorbanzadeh, T. Blaschke, K. Gholamnia, S. R. Meena, D. Tiede, and J. Aryal, "Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection," *Remote Sens.*, vol. 11, no. 2, Jan. 2019, doi: 10.3390/rs11020196.
[3] E. Ghaderpour and T. Vujadinovic, "Change detection within remotely sensed satellite image time series via spectral analysis," *Remote Sens.*, vol. 12, no. 23, pp. 1–27, Dec. 2020, doi: 10.3390/rs12234001.
[4] S. Kumar, S. Shwetank, and K. Jain, "Development of spectral signature of land cover and feature extraction using artificial neural network model," *Proc. - IEEE 2021 Int. Conf. Comput. Commun. Intell. Syst. ICCCIS 2021*, pp. 113–118, 2021, doi: 10.1109/ICCCIS51004.2021.9397172.
[5] I. Abed, "Lung Cancer Detection from X-ray images by combined Backpropagation Neural Network and PCA," *Eng. Technol. J.*, vol. 37, no. 5A, pp. 166–171, 2019, doi: 10.30684/etj.37.5a.3.
[6] S. Li and G. Guo, "The application of improved HSV color space model in image processing,"

Proc. 2010 2nd Int. Conf. Futur. Comput. Commun. ICFCC 2010, vol. 2, pp. 10–13, 2010, doi: 10.1109/ICFCC.2010.5497299.

[7] D. V. R. Ganesan P, "VALUE BASED SEMI AUTOMATIC SEGMENTAION OF SATELLITE IMAGES USING HSV COLOR SPACE, HISTOGRAM EQUALIZATIONAND MODIFIED FCM CLUSTERING ALGORITHM," in *2013 International Conference on Green Computing, Communication and Conservation ofEnergy (ICGCE)*, 2013, vol. 8, pp. 77–82.

[8] I. S. A.-J. Aliris, "Face Recognition in Digital Video by Using Artificial Intelligence Techniques," University of Technology, Department of Computer Science, 2015.

[9] A. R. Abbas and A. R. Kareem, "Age estimation using support vector machine," *Iraqi J. Sci.*, vol. 59, no. 3, pp. 1746–1756, 2018, doi: 10.24996/IJS.2018.59.3C.19.

[10] M. Kaur, J. Kaur, and J. Kaur, "Survey of Contrast Enhancement Techniques based on Histogram Equalization," *Int. J. Adv. Comput. Sci. Appl.*, vol. 2, no. 7, 2011, doi: 10.14569/ijacsa.2011.020721.

[11] W. Zhang, W. Quan, and L. Guo, "Blurred star image processing for star sensors under dynamic conditions," *Sensors (Switzerland)*, vol. 12, no. 5, pp. 6712–6726, 2012, doi: 10.3390/s120506712.

[12] H. A. Jaber and M. T. Rashid, "HD-sEMG Gestures Recognition by SVM Classifier for Controlling Prosthesis," *Iraqi J. Comput. Commun. Control Syst. Eng.*, pp. 10–19, 2019, doi: 10.33103/uot.ijccce.19.1.2.

[13] E. O. Brigham and R. E. Morrow, "The fast Fourier transform," *IEEE Spectr.*, vol. 4, no. 12, pp. 63–70, 1967, doi: 10.1109/MSPEC.1967.5217220.

[14] F. K. Soong, A. E. Rosenberg, L. R. Rabiner, and B. H. Juang, "Vector Quantization Approach To Speaker Recognition.," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, no. May, pp. 387–390, 1985, doi: 10.1109/icassp.1985.1168412.

[15] R. A. Mohammed and N. F. Hassan, "Proposal Personal Identification Model based on Multi Biometric Features," University of Technology, Department of Computer Science, 2020.

[16] N. H. Marza, M. E. Manaa, and H. A. Lafta, "Classification of Spam Emails using Deep learning," pp. 63–68, 2021, doi: 10.1109/bicits51482.2021.9509909.

[17] M. M. Mahmoud and A. R. Nasser, "Dual Architecture Deep Learning Based Object Detection System for Autonomous Driving," *Iraqi J. Comput. Commun. Control Syst. Eng*, vol. 21, no. 2, pp. 36–43, 2021.

[18] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep Learning for Computer Vision: A Brief Review," *Comput. Intell. Neurosci.*, vol. 2018, 2018, doi: 10.1155/2018/7068349.

[19] S. Sumathi and G. K. Pugalendhi, "Cognition based spam mail text analysis using combined approach of deep neural network classifier and random forest," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 6, pp. 5721–5731, 2021, doi: 10.1007/s12652-020-02087-8.

[20] E. Alpaydin, "Neural Networks and Deep Learning," *Mach. Learn.*, 2021, doi: 10.7551/mitpress/13811.003.0007.

[21] R. M. J. Al-Akkam and M. S. M. Altaei, "Plants Leaf Diseases Detection Using Deep Learning," *Iraqi J. Sci.*, vol. 63, no. 2, pp. 801–816, 2022, doi: 10.24996/ijs.2022.63.2.34.

[22] B. Endale, A. Tullu, H. Shi, and B. S. Kang, "Robust approach to supervised deep neural network training for real-time object classification in cluttered indoor environment," *Appl. Sci.*, vol. 11, no. 15, 2021, doi: 10.3390/app11157148.

[23] "https://www.kaggle.com/mahmoudreda55/satellite-image-classification."

[24] Z. A. Mohammed, M. N. Abdullah, and I. H. Al-hussaini, "Predicting Incident Duration Based on Machine Learning Methods," *Iraqi J. Comput. Commun. Control Syst. Eng.*, pp. 1–15, 2021, doi: 10.33103/uot.ijccce.21.1.1.