Machine Vision Based Object Detection Using Deep Learning Techniques

1st Garapati..Deva ram ganesh 2nd P.Vidyullatha 3rd Maddipati.Ravi krishna

Computer Science and Engineering Computer Science and Engineering Computer Science and Engineering

KLEF, Vaddeswaram, Greenfields, KLEF, Vaddeswaram, Greenfields, KLEF, Vaddeswaram, Greenfields,

Guntur, AP, India. Guntur, AP, India Guntur, AP, India.

[devaramgarapati@gmail.com](mailto:devaramgarapati@gmail.com) [latha22pellakuri@gmail.com](mailto:latha22pellakuri@gmail.com) [ravikrishna7755@gmail.com](mailto:ravikrishna7755@gmail.com)

4th S.Thanooj prapulla 5thA. Pavan Saran 6th Puppala Ramya

Computer Science and Engineering Computer Science and Engineering Computer Science and Engineering

KLEF, Vaddeswaram, Greenfields, KLEF, Vaddeswaram, Greenfields, KLEF, Vaddeswaram, Greenfields

Guntur, AP, India. Guntur, AP, India. Guntur, AP, India.

[sanapabittu2002@gmail.com](mailto:sanapabittu2002@gmail.com) [pavansaran2002@gmail.com](mailto:pavansaran2002@gmail.com) mothy274@kluniversity.in

*Abstract*— **The identification of items on the surface of the earth is widely known to be possible using hyperspectral images. To do classification and identify the various items on the image, the majority of classifiers just take into account spectral information. In this study, a neural network convolutional is used to classify the hyperspectral picture based on spectral and spatial properties (CNN). There are only a few areas in the hyperspectral picture. The multilayer perceptron aids in the categorization of visual characteristics into many classes while CNN builds the upper categorical level of strategic spectral and spatial aspects in each of the patch. The patch size of 13× 13 is found to be sufficient to attain the best accuracy. Compared to other classifiers, CNN requires greater computing time for training and testing. In comparison to other classifiers, simulation findings indicate that CNN stores the hyperspectral picture with the best classification accuracy.**

Keywords— *Hyperspectral image, CNN model, Spatial data, Pattern Recognition, ReLu.*

I**. INTRODUCTION**

A hyperspectral image capture sensor with improved spectral as well as spatial resolution has been created using recent advancements in optics and photonics. It is effectively used to be identify the objects and recognize the things on the earth's surface to use the spectral and spatialized information [1,2]. The spectrally fingerprints are modelled such that it will allow them to distinguish between the various objectified images. It is reasonable to think of the categorization of picture pixels according to their spectral properties as a task that identifies various, Based on their reflectance, materials, items, and surface. ground cover classifications qualities [3,4]. Applications for classification of hyperspectral imaging include astronomy, surveillance, biomedical imaging, agriculture, environmental science, and environmental science. The identification of the hyperspectral pictures, however, has its own distinct, the limited sample count that have been labelled and the Large spatialized diversity of spectral fingerprints are also factors [5,6]. A large portion of current research aim is on hyperspectral type of classified data that uses the conventional pattern of recognition approach, which entails 2 distinct steps: Second, employ k-nearest neighbours, minimal distance, logistic regression, maximum likelihood, parallelepiped classification, support vector machines (SVM), neural networks (NN), and to learn. The detailed kind of handmade dimensions are retrieved from the raw data input.. Most of the methods listed above are subject to the "curse of dimensionality". It has been recommended to maintain the drawbacks of dimensionality and the small number of hyperspectral data training samples available when utilising dimensionality reduction-based classification methods [7,8]. Other techniques for dealing with dimensionality include band selection and transformation [9,10]. Numerous problems have been solved using statistical learning methods in general. Support vector machine is one of the top classification method for hyperspectral type of data that is presented [11,12]. SVM is a poor learner due to its resistance to the Hughes effect and sensitive to the curse of dimensionality. In certain situations, in terms of“ classification accuracy, SVM-based classifiers outperform other frequently employed pattern recognition techniques [13]. For a very longest period, these classifiers were state-of-the-art technical devices [14]. The previous years, the classification of hyperspectral data has greatly benefited using spatial information. Spatial-spectral classification techniques provide significant efficiency advantages [15]. To address the geographical heterogeneity of spectral signatures, several innovative techniques try to take spatial data into consideration. Using SVM as well as a transmission map filter, presents a method for categorizing hyperspectral pictures. The spatial features are included into the Classification model using the guided image filter [16]. In order to provide spatial characteristics to the SVM classifier, edge preserving filters like the bidirectional filter and sparse representation filter are incorporated [17]. Given the wide variety of represented materials, it is crucial to understand which traits are most important for the classification algorithms. For classification applications, multiple models for deep learning have been created [18]. These models are developed using a range of feature levels. To build the high-level features required to train the model, low-level features are employed. Compared to any convolution image face detector, these methods streamline the feature extraction for every challenge. In addition, larger datasets and pictures having higher spectral and spatial resolutions likely to match and solve the classification problem more successfully for deep learning systems. Deep learning techniques have previously demonstrated promising results for categorizing and recognizing genuine objects, like dude objects or vehicles. Used a deep learning method to classify hyperspectral data correctly in more detail. To be more specific, the auto encoders aid in the development of the deep architecture, which collects highest level spatialized information from the picture in a hierarchical manner in order to classify each pixel. In a further stage, spectral properties were combined with spatially dominant data and provided as input to a logistic regression classifier. In a similar vein, we offer a deep education process for the categorization of spectral images into several classifications. Our approach, however a coherent kind of building which unifies spectral and spatial data in a single step while simultaneously producing high-level spectral-spatial properties. In particular, we recommend the employment of a Multi-Layer Perceptron, that helps to utilized for the categorization the images and the M-CNN, or High-level feature generation is carried out through a modified convolutional neural network. Due to the presence of the feed forward networks in CNNs & MLPs, the improved framework makes building spectral-spatial features while making instantaneous predictions of different classes in the picture is possible with this kind of architecture.

**II. METHEDOLOGY**

Diagram

Description automatically generated with medium confidence The most popular deep learning approach for image categorization and identification tasks is the convolution neural network (CNN). In addition to these uses, CNN is frequently used to categorize objects and identify human faces. In the learning phase of the CNN algorithm, features are extracted from the input picture and used to categorize it into multiple groups or categories. The resolution of the input picture will affect how the programmed perceives it as an array of a matrix. About the grayscale and RGB images, CNN considers the matrix. Bypassing input picture series through several convolution layers employing pooling, fully linked layers, and ultimately the soft-max function to identify the objects are the layers that use using kernels or filters to extract characteristics. probabilistically, one may theoretically develop a CNN model. *Convolution Layer-* A convolution layer is used to extract the features present in the input pictures*.* During learning phase, a tiny square of input data is taken, preserving the connection between the image's pixels. A kernel or filter with an image matrix are multiplied mathematically to produce the convolution. Figure 2 depicts the image's convolution process. Consider a 5 x 5 matrix with the values 0 and 1 for an image pixel as well as a 3 x 3 filter matrix or kernel. The image matrix is multiplied by a filter matrix in the convolution layer to create the convolved feature or feature map. In Figures 3 and 4, the same convolved procedures are depicted. In most situations, the chosen kernel or filter is not suitable for the type of picture input. The situations for the above has followed the options used to balance or match the input to visuals. To accommodate the input pictures, zero paddings are applied to the input image. Reduce the image's pixel count to fit the supplied pictures. ReLU, or the Rectified Linear Unit, operates on nonlinear data which is selected for Max(0,input) of the convolution network computation the results of ReLU include . RELU, or Rectified Linear Unit carries out non-linear operations. For each feature value that has been convolved, the average pooling selects the mean value. The addition of each component The value of the convolved feature is chosen using the sum pooling. *Fully Connected Layer (FCL)*. The pooling layer's two-dimensional matrix is changed into a one-dimensional matrix. A totally connected layer receives this single layer vector as input, much like a neural network. Figure 1 depicts how a completely linked layer functions. A vector representation of the output feature matrix is created Learning the goal of introducing nonlinearity to the convolution network is to produce non-negative linear values to ReLU. The CNN ReLU procedure is depicted in image. Non-linear functions that can be substituted with ReLU are Tanh and sigmoid. However, ReLU performs better than a pair of non-linear processes. *The convolution layer's output pixels are in the pooling layer* are decreased by the pooling layer, which also lessens the complexity of huge pictures. By keeping most of the image's information, the spatial pooling, sometimes referred to as pixel down sampling or subsampling as. reduces the image's spatial dimension. These characteristics are used to build a model having a layer that is fully linked, and the outputs are then classified using the soft-max or sigmoid function. The implementation process for CNN-based hyperspectral image categorization. The definition of a hyperspectral picture is a 3D the spectral channel is indicated by the image's height and breadth, in this size matrix. The many channels of hyperspectral picture enhance the training and prediction systems' computing capacity and memory resources. process. Yet, statistical research reveals that the spectral response variance for each type of pixel is minimal. Consequently, similar spectral values are shared by pixels with the same class labels. for each channel, but the spectral values of pixels with various class designations are distinct. Based on these characteristics, The input data's dimensionality might be reduced using a dimensionality reduction strategy. and improve the procedures for categorization and training. The decrease in dimensions approach known as PCA decreases the spectral dimensions of the hyperspectral picture without sacrificing any image data. Split the hyperspectral picture into tiny patches after dimensionality reduction to make it compatible with CNN's fundamental structure

