

Semantic Annotation of Land Cover Remote Sensing Images Using Fuzzy CNN

K. Saranya^{1,*} and K. Selva Bhuvaneshwari²

¹Department of Electronics and Communication Engineering, University College of Engineering Kanchipuram, Kancheepuram, 631552, India

²Department of Computer Science and Engineering, University College of Engineering Kanchipuram, Kancheepuram, 631552, India

*Corresponding Author: K. Saranya. Email: yasaran84@gmail.com

Received: 29 August 2021; Accepted: 16 November 2021

Abstract: This paper presents a novel fuzzy logic based Convolution Neural Network intelligent classifier for accurate image classification. The proposed approach employs a semantic class label model that classifies the input land cover images into a set of semantic categories and classes depending on the content. The intelligent feature selection algorithm selects the prominent attributes from the given data set using weighted attribute functions and uses fuzzy logic to build the rules based on the membership values. To annotate remote sensing images, the CNN method effectively creates semantics and categorises images. The decision manager then integrates the fuzzy logic rules with the CNN algorithm to achieve accurate classification. The proposed approach achieves a classification accuracy of 90.46% when used with various training and test images, and the three class labels for vegetation (84%), buildings (90%), and roads (90%) provide a higher classification accuracy than other existing algorithms. On the basis of true positive rate, false positive rate, and accuracy of picture classification, the suggested approach outperforms the existing methods.

Keywords: Land cover; high resolution; annotation; CNN; fuzzy logic

1 Introduction

Remote sensing is a technique for monitoring and detecting the earth's surface area without physical contact, utilizing specialized sensing devices such as high-resolution cameras and satellite images [1]. The benefits of remote sensing are that it enables remote monitoring of an unattended environment on a continual dynamic basis [2]. It also offers information about changes that occur in the environment, which is beneficial. With the advancement of remote sensing technology, research on remote sensing has shifted its focus to precise image classification through the application of image processing and machine learning algorithms [3]. Image annotation is a popular machine learning technique that makes use of artificial intelligence to annotate images depending on their context. The majority of existing image annotation systems classify the image using a single class label in order to provide an overall understanding of the image. Unfortunately, image annotation based on a single class label gives inadequate information to categorise and annotate images more accurately in nature. Segmentation-based semantics is a widely used technique that classifies objects in a scene based on their image pixels in order



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

to extract information [4]. Semantic image segmentation based on a single class label, on the other hand, is a difficult and time-consuming process. Additionally, this procedure incurs computational overhead, which might deplete system resources, and the annotation of the image classification is imprecise by nature [5]. To address the limitation of a single class label classifier, a classifier based on multiple class labels has been developed to offer adequate information about the scene to annotate multi label and categorize the image with greater accuracy [6]. When compared to single-label remote sensing image classification, multi-label remote sensing image classification is a more realistic challenge. The purpose of multi-label annotation is to predict numerous semantic labels that will be used to characterize a remote sensing image scene. Due to its higher descriptive capacity, multi label may be used in numerous disciplines, like image annotation [7,8] and image retrieval [9–11]. However, the multi-class labels demonstrated limited performance as the image annotations are classified using handcraft features from the given images and the images in high-level semantics are not displayed to ensure a precise classification and annotation of the images in a more accurate way. To overcome the limitations of existing systems, this study proposes an efficient fuzzy logic-based CNN Intelligent Semantic Multi-label annotation technique for more precisely classifying and annotating Land-cover high-resolution remote sensing images. The proposed intelligent classifier makes use of a semantic multi-label model in which the image is represented using high-level semantics. Further, the presented intelligent classifier classifies the image into a series of semantic categories, each with a unique set of classes based on the content of the remote sensing image. Furthermore, it applies intelligent feature selection, in which the prominent characteristics are chosen based on the weightage attribute functions utilizing the information gain ratio.

2 State of the Art

In the realm of remote sensing, it is important to annotate scene images with multiple labels in order to comprehend the images [12,13]. Qi et al. (2020) constructed a multi-label high spatial resolution dataset to understand well about semantic scene images with deep learning approach from the overhead perspective. Their suggested method enables the classification and retrieval of multi-label images via deep learning. This strategy outperformed previous methods in multi-label image classification and retrieval tasks. To evaluate the performance of image classification, investigators employed mean average precision, average F1 score and precision at number of retrieved images, as well as average normalised modified retrieval rank, mean average precision, and precision at number of retrieved images. Zhu et al. (2020) proposed a deep learning framework for multi-label annotation of remote sensing images. One of the primary features of this system is the use of convolutional neural networks to learn features from dual-level semantic ideas. One problem of this approach is that it neglects to include the dependence of the label at the object level and label relationships between the scene level and the object level. Vanegas et al. (2019) presented a kernel matrix factorization based semi-supervised online learning approach for automatic multi-label annotation. The proposed method worked with large datasets, which addresses one of the primary shortcomings of kernel-based methods, namely their inability to scale. Also, this method is ideal for non-linear complicated relationships and significantly reduces the amount of memory and calculation time required for multi-label annotation tasks.

Hu et al. (2013) proposed a multi-level max-margin discriminative analysis for the annotation of high-resolution images. To create discriminative features, the algorithm use the maximum entropy discrimination latent Dirichlet Allocation technique. It utilises the bag-of-words to incorporate both word-level and topic-level elements in order to increase annotation performance in multi-level semantics and contextual information. Jeppesen et al. (2019) introduced a remote sensing network which is a deep learning model to detection cloud free images in optical satellite imagery. This model was trained and evaluated using Landsat 8 Biome and SPARCS dataset over biomes with cloud over snowy and icy region images. Further, the model treated the noisy data and increases performance over cloud masking method.

