Journal of Intelligent Systems and Internet of Things (JISIoT)
Vol. 09, No. 01, PP. 36-48, 2023
DOI: https://doi.org/10.54216/JISIoT.090103

Intelligent Multi-level Feature Fusion using Remote Sensing and CNN Image Classification Algorithm

Mustafa Altaee[1]*, A. Talib[2], M. A. Jalil[3], J. Ali[4,5], Thamer A. Alalwani[6]

[1]Department of Medical Instruments Engineering Techniques, Al-farahidi University, Baghdad, Iraq, [2]Department of Computer Communications Engineering, National University of Science and Technology, Thi Qar, Iraq, [3]Department of Computer Engineering Techniques, Al-turath University College, Baghdad, Iraq, [4]Department of Computer Engineering Techniques, Mazaya University College, Thi Qar, Iraq, [5]MEU Research Unit, Middle East University, Amman, Jordan, [6]Department of Radiological Techniques, Al- Mustaqbal University College, 51001 Hilla, Iraq
Emails: m.altaee@alfarahidiuc.edu.iq; ali.j@nust.edu.iq; mohammed.jalil@turath.edu.iq; a.jawad@mpu.edu.iq; thamerabdalhamza@uomus.edu.iq

Abstract

The collection of features in both multispectral and hyperspectral domains is possible with hyperspectral image (HSI). It comprises a vast array of multispectral bands with functional relationships. However, they become more complex when dealing with small samples. To tackle this issue, researchers employed a deep learning convolutionary neural network system (DL-CNN) and implemented a temporal abstraction strategy to grade HSI. This approach is an intelligent multi-level feature fusion that combines the temporal abstraction strategy and DL-CNN for HSI grading. Researchers have introduced the impact of seven separate classifiers in implementing the location estimation on our broad convolutionary neural network system (CNN) framework, which plays the shallow CNN model's main training phase. PSO, Adagrad, Plans to implement, Alexnet, Adam, Environmental benefits, and Nadam are the seven distinct remained significantly included in this analysis. A detailed study of the four multispectral remote sensing technique sets showed the deep CNN system's supremacy followed with the HSI identification AlexNet Optimizer.

Received: August 17, 2022 Revised: January 01,2023 Accepted: February 19, 2023

Keywords: Convolutional Neural Network; Image Classification; Intelligent Multi-level Feature Fusion; Remote Sensing; Deep Learning

1. Introduction

Mass networking takes the shape of visual photographs in the modern field of internet media. There are various photos from media platforms to the traditional web pages that have to be categorized by intelligent and powerful software or strategies [1]. Quality monitoring, optical character recognition technologies, and satellite imagery are currently used for classification methods. Despite variations in posture, stance, or perspective, the naked eye may distinguish a family member from hundreds of unknown persons [2]. The human heart and mind categorize a picture using perceptual interpreting components; data mining techniques may identify a photograph. A recognition system can understand an image by extracting features such as structure, chronology, and softness [3].

Support Vector Machine (SVM) is a highly effective and unique classification. The controlled method of segmentation of tweets was integrated with the learning algorithm from image processing techniques in the study. For the variety of pictures, SVM classification is employed. The ensemble classifier is a methodology of choice or a model used in different fields, for example, action recognition, categorization of papers, etc. [4]. In NLP (machine learning) and in

the compilation of knowledge (in which a text or statement is viewed as a similarity measure of his phrase), the term bag is often commonly used in the template (BoW prototype) [5]. In several object detection and natural language processing topics, convolutionary neural network system (CNN) has proved to turn out well. CNN provides image processing for many uses. Photos are taken as artifacts the picture represents and categorized the identification of the subject's type [6].

In the 2000s, the idea suggested and continued the development of ground-breaking machine learning. CNN can investigate feature representations of remote imaging in tandem with the rise of deep computing to address challenges facing conventional supervised learning. Through using networks consisting of several feature maps, CNN can quantify theoretically resilient functions [7]. The CNN architecture can change its connection weights such that artifacts in the picture are defined more robustly. In other sentences, CNN can derive the characteristics at low levels and the elements at top standard. Conversely, CNN model specifications are periodical manner on a required measurement and solved very slowly [8].

For CNN, the issues with conventional pixel-based identification are serious, along with salt and pepper impact and the boundaries fluctuating effect and within the category [9]. CNN often needs several storage stream processors, which consume lots of space and resources on the hard disc. For medium-resolution remote sensing image processing, patched-based CNN identification has been suggested, and different patches are not standardized in scale.

In the other direction, deep learning CNN (DL-CNN), replacing small fixtures of knowledge, is a way to divide distant location images into concrete object classes employing machine learning and their properties by geographical, temporal, and sequential structures [10]. Transform-based categorization, pre-estimation, object collection, picture description, etc., have taken much advancement with GEOBIA.

A DL-CNN provides a plausible way to deal with serious learning issues. The pixel values are compared with theoretical CNN. Second, no customer computer vision technologies are needed [11]. For DL-CNN, object detection sizes vary considerably with machine learning, which is inappropriate as CNN feedback.

The rest of the research work as follows. Section 2 deals with the literature and the background of the CNN image classification methods. The proposed DL-CNN is designed, and implementation is done in section 3. Software implementation and comparison are discussed in Section 4. Section 5 illustrates the proposed system's conclusion and future scope.

2. Related Works

Specific classification techniques, including KNN, maximal probability of incidence, minimal distance, regression analysis (LR), adaptive intelligence grid, and supporting machine learning algorithm, have been implemented in the first stage of identifying the Hyperspectral Image (HSI) framework (SVM) [12]. In the HSI classification stage, the proposed method has been quite good. These classifications were the cutting-edge tools for the identification of HSI for a good while longer. In the past months, Deeper Training, which involves image detection, linguistic analysis, voice control, and many others, has main attraction success in computer training and has been used extensively from almost all implementations [13].

In image processing, the DL approach was recently developed by Bera and Shrivastava [14]. The DL method will automate the learning of representation functionalities from the information, including auto-encoder (SAE), deep-belief System (DBN), recurring neural network, and genetic algorithms of quantization (CNN) [15]. Due to its protection of local connections and weight exchanging, Deep CNN has already shown existing members and usefulness among different DL measures for HSI removal. Li *et al*. suggested that 1D, 2D, or 3D-CNN derives synthetic, temporal, and spectroscopic spatial characteristics individually [16]. Hamouda, Ettabaa, and Bouhlel introduced a Clustering algorithm of the adjustable kernel and display settings. Shao and Cai suggested the spectrum and multi-scale Neural network model [17].

Martins *et al*. suggested CNN triple technology to use comparable spectral-spatial details. A combination of low-level characteristics with comprehensive information and system requirements attributes, and semantic knowledge, offers complimentary sensory features [18]. A Bavarian structure based on CNN and conditional random field was suggested by Sujitha *et al*. (MRF) [19]. The SVM classifier should be used for obtaining spectral-space characteristics, and for using location data, the transparent MRF is provided in advance. A convolution layer with incorporated semantic features was introduced by Ye *et al*. [20].

Optimizers play an essential part in the preparation of the algorithm in computer vision. Conversely, several profound schooling systems have been optimized using linear regression (SGD) [21]. However, the backpropagation in SGD is very easy to alter since the scales of the varying factors vary considerably. Optimized SGD versions such as Adagrad, Support vector, Adam, Environmental benefits, and Nadam have been suggested to solve such challenges [22]. Between these adaptive optimization methods, Adam Optimizer was used mainly in a few HSI identification literature and remainder optimization methods were not investigated to the maximum of our expertise when implementing HSI identification [23]. In comparison, a few courses are written in which quantitative study of different optimizing compilers has been provided concerning the profound CNN template's success.

Neither one of them treated the multispectral data collection for remotely sensed. Consequently, researchers have examined the seven optimization algorithms listed above and provided a proper assessment of the profound CNN models success in the HSI sample group category [24]. In brief, this study's significant strengths are: (1) Designing a more in-depth Classification algorithm with various HSI identification implementations, (2) contrasting the specific deep CNN models for Lesion segmentation to a wide variety of simulation technologies.

3. Proposed DL-CNN

Two hundred images of five categories, namely, binoculars, planes, faces of people, watches, and motorbikes, have been taken from remote sensing image archive known. The Caltech 256 contains 30,607 images of 256 objects. Figure 1 shows some of the photos used in the experiment.

3.1. Feature extraction

During the first phase, all pictures must be removed. This system is achieved by constructing an entity in the picture store containing all the images and their names.

3.1.1. Word bag (boW)

One of the several separation processes used is BoW to remove image characteristics from each collection. A boW can be described as "pixel interpretation based on the subjective characteristics" boW is a three-stage image: Identification, role definition, and codebooks production.

The boW would display these characteristics in a vector, referred to as a function description, after identification. To identify and improve the function predictor Speed-Up Robust Features (SURF) functionality, BoW utilizes the SURF function applies. SURF operates in three stages: Finding highlights, describing the local district, and balancing. To classify connected components, SURF utilizes a blob sensor dependent on the Configuration space's leading determining factor.

The local district's definition generates the representation solely based on its features, for example, pixel strength compared to the feature vectors. The derived types are classified by cluster analysis with K-means. K-means that each measurement is split into a group with a medium close to the height. This system is done through two stages of interchanging.

Figure 1 shows the proposed remote sensing image classification DL-CNN. Evaluate phase where the constellation with a medium ultimately decreases square cluster total amount (WCSS), and a digital paradigm update process has been assigned to each analysis. Application phase: Eventually, these characteristics are used in the SVM classification Linear and Wave equation.

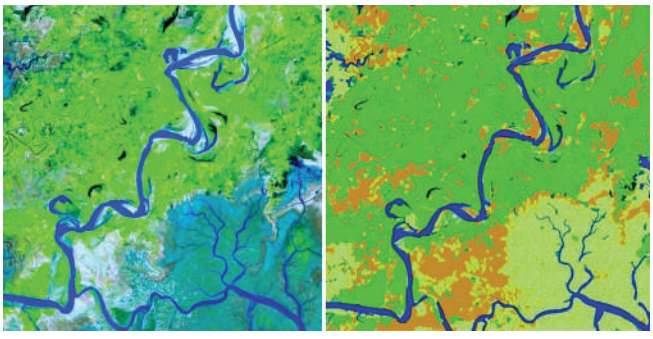

Figure 1: Remote sensing image classification method of the proposed deep learning convolutionary neural network system.

3.1.2. CNN feature extraction

The other method used for removing information is CNN. It utilizes strength shared and regional convolution layers. CNN has two main properties of CNN. Optimally, CNN is fitted with three kinds of surface topology, pool surface, and overlay linked. The principal aim of fully connected layers is to data comprises and to model the softmax function. The ultimately linked layer learns from the properties of the convolution operation.

The AlexNet Template is used here as a quick post-CNN. Pictures in 2000 classes can condition the template. However, AlexNet is used in this article to derive picture functions. AlexNet's following multiple levels are fully associated layers, and the final surface generates a 2000 class mark delivery. A photo of kernel function 228 × 228 requires the first fully connected layers sheet. Both photographs are then updated to the appropriate size.

Per layer of 2048 electrodes is associated. Data rises and decreases in AlexNet are minimized in multiple forms. Feature vectors are extended dynamically with modifications that retain the mark and produce translated pictures with very few computations. Dropdown, with 0.3 chance, the invisible neuronal is considered to be zero. This system eliminates through and backward spread and helps nodes acquire resilient characteristics, as they do not rely much on other memories.

Googlenet has several levels, and not all strands are appropriate for removing functions. Properties, including blobs and corners, are extracted and on the first surface. There are, therefore, more significant distinct functions using a more prominent chance. Thus, fc6 is required to data comprises before identification sheet. These functions are then used for teaching and evaluating SVM classification systems in standard and polynomial formats.

3.2. Remote sensing classifiers

Predictive methods which can classify new findings into an appropriate folder are models. In the master training course, the paradigm that knows about the correct area before classifying an unfamiliar object or experience is called the training set.

Figure 2 shows the remote sensing optimization techniques of the proposed DL-CNN. The detected feature is clustered, then visual words are displayed from the vocabulary in the database. This analysis splits the picture sets into the classification model, and trials of the 100 samples, 125 are included for classification training, and 30 are required for classification research. In addition to this configuration, the 250 cameras allow for 10-time-10-fold testing data.

3.2.1. SVM vector

SVM utilizes a Decision boundary that maximizes the margins, that is, the length from the closest point of practice.

3.2.2. SVM square

Kernel Function, which requires a variational distance matrix, the hyperplane converts to an input image.

3.2.3. CNN

The visual perception structure is predicated on neuroscience and can effectively detect and execute instructions. The machine-learning author studied many data analysis techniques for the recognition and characterization of objects using such principles. In the last few years, CNN is one of the most appealing data mining techniques for this theory.

Figure 3 shows the neural network analysis of the proposed DL-CNN. In particular, due to a shared network and weight-sharing process, CNN is defined by a few mathematical variables. A complete network architecture with an overall framework, weight matrix, and entirely linked layer is created.

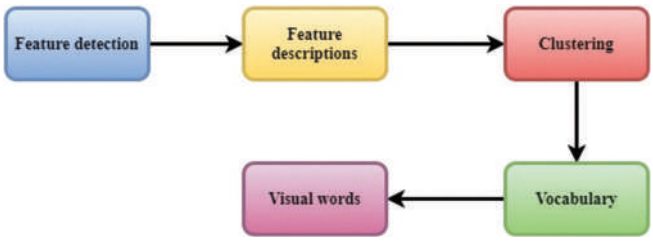

Figure 2: Remote sensing Optimization techniques of the proposed deep learning convolutionary neural network system based on intelligent multi-level feature fusion.

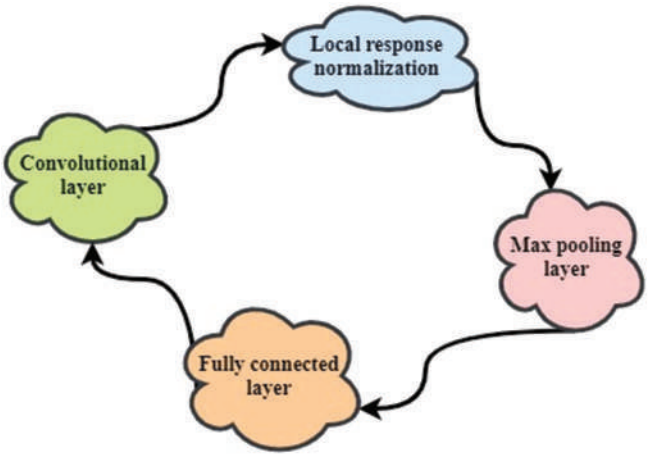

Figure 3: Neural network analysis of the proposed deep learning convolutionary neural network system.

There are two critical elements: Paying, one buffer (or kernels), and the second. The moderators have a limited dynamic range and are qualified to obtain those image characteristics. The neural network demonstrates that this function is expressed in Equation (1)

$$x_n^{'} = a\left(\sum_{m=1}^{K} x_m^{r-1} * w_{mn}^{r} + b_m^{r}\right) \tag{1}$$

X'n describes the new r^{th} layer of n^{th} input neuron, x_m^{r-1} is the former $(r-1)^{th}$ function representation, and K is the number of masks of the current amplification. Synaptic weights variables are w_{mn}^{r} and b_m^{r} The * generator can be used for matrices, and α corresponds to the variable induction. After using the initialization feature from every input vector, the configuration mapping is then sent to the convolutional layers.

The combining system includes the separability of the localization by reducing the accuracy of the triggering chart. The *dxd* trigger map interface produces the trigger layer of both the collection. Accumulation is by far the most common form. Finally, to construct the required volume, the ultimately linked layer integrates information from all input images in the preceding stage.

3.2.3. Operator

The profound CNN is learned by upgrading the specifications of all stages of the system and maximizing iterates. In the human brain, the loss function technique is a convenient addition.

The specifications are modified in the opposite way of the optimization problem regression fixed-function concerning the variables to reduce the optimal solution $J(\theta)$, $\nabla_{\theta} J(\theta)$ function. Any evaluation compares the required outcome and the expected input and reverses error. Territorial is one of the most common calculating output indicators. When the preferred result and projected performance are almost equal, the bridge value is nearly null; this is the crucial goal of any appropriate objectives.

The following are the various evaluation approaches.

SGD: It is a strategy for optimizing function θ to be modified at each period t. The Equation SGD obtains the following Equation (2).

$$\theta_{t+1} = \theta_t - \eta d_t \tag{2}$$

Where d_t denotes an empirical feature regression dependent on θ the period phase t of η reflects the training set.

(2) Adagrad: It is an interactive learning rate-adapting technology with all the variables modified. The mass related to the changing for Rmsprop is expressed in Equation (3).

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{D_t + \varepsilon}} d_t \tag{3}$$

Any triangular factor is denoted as D_t and is the amount of the approximate quadruple differential concerning θ at period t, whereas ε is a relatively small percentage variable to prevent the differential from dividing by zero.

(3) Adadelta: It is an amendment to Adagrad that aims to reduce its cognitive development, decreasing monotonously. Adagrad builds up all previous stable atom for the rate update in typical instances, while Adadelta uses the specified number's openings to build up previous patterns. In this case, the number of brush strokes is interpreted as a decreasing aggregate of ever. The running mean $R[d^2]_t$ value at stage t depending, as shown in formula, on the later mean and current trajectory is expressed in Equation (4).

$$R[d^2]_t = \gamma R[d^2]_{t-1} + (1-\gamma)d^2{}_t \tag{4}$$

Here γ is a continuous decomposition of 0.8 worth. Researchers can now reinterpret the Equation (3) download Adagrad while using the variable $\Delta\theta_t$ as displayed in equation (5)

$$\Delta\theta_t = -\frac{\eta}{\sqrt{D_t + \varepsilon}} d_t \tag{5}$$

The median deterioration over past triangular laplacian $R[d^2]_t$ as seen in Equations was substituted by D_t triangular in delta form that is expressed in Equation (6). This concept is defined as the root average differential error (RMS), as seen in Equation (7). The training rates η in Equation (7), as shown in Equation, are supplemented by the base means squares mistakes of the change variable $RMS\,\Delta\theta_t$ is expressed in Equation (8).

$$\Delta\theta_t = -\frac{\eta}{\sqrt{R[d^2]_t + \varepsilon}} d_t \tag{6}$$

$$\Delta\theta_t = -\frac{\eta}{RMS[d]_t} d_t \tag{7}$$

$$RMS\Delta\theta_t = \sqrt{R[\Delta\theta^2]_t + \varepsilon} \tag{8}$$

Any triangular factor is denoted as D_t and is the amount of the approximate quadruple differential concerning θ at period t, whereas ε is a relatively small percentage variable to prevent the differential from dividing by zero where $R[\Delta\theta^2]_t$. The rounded notifications and the size specified for the Equation by the packed variable are expressed in Equation (9).

$$R[\Delta\theta^2]_t = R\gamma[\Delta\theta^2]_{t-1} + (1-\gamma)\Delta\theta^2{}_t \tag{9}$$

As $RMS\,\Delta\theta_t$ is not known for the actual time, the previously $RMS\,\Delta\theta_{t-1}$ upgrade rule in Equation (10) and the last Adadelta rule in the Formula is assumed to be generated in the formula Adadelta upgrade term is expressed in Equation (11).

$$\Delta\theta_t = -\frac{RMS\Delta\theta_{t-1}}{RMS[d]_t} d_t \tag{10}$$

$$\theta_{t+1} = \theta_t + \Delta\theta_t \tag{11}$$

(4) RMSprop: Another adaptive optimism strategy to reduce the monotonous decay of backpropagation of the Adagrad. The additional mass services of RMSprop distinguish the information gain between the inversely proportional raw root of the mean continuum $R[d^2]_t$. The Equation gives the mass change is expressed in Equation (12).

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{R[d^2]_t + \varepsilon}} d_t \tag{12}$$

(5) Adam: It holds an infinitely long aggregate of current (a_t) gradients that is the initial (*mean*) time and previous (u_t), the third attempt (*variance)*. The following was calculated a_t and u_t is expressed in Equations (13) and (14)

$$a_t = \beta_1 a_{t-1} + (1-\beta_1)d_t \tag{13}$$

$$u_t = \beta_2 u_{t-1} + (1-\beta_2)d_t{}^2 \tag{14}$$

It holds an infinitely long aggregate of current (a_t) gradients that is the initial (*mean*) time and previous (u_t), the third attempt (*variance*). When the decline rate is minimal (i.e., β_1 and β_2 are close to 0, a_t and u_t). For the first and successive, instants with the skewed word are evaluated to solve this circumstance that is expressed in Equations (15) and (16)

$$\widehat{a_t} = \frac{a_t}{1-\beta_1^t} \tag{15}$$

$$\widehat{u_t} = \frac{u_t}{1-\beta_2^t} \tag{16}$$

It holds an infinitely long aggregate of current (a_t) gradients that is the initial (*mean*) time and previous (u_t), the third attempt (*variance*). When the decline rate is minimal (i.e., β_1 and β_2 are close to 0, a_t and u_t). The system relates the Adam weighted modification is expressed in Equation (17).

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\widehat{u_t} + \varepsilon}} \widehat{a_t} \tag{17}$$

Any triangular factor is denoted as D_t *and* is the amount of the approximate quadruple differential concerning θ at period t, whereas is a relatively small percentage variable to prevent the differential from dividing by zero. It holds an infinitely long aggregate of current (a_t) gradients that is the initial (*mean*) time and previous (u_t), the third attempt (*variance*).

Figure 4 shows the representation of θ_{t+1}. Any triangular factor is denoted as D_t and is the amount of the approximate quadruple differential concerning θ at period t, whereas is a relatively small percentage variable to prevent the differential from dividing by zero. It holds an infinitely long aggregate of current (a_t) gradients that is the initial (*mean*) time and previous (u_t), the third attempt (*variance*).

(6) AdaMax: The AdaMax mass upgrade procedure ensures a substitution by the upper bound indefinitely standard (v_t) for the word oscillating in $\sqrt{\widehat{u_t} + \varepsilon}$ as seen in Equation (17) in Adam's error propagation principle is expressed in Equation (18)

$$\theta_{t+1} = \theta_t - \frac{\eta}{v_t} \widehat{a_t} \tag{18}$$

Where v_t is expressed in Equation (19)

$$v_t = \beta_2 u_{t-1} + (1-\beta_2) d_t = (\beta_2 . u_{t-1}, d_t) \tag{19}$$

Any triangular factor is denoted as D_t *and* is the amount of the approximate quadruple differential concerning θ at period t, whereas ε is a relatively small percentage variable to prevent the differential from dividing by zero [25,26].

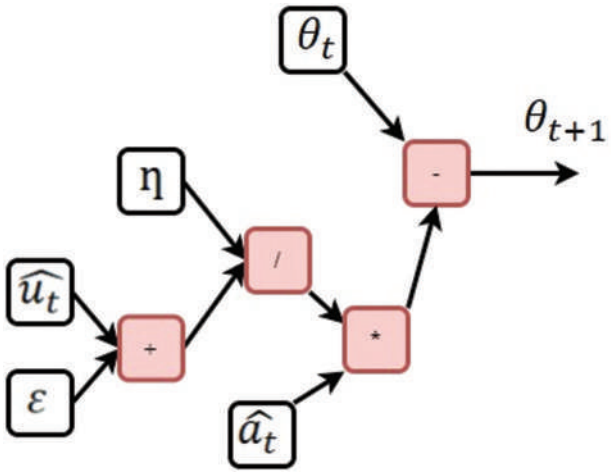

Figure 4: Representation of θ_{t+1}.

(7) Nadam: Innovating is an error propagation law option too. The function θ_{t+1} in Equation (17) with the description a_t and $\hat{a}s$ provided in the model is determined in Equation (20).

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\widehat{u_t} + \varepsilon}} \left(\frac{\beta_1 a_{t-1}}{1-\beta_1{}^t} + \frac{1-\beta_1}{1-\beta_1{}^t} d_t \right) \tag{20}$$

The variable amplitude of the preceding time stage prejudice calculation is denoted as $\frac{\beta_1 a_{t-1}}{1-\beta_1{}^t}$. So it's $\widehat{a_{t-1}}$ substitutable is expressed in Equation (21)

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\widehat{u_t} + \varepsilon}} \left(\beta_1 \widehat{a_{t-1}} + \frac{1-\beta_1}{1-\beta_1{}^t} d_t \right) \tag{21}$$

The Nadam addition to the structural is then given by substituting the biosecure approximation of the former vector propagation $\widehat{a_{t-1}}$ for the biosecure calculation of the existing vector; magnitude is denoted as $\widehat{a_t}$ is expressed in Equation (22)

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\widehat{u_t} + \varepsilon}} \left(\beta_1 \widehat{a_t} + \frac{1-\beta_1}{1-\beta_1{}^t} d_t \right) \tag{22}$$

Figure 5 shows the pictorial representation of θ_{t+1}. For geographic abstraction and HSI identification, our neural network template has been used. In the publication, we found that remote sensing characteristics play an essential role in the characterization of HSI. Conversely, the numerical volume increases in both hyperspectral features. Moreover, in terms of process precision and measurement difficulty, researchers have noticed from many publications that the influence of extracted relationships alone is more substantial than critical specs. This report has addressed only geographical characteristics.

3.2.4. HSI identification deep CNN template

Until submitting to the primary analysis method, the HSI information collection is standardized (PCA). Instead, a PCA was used to minimize complexity and use the initial vital element with maximal variation as an image space to manipulate further. From the 2D picture, researchers retrieved the size scratches (26 to 26).

Figure 6 shows the workflow of the proposed DL-CNN. Researchers used three levels of matrix multiplication, two layers of merging, and one convolution layer. There were 38, 68 and 256 filters of initial, second, or third layers of modulation, which produced 38, 68, and 256 layouts, including both. Again when, and use these flaws as activation function and the central-pixel labels, the neural network classifier is constructed thru many optimizers as actual data. Finally, the model suggests the sensor packaging of the research results set and produces the required volume.

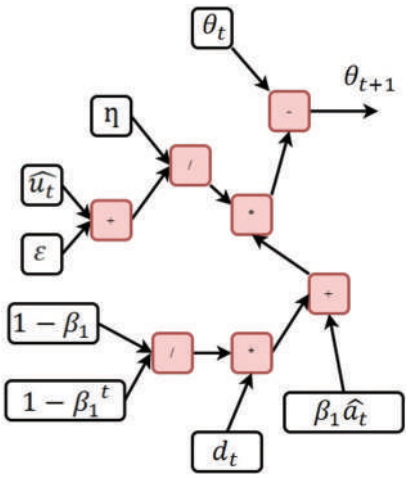

Figure 5: Pictorial representation of θ_{t+1}.

The full quantization is used in convolutional layers with a sliding window of 2–2 with phase 2. Finally, a normality test (LR) clustering algorithm employs soft-max induction in the convolutional layers. Moreover, L2 convolution and high school senior methodologies have been used to reduce dimensionality.

4. Software Analysis and Performance Comparison

The proposed DL-CNN is designed and implemented in this section. The different software performance parameters such as accuracy, loss, classification error, test, and validation analysis are done.

Figure 7a and b show the validation and test analysis of the proposed DL-CNN. The different simulation parameters such as linear SVM, quadratic SVM, linear Alexnet, and the quadratic Alexnet are considered. The proposed DL-CNN is analyzed, and the results are plotted in the above graph. The results show that the proposed DL-CNN has the highest performance in all the scenarios.

Figure 8 shows the remote sensed image classification and identification of the proposed DL-CNN. The first image shows the particular area's remotely sensed image, and the second image shows the zoomed portion of the small part of the first image. The last picture shows the RGB-converted image of the second image. From the third image, the fields and crops are analyzed with the diseases infected level from the size and the color of the crops and land. The proposed DL-CNN identifies well the crops.

Table 1 shows the confusion matrix analysis of the proposed DL-CNN. The confusion matrix is analyzed for the proposed DL-CNN. The different remote sensed images such as tomatoes, wheat, grass, water, and barley images are identified, and the result is tabulated in the above table. The results show that the proposed DL-CNN has the highest performance in all the scenarios. As the number of iterations increases, the accuracy increases.

Figure 9a and b show the classification percentage analysis of the proposed DL-CNN and the existing ANN method, respectively. The different types of images considered for the classification input include a binocular picture, watch image, motorbike images, and airplane images. The input images are analyzed and classified in both the proposed DL-CNN and the existing ANN method. The results show that the proposed DL-CNN highest accuracy in all the scenarios.

Table 2 shows the classification percentage analysis of the proposed DL-CNN. The different types of images considered for the classification input include a binocular picture, watch image, motorbike images, and airplane images. The input images are analyzed and classified in both the proposed DL-CNN and the existing ANN method. The simulation

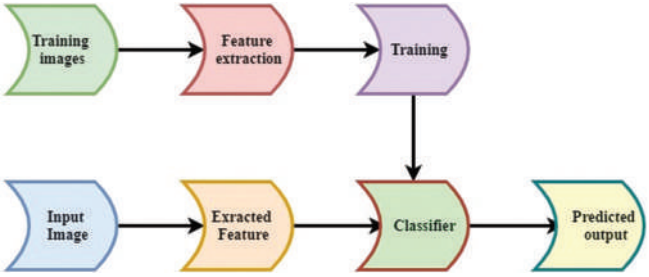

Figure 6: The workflow of the proposed deep learning convolutionary neural network system.

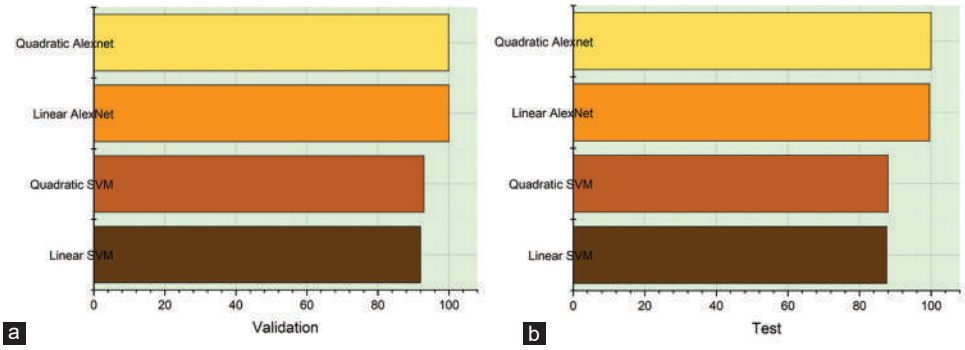

Figure 7: (a) Validation analysis of the proposed deep learning convolutionary neural network system. (b) Test analysis of the proposed deep learning convolutionary neural network system.

Table 1: Confusion matrix analysis of the proposed deep learning convolutionary neural network system

Ground item	Tomatoes	Trees	Wheat	Grass	Water	Barley
Tomatoes	1092	0	1	0	1	0
Trees	0	2475	0	3	2	2
Wheat	2	1	2854	2	0	4
Grass	4	0	2	1242	1	0
Water	8	3	4	4	124	2
Barley	9	2	0	6	4	321

Table 2: Classification percentage analysis of the proposed deep learning convolutionary neural network system

Methods/iterations	Building	Road	Parking	Water	Shadow
10	0.72	0.82	0.81	0.94	0.77
20	0.82	0.73	0.86	0.97	0.79
30	0.86	0.69	0.78	0.96	0.72
40	0.84	0.71	0.76	0.97	0.67
50	0.89	0.68	0.68	0.94	0.57
60	0.84	0.64	0.72	0.96	0.82

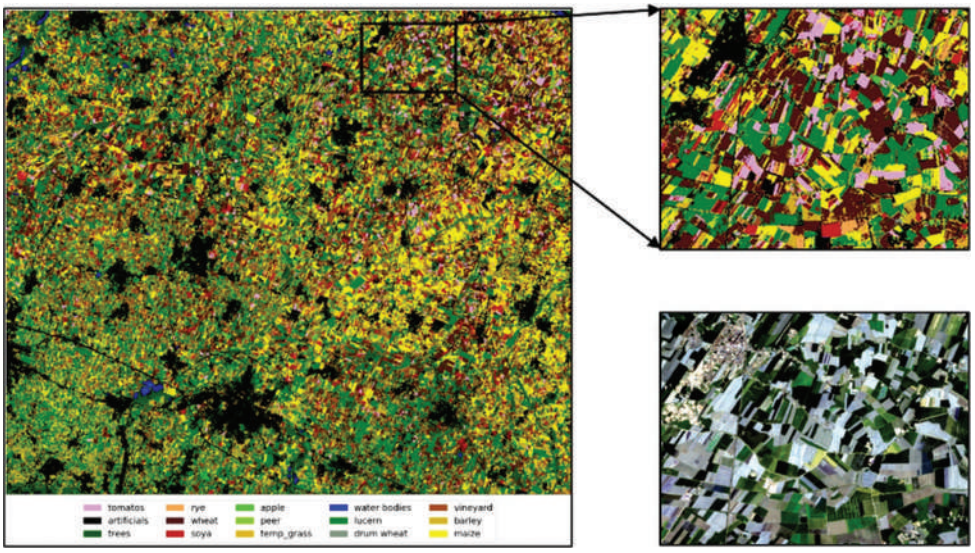

Figure 8: Remotely sensed image classification and identification of the proposed deep learning convolutionary neural network system.

iterations are varied from 10 to 60 with a step size of 10; as the number of iteration increases, the classification percentage increases. The results show that the proposed DL-CNN has the highest efficiency.

Figure 10a and b show the loss analysis of the existing ANN system and the proposed DL-CNN, respectively. The epochs are varied from 0 to 30 with a step size of 5 linearly. The respective loss percentage of the input CNN image is analyzed, and the result is plotted in the above graph. As the epochs increases, the loss percentage of the CNN image decreases. The results show that the proposed DL-CNN has the highest performance in all the scenarios.

The proposed DL-CNN is designed and implemented in this section. The different software performance parameters such as accuracy, loss, classification error, test, and validation analysis are done. The results show that the proposed DL-CNN has the highest performance in all the scenarios.

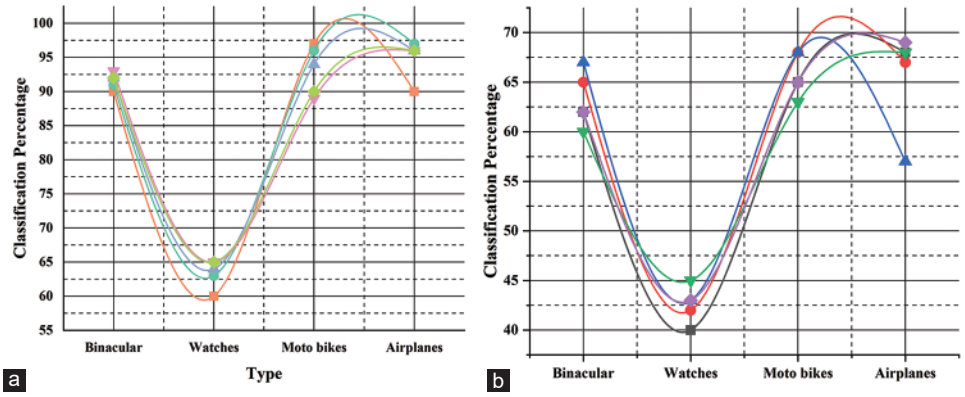

Figure 9: (a) Classification percentage analysis of the proposed deep learning convolutionary neural network system. (b) Classification percentage analysis of the existing ANN method.

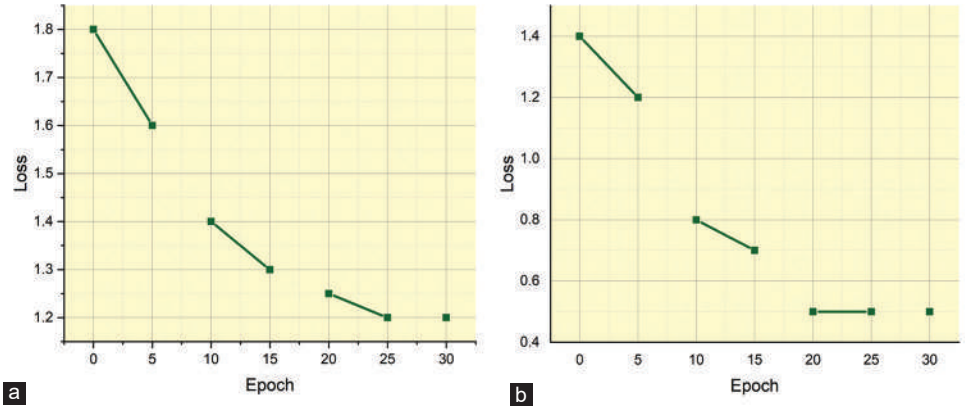

Figure 10: (a) Loss analysis of the existing ANN system. (b) Loss analysis of the proposed deep learning convolutionary neural network system.

5. Conclusion

The comparative study between the BoW and Mobilenet CNN function extraction was performed in this document. A DL-CNN is proposed in this research work. A pre-trained CNN was found to be able to retrieve stronger descriptors in comparison to BoW. Therefore, both the categories, that is, Sequential SVM and Quaternion SVM, can effectively evaluate the picture in the right order. There has been a comparable performance between the two classifications, and Wave equation SVM has been better overall than immunomodulatory every time. With the 20 times cross-standardized protocol, BoW cannot perform possibly the best. Its precision has fallen from 87% to 75.8% for QuadraticSVM and from 89.6% to 54.9% for LinearSVM. Precision has lessened by 0.1% in 12-fold to 20-fold bridge contexts and an increment of 0.2% for Linear SVM.

6. Funding

"This research received no external funding."

7. Conflicts of Interest

"The authors declare no conflicts of interest."

References

[1] D. L. Vu, T. K. Nguyen, T. V. Nguyen, T. N. Nguyen, F. Massacci, and P. H. Phung, "HIT4Mal: Hybrid Image Transformation for Malware Classification," *Transactions on Emerging Telecommunications Technologies*, vol. 31, no. 11, p. e3789, 2020.

[2] S. Bhattacharya, P. K. R. Maddikunta, Q. V. Pham, T. R. Gadekallu, C. L. Chowdhary, M. Alazab, and M. J. Piran, "Deep Learning and Medical Image Processing for Coronavirus (COVID-19) Pandemic: A Survey,"

Sustainable Cities and Society, vol. 65, p. 102589, 2021, doi: 10.1016/j.scs.2020.102589

[3] J. Khan, G. A. Khan, J. P. Li, M. F. AlAjmi, A. U. Haq, S. Khan, N. Ahmad, S. Parveen, M. Shahid, S. Ahmad, and M. Raji, "Secure Smart Healthcare Monitoring in Industrial Internet of Things (iiot) Ecosystem with Cosine Function Hybrid Chaotic Map Encryption," *Scientific Programming*, vol. 2022, p. 8853448, 2022, doi: 10.1155/2022/8853448

[4] P. Ramasamy, V. Ranganathan, S. Kadry, R. Damaševičius, and T. Blažauskas, "An Image Encryption Scheme Based on Block Scrambling, Modified Zigzag Transformation and Key Generation Using Enhanced Logistic-Tent Map," *Entropy*, vol. 21, no. 7, p. 656, 2019, doi: 10.3390/e21070656

[5] U. Maheshwari, and Silingam, K, "Multimodal Image Fusion in Biometric Authentication," *Fusion: Practice and Applications*, vol. 1, no. 2, p. 7991, 2020.

[6] D. Datta, P. K. Mallick, A. V. Reddy, M. A. Mohammed, M. M. Jaber, A. S. Alghawli, and M. A. Al-qaness, "A Hybrid Classification of Imbalanced Hyperspectral Images Using ADASYN and Enhanced Deep Subsampled Multi-grained Cascaded Forest," *Remote Sensing*, vol. 14, no. 19, p. 4853, 2022, doi: 10.3390/rs14194853

[7] Y. Liu, X. Chen, Z. Wang, Z. J. Wang, R. K. Ward, and X. Wang, "Deep Learning for Pixel-level Image Fusion: Recent Advances and Future Prospects," *Information Fusion*, 42, pp. 158-173, 2018, doi: 10.1016/j.inffus.2017.10.007

[8] Z. H. Arif, and K. Cengiz, "Severity Classification for COVID-19 Infections Based on Lasso-Logistic Regression Model," *International Journal of Mathematics, Statistics, and Computer Science*, vol. 1, pp. 25-32, 2022, doi: 10.59543/ijmscs.v1i.7715

[9] X. Xu, Y. Chen, J. Zhang, Y. Chen, P. Anandhan, and A. Manickam, "A Novel Approach for Scene Classification from Remote Sensing Images Using Deep Learning Methods," *European Journal of Remote Sensing*, vol. 54. pp. 1-13, 2020, doi: 10.1080/22797254.2020.1790995

[10] V. E. Sathishkumar, U. Moorthy, J. Park, C. Shin, and Y. Cho, "Internet Role in Remote Sensing and Geo Informatics System", *International Journal of Innovative Technology and Exploring Engineerin*g, vol. 9, pp. 57-64, 2019, doi: 10.35940/ijitee.A4859.129219

[11] D. Ezhilmaran, and M. Adhiyaman, "Fuzzy Approaches and Analysis in Image Processing", In: *Advanced Image Processing Techniques and Applications*", IGI Global, United States, pp. 1-31, 2017.

[12] D. Ezhilmaran, and M. Adhiyaman, "A Review Study on Fingerprint Image Enhancement Techniques", *International Journal of Computer Science and Engineering Technology*, vol. 5, pp. 2229-3345, 2014.

[13] M. Anbarasan, B. Muthu, C. B. Sivaparthipan, R. Sundarasekar, S. Kadry, S. Krishnamoorthy, and A. A. Dasel, "Detection of Flood Disaster System Based on IoT, Big Data and Convolutional Deep Neural Network," *Computer Communications*, vol. 150, pp. 150-157, 2020, doi: 10.1016/j.comcom.2019.11.022

[14] S. Bera, and V. K. Shrivastava, "Analysis of Various Optimizers on Deep Convolutional Neural Network Model in the Application of Hyperspectral Remote Sensing Image Classification," *International Journal of Remote Sensing*, 41(7), pp. 2664-2683, 2020, doi: 10.1080/01431161.2019.1694725

[15] M. Zhou, G. Fortino, W. Shen, J. Mitsugi, J. Jobin, and R. Bhattacharyya, (2016), "Guest Editorial Special Section on Advances and Applications of Internet of Things for Smart Automated Systems," in: *IEEE Transactions on Automation Science and Engineering*, vol. 13, no. 3, pp. 1225-1229, 2016, doi: 10.1109/TASE.2016.2579538

[16] Y. Liu, J. Nie, X. Li, S. H. Ahmed, W. Y. B. Lim, and C. Miao, "Federated Learning in the Sky: Aerial-Ground Air Quality Sensing Framework with UAV Swarms," in: *IEEE Internet of Things Journal*, vol. 8, pp. 9827-9837, 2020.

[17] Z. Shao, and J. Cai, "Remote Sensing Image Fusion with Deep Convolutional Neural Network," in: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 5, p. 1656-1669, 2018, doi: 10.1109/JSTARS.2018.2805923

[18] V. S. Martins, A. L. Kaleita, B. K. Gelder, H. L. da Silveira, and C. A. Abe, "Exploring Multi Scale Object-based Convolutional Neural Network (Multi-OCNN) for Remote Sensing Image Classification at High Spatial Resolution," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 168, pp. 56-73, 2020, doi: 10.1016/j.isprsjprs.2020.08.004

[19] B. Sujitha, V. S. Parvathy, E. L. Lydia, P. Rani, Z. Polkowski, and K. Shankar, "Optimal Deep Learning Based Image Compression Technique for Data Transmission on Industrial Internet of Things Applications," *Transactions on Emerging Telecommunications Technologies*, vol. 32, p. e3976, 2020, doi: 10.1002/ett.3976

[20] F. Ye, Y. Su, H. Xiao, X. Zhao, and W. Min, "Remote Sensing Image Registration Using Convolutional Neural Network Features", in: *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 2, pp. 232-236,

2018, doi: 10.1109/LGRS.2017.2781741

[21] B. Petrovska, E. Zdravevski, P. Lameski, R. Corizzo, I. Štajduhar, and J. Lerga, "Deep Learning for Feature Extraction in Remote Sensing: A Case-Study of Aerial Scene Classification," *Sensors* (*Basel*), vol. 20, no. 14, pp. 3906, 2020, doi: 10.3390/s20143906

[22] Q. Liu, X. Xiang, Y. Wang, Z. Luo, and F. Fang. "Aircraft Detection in Remote Sensing Image Based on Corner Clustering and Deep Learning," *Engineering Applications of Artificial Intelligence*, vol. 87, p. 103333, 2020, doi: 10.1016/j.engappai.2019.103333

[23] W. Han, R. Feng, L. Wang, and Y. Cheng, "A Semi-supervised Generative Framework with Deep Learning Features for High-resolution Remote Sensing Image Scene Classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 145, pp. 23-43, 2018, doi: 10.1016/j.isprsjprs.2017.11.004

[24] Z. Y. Algamal, M. R. Abonazel, and A. F. Lukman, "Modified Jackknife Ridge Estimator for Beta Regression Model with Application to Chemical Data", *International Journal of Mathematics, Statistics, and Computer Science*, vol. 1, pp. 15-24, 2023, doi: 10.59543/ijmscs.v1i.7713

[25] M. M. Jaber, M. H. Ali, S. K. Abd, M. M. Jassim, A. Alkhayyat, R. S. Bader, and A. R. Alkhuwaylidee, "Application of Image Encryption Based Improved Chaotic Sequence Complexity Algorithm in the Area of Ubiquitous Wireless Technologies", *Wireless Networks*, vol. 2022, pp. 1-14, 2022, doi: 10.1007/s11276-022-03162-y

[26] M. H. Ali, M. M. Jaber, S. K. Abd, A. Alkhayyat, and H. A. Jameel, "Model for Wireless Image Correlation Assisted by Sensors Based on 3D Display Technology," *Optik*, vol. 268, p. 169794, 2022, doi: 10.1016/j.ijleo.2022.169794