Figure 1: Fully Connected Layer using CNN model

A single pixel's spectral and spatial properties are included in each produced patch. Square patch, more specifically. A family of Convolutional neural networks (CNN/ConvNet), also known as deep neural networks, are commonly used to interpret visual vision in deep learning. ConvNet does not use matrix multiplications, contrary to what we typically think of when we think about neural networks. Convolution is a unique method that is employed. A mathematical procedure called convolution on two type of functions which uses third activation function that expresses how the form of one is changed by the other. Around the 1980s, CNNs were first created and put to use. At the time, a CNN could only identify handwritten numbers to a certain extent. To read zip codes, pin numbers, etc., it was mostly utilized in the postal industry. The most crucial thing to keep in mind about any deep learning model is that it needs a lot of computational power and data to train. The dimensionality reduction approach known as PCA decreases the spectral dimensions of the hyperspectral picture without sacrificing any image data. Split the hyperspectral picture into tiny patches after dimensionality reduction to make it A picture containing shape

Description automatically generatedcompatible with CNN's fundamental structure.

Figure 2: Image Convolution Operation Using Filter Kernel

III. RESULTS AND DISCUSSIONS

The proposed method is imitate utilizing the hyperspectral picture. The Indian Pines dataset, a hyperspectral picture taken using AVIRIS sensors at a test location in northwest Indiana. The following criteria are used to assess how well the present approach performs: overall training components of the accuracy, test data loss, precision, recall, and f1-score. The spectral bands in the simulation are narrowed using PCA, and the classification is done using 30 principal components. Each patch includes dimensions after dimensionality reduction. The patch size is set to 5 during the simulation to take each pixel's 24 closest neighbors into account. The CNN architecture receives each patch for categorization purposes. The 5 × 5 patch's classification accuracy was 84.12%, and the simulation took 1142.26 seconds. When the picture is 145 x 145 pixels in size, and there are 220 spectral bands and a spectral range is form (0.4-2.5) µm, and the number of classes is 10. When the patch size is 5 x 5 then overall accuracy (%) is 8.12 and computational time/sec is 1142.2. When the patch size is 7 x 7 then overall accuracy (%) is 91.10 and computational time/sec is 2111.512. When the patch size is 9 x 9 then overall accuracy (%) is 95.00 and computational time/sec is 3047.524. when the patch size is 11 x 11 then overall accuracy (%) is 97.42 and computational time/sec is 4263.925. When the patch size is 13 x 13 then overall accuracy is 98.28 and computational time/sec is 5002.114. When the model is PCA then Overall Accuracy is 88.1 and Computational Time (sec) is 568.1.when the model is LCA then Overall Accuracy(%) is 86.53 and Computational Time(sec) is 423.58. when the model is CNN then Overall Accuracy(%) is 98.28 and Computational Time(sec) is 5002.114. The simulation duration goes up when the patch size is increased, improving classification accuracy. displays the outcomes of CNN classification using various patch sizes. Figure 3 provides information on the CNN classification's accuracy and calculation time. The table shows that CNN performs best for the patch size 13 × 13 in terms of classification accuracy and calculation time. No new developments are noted in When the patch size exceeds 13, the classifier's accuracy actually degrades and the amount of computing work required rises. In support vector machine classifier is contrasted with the suggested CNN method for the classification of the hyperspectral picture. Support vector machines (SVMs) employ spectral information for classification, and guided image filters and bilateral filters add spatial features to the output of the SVMs. PCA type and LDA’s are also used to minimize the dimensionality of the spectral features together with SVM. The SVM algorithm, when used in conjunction with bilateral and directed image filters, improves classification accuracy while shortening calculation time. Without employing bilateral and directed image filters, the CNN algorithm simultaneously incorporates spectral and spatial characteristics. In comparison to SVM classification accuracy, CNN method archives the maximum classification accuracy with rising calculation time. In comparison to the CNN method, the SVM algorithm requires less processing time.

Shape

Description automatically generated with low confidence

Figure 3: 5 by 5 & 3 by 3 Matrix of striding

Table

Description automatically generated

Figure 4: Max Pooling Operations 4 by 4 image

A picture containing table

Description automatically generated

A picture containing table

Description automatically generated

Figure 5: Average Pooling Operations 4 by 4 image

Table

Description automatically generatedFigure:6 Transfer ReLU Operation in CNN.

Figure:7 Performance parameters for deep learning techniques.

**IV CONCLUSION**

The hyper-spatialized images are classified using a convolutional neural networks.. The classification of the picture takes into account both spectral and spatial information. The includes spatial components categorization to enhance the categorization. Among support vector machine classifier and other approaches, CNN records the greatest classification training accuracy of 95.28% and testing accuracy of 96.2% where the patch size taken into consideration for the categorization will affect CNN's accuracy. The number of spatial features taken into account for categorization is indicated by patch size. The patch size of 13× 13 is found to be sufficient to attain the best accuracy. Compared to other classifiers, CNN requires greater computing time for training and testing. Edge preserving filters are not used in the proposed technique to include spatial information in the classification.

**REFERENCES:**

1. Deep Learning for Hyperspectral Image Classification: An Overview, S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson, IEEE Transactions on Geoscience and Remote Sensing, vol. 57(9), pp.. 66906709, 2019.
2. [Sagar Imambi, S.](https://www.scopus.com/authid/detail.uri?authorId=57202029290), [Mandhala, V.N.](https://www.scopus.com/authid/detail.uri?authorId=56195835600), [Naaz, M.A.](https://www.scopus.com/authid/detail.uri?authorId=57216590948)” [The Amalgamation of Machine Learning and LSTM Techniques for Pharmacovigilance](https://www.scopus.com/record/display.uri?eid=2-s2.0-85084003511&origin=resultslist&sort=plf-f)”, [Lecture Notes in Networks and Systems](https://www.scopus.com/sourceid/21100901469?origin=resultslist), 2020, 105, pp. 123–132.
3. B. Tu, J. Wang, X. Kang, G. Zhang, X. Ou and L. Guo, "KNN-Based Representation of Superpixels for Hyperspectral Image Classification," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 11, pp. 4032-4047, Nov. 2018.

1. IEEE China Summit & International Conference on Signal and Information Processing, Xi'an, 2014, pp. W. Li, S. Prasad, E. W. Tramel, J. E. Fowler, and Q. Du, "Decision fusion for hyperspectral image classification based on minimumdistance classifiers in thewavelet domain,". 162165.
2. S. Bajpai, H. V. Singh and N. R. Kidwai, "Feature extraction and classification of hyperspectral images using singular spectrum analysis and multinomial logistic regression classifiers," 2017 International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT), Aligarh, 2017, pp. 97-100. Using principal component analysis and deep learning, Q. Sun, X. Liu, and M. Fu classified a hyperspectral picture, International Conference on Electronics Information and Emergency Communication, 2017, pp. 356-359.
3. [Jammalamadaka, K.](https://www.scopus.com/authid/detail.uri?authorId=57205031541), [Parveen, N.](https://www.scopus.com/authid/detail.uri?authorId=57219789147) “[Testing coverage criteria for optimized deep belief network with search and rescue](https://www.scopus.com/record/display.uri?eid=2-s2.0-85104725442&origin=resultslist&sort=plf-f)”, Journal of Big Data, 2021, 8(1), 59.
4. S. Zhong, C. Chang and Y. Zhang, "Iterative Support Vector Machine for Hyperspectral Image Classification," International Conference on Image Processing, 2018, pp. 3309-3312.
5. Few-Shot Hyperspectral Image Classification With Unknown Classes Using Multitask Deep Learning, IEEE on Geoscience and Remote Sensing, 2020; S.Liu, Q. Shi, and L. Zhang.
6. Shambulinga M and Sadashivappa G., "Hyperspectral Image Classification Support Vector Machine with Guided Image Filter" International Journal of Advanced Computer Science and Applications, vol. 10(10), 2019.
7. Semisupervised Hyperspectral Image Classification Using Deep Features, M. S. Aydemir and G. Bilgin, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 12(9), pp. 3615–3622.
8. "Supervised Hyperspectral Image Classification Using SVM and Linear Discriminant Analysis" by M. Shambulinga and G. Sadashivappa was published in the International Journal of Advanced Computer Science and Applications on October 10, 2020..
9. I.Bidari, S.Chickerur, H.Ranmale, S.Talawar, H.Ramadurg and R.Talikoti, "Hyperspectral Imagery Classification Using Deep Learning," Conference on Smart Trends in Systems, Security and Sustainability, 2020, pp. 672-676.
10. [Gayatri, N.](https://www.scopus.com/authid/detail.uri?authorId=35292092500), [Vamsi, B.](https://www.scopus.com/authid/detail.uri?authorId=58019037100), [Vidyullatha, P.](https://www.scopus.com/authid/detail.uri?authorId=56586053000)” [Deep Learning LSTM Approach on Hyperspectral Images using Keras Framework](https://www.scopus.com/record/display.uri?eid=2-s2.0-85130093462&origin=resultslist&sort=plf-f)”, International Conference on Sustainable Computing and Data Communication Systems, ICSCDS 2022 - Proceedings, 2022, pp. 1722–1726.
11. [Jadhav, A.D.](https://www.scopus.com/authid/detail.uri?authorId=57218176005), [Pellakuri, V.](https://www.scopus.com/authid/detail.uri?authorId=56586053000), ”[Highly accurate and efficient two phase-intrusion detection system (TP-IDS) using distributed processing of HADOOP and machine learning techniques](https://www.scopus.com/record/display.uri?eid=2-s2.0-85117300922&origin=resultslist&sort=plf-f)”, Journal of Big Data, 2021, 8(1), 131.
12. [Jammalamadaka, K.](https://www.scopus.com/authid/detail.uri?authorId=57205031541), [Parveen, N.](https://www.scopus.com/authid/detail.uri?authorId=57219789147), "[Equivalent mutant identification using hybrid wavelet convolutional rain optimization](https://www.scopus.com/record/display.uri?eid=2-s2.0-85116261878&origin=resultslist&sort=plf-f)”, [Software - Practice and Experience](https://www.scopus.com/sourceid/20007?origin=resultslist), 2022, 52(2), pp. 576–593.
13. [Prassanna, P.L.](https://www.scopus.com/authid/detail.uri?authorId=56461364700), [Sandeep, S.](https://www.scopus.com/authid/detail.uri?authorId=57220899666), [Rao, K.](https://www.scopus.com/authid/detail.uri?authorId=57218179592), ...[Mounika, P.](https://www.scopus.com/authid/detail.uri?authorId=56285125000), [Govardhani, U.](https://www.scopus.com/authid/detail.uri?authorId=57226566724) “[Image Classification Based on Convolutional Neural Network](https://www.scopus.com/record/display.uri?eid=2-s2.0-85123348662&origin=resultslist&sort=plf-f)”, [Lecture Notes on Data Engineering and Communications Technologies](https://www.scopus.com/sourceid/21100975545?origin=resultslist), 2022, 93, pp. 833–842.
14. [Vineethmohan, M.](https://www.scopus.com/authid/detail.uri?authorId=57217085571), [Hemanth, P.](https://www.scopus.com/authid/detail.uri?authorId=57219614066), [Mounica, M.](https://www.scopus.com/authid/detail.uri?authorId=57217084826), [Lakshmi Prasanna, P.](https://www.scopus.com/authid/detail.uri?authorId=56461364700) ”[Image classification using deep learning](https://www.scopus.com/record/display.uri?eid=2-s2.0-85086110999&origin=resultslist&sort=plf-f)”, Journal of Advanced Research in Dynamical and Control Systems, 2020, 12(6), pp. 1–10.
15. [Sakhare, N.N.](https://www.scopus.com/authid/detail.uri?authorId=56857608200), [Sagar Imambi, S.](https://www.scopus.com/authid/detail.uri?authorId=57202029290) “[Technical Analysis Based Prediction of Stock Market Trading Strategies Using Deep Learning and Machine Learning Algorithms](https://www.scopus.com/record/display.uri?eid=2-s2.0-85142122988&origin=resultslist&sort=plf-f)”, International Journal of Intelligent Systems and Applications in Engineering, 2022, 10(3),