Kadhim et al. (2019) proposed an effective method for deep learning and CNN based satellite image classification technique for feature extraction. Four effective ways for improving the performance of satellite image classification were presented. Cao et al. (2020) presented an automatic image annotation technique based on CNN with threshold optimization to address the problem of over- or under-labeling in multi-label image annotation. Hoxha et al. (2020) proposed a remote sensing image retrieval system capable of generating and utilizing textual descriptions that characterise the relationship between objects and their associated attributes in remote sensing images. Xia et al. (2021) suggested a stacked ensemble method for improving the pairwise label correlation and weight learning processes. Additionally, they created an optimization approach to achieve an ideal ensemble solution that is both efficient and optimal. Markatopoulou et al. (2019) addressed deep convolutional neural network architecture that taking the problem of multi-label video/image annotation by exploiting multi-task learning to find the relation between targets and structured output learning to find the correlation between the concepts. Both models are built using standard layers that may be trained using back propagation to increase the accuracy of annotations. Wanga et al. (2019) experimented with an automatic image annotation technique based on a multiclass label selection algorithm. Using a convolutional neural network, this technique improves annotation performance. Alshehri (2020) discussed a technique for extracting image features using principal component analysis and the wavelet transform. Moreover, the author suggested a prediction technique based on neural networks for image classification of retrieved data. Jabari et al. (2013) proposed a classification method in high resolution urban satellite images using fuzzy logic. Fuzzy logic is used for satellite image to handle the main problem such as uncertainty in the position of object borders in high resolution image classification. Li et al. (2017) investigated a method for extracting visual attention features using a multi-scale procedure. Further, researchers created a fuzzy classification method for classifying high-resolution remote sensing scene images. This approach allows for an accurate classification rather than other measurements of quantitative accuracy. Gheshlaghi et al. (2017) proposed an analytical network process and fuzzy based decision making system for detecting landslides problems. Bharti and Kurmi (2017) described a novel approach for classifying high-resolution urban satellite images into three categories using fuzzy logic: road, building, and vegetation. Ma et al. (2017) discussed the use of remote sensing imagery to classify land cover images using an object-based approach. According to the literature review, the majority of existing image classification algorithms are ineffective at accurately detecting class labels and at semantically annotating the multiple class labels. Motivated by these findings, a unique intelligent classification technique is suggested in this work, which leverages intelligent fuzzy rules in conjunction with the CNN algorithm to categorise image class labels with more accuracy. According to high level semantics, proposed intelligent classifier combines CNN algorithm with convolutional layer, min-max pooling layer, and decision manager to efficiently classify images into different types of label classes. Finally, the decision manager decides on image annotation by integrating intelligently produced fuzzy rules and CNN classification.

The contributions of the proposed system are

1. The Multi label semantics where the images are annotated with single-class label and represent them in high level semantics.
2. The proposed model provides intelligent feature selection algorithm where the prominent features are selected from the class label.
3. The proposed system incorporates an intelligent classifier, which utilises intelligent fuzzy rules and a CNN classifier to appropriately annotate images using the retrieved feature set.

3 Proposed System Architecture

The proposed system's architecture is depicted in Fig. 1, and it is composed of eight modules: an image dataset module, a semantic analysis module, a class label classification module, an intelligent feature extraction module, a CNN classification module, a fuzzy rule generator module, a fuzzy inference module, a knowledge base module, and a decision manager module.

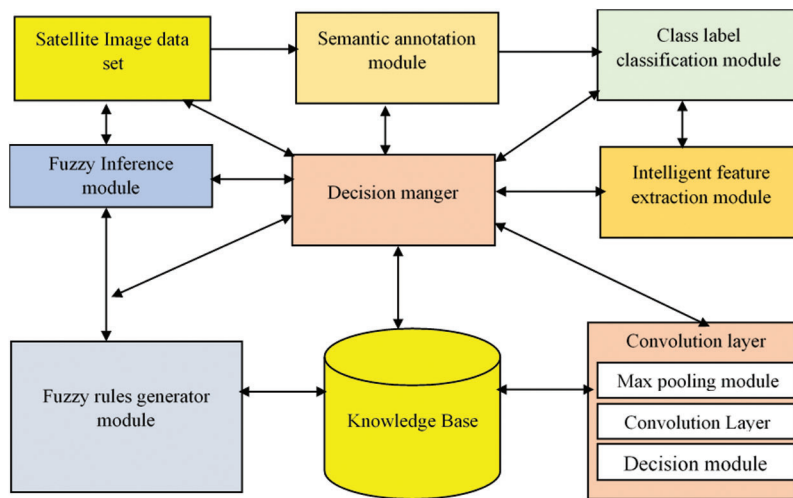


Figure 1: Architecture of annotation for land-cover remote sensing images

Image data set module is the first module of a proposed system. It allows use of the UC Merced data set, with 70% of the images being used for training and 30% for testing the proposed system. The second module is devoted to semantic analysis. The semantic analysis module's primary function is to provide a higher-level knowledge of the given scenery and to annotate the given image with various class labels [14,15]. The following module is for intelligent feature extraction. The main role of this module is to identify and extract the prominent features from the annotated class labels. The next module is CNN classification. This module is further divided into three submodules: max pooling, convolutional, and decision. The following module is fuzzy inference, which employs fuzzy rules. The accompanying module is the fuzzy rule generator. The fuzzy rule generator module is further divided into four submodules: fuzzification, rule creation, rule firing and matching, and rule execution. The succeeding module is the knowledge base, which stores the created fuzzy rules. Decision manager is the last module of the system proposed. The main role of the decision manger is to take the decisions and control and coordinate the other modules present in the system.

4 Proposed System

The semantic annotation phase is the initial phase of the proposed system. The data collection UC Merced is used as an input, and it contains high-resolution remote sensing images of land cover, and some sample images are shown in Fig. 2. The most important job of the semantic annotation phase is to carry out high-level thinking and assign different class labels to the image [16]. The proposed approach employs multiple label annotation to more precisely classify the images. Algorithm 1 details the algorithm for the semantic annotation module.

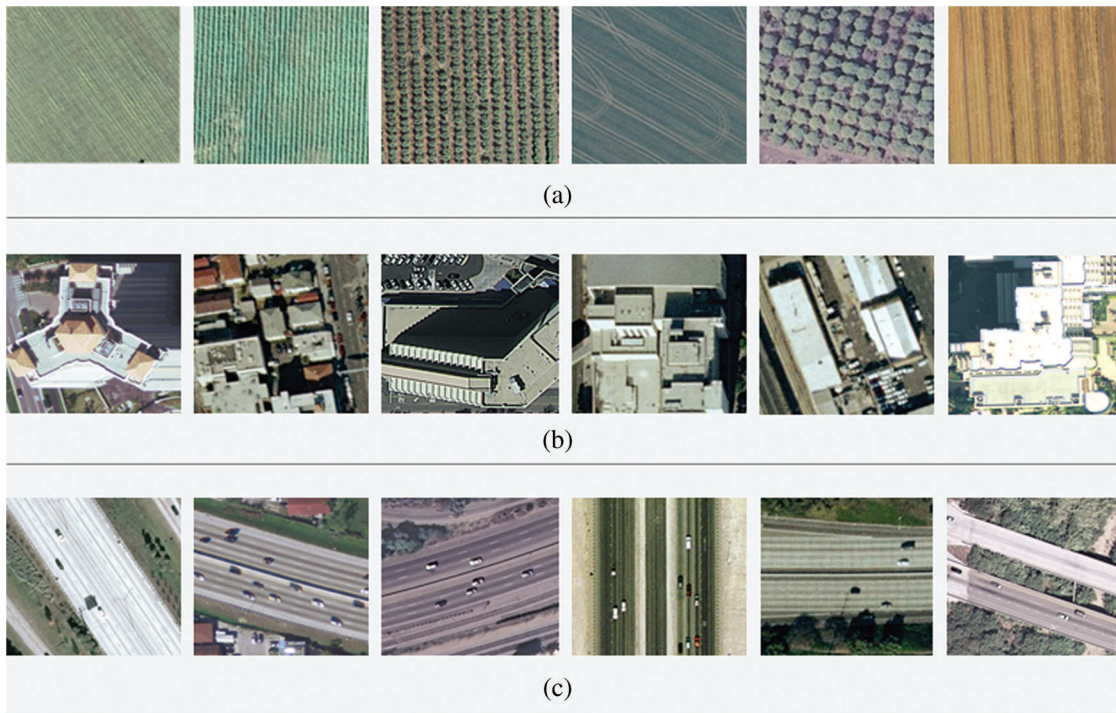


Figure 2: Sample images from the UC-Merced data set (a) Agriculture (or) Vegetation area; (b) Building; (c) Roads (or) Freeways

4.1 Initial Level Image Segmentation Phase

In this step, the images are segmented into homogenous non-overlapping discrete sections based on their attributes such as values of their grey pixel, texture and auxiliary data. The proposed systems segment the image using a multi-resolution segmentation technique. The proposed system takes into account three critical parameters: scale, shape, and compactness. Based on the results of initial segmentations the class label items such as shadows, vegetation area and roadways can be defined. However, for the accurate detection of buildings, the second level image segmentation is essential.

4.2 Intelligent Fuzzy Based Image Classification

In the fuzzy image classification, the segments are classified on the basis of the specific values defined in the membership functions instead of applying a decision based upon the binary values. The membership functions based on fuzzy logic have values ranging from 0 to 1. Where 0 indicates that the object is not a member of the class and 1 indicates that the object is a member of the class. A triangle membership function is used in the proposed system [17–20]. The approach utilizes three variables: low, medium, and high. The fuzzy inference system generates intelligent fuzzy rules based on the linguistic variables. The decision manager makes the decision based on the generated fuzzy rules. The suggested system tests all object classes by classifying each image segment using intelligent fuzzy rules. The parameters used for the object based image classification are explained as follows.

4.2.1 Segment Shadowing

In the object based image classification, the Segment Shadowing is used to identify the objects which are elevated. The proposed system uses two parameters namely segment brightness and segment density to determine the shadow of the image segments.

4.2.2 Brightness

The brightness is defined as amount of mean value of each pixel or segment present in all the bands of the image. The brightness of the segments of the image can be computed by using Eq. (1) [21]:

$$\text{Segment_brightness} = \text{red_colour} + \text{blue_colour} + \text{green_colour} + \text{NIR}/4 \quad (1)$$

In object based image classification, the shadow objects in the image segments contains low brightness values. Moreover, in order to improve the accuracy of the brightness the proposed system employs a fuzzy logic based K-means clustering approach to detect the cluster with darkest values. Finally, the mean and Standard Deviation (SD) of the darkest cluster is computed. To build intelligent fuzzy rules, the computed darkest cluster is used as a linguistic variable for the shadow brightness. The two parameters NIR ratio [21] and NDVI [21] are considered in the proposed approach to calculate the vegetation area for the given image segment. Eq. (2) contains the formula for calculating the NIR ratio and follows:

$$\text{NIRRatio} = \frac{\text{NIR}}{\text{NIR} + \text{R} + \text{G} + \text{B}} \quad (2)$$

where NIR ratio is Near Infrared imaging and NDVI is the difference vegetation index which is used to identify the vegetation area from the given image segment. Two important metrics, L_{cm} [21] and L_e [21], are taken into account while identifying the road from an image segment. The L_{cm} and L_e are computed by using Eqs. (3) and (4):

$$L_{CM} = \frac{\sqrt[2]{\pi \text{area}(\text{object})}}{\text{perimeter}(\text{object})} \quad (3)$$

$$L_e = \frac{\text{Area}(\text{object})}{\text{Length}(\text{object})^2} \quad (4)$$

The intelligent fuzzy rules can be used to identify the road class label from a given image segment based on these criteria. To determine the building class labels from a given image segment, three critical characteristics must be considered: the elliptical fit, the rectangular fit, and the shadow position within the given image segment. Rectangular fit is defined as the degree to which objects (buildings) fit within a rectangle. The number 0 indicates that the objects do not fit within the rectangle, whereas the value 1 shows that the objects do fit within the rectangle. The term “elliptical fit” refers to the degree to which objects fit into an elliptical structure. The value 0 indicates that the objects do not fit within the elliptical structure, whereas the value 1 shows that the things do fit within the elliptical framework. The shadow position is used to denote the locations of buildings in an image segment. The most frequently used position of the shadow for identifying buildings in a given image segment is on the southern or western side. Intelligent fuzzy rules are formed based on all of these computed parameters [22–24].

5 Semantic Annotation Module

The algorithm for semantic annotation module is explained in Algorithm 1. In this algorithm, the UC Merced image data set is taken as input and the output of this algorithm is Set of annotated class labels [25]. Initially, in this algorithm Image Set (IS) is defined as set of images ranging from IS_1 to IS_n . The next step is store the elements of the IS into an array. The Class Label (CL) set is defined as set of class labels ranging from CL_1 to CL_n and these elements are stored in an array. The next phase is image segmentation phase. The image is divided into equal non-overlapping segments S_1 to S_n in this step using the Image Set (IS).

Algorithm 1: Semantic annotation feature extraction module

Input: UC Merced image data set

Output: Set of annotated class labels

- (1) Define image set $IS = \{IS_1, IS_2, IS_3 \dots IS_n\}$
 - (2) Store the IS into an array, i.e $IS = [IS_1- IS_n]$
 - (3) Define class labels set $CL = \{CL_1, CL_2, CL_3 \dots CL_n\}$
 - (4) Store the CL into an array i.e., $CL = [CL_1- CL_n]$
 - (5) For each IS do the following
 - Partition $P(IS) \rightarrow \{S_1, S_2, S_3 \dots S_n\}$ Where S_1, S_2, S_n are the non-overlapping segments IS
 - Repeat the process until $P(IS) = \{\phi\}$
 - End for
 - (6) For each P(IS) do the following Class label detection probability (CLD_{pm}) = classify $p(IS_1)$
 - \rightarrow MedLDA algorithm to identify the Class Labels in the given $p(IS_1)$
 - Repeat $p(IS_1) = \{\phi\}$
 - End For
 - Class label detection probability (CLD_{ps}) = classify $p(IS_1) \rightarrow$ SVM algorithm to identify the Class Labels in the given $p(IS_1)$
 - Repeat $p(IS_1) = \{\phi\}$
 - (7) Compute the Total class label detection probability $TCLP_n = CLD_{pm} \cup CLD_p$
 - (8) For each ($TCLP_n$) \rightarrow Assign (CL) // Assign the class labels to each image partition segments
 - Repeat the process until $TCLP_n = \{\phi\}$
 - End For
-

Calculate the class label detection probability for each portioned image set using the MedLDA and CNN algorithms. The combined class label detection probability of the MedLDA and CNN algorithms is used to get the total class label detection probability. Class labels are assigned to each image segment based on the estimated total class label detection probability, and annotation is performed using the class labels.

5.1 Intelligent Class Labels Extraction Phase

In this phase, features are extracted from the annotated class labels in order to achieve more accurate classification of the target images. This step extracts features using machine learning algorithms. Algorithm 2 explains the stages involved in the collection of Intelligent class labels. This method takes a set of annotated class labels as input and extracts intelligent features from the set of class labels as output. The class labels are initially loaded and stored in an array. Each class label undergoes pre-processing. The proposed algorithm extracts intelligent features from a set of class labels using an optimized VGG16 model and a RESNET model [21,22]. Finally, for improved classification, the retrieved features are assigned to the appropriate classes.

Algorithm 2: Intelligent Class labels extraction phase

Input – Set of class labels.

Output – Intelligent features from the set of class labels.

- (1) Begin
 - (2) Load the class labels and store them into an array CL_1 to CL_n // where CL is the set of class labels.
 For each CL_1 to CL_n
 DO
 - (3) Perform pre-processing to the set of class labels
 Convert the CL \rightarrow RGB to grey scale
 Divide each class labels into various pixels of P_1 to P_n
 Repeat until $CL = \{\phi\}$
 End For
 - For each pre-processed class label
 Do
 - (4) Extract the features using optimised VGG16 Model to all the class labels.
 $F_{Ext1} = \text{Extract Features (CL)} \rightarrow \text{VGG16 Model}$
 - (5) Assign the extracted class label features to the corresponding class labels such that
 $CL_1 = \{F_{Ext1} \dots F_{Extn}\}$ // where F_{Ext} is the extracted class label features.
 Repeat until $CL = \{\phi\}$
 End For
 - For each pre-processed class label
 Do
 - (6) Extract the features using optimised ResNet Model to all the class labels.
 $F_{Ext1} = \text{Extract Features (CL)} \rightarrow \text{ResNet Model}$
 - (7) Assign the extracted class label features to the corresponding class labels such that
 $CL_2 = \{F_{Ext1} \dots F_{Extn}\}$ // where F_{Ext} is the extracted class label features.
 Repeat until $CL = \{\phi\}$
 End for
 Compute $FF_{Ext} = F_{Ext1} + F_{Ext2}$ // where FF_{Ext} is fusion feature extraction
 - End for
 For each class labels CL_1 to CL_n
 Do
 - (8) Assign FF_{Ext} to the corresponding class labels and store in the database
 End for
-

In this step, land cover satellite images were intelligently classified using fuzzy logic and the CNN algorithm. This step takes the annotated class labels and extracted intelligent features as input. Tab. 1 has annotations for the possible land cover images. Intelligent fuzzy rules are built using these annotated class labels and the collected intelligent features [26–28]. The object shape, annotated class labels, and its features are referred to as membership functions in the proposed system. Algorithm 3 provides the intelligent fuzzy rules for the proposed system for improving land cover image classification [21]. The proposed intelligent-based fuzzy classification algorithm makes use of a triangular membership function that is more appropriate for mamdami than sugeno models. As a result, the mamdami model is favored above the sugeno model in the proposed model. The algorithm illustrates the classification of vegetation areas using clever fuzzy rules.

Algorithm 3: Intelligent fuzzy rules for image classification (Classification of Vegetation)

Input – Set of class annotated class labels, intelligent features

Output – Intelligent fuzzy rules

One of the Fuzzy rules -**Classification of Vegetation**

“*NIR_Ratio == Low & NDVI_Value == Low => Image_Segment_Classification = Not_a_vegetation_area*”

“*NIR_Ratio == Low & NDVI_Value == Medium => Image_Segment_Classification = May_be_a_vegetation_area*”

“*NIR_Ratio == Low & NDVI_Value == High => Image_Segment_Classification = May_be_a_vegetation_area*”

“*NIR_Ratio == Medium & NDVI_Value == Low => Image_Segment_Classification = Not_a_vegetation_area*”

“*NIR_Ratio == Medium & NDVI_Value == Medium => Image_Segment_Classification = May_be_a_vegetation*”

“*NIR_Ratio == Medium & NDVI_Value == High => Image_Segment_Classification = Medium_Vegetation_area*”

“*NIR_Ratio == High & NDVI_Value == Low => Image_Segment_Classification = May_be_a_vegetation_area*”

“*NIR_Ratio == High & NDVI_Value == Medium => Image_Segment_Classification = Highly_vegetation_area*”

“*NIR_Ratio == High & NDVI_Value == High => Image_Segment_Classification = Strong_Vegetation_area*”

Table 1: Classification and efficiency results using binary cross entropy algorithm

Class label	Training images	Testing images	TP	TN	FP	CA
Vegetation area	350	150	73	26	1	73
	400	200	74	24	2	74
	450	250	75	22	3	75
	500	300	77	22.5	1.5	76.23
	550	350	79	20.5	0.5	78.88

(Continued)

Table 1 (continued)						
Class label	Training images	Testing images	TP	TN	FP	CA
Buildings	350	150	74.5	18.5	7	74.5
	400	200	76.5	19.5	4	76.5
	450	250	77.5	20	3.5	76.73
	500	300	79.6	17.5	2.9	79.6
	550	350	80.5	17.7	1.8	80.5
Roads	350	150	76.5	14.4	9.1	76.5
	400	200	78.4	14.4	7.2	78.4
	450	250	79.4	13.5	7.1	79.4
	500	300	80.1	14.9	5	80.1
	550	350	81	13.6	4.8	81.4

Fig. 3a illustrates a mamdani-based fuzzy inference system model for classifying vegetation areas from image segments. This model considers two membership functions, namely the NRI ratio and NDVI values, with weightings of low, medium, and high. Nine intelligent fuzzy rules are generated based on these membership functions. To perform the fuzzification process, the intelligent fuzzy rules are triggered and the rules are executed. The decision manager performs defuzzification and generates three alternative results from the performed intelligent fuzzy rules, namely strong vegetation area (3), high vegetation area (2), may be vegetation area (1), and not vegetation area (0). The membership function for the input variable NIR Ratio is shown in Fig. 3b. Fig. 3c gives Membership function for an input variable NDVI value is shown in Fig. 3c, and the Membership function for an output variable of image segment classification for vegetation area is shown in Fig. 3d. Similarly, the fuzzy inference systems are built for road and building classifications.

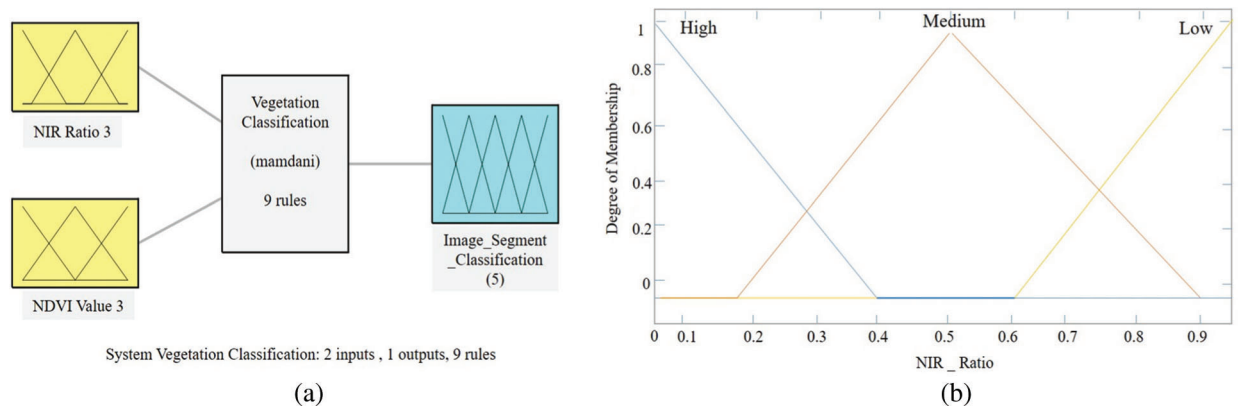


Figure 3: Continued

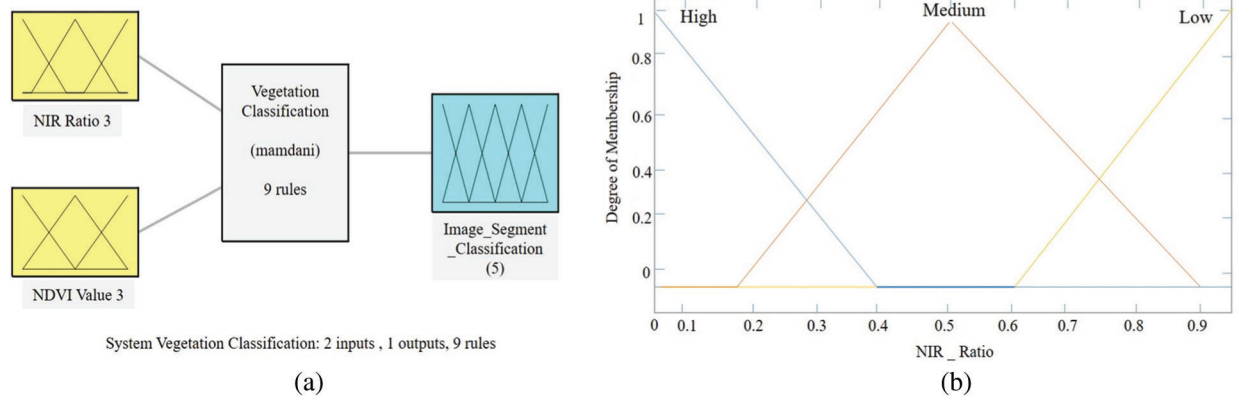


Figure 3: (a) Fuzzy inference system for vegetation area classification; (b) Membership function for an input variable NIR_Ratio; (c) Membership function for an input variable NDVI_vale; (d) Membership function for an output variable of image_segment_classification (NV- Not_a_Vegetation_area, MBV–May_Be_a_Vegetation, MVA–Medium_Vegetation_Area, HVA–Highly_Vegetation_Area, SVA–Strongly_Vegetation_Area)

6 CNN Based Classification Algorithm

By extending the CNN algorithm with intelligent fuzzy rules, the proposed system develops a novel intelligent CNN-based image classification algorithm. The proposed algorithm performs the convolution operations for two functions x and y for the operator using the integral given in Eq. (5) as follows.

$$(x * y)(t)^{def} = \int_{-\infty}^{\infty} x(\tau)y(t - \tau)d\tau = \int_{-\infty}^{\infty} x(t - \tau)y(\tau)d\tau \tag{5}$$

The Eq. (5) provides the convolution operations for two functions x and y for the operator using the integral. By employing this equation, the proposed system develops nine maximum pooling layers and ten convolutional layers for the two functions x and y , which are used to conduct image classification. Moreover, the proposed system employs the sigmoidal function as a activation function and we use the bias function defined by $f(x) = x + 1/x$ as a bias function along with the CNN. The proposed algorithm employs nine max pooling layers and ten convolution layers for performing the classification of the image. All of these layers operate on the image data set and provide a set of features such as NIR ratio, NDV ratio, LCM value, LE value, Std value, rectangular fit, and elliptical fit that can be used to classify the images in the given data set. In this proposed model, we use the sigmoidal function as a activation function and we use the bias function defined by $f(x) = x + 1/x$ as a bias function along with the CNN. By comparing them to the features selected by the feature selection algorithm, the CNN applies fuzzy rules to obtain feedback on the selected features. If both are identical, the classification process is initiated. In the event of a mismatch, it asks the decision manager for guidance on the qualities to employ depending on their sensitivity. Errors are communicated in reverse order and are minimized during the classification process. The intelligent fuzzy CNN proposed in this paper performs multiclass classification on a variety of distinct class labels, including buildings, vegetation, land, roads, and vehicles.

7 Experimental Setup and Results

The proposed intelligent classification model is implemented using the MATLAB 2013a software. The proposed model generates intelligent fuzzy rules using a mamdami model with triangle membership

functions. The proposed classification model is compared to previously published models using performance criteria such as True Positive Rating (TPR), True Negative Rating (TNR), False Positive Rating (FPR), and Classification Accuracy (CA). The TPR, TNR, FPR, and CA are calculated in the following manner.

$$TPR = \frac{TP}{TP + FN} \quad (6)$$

$$TNR = \frac{TN}{TP + FP} \quad (7)$$

$$FPR = \frac{FP}{FP + TN} \quad (8)$$

$$CA = \frac{TPR}{TPR + TNR + FPR} \quad (9)$$

Tab. 1 gives the classification detection accuracy of various class labels such as vegetation, road, and buildings for binary cross entropy algorithm. In **Tab. 1**, three class labels namely vegetation area, building and roads are considered. For each class labels different sets of training and testing images are given as input to the buildings for binary cross entropy algorithm.

As shown in **Tab. 1**, the average true positive value for the vegetation area class label is 75.6%, the average true negative value is 23%, the average false positive value is 1.6%, and the classification accuracy for the binary cross entropy algorithm is 75.42 percent when training and testing images are varied. For building class label the average true positive value is 77.72%, average true negative value is 18.64%, average false positive value is 3.84% and classification accuracy of for binary cross entropy algorithm is 77.56% for varying training and testing images. For road class label the average true positive value is 79.08%, average true negative value is 14.16%, average false positive value is 6.64% and classification accuracy of for binary cross entropy algorithm is 79.16% for varying training and testing images. **Tab. 2** exhibits the classification detection accuracy for CNN utilizing the RNN algorithm for various class labels such as vegetation, road, and building. In **Tab. 2**, three class labels are considered: vegetative area, building, and road. Different sets of training and testing images are fed into the CNN using the RNN algorithm for each class label. From the **Tab. 2**, it is clear that for vegetation area class label the average true positive value is 79%, average true negative value is 15.3%, average false positive value is 1.2% and classification accuracy of for binary cross entropy algorithm is 79.14% for varying training and testing images. For building class label the average true positive value is 82.4%, average true negative value is 15.66%, average false positive value is 2.54% and classification accuracy of for CNN using RNN algorithm is 83% for varying training and testing images. For road class label the average true positive value is 79.08%, average true negative value is 12.3%, average false positive value is 4.5% and classification accuracy of for CNN using RNN algorithm is 83.16% for varying training and testing images.

Table 2: Classification and efficiency results by CNN using RNN algorithm

Class label	Training images	Testing images	TP	TN	FP	CA
Vegetation area	350	150	74	2.5	1	74.7
	400	200	77	22.5	0.5	77
	450	250	79	20	1	79
	500	300	82	16	2	82
	550	350	83	15.5	1.5	83

(Continued)

Table 2 (continued)						
Class label	Training images	Testing images	TP	TN	FP	CA
Buildings	350	150	80.5	16.5	3	80.5
	400	200	81.6	17.1	2.3	80.7
	450	250	82.7	15.3	2	82.7
	500	300	83.1	14.8	3.1	82.27
	550	350	84.1	14.6	2.3	83.26
Roads	350	150	81.4	12.6	6	81.4
	400	200	82.4	12.7	4.9	82.4
	450	250	82.6	11.8	4.6	83.43
	500	300	84.1	12.6	3.3	84.1
	550	350	84.5	11.8	3.7	84.5

Tab. 3 summarizes the classification detection accuracy for the proposed intelligent classification algorithm for various class labels such as vegetation, road, and building. Tab. 3 considers three class labels: vegetative area, building, and road. Different sets of training and testing images are given into the buildings for the proposed intelligent classification algorithm for each class label. From the table it is clear that for vegetation area class label the average true positive value is 88.42%, average true negative value is 11%, average false positive value is 1.38% and classification accuracy of for binary cross entropy algorithm is 87.68% for varying training and testing images. For building class label the average true positive value is 90.96%, average true negative value is 7.58%, average false positive value is 1.66% and classification accuracy of for proposed intelligent classification algorithm is 90.68% for varying training and testing images. For road class label the average true positive value is 90.5%, average true negative value is 7.44%, average false positive value is 2.06% and classification accuracy of the proposed intelligent classification algorithm is 90.46% for varying training and testing images.

Table 3: Classification and efficiency results for the proposed intelligence classification algorithm

Class label	Training images	Testing images	TP	TN	FP	CA
Vegetation area	350	150	85	14	1	85
	400	200	87	12	2	86.13
	450	250	88.5	11	1.5	87.6
	500	300	89.9	10	1.1	89
	550	350	91.7	8	1.3	90.7
Buildings	350	150	86.1	12.8	2.1	85.24
	400	200	88.9	10	1.1	88.9
	450	250	92.1	6.2	1.7	92.1
	500	300	93.6	5.1	1.3	93.1
	550	350	94.1	3.8	2.1	94.1

(Continued)

Class label	Training images	Testing images	TP	TN	FP	CA
Roads	350	150	86.5	10.1	3.4	86.5
	400	200	89.8	8.2	2	89.5
	450	250	90.6	8	1.4	90.6
	500	300	91.8	6.1	2.1	91.8
	550	350	93.8	4.8	1.4	93.8

Fig. 4 illustrates the classification accuracy of three classification algorithms: CNN with binary entropy, CNN with RNN algorithm, and the proposed intelligent classification algorithm for three image class labels: vegetative area, buildings, and roads. From the Fig. 4, it is observed that the proposed intelligent classification algorithm has better classification accuracy of three class labels for vegetation class label (84%), building (90%) and roads (90%) when it is compared with other existing classification algorithms such as CNN with binary entropy for vegetation class label (75%), building class label (76%) and road class label (79%) and CNN with RNN algorithm with classification accuracy of various class labels vegetation class label (79%), building class label (80%) and road class label (83%). The proposed intelligent classification technique achieves a higher classification accuracy because it combines semantic analysis with single label image segmentation.

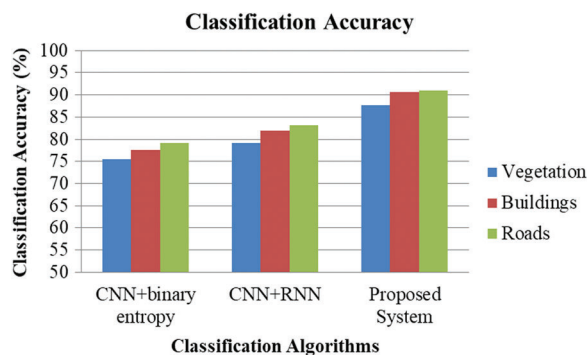


Figure 4: Classification accuracy

Moreover, when compared to other existing classification algorithms, the proposed intelligent classification algorithm employs intelligent fuzzy rules in conjunction with the CNN algorithm to accurately classify the selected class labels vegetation area, buildings, and roads. Furthermore, the proposed intelligent classification method has a higher proportion of genuine positives and a lower proportion of true negatives, as well as a lower false positive rate. As a result, the proposed intelligent classification method outperforms other existing classification algorithms in terms of class label accuracy.

8 Conclusion and Future Work

A novel intelligent Fuzzy based CNN image classification model has been proposed in this paper. This strategy is advantageous for enriching the target information and outperforms manual image labeling by collecting semantic descriptors from the images automatically. The experimental results from three remote

sensing image datasets demonstrate that the proposed framework significantly improves the performance of Multi Label annotation when compared to alternative annotation approaches. In comparison to previous methods, the improvised algorithm adaptively decides the number of semantic classifications within class labels during annotation. The proposed intelligent classifier overcomes the least probability retrieval error during classification. This approach produces more true positives and fewer true negatives, as well as lower false positive rates. The future research will entail modifying the similarity measurements in order to generate more semantically related scenes using enhanced metric learning approaches. Additionally, it focuses on developing Fuzzy-CNN to operate on many classes and incorporate methods for classification judgments, incorporating Multi Label and Multi Class output models into land cover remote sensing images.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] R. F. Berriel, A. T. Lopes, A. F. Souza and T. Oliveira-Santos, "Deep learning-based large-scale automatic satellite crosswalk classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 9, pp. 1513–1517, 2017.
- [2] J. Vanegas, H. Escalante and F. González, "Scalable multi-label annotation via semi-supervised kernel semantic embedding," *Pattern Recognition Letters*, vol. 123, pp. 97–103, 2019.
- [3] Y. Xia, K. Chen and Y. Yang, "Multi-label classification with weighted classifier selection and stacked ensemble," *Information Sciences*, vol. 557, pp. 421–442, 2021.
- [4] F. Markatopoulou, V. Mezaris and I. Patras, "Implicit and explicit concept relations in deep neural networks for multi-label video/image annotation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 6, pp. 1631–1644, 2019.
- [5] H. Kim, J. Park, D. Kim and J. Lee, "Multilabel naïve Bayes classification considering label dependence," *Pattern Recognition Letters*, vol. 136, pp. 279–285, 2020.
- [6] L. Wang, T. Zhou, Y. Lee, K. Cheoi, K. Ryu *et al.*, "An efficient refinement algorithm for multi-label image annotation with correlation model," *Telecommunication Systems*, vol. 60, no. 2, pp. 285–301, 2015.
- [7] P. Zhu, Y. Tan, L. Zhang, Y. Wang, J. Mei *et al.*, "Deep learning for multilabel remote sensing image annotation with dual-level semantic concepts," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 6, pp. 4047–4060, 2020.
- [8] G. Xia, Z. Wang, C. Xiong and L. Zhang, "Accurate annotation of remote sensing images via active spectral clustering with little expert knowledge," *Remote Sensing*, vol. 7, no. 11, pp. 15014–15045, 2015.
- [9] F. Hu, W. Yang, J. Chen and H. Sun, "Tile-level annotation of satellite images using multi-level max-margin discriminative random field," *Remote Sensing*, vol. 5, no. 5, pp. 2275–2291, 2013.
- [10] J. Joshua Bapu and D. Jemi Florinabel, "Automatic annotation of satellite images with multi class support vector machine," *Earth Science Informatics*, vol. 13, no. 3, pp. 811–819, 2020.
- [11] J. Jeppesen, R. Jacobsen, F. Inceoglu and T. Toftegaard, "A cloud detection algorithm for satellite imagery based on deep learning," *Remote Sensing of Environment*, vol. 229, pp. 247–259, 2019.
- [12] M. A. Kadhim and M. H. Abed, "Convolutional neural network for satellite image classification," in *Asian Conf. on Intelligent Information and Database Systems, Intelligent Information and Database Systems: Recent Developments*, Indonesia, pp. 165–178, 2019.
- [13] J. Cao, A. Zhao and Z. Zhang, "Automatic image annotation method based on a convolutional neural network with threshold optimization," *PLoS One*, vol. 15, no. 9, pp. 1–21, 2020.
- [14] G. Hoxha, F. Melgani and B. Demir, "Toward remote sensing image retrieval under a deep image captioning perspective," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 4462–4475, 2020.

- [15] S. Heng, S. Li, W. Li and M. Xiaoyong, "Fuzzy semantic retrieval of distributed remote sensing images," in *Int. Conf. on Computational Intelligence and Security*, China, pp. 1435–1441, 2006.
- [16] S. Ghosh, D. Biswas, S. Biswas, D. Chandasarkar, P. Sarkar *et al.*, "Soil classification from large imagery databases using a neuro-fuzzy classifier," *Canadian Journal of Electrical and Computer Engineering*, vol. 39, no. 4, pp. 333–343, 2016.
- [17] M. Ivacic-Kos, M. Pobar and S. Ribaric, "Two-tier image annotation model based on a multi-label classifier and fuzzy-knowledge representation scheme," *Pattern Recognition*, vol. 52, pp. 287–305, 2016.
- [18] M. Alshehri, "A content-based image retrieval method using neural network-based prediction technique," *Arabian Journal for Science and Engineering*, vol. 45, no. 4, pp. 2957–2973, 2019.
- [19] L. Wang, A. Zhang, P. Wang and Y. Dong, "Automatic image annotation using model fusion and multi-label selection algorithm," *Journal of Intelligent & Fuzzy Systems*, vol. 37, no. 4, pp. 4999–5008, 2019.
- [20] A. Alzubi, A. Amira and N. Ramzan, "Semantic content-based image retrieval: A comprehensive study," *Journal of Visual Communication and Image Representation*, vol. 32, pp. 20–54, 2015.
- [21] S. Jabari and Y. Zhang, "Very high resolution satellite image classification using fuzzy rule-based systems," *Algorithms*, vol. 6, no. 4, pp. 762–781, 2013.
- [22] L. Li, T. Xu and Y. Chen, "Fuzzy classification of high resolution remote sensing scenes using visual attention features," *Computational Intelligence and Neuroscience*, vol. 2017, pp. 1–9, 2017.
- [23] B. Feizizadeh, T. Blaschke, D. Tiede and M. Moghaddam, "Evaluating fuzzy operators of an object-based image analysis for detecting landslides and their changes," *Geomorphology*, vol. 293, pp. 240–254, 2017.
- [24] H. A. Gheshlaghi and B. Feizizadeh, "An integrated approach of analytical network process and fuzzy based spatial decision making systems applied to landslide risk mapping," *Journal of African Earth Sciences*, vol. 133, pp. 15–24, 2017.
- [25] R. Bharti and J. Kurmi, "A survey of satellite high resolution image classification," *International Journal of Computer Applications*, vol. 164, no. 1, pp. 26–28, 2017.
- [26] L. Ma, M. Li, X. Ma, L. Cheng, P. Du *et al.*, "A review of supervised object-based land-cover image classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 130, pp. 277–293, 2017.
- [27] M. Sameen and B. Pradhan, "A two-stage optimization strategy for fuzzy object-based analysis using airborne LiDAR and high-resolution ortho photos for urban road extraction," *Journal of Sensors*, vol. 2017, pp. 1–17, 2017.
- [28] D. Hou, Z. Miao, H. Xing and H. Wu, "V-RSIR: An open access web-based image annotation tool for remote sensing image retrieval," *IEEE Access*, vol. 7, pp. 83852–83862, 2019.