CLOUD DETECTION IN REMOTE SENSING IMAGES USING DEEP LEARNING

By

Maria Amjad

NATIONAL UNIVERSITY OF MODERN LANGUAGES

ISLAMABAD

January, 2023

Cloud Detection in Remote Sensing Images Using Deep Learning

By

MARIA AMJAD

BSCS, PMAS Arid Agriculture University, Rawalpindi, 2016

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

In **Computer Science**

To

FACULTY OF ENGINEERING & COMPUTER SCIENCE

NATIONAL UNIVERSITY OF MODERN LANGUAGES ISLAMABAD

NATIONAL UNIVERSITY OF MODERN LANGUAGES   FACULTY OF ENGINEERING & COMPUTER SCIENCE

THESIS AND DEFENSE APPROVAL FORM

The undersigned certify that they have read the following thesis, examined the defense, are satisfied with overall exam performance, and recommend the thesis to the Faculty of Engineering and Computer Sciences for acceptance.

Thesis Title: Cloud Detection in Remote Sensing Images Using Deep Learning

Submitted By: Maria Amjad   **Registration #:** 34 MS/CS/S20

Master of Science in Computer Science (MSCS)
Title of the Degree

Computer Science
Name of Discipline

Dr. Fazli Subhan
Name of Research Supervisor   ____________________ Signature of Research Supervisor

Dr. Sajid Saleem
Name of Research Co-Supervisor   ____________________ Signature of Research Co-Supervisor

Dr. Muhammad Noman Malik
Name of Dean (FE&CS)   ____________________ Signature of Dean (FE&CS)

Brig. Syed Nadir Ali
Name of Director General   ____________________ Signature of Director General

January 12th, 2023

AUTHOR'S DECLARATION

I Maria Amjad

Daughter of Muhammad Amjad

Registration # 34 MS/CS/S20

Discipline Computer Science

Candidate of **Master of Science in Computer Science (MSCS)** at the National University of Modern Languages do hereby declare that the thesis **Cloud Detection in Remote Sensing Images Using Deep Learning** submitted by me in partial fulfillment of MSCS degree, is my original work, and has not been submitted or published earlier. I also solemnly declare that it shall not, in future, be submitted by me for obtaining any other degree from this or any other university or institution. I also understand that if evidence of plagiarism is found in my thesis/dissertation at any stage, even after the award of a degree, the work may be cancelled and the degree revoked.

____
Signature of Candidate

Maria Amjad
Name of Candidate

12th January, 2023
Date

ABSTRACT

Title: Cloud Detection in Remote Sensing Images Using Deep Learning

Remote sensing images play a vital role in the analysis of the earth's surface. The earth analysis is productive if the sky is clear because clouds obscure the earth surface and create problems for remote sensing applications such as change detection, agriculture, surveillance, urban and rural planning. Various methods have been proposed for the detection of clouds, which vary from pixel intensity transformation based methods to deep learning methods. The intensity transformation methods are generally fast but they are susceptible to the variation in the pixel intensities, illumination changes and noise. On the other hand, the deep learning methods are efficient and accurate but require training on dataset(s) prior to cloud detection. In this thesis, You Only Look Once (Yolo) algorithm is investigated for the cloud detection. Yolo has been successfully applied for the detection and recognition of real life objects in indoor and outdoor images. In this thesis, the Yolo algorithm is combined with other three state of the art deep learning algorithms in order to improve its accuracy for cloud detection. These algorithms are Practical Portrait Human Segmentation Lite (PP-HumanSeg_Lite), Deep Dual Resolution Network (DDRNet) and Disentangled Non-Local Network (DNLNet). All these algorithms have been used for the semantic segmentation of the real life objects. The combination of Yolo with PP-HumanSeg_Lite, DDRNet and DNLNet is done through an ensemble learning method where the responses of Yolo and the other algorithms are combined and provided to Random Forest for accurate cloud detection. Experiments are performed on two different cloud datasets which are High Resolution Cloud Detection (HRCD) Dataset and 38-Cloud Segmentation datasets. The experimental result shows that Yolo+PP-HumanSeg_Lite give the best results and achieves accuracy of 96% and 93% on HRCD and 38-Cloud datasets, respectively. Whereas, Yolo achieves 91.2% and 81.5% accuracy, respectively.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANN	-	Artificial Neural Network
CNN	-	Convolutional Neural Network
DAFNet	-	Dense Attention Fluid Network
DDRNet	-	Deep Dual-resolution Network
DNLNet	-	Disentangled Non Local Neural Network
FCN	-	Fully Convolutional Networks
GF-1	-	Gaofen-1
HRCD	-	High Resolution Cloud Detection
IOU	-	Intersection Over Union
KNN	-	K-Nearest Neighbour
LCD	-	Light-Weight Cloud Detection
PP-HumanSeg_Lite	-	Practical Portrait Human Segmentation Lite
R-CNN	-	Regions with Convolutional Neural Network
ResNet	-	Residual Networks
RF	-	Random Forest
RGB	-	Red Green Blue
RNN	-	Recurrent Neural Network
U-Net	-	U shaped Convolutional Neural Network
ZY-3	-	Ziyuan-3, 'Resource-3'

ACKNOWLEDGMENT

First of all, I wish to express my gratitude and deep appreciation to Almighty Allah, who made this study possible and successful. This study would not be accomplished unless the honest espousal that was extended from several sources for which I would like to express my sincere thankfulness and gratitude. Yet, there were significant contributors for my attained success and I cannot forget their input, especially my research supervisors, Dr. Fazli Subhan and Dr. Sajid Saleem, who did not leave any stone unturned to guide me during my research journey.

I shall also acknowledge the extended assistance from the administrations of Department of Computer Sciences who supported me all through my research experience and simplified the challenges I faced. For all whom I did not mention but I shall not neglect their significant contribution, thanks for everything.

DEDICATION

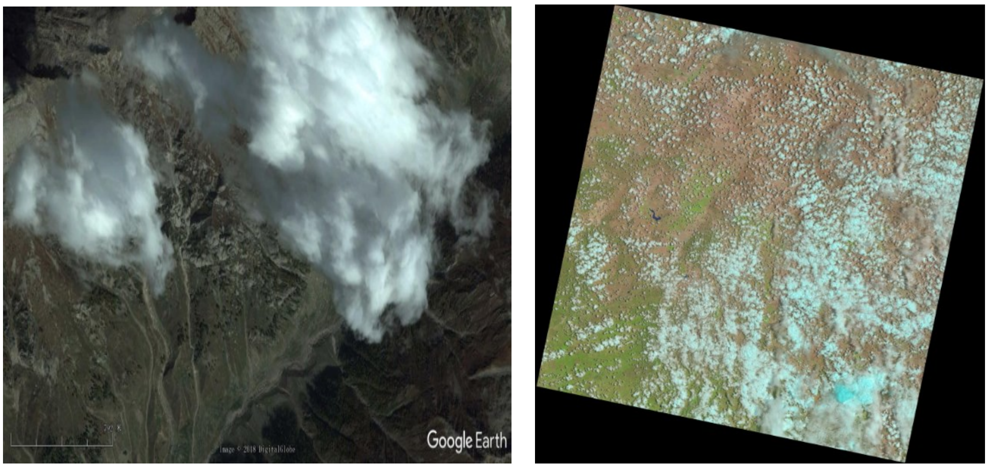
This thesis work is dedicated to my parents and my teachers throughout my education career who have not only loved me unconditionally but whose good examples have taught me to work hard for the things that I aspire to achieve.

CHAPTER 1

INTRODUCTION

1.1 Overview

Remote sensing images are widely used in the analysis of earth surface [1]. The presence of clouds in the remote sensing images, obscure the objects and other information on the earth surface and creates problems for applications such as change detection, agriculture, surveillance, urban and rural planning [2]. Example of remote sensing images containing clouds is shown in Figure 1.1.



(a) (b)

Figure 1.1: Example images containing clouds taken from datasets (a) HR Cloud Detection [3] (b) 38-Cloud Segmentation [4]

Over the past few years the cloud detection has been shifted from classical methods to machine/deep learning methods [2]. The classical methods are pixel intensity thresholding based methods [5][6]. Such methods have limitations like pixel intensity is susceptible to noise and other illumination variations. Also clouds are of different types like thin and dense clouds while create problems for the intensity thresholding based methods. Also these methods fails to detect different types of clouds in remote sensing images of different spectral bands [7]. In contrast the machine and deep learning methods are accurate and efficient. They are first trained on datasets to learn about the clouds and then detect the cloud efficiently [8].

Thin cloud detection is a challenging problem, a method for thin cloud detection is proposed which extracts the dark pixel [9] of the image. Then, it extracts the scattered thin regions of the image and Thiessen polygon parameter is used to calculate the density of these pixels. As it is an effective method but still it fails to detect thin clouds. High resolution satellite images are used and taken from the urban areas for cloud detection and the method detects the clouds by using its color and edge features [10]. Experiment has been performed and accuracy is achieved but it is unable to detect some thin cloud regions.

Deep learning which is the subset of machine learning, multiple deep learning methods applied on multiple types of input images for cloud or no cloud detection [11]. Input types of the images contain pixel values, features and patches. After performing an experiment convolution neural network gives the accurate results as it gives 86% accuracy only by taking input type of images as patches.

Deep convolution network is used to extract the features of the clouds from satellite images and adaptive simple linear iterative clustering algorithm is used for pixel wise detection of clouds [12]. This algorithm gives the accuracy up to 94% but is inefficient for thin cloud detection. Proba-V which is a small satellite from which images are taken and it is difficult to identify the location of clouds from these images.

Different machine learning models are used from which artificial neural network is used for dataset training and gives the accurate result by identifying thin clouds, cloud borders and bright surfaces [13]. This method gives the accuracy of 95% of predicted clouds. From the comparison of these classical and machine learning algorithm it is observed that machine learning algorithms gives high accuracy and hybrid methods combining the artificial neural networks can overcome the drawbacks of existing algorithms.

For thin cloud detection modified simple linear iterative clustering algorithm is used in which natural scene statistics model detect the cloud and snow regions of the image [14]. GF-1 satellite images are used and the algorithm gives 93% accuracy which is the high accuracy as compared to other methods.

Deep matting process separates the foreground image from background to detect the clouds from the image [15]. A method is proposed for cloud detection and removal in remote sensing images which improve the performance of remote sensing applications. Thick and thin clouds are detected by weakly supervised learning [16]. Deep learning method is proposed for cloud and shadow detection from Gaofen-1 satellite images [17]. Multi-scale deep learning method is introduced to detect cloud and shadow using Gaofen-1 satellite images which provide accuracy but cloud detection is also needed for real time remote sensing applications.

In past years, a lot of work is done on natural images which give an accurate result for object detection, aerial images are different from natural images as they produce visual spectrum. DOTA dataset is used for object detection from aerial images, images taken from the Earth vision [18]. Deep convolution neural network classifies the cloudy and non-cloudy areas, proposed haze or cloud detection method for airborne videos which do not produce the effective results [19].

Deep learning approach gives highly accurate results from recent years. Deep learning methods will be trained for shape based cloud detection. Yolo algorithm is fast, real time object detection algorithm. It is highly generalized algorithm, trained on a loss function that directly corresponds to detection performance and trained directly on full images [20]. We will combine Yolo algorithm with other deep learning methods to improve the accuracy, experiment will be performed which will produce effective results. Figure no. 2 of the article [20] shows the real time object detection from Yolo algorithm.

1.2 Motivation

From past years, cloud detection is increasingly used in remote sensing image. Many methods have been proposed for cloud detection but still there some challenges exist, to address especially these using deep learning techniques. Recent studies use the machine learning and deep learning methods for cloud detection which need further improvement, enhancement and investigation [17].

1.3 Problem Background

Cloud detection containing faces challenges due to nature of clouds which appear thin, thick cloud and similar to snow and create problems for classical and machine learning algorithms [2]. There is a need to improve the accuracy of cloud detection by combining different neural network based methods to remove the shortcomings in the existing algorithms. Deep matting has been used to perform cloud detection under different cloud covers on remote sensing images [15]. But still accuracy suffers due to the image quality and other degradation in the acquisition process. Deep convolution neural network method is used for the cloud or haze detection, effective results has been drawn [19]. But there is a need to find the exact boundaries of the cloud and to improve the quality detection. Cloud detection is performed on high resolution remote sensing images under weakly supervision [16]. But there is a need to use the advance deep network architecture and to improve the performance of the method.

1.4 Problem Statement

Real time irregular shape based cloud detection is required from remote sensing images [17]. Deep learning methods with the passage of time are improving and require hybrid method that combines the strength of different deep learning methods. Yolo which is accurate in real time object detection requires to be used for cloud detection with hybridization of other deep learning methods. So that pixel wise detection and accurate results can be obtained.

1.5 Research Questions

Q.1. How to increase the accuracy for cloud detection using Yolo and other deep learning methods?
Q.2. How the combination of Yolo with other deep learning methods reduced the number of missed dense and thin clouds in cloud detection?

1.6 Aim of Research

The aim of this research is to do cloud detection on remote sensing images. There are various applications which rely on remote sensing which should be cloud free so that measures related to change detection, urban planning, and agriculture can be made accurately

1.7 Research Objectives

- To increase the cloud detection accuracy using Yolo and other deep learning methods.
- To combine Yolo with other deep learning algorithms to reduce the number of missed dense and thin clouds.

1.8 Scope of the Research Work

Different techniques are used for cloud detection but there are still some challenges like real time detection or to increase the performance of the model. The aim is to design the rich model which is used for remote sensing applications. The model is helpful for remote sensing environment like for land monitoring and will produce the better results.

1.9 Thesis Organization

The rest of the thesis is organized as follow:

Chapter 2 provides background on machine and deep learning algorithms used for the cloud detection. Some of the methodologies and the algorithms are discussed. The issues and the challenges are highlighted.

Chapter 3 presents detail of deep learning algorithms applied on real life dataset which contains indoor and outdoor images. Semantic segmentation is done for the object detection. The methodology of the methods is discussed with their architecture diagram.

Chapter 4 presents the proposed methodology applied on two datasets one is high resolution dataset. The performance is measured in terms of accuracy. The ensemble technique and the training model are discussed in this chapter.

Chapter 5 provides the results and analysis of the proposed method. The comparison between the methods is shown in form of graphs and table. The accuracy achieved on both datasets is also discussed.

Chapter 6 gives the summary of the contributions in detail-along with conclusion.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

In this chapter, cloud detection schemes are discussed. Thick/thin clouds are detected by different techniques. Some traditional classical algorithms are discussed for the detection and removal of clouds. Taxonomy of cloud detection is presented which explores the different cloud detection techniques for remote sensing images. Machine learning and deep learning methodologies used for object and pixel-wise cloud detection schemes are discussed. Some image processing techniques are reviewed. Comparison of techniques, their strength and limitations are also discussed.

Cloud is an essential part because due to this rain occur. Cloud detection is important for the study of Earth surface. A survey is performed on the cloud detection in which there is a discussion about how different methodologies are applied on satellite images having various forms of clouds and also their results or findings. Many researchers that worked on the threshold-based classification algorithms and machine learning techniques with the accuracy is also discussed and after reviewing these methodologies with their results, it is found that cloud detection accuracy could be improved by different time and environment and the hybrid methods of artificial neural networks can remove the drawbacks of existing algorithms [2].

Classical algorithm approach is threshold-based algorithm; cloud detection methods that are based on threshold algorithm give poor quality or results because the results of classical algorithm approach vary due to change of climate and time. Machine learning techniques are easy to understand, their results are not consistent, the training depends on the input data but the hybrid technique which is the combination of machine learning techniques and classical algorithm give the accurate results. This survey shows that more work is needed which provide more accurate results for cloud detection methods. The table no. 3 of the article [2] shows the strength and limitations of classical and machine learning algorithm including their methods. In past years, pixel-wise image classification for cloud detection is a problem as it is difficult to detect the thin clouds. Deep cloud matting method in which matting is used in image processing which separates the foreground objects from the background in the given image, also deals with these three tasks which are cloud detection, cloud cover assessment and cloud removal [15].

When the satellite or any aircraft fly above the clouds mixed energy received by the sensors, the model is proposed for this energy of cloud images named mixed energy imaging model. Network based on convolution neural network is created and the network is trained with multitask loss function. Experiment performed on the remote sensing dataset having 328 images captured by the satellite, 72 images are for training and 256 for testing and images contained either thick or thin clouds. 38 background images and 34 thick-cloud images are selected from training set to train cloud matting network and for cloud detection and removal, 23 background images and 43 thick cloud images are selected from testing set. Highest accuracy is found on detecting the thin clouds. The table III of the article [15] shows the results.

High resolution remote sensing images are taken for the classification of cloud or non-cloud images in which the learning of deep network whose architecture is similar to convolution neural network, done under weak supervised learning [16]. New framework is proposed which train networks with block-level labels to detect the image contain the clouds or not. Iterative backward propagations applied for further computations. The proposed method gives the desired results, but there is a need to improve its performance. China National Space Administration developed a Gaofen-1 satellite mission for near real time earth observation [17]. Images are taken from Gaofen-1 satellite for cloud and non-cloud segmentation. The new algorithm proposed in deep learning architecture which extracts the features during learning or training process. The new method proposed and compared with other methods to evaluate the accuracy of the methods in cloud and cloud shadow segmentation. The proposed method has a good performance even facing the small clouds than the other methods. The figure no. 5 of the article [17] shows the comparison of these methods.

DOTA which is the largest dataset introduced for object detection, contains the aerial images from different sensors and platform [18]. Ten algorithms have been evaluated for guidance of design of object detection algorithms. It is challengeable for the researchers to identify the objects in aerial images. With the advancement of deep learning, many researchers adapt deep object detection from natural to aerial images. Code library is built by integrating some algorithms and multiple experiments performed for the object detection in aerial images. The figure no. 9 of the article [18] shows the results of DOTA dataset detecting the objects in aerial images. Cloud/haze detection done on airborne videos in which dataset of multiple videos are taken, videos are captured by an airplane or UAV [19].

Frames from the video are taken in a sequence for the detection as this method worked on the sequence of frames. These frames are considered as a set and further divided into patches. Deep convolution neural network is proposed having three convolution layers each followed by a pooling layer. Network is trained by using patch sets and labeled them 1 if there is cloud or haze and 0 for non-cloud image. The figure no. 8 of the article [19] shows an effective results of their proposed solution. Hybrid method is proposed for the cloud detection which is the combination of threshold and deep learning method [4]. Convolution neural network which is a deep learning model is used and dataset contains the Landsat 8 images. The method detects the clouds from the image and removes the snow and ice region while training the dataset. Pixel wise identification is performed and the accuracy is achieved. There is a need to improve the network to identify more cloud context. Deep learning methods or algorithms are applied on multiple satellite data [21] which extract the clouds and clear sky pixels. Back propagation neural network algorithm which is based on Keras deep learning framework platform is also used for cloud detection from other satellite data. The results provide more than 90% accuracy but there are some issues related to original and simulated data and other errors also occurs during cloud detection while using the trained dataset.

Cloud detection is also done on remote sensing images using machine learning models [22] and these models are also compared to determine the highest accuracy of the model. SVM which is the machine learning model gives the highest accuracy and valid for cloud detection but there is a need to compare the performance of sentinel image dataset with Landsat or other satellite data. Pixel level and object level cloud detection schemes are discussed in detail.

2.2 Pixel based Cloud Detection Schemes

Different machine and deep learning techniques are used to detect clouds from remote sensing images by using different parameters, like cloud shape, density, shadow or clear sky background difference [2]. Classical algorithms [9],[23],[24] contains threshold approach in which there is a variation in the cloud detection results of different areas due to climate or season change. Previous researchers used the spectral features of the cloud based on threshold method for detecting clouds. Radiative transfer equation based algorithm [23] is applied on Landsat-8 images to obtain the band color for thin cloud detection. Coastal and blue band shows the non-cloud and cloud cover area. Effective algorithm is proposed to extract two images having cloud and clear sky with satisfactory result.

Unified method, state of the art method for the removal of thin clouds [24] from remote sensing images. Spectral technique is used to extract cloud pixels and make their clusters. Method is used to extract multiple spectra of images to calculate the presence of cloud. For the classification of pixels, thresholding technique is applied. Thin clouds are detected by the proposed method and removed but cannot fully detected the thick clouds. Thin cloud masking is still challenging for the researchers detected the clouds by the classical algorithms. Dark pixel region is detected by the path radiance [9] which detects that the number of dark pixels are less than the clear pixels. Path radiance is added due to the scattering effect of thin clouds. For the density calculation of thin clouds, window size parameter unable to give accurate result due to irregular shape and different size of clouds. Thiessen polygon parameter is used to measure the density calculation of dark pixels. The multiband image is converted into single band through transformation and the cloud is segmented to obtain the thin cloud masks. Some missing and false detection is also observed.

J. Zhang has proposed the minimum cross entropy threshold method [25] for high resolution remote sensing images. Firstly, the objects are extracted from the images using the threshold method. Transformation filter is applied to decompose the image to extract its detailed information. Feature detail binary image is obtained by the minimum cross entropy threshold method for image segmentation. Regular shaped objects like circle or rectangle are removed to improve the performance. Boundary tracking method is used to extract the areas. Further refining is done for the efficient cloud detection performance. The method is also compared with previous researches including supervised learning techniques. Method is suitable for complex ground objects but the spectral data is required to overcome the false cloud detection. Many classical methods are proposed for not only the detection but the removal of the clouds.

Image processing is applied for the enhancement of the image which is the combination of block-based processing [26] to accurately detect the clouds. Inpainting technique is used as a threshold method for the binarization of an image. The proposed algorithm is based on image preprocessing and compared with previous classical method achieved better results. Another method used the multiple spectral features [27] of remote sensing images to detect the clouds in crop areas. Thresholding is applied as the dataset contains images having all kind of climate change and the method is efficient for agriculture applications. Figure 2.1 shows the taxonomy of machine learning and deep learning techniques for the cloud detection.

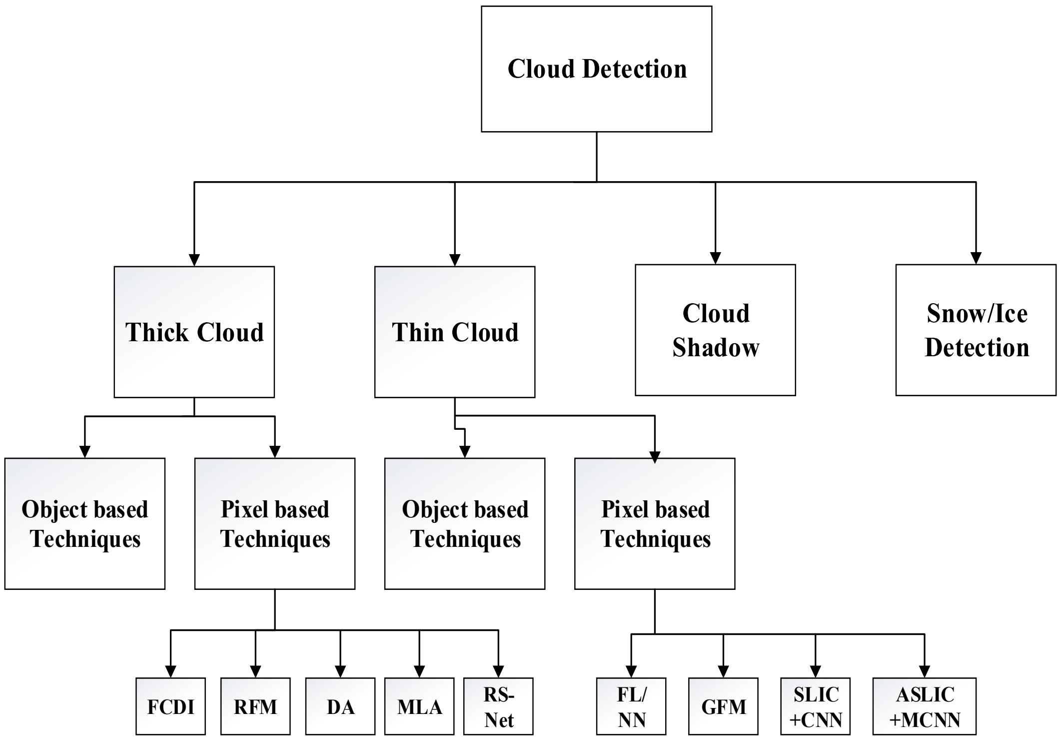


Figure 2.1: Taxonomy for Cloud Detection Schemes

2.2.1 Thick/Thin Cloud Detection Schemes

Machine learning techniques are applied in previous studies in which training and testing of data is done on remote sensing images to detect clouds. MSG images [28] trained on three algorithms fuzzy logic, multi-layer perceptron and neural network in which the data is labeled by extracting the spectral features of the images. Three features are extracted and the application is used to detect the non-cloud areas in which the algorithm is based on thresholding technique. The cloud mask is obtained which contains the binary image. The cloud detection methodology fuzzy logic is applied on the training dataset to check the pixel whether it is cloudy or not and then it classifies the result. Multi-layer perceptron is used for image processing of fuzzy logic to train the data. It contains the input, hidden and output layer and the hidden layer detect the clouds in less time. The proposed method detected the thick and thin cloud both but it is limited only for the day time cloud detection. Small and thin cloud detection gives the least accuracy as they are challenging to detect. Due to the irregular shapes of clouds, it is difficult to detect the small clouds, for this, patch based methods are applied in which the image is segmented and then pixels are combined to make a patch. Ground areas are complex to analyze; Gabor feature filter [14] is applied to differentiate cloud and snow by extracting the features of the image. The data is trained on support vector machine (SVM) model to classify cloud and snow. The method is compared with other algorithms.

Many image processing techniques [29] has been discussed for the processing of remote sensing images. Histogram and optimization based methods are also used to enhance the images. It is necessary to enhance the image to get its all detailed information. Differential evolution algorithm [30] is applied to the remote sensing images by defining the fitness function. Method is applied to the natural images to enhance them by expanding its contrast. The computations are complex but it preserves the minor details of the image. Many multi-level threshold methods are proposed for image segmentation and these methods are expensive. Feature extraction used for the classification of the image and image segmentation used to categorize the pixels of the image are the image processing techniques. PSO, DPSO and ABC [31] are the three methods containing the threshold methods having the time complexity issues. The proposed method with Otsu threshold method has given the better segmentation results. CSMcCulloch algorithm combined the functions of these three threshold methods by giving the better computation but with lowering the threshold quality.

To classify the urban scene [32] image fusion is done. When two or more images are combine to create a new image is the image fusion process. The proposed method extracted the characteristics and features of urban scene at pixel and feature level. ZY-3 dataset is used on which the image is segmented by combining the pixel. Detailed urban classification can be done with the results. The images are further refined to remove the effects of salt and pepper and method is also compared with spectral bands. Angular features are extracted to describe the angular information but the method is not efficient for multi-feature classification.

To identify different types of clouds is challenging, different machine learning methods are proposed. Full spectra resolution data [33] is used for cloud detection. The two channels are selected for pairing to collect the cloud information and check the temperature at different areas. The new model is proposed to categorize the areas in terms of land, sea, latitude and longitude etc. The model is used to differentiate the cloud, its properties, size and other characteristics. The method is used to quantitatively detecting the clouds and to improve cloud detection classification. Machine learning method extreme learning is applied on the model to identify and classify the clouds. The method is improved with more classification details.

For the proposed model [34] standard deviation is removed and for pairing channel functions and different cloud levels are selected. Threshold method is used for the pairing of new channel. Labeling is done using the proposed model. The proposed model is improved and it can recognize the clouds in different atmosphere and areas. Extreme learning and support vector machine methods are inefficient as the good results are not achieved. Multi-layer perceptron is applied on the proposed model with different activation functions to train the data and better results are obtained. Many machine learning techniques used the ground truth values applied on the data of the sensors to classify the clouds. The method could be expensive for all the sensors. Data of two sensors is used and histogram is used to remove the image properties differences of these sensors.

Image is transformed at pixel level for which domain adaptation framework [35] is proposed. Only one dataset Landsat-8 is labeled for training by calculating its ground truth value. Transformation is applied to upscale the dataset. Semantic segmentation is applied on the training dataset and then the dataset is manually labeled and trained on cloud detection algorithm. Cloud detection form multiple images is time consuming process. As the cloud mask has to be firstly prepared from multiple images which makes the method more complex. For this, many methods are proposed to detect the clouds from single image.

A method is proposed for automated cloud detection based on single image in which cloud atoms are selected using random fractal model [36]. The number of pixels is converted into patches and the histogram value of each patch is calculated based on their RGB color value. Another model is applied to match the patches value with histogram value as it is the threshold method and the values are cloud atom. There is no need of dataset labeling in this approach. The proposed algorithm is applied for the selection of cloud atoms and the efficiently clouds are detected and gives the better performance than the other machine learning algorithms. Multilayer cloud detection is another challenging problem; many techniques are proposed to detect the multilayer clouds.

Four machine learning algorithms [37] are applied on the dataset containing day and night images to extract the multilayer pixels. Multilayer pixels are selected from the dataset of different satellites using the properties of their instruments. Using the model parameters of machine learning algorithms (ANN, KNN, RF and SVM) dataset is trained and the new model is proposed. Dataset is classified into single layer and multilayer cloud pixels. Two models are proposed, one for day time dataset and the other for night time dataset. The algorithm gives the best results for day time dataset multilayer cloud detection but it misclassifies the thin clouds. Random forest machine learning model gives the best results while comparing with other machine learning models.

Although it is difficult to detect various forms of cloud, multiple features of cloud [38] are combined and trained on machine learning model. Spectral feature is used to extract the color information of the cloud; statistical properties of the clouds are calculated to identify the gray pixels. Cloud texture details are extracted by applying the Gabor filter and structural features are used to check the similarity of some part of the image with whole. These features are combined to propose a new algorithm and the cloud features are extracted using this algorithm. Support vector machine model is applied to train the dataset and classify the image having cloud or not.

Many spectral and spatial based cloud detection algorithms are proposed, pixel-level information [39] is used for the cloud detection. Random forest model is used to calculate the ground truth value of remote sensing images. RF model is used to detect pixels of different types of surface. Two RF models are trained and pixels are labeled using spectral features of images. Thermodynamic algorithm is proposed for the classification of clouds. Ground truth values or reference masks and classification of clouds both are obtained from RF models. The RF models advantages and disadvantages shows that both models can train the dataset efficiently and gives the better performance but there is a need to use the probability with more quantity. To detect broken or thin cloud from complex scene another scheme RFmask algorithm [40] is applied. Image Segmentation is done on Landsat-8 images combined with machine learning RF model. Pixels of an image having spectral features are extracted making a superpixel. Through this database is created having labeled images. Cloud and non-cloud pixels of the image are stored in a database for training. Pixels are passed to the RF model for training. Spectral feature property of an image is selected to train the data. Algorithm is proposed for pixel level detail and image is segmented by super pixel. RFmask algorithm is proposed in which RF model and algorithm results are combined for classification of an image. The architecture diagram of the Random forest model is shown in Figure 2.2.

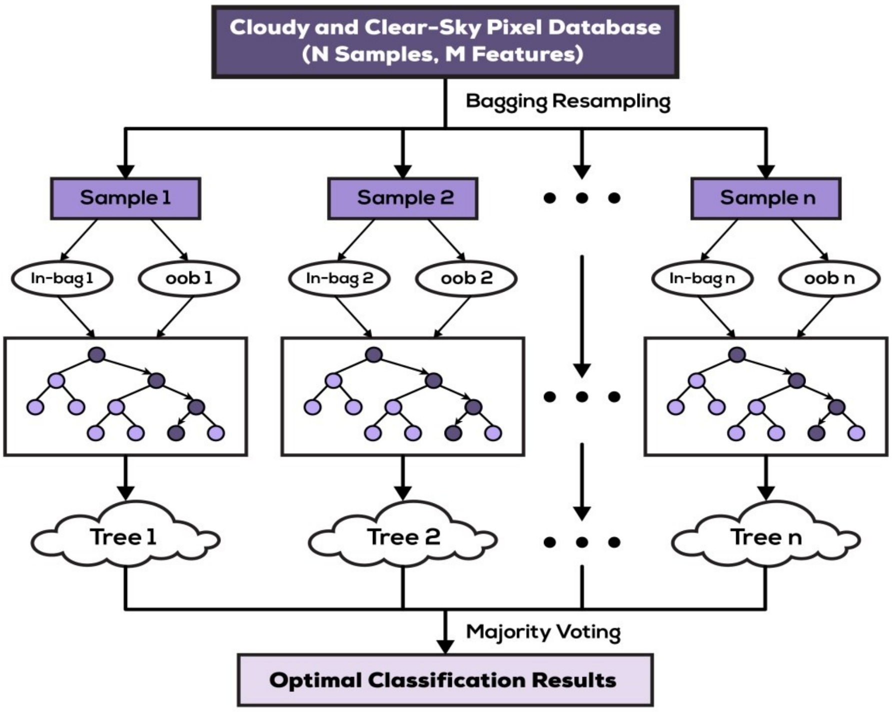


Figure 2.2: Architecture of Random Forest Model [40]

Many deep learning algorithms applied by different researchers for cloud detection are discussed in detail [41] including different types of cloud, their labeling and image processing. How to create accurate cloud mask from dataset, their use, existing datasets are discussed; and suggestion for the creation of new dataset but it could be expensive. Many pixel wise and object wise cloud detection algorithms are discussed in detail. For object wise detection, there is a need to generalize the cloud detection model for accuracy. Light weight models are more efficient as they reduce the calculations and more efficient convolution network is designed. Many traditional algorithms and machine learning algorithms are compared with deep learning algorithms in which many CNN network based deep learning techniques are analyzed.

Deep learning algorithms are applied in many applications in which pixel wise algorithms or techniques are used. Hyperspectral dataset [21] is used for training and applied on the dataset of different sensors. Cloud and no cloud spectral pixels are selected from different images having different surface types. To train the remote sensing images, multiple channel bands are used to detect cloud pixels. Back propagation algorithm based on keras network is applied to train the dataset. Each neuron layer is calculated by forward propagation and in backward feedback there is a need to correct the neuron weights and threshold. Another model is used for data classification and parameter correction. The proposed method gives the better performance and results. Another scheme [42] , which is based on deep learning and machine learning method, is used. Dataset contain RGB images converted into gray scale and extra part of the image is removed. Image is filtered to remove the noise. Image is normalized and transformed by mapping. Intensity values of the pixels are changed by the process of feature extraction. Image is preprocessed and database created from which feature extraction is done, dataset is trained on recurrent neural network (RNN) model. After feature extraction image filtration is done by evaluating filter value and random forest (RF) is applied for classification. By the combination of both algorithms optimum results are obtained. The proposed method is useful for the land monitoring applications.

2.2.2 CNN based Models Schemes

CNN based attention mechanism [43] is proposed for cloud detection. Dataset is taken from specific sensors is preprocessed by rotating and cropping the images. RGB channels of images are fused and each image is divided into blocks and some blocks are used for training. New features are added to the FCN and ResNet model to create the new architecture. Semantic segmentation techniques are proposed for the detection of clouds. Basically cloud detection is a preprocessing step in many remote sensing applications. Many techniques are proposed to extract the features of clouds from pixel based images. Another DCNN method [44] is proposed which is deep convolutional neural network method whose architecture is based on encoder decoder network. Multiple operations are performed to encode the images by passing them form convolution layer then another layer is used to extract the detailed information of the image by replacing the pooling layer. Then this information is decoded which is the reverse of encoding image. FCN is used for the large size remote sensing image. Encoding do not loss the detail of the image and decoding retrieves the loss information; the method is used to recover the loss details of the image. The results give the higher achievement than the traditional method.

CNN algorithm is developed for the detection of cloud and cloud shadow which extracts the clouds features using its spectral and spatial [45] information using infrared and RGB channels. Dataset is taken from two different satellites. CNN architecture is used to obtain the cloud mask or ground truth value of each pixel of the image patch. Large patch size of image with four channels is used and many filters are applied on each convolution and pooling layer. To classify the image labeling is done on each pixel. CNN architecture is modified with new threshold for the better performance of the algorithm. Algorithm performs well on both dataset and image is multi classified as the algorithm is able to detect both cloud and cloud shadow. The algorithm is applied for limited set of channels and only on two dataset. There is a need to generalize the algorithm for the use of global dataset. There is a need for texture feature extraction from the cloud dataset. Cloud detection is a need to remove the clouds from the Landsat images used in remote sensing applications to monitor the land. CNN is used for the detection and removal of thick clouds from Landsat images.

In the proposed CNN architecture multiple convolutional and pooling layers are used [46] to produce new cloud features. In this scheme, CNN architecture is modified by replacing fully connected layer with global average pooling layer. Three networks are presented for the detection and removal of clouds. Patches of images with their three channel values are obtained to train the dataset. The three networks are used for the detailed information of images. The method is able to not only detect but also remove the thick and thin clouds from the remote sensing images. CNN architecture is used for the detection of clouds and the architecture is based on three network structures for the removal of clouds. This scheme is not applicable if the land covered area may change so there are some limitations with this method. In another CNN based scheme, multi-feature CNN method [47] is used to detect thick, thin and non-cloud parts from the images. It is difficult to detect thin cloud; the proposed model identified the thick and thin clouds by extracting the global features of the images. The convolutional layers are combined to obtain multi-scale information of image and the spectral information of image is used to train the proposed model. Two datasets are used and taken from different earth surface which is complex and the ground truth values of the pixels are manually calculated. Some cloud and non-cloud pixels are misclassified due to complex earth surface. Performance matrix of different methods which are the previous methods is calculated in order to compare with this method and efficient results are obtained. Performance matrix is improved by comparing with other methods. Mostly the cloud detection techniques are applied on daytime image dataset or nighttime image dataset.

In this scheme [48], the technique is applied on both time day and night time dataset. Deep learning architecture is proposed which is based on three layers to identify the pixel either it is sky or cloud. Light weight architecture is applied on two public datasets. The three channel image is passed to the convolutional layer for encoding. After passing through the pooling layer the image is decoded and then the cloud masks are obtained. These ground truths are manually obtained and the threshold is also calculated. Dataset containing daytime and nighttime images is trained using binary labeling of images. Different augmentation techniques are applied to extract the features for efficient cloud segmentation. Performance matrix is calculated and compared with daytime, nighttime and, daytime and nighttime dataset. Better accuracy is achieved but there is a need to apply the model on the larger dataset to obtain better results. Multi-scale convolutional feature fusion architecture [3] is proposed for different type of satellite images. Dataset is preprocessed for the computation of cloud masks. The cloud and cloud shadow pixels are labeled by applying normalization method. Image is further divided into patches for training. Fully connected network is used for the classification of pixel. Encoder-decoder structure is used for the extraction of features of the image by passing through multiple convolutional and pooling layers of the architecture. Then the multi-scale features of the images are extracted. The dataset is trained on the architecture but good accuracy is achieved on medium and high resolution images. Clouds are detected on ten different types of images but the cloud shadow detection gives the low accuracy. The proposed architecture efficiently detects the clouds from different satellite images and best for remote sensing application. The method is also compared with other methods by taking different types of images.

In the scheme [49], a new framework is proposed in which two methods (Fmask and MSCFF) are used to compute the cloud masks for cloud and cloud shadow detection and removal. Detailed information is needed to check the position of the cloud and shadow to get the cloud masks. The image is divided into patches which are the group of the pixels and it eliminates the invalid information. It reduces the computation complexity and for the restoration of these patches, multi-scale and contextual information of feature is needed to locate the cloudy areas. Then the final results or patches are restored using aggregation iteration. The experiment is performed on two datasets effective results are produced with some limitations. The cloudy regions should be separately constructed to overcome the limitations. U-Net architecture [50] is used for cloud detection which contains the encoder decoder architecture. The method is proposed for semantic segmentation of the images. In this technique, encoder uses convolutional and pooling layer. The encoder part is used for the extraction of pixel feature and the feature representation. Decoder is used to segment these features. Encoder used the attention technique to learn the important features of the image and it ignores the incorrect information. The architecture is modified by using the attention gate which represents the features and get the useful information. Loss function is used to train the model. The proposed model modified with attention mechanism is also compared with U-Net architecture. Attention gate is also combined with the architecture for better results. Excellent results with better performance of the model are achieved.

MCNet model [51] is proposed for the extraction of multi-scale features. U-Net architecture based model is proposed with no connections in which encoder contains the convolutional layers to represent the image features on multiple channel RGB (red, green, blue) and near infrared using pyramid convolutional layer. Decoder is used to extract the semantic details of the features. It improves the cloud masks. It is not possible to extract the features from multiple scales using single convolution layer. Pyramid layer is used which contains multiple layers and other module is used for the combination of multiple channel features. More attractive features are extracted as compared to other architectures. The method is also based on attention mechanism. 38-cloud dataset is used for training and good performance metrics is achieved. Previous contains weak textual features and the results are not effective. For the extraction of texture features the model is proposed which contains convolutional neural network based on encoder decoder architecture. The proposed model [52] extract the texture features of the cloud pixel and automatically detects the clouds. Up down blocks are used with encoder decoder architecture to recover the loss of spatial information. Threshold value is used for the classification of image. Texture features are used for the segmentation of the image as more detailed texture features are extracted using down block. It reduces the loss of information. For the detail information of the object up block is used which contains the spatial and semantic detail of the features to restore the image. Pixel level classification is done and decoder allows mapping the feature from low to high resolution. The proposed method is efficient for the detection of thin clouds. The model is also compared with other architectures for the performance evaluation and the effective results are achieved.

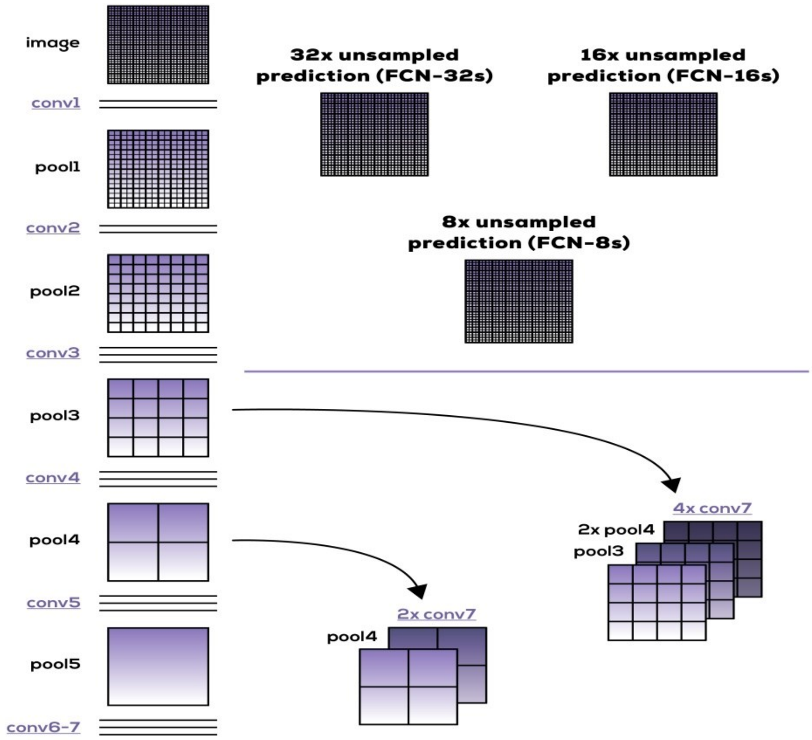


Figure 2.3: CNN Architecture [53]

CNN models achieved good accuracy for segmentation of natural images but it could not perform well for remote sensing images due to low amount of ground truth labels. In [54], cropland segmentation exploiting two labels commonly used in remote sensing datasets as weak observations. First are labels composed of individual geo-tagged points and second are image-level labels. It used the U-Net on trained label which outer performed the benchmark models such as logistic regression, support vector machine and random forest. However, neural networks performed better for larger datasets. Hence, it could be hybrid with U-net for better classification and efficient use of labels to get pixel level labels from image labels.

In [55], discussion about the research performance of rainy cloud detection and convective precipitation on GOES-ABI data with DNN had done. It performed well in terms of accuracy. However, it noticed the overestimation in the identification of convective precipitation, especially over land areas. Due to the limitations of passive instrumentation, one of the recognized weaknesses in deriving cloud information from brightness temperature was that it only provided the vertical column-integrated cloud information. Although the spectral parameters adopted in this research were effective in reflecting LWP information using differences in BT, they do not measure cloud raindrops directly as PMW data do. Therefore, the IR-BTD based method did not perform well enough in rain rate decision making. Areas of convective rain overestimated in validations and case studies. However, ABI data outperformed PMW data in both spatial and temporal resolution, providing real-time monitoring and detection. The time resolution of ABI was 5 minutes, and it produced classification results at the same interval after the built of the system. Conversely, IMERG's time resolution is 30 minutes, six times less often than ABI. The system provided rain cloud classification results at a spatial resolution of 2 km, which was five times greater than the IMERG data. It worked in predicting convective precipitation for both emergency precipitation hazards and routine weather forecasts, and was an alternative when PMW data are not available. Radar-based QPE offers a high-quality precipitation estimate. However, it deployed ground-based radar sparsely and was not available over ocean that has a high percentage of heavy precipitation (i.e., Open Ocean) and areas of most significant economic impact (i.e., coasts). GOES-ABI had full and continuous coverage of North and South America and large areas of the ocean.

In [56], the research problem was cloud masking. It used High resolution satellite images for detection of clouds in visible and multispectral imagery. Machine learning was promising domain for cloud detection; therefore, it proposed CloudFCN based on CNN based U-net architecture as it was standard for image detection. It connected the shallowest and deepest layers of the network, thereby directing low-level visible content to its deepest layers. It also offered a wide range of experiments on this topic, including data from two high-resolution sensors – Carbonite-2 and Landsat 8 – and several complementary tests. With various design options and performance-enhancing training techniques, it exhibited peak performance that is comparable to other methods, high speed, and robustness to many different terrains and sensor types.

CNN based method DABNet [57] is proposed for less computation and parameters. It is a light weight model which contain back propagation algorithm to prevent the loss of spatial information. Feature extraction is done using convolution layers containing normalization layers. By capturing long range dependencies the other module of the network is used to represent the spatial features of the image. Restoring network is used to restore the resolution by combining the spatial and semantic details of the features. Loss function is used to predict the pixel value either the pixel is cloud or not. Multi-scale features are extracted to achieve the information of cloud boundary. The method is proposed for the detection of complex clouds. High accuracy is achieved after applying on the Landsat dataset with better performance. The spectral library [58] is organized which contains the spectral features of clouds and surface features. The library contains the clouds and other ground objects as a training dataset. The surface features are removed; only cloud spectral features are organized in another new library for cloud detection. Each pixel contains the spectral details including all channels near infrared also. So, the new library contains the multi spectral details and the three remote sensing satellite datasets. Pixel level extraction is performed to get the details and the method is proposed in which there are two modules. One is to extract the hidden features and it extracts the limited information. Other is used for feature representation and it work as a classifier. Method is a combination of spectral library and the proposed model. It is compared with other networks and accurate results are obtained. The method is proposed for automatic training and automatic cloud detection.

As semantic segmentation is discussed in different studies, still there are many issues for optimum method. Due to manual labeling of clouds thin clouds are ignored. It affects the accuracy of detection method. The model is proposed which is the combination of CNN and transformer named Cloudformer [59]. In CNN module there is pyramid module for the extraction of spatial properties. The image is passed to encoder decoder architecture for feature mapping. Pyramid layer extract the multi-scale features and to determine the hidden relationship between the pixels. It enhances the features and decoder is used for the pixel restoration. The computation is performed for the spatial and shallow features. Transformer module is used for the classification and to obtain the ground truth of the image. For mask segmentation, pixel level approach is used which is the combination of semantic and contextual information. After the preprocessing of the dataset, the method is applied and also compared with other methods and efficient results are achieved. In pixel wise labeling, the other challenging problem that the labeling have to be done on limited number of the images is occurred and to resolve this issue another scheme is proposed. In this scheme [60], traditional approach is used in which the dataset is trained containing the large number of data unlabeled and few data with labeled sample. To generalize the segmentation network, the method is proposed to reduce the distribution gaps by representing the features at feature level and output level. For the classification, features are classified for selection and alignment and the predicated results are aligned indirectly. The method is proposed to reduce the distribution gaps between the labeled and unlabeled dataset. The model shows the good performance using limited number of datasets but there is a need for loss function and threshold value to make the performance more efficient. The scheme gives an effective method for cloud detection.

From existing studies, the previous methods require a lot of resources as deep learning methods need more computation power. Light weight models require less computation power. For this, LCD network [61] is developed which is applied on cloud and shadow dataset. The method contains the encoder decoder structure containing attention gate module. In encoding part, the multi-scale information of features is extracted and attention mechanism is used between encoder and decoder for thin cloud detection. Decoder is used to extract the spatial information of the cloud layer. It improves the accuracy and distinguished the ground and object pixels. The proposed method reduces the number of parameters and it also requires less computation. It segments the cloud pixels and improves the performance of the network. The proposed method gives high accuracy with low cost for cloud and shadow segmentation.

2.3 Object based Cloud Detection Schemes

Researchers proposed various methods for object detection in remote sensing images. It is more challenging to detect the objects in scene classification dataset as these are multi-scale natural images. Small objects are difficult to detect and contain few pixels in high resolution images. Different deep learning methods are proposed for object detection. New benchmark RSSOD dataset [62] is used in which main focus is to detect the small images. The new network is proposed containing convolutional layer to extract the feature of the image. The features are combined from the blocks and the other cyclic network is used. Cyclic network is used to generate the low resolution image from the high resolution. The proposed model is the combination of three networks including Yolo. Loss function is used for the selection of random images. Some quality metrics are used to improve the quality of the objects. Satisfactory results are obtained for different object detection in the same image. The proposed method with Yolo architecture gives the high accuracy and easily applied on other datasets for object detection. U-Net based architecture is proposed in the scheme [63], for remote sensing dataset. Spatial feature details are extracted and the single spectral band is used for the classification of the objects. An algorithm is proposed in which fully convolution neural network is used. This is a simple approach in which RGB channel is used for learning. The object annotation is done using Fmask algorithm. Noisy dataset is taken for cloud detection to check the performance of the model. Algorithm is used to detect the clouds using the spatial and spectral feature combination. High performance is achieved using RBG channel band. Model successfully trained the noisy dataset but there is a need for the improvement of the model. New architectures could be used for more good performance other than the U-Net. There is a need to improve the cloud detection accuracy for this purpose multiple models are proposed. In this scheme [64], encoder decoder structure based cloud detection network is proposed in which ResNet architecture is the backbone of the network to extract the features. ResNet is modified by using filters which extracts the features by increasing the performance of the network using convolutional layers. Pooling layer is also used and pyramid module helps to capture the contextual information of the image. Another module is used for the refinement of object boundaries by using fully convolutional neural network. For the classification of object, softmax function is used which extracts the classified features of the image. Experiment is performed on multiple Landsat datasets good results are achieved but there is still need to improve the model as the accuracy for thin clouds is low which needs to be high.

For land cover information, different deep learning methods are applied on high resolution remote sensing images for image classification and object detection. Different researches are proposed for object level detection [65] using deep learning methods. CNN based architecture is proposed to extract the feature representation to identify the boundary of the object. Then the features are captured from these representations for classification. Intersection over union (IOU) is used as performance metrics for thresholding. Method proposed for real time object detection which is based on Yolo architecture it directly predicts the feature maps and classifies the objects. High spatial details of the image helps in land cover classification using deep learning approaches. Google Earth images are used as a dataset for land cover object detection in which RGB bands are used to create box for the identification of the objects. Pixel level detection which is semantic segmentation gives better results than the object level using deep learning approaches. But the deep learning methods give satisfying results than the traditional approaches for object detection. High resolution image classification is still problem due to the complex spatial information of image. CNN based architecture for object detection is used in the scheme [66] resolve the pixel level CNN issues. OCNN model which is object-based contains two images as an input, one is labeled and other is context patch. Then the image is cropped taking the object with its background containing large pixels of objects in an image. The geometric characteristics of the object are extracted to discriminate two different objects. To recover the loss details of geometric characteristics object deformation is done. Deep features of the object are extracted which helps to predict the object boundaries. The network contains the stacked blocks including convolutional and pooling layers of CNN. The contextual information of the object is used for classification and activation function is used. The method is compared with other benchmark methods and high performance is achieved with low computational cost. Pixel level requires more time than the object level. But it is difficult to detect the small objects using geometric characteristics. K. Itakura used Yolo algorithm for automatic tree detection in the scheme [67]. The method is proposed for tree identification, its location and counting details. Firstly, the dataset is preprocessed for tree trunk diameter and height estimation measurement. The images of the datasets are divided into grid of cells. Each cell is computed to check its probability whether it contains the object or not. The images are classified as each cell consists of different colors. The boundary boxes are created outside the color using Yolo v2 method. The bounding box is calculated using class probability and intersection over union score. The Figure 2.4 shows how to obtain the boundary boxes using Yolo detector [67]. Yolo is also deep learning model used to extract the features. Image classification is done on ImageNet dataset using ResNet model. Dataset is trained on ResNet architecture and the class and location of the object is computed. Anchor box is used to calculate the exact location of the boundaries. It calculates the width and height of the box. The number of grid cells and size of feature map both are calculated. Dataset in the form of images are taken from video frames as they are 3D images. High performance is achieved by applying the method on the dataset the objects are detected. The method helps in investigating the tree diseases by observing the leaves and their classification. The method should be tested with more tree images for more high accuracy and better performance.

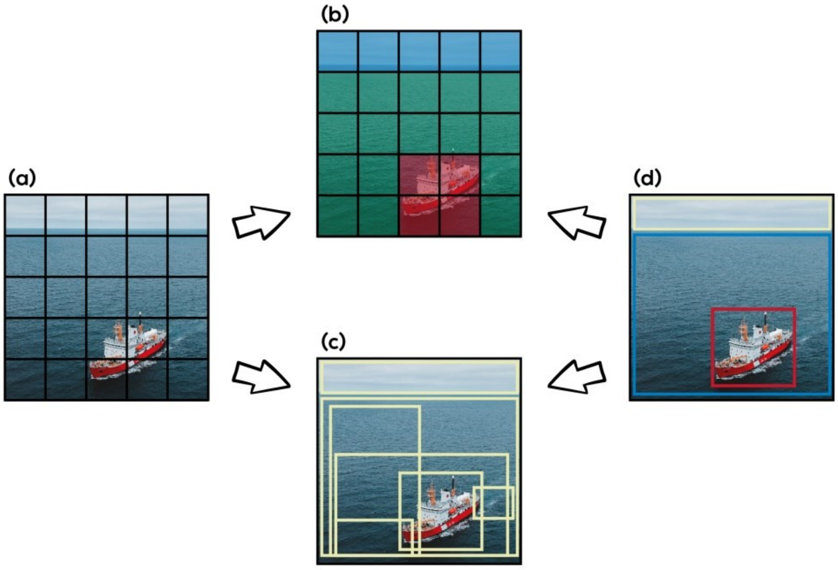


Figure 2.4: Yolo algorithm detection [67]

The new version of Yolo is proposed for object detection which is Yolo v3 [68]. The architecture is based on DarkNet model which contains the number of parameters for calculation. The structure is divided into different blocks and the channel bands are combined using connecting layers to extract the features that are again used. The back propagation algorithm is used for enhancement. The network is used for feature representation and extraction. The algorithm is proposed for the multi-scale feature of the object in which image is divided into grid of cells for feature mapping. The detecting layer is used for the detection of the objects. Remote sensing images contain large number of objects. For the prediction of object boundaries, anchor boxes are used in Yolo. K-means algorithm is used for anchor box creation and the ground truth IOU has been calculated which results with large values. The boundary box values are calculated and several values are obtained. The algorithm compared the value of boundary box with other values to detect the required object. The experiment is performed on the datasets and high performance is achieved. Better accuracy is obtained by detecting the small objects. Yolo give the low performance for complex scenes but the proposed method performs well with Yolo new version. DAFNet architecture is used in the scheme [69] which is based on encoder decoder architecture. Encoder is used for feature extraction global attention map is used instead of convolutional blocks. Max pooling layer is used in the network and feature learning is done. The module is used in which global features are captured for feature alignment and combining the global semantic information to check the relationship between the pixels and it makes the feature mapping easier. Feature decoder is used to get multiple outputs. It is also used spatial details and channel bands for feature representation. These details also linked the semantic information. These feature details are restored in decoder phase. Loss function is used to capture the multi-scale information. There are also some challenges in dataset including small objects. Dataset is constructed and attention mechanism is used for feature learning. It maps low level features to high resolution and avoiding the global contextual information. Benchmark dataset is used in which image scenes are complex; containing ships and clouds objects. The scenes are also distorted because of mixed objects. The dataset is larger and more challenging due to these issues but the deep learning method is applied on it. Better results are achieved with good performance. Different CNN based deep learning methods are used for object detection and efficient results are achieved. Method is proposed named few-shot learning [70] in which few labeled samples are used for unseen object classes. Object detection is challenging as the method needs to identify the location of the object first. Meta-feature extractor is used to extract the single-scale features of the object. Remote sensing images contain the objects of different sizes and for this multi-scale feature extraction is used. So, the DarkNet model is used for this purpose. The light weight CNN model is used to map the features on each scale. Three groups are created for feature map for object detection. These feature mapping calculate the object location and also the boundary boxes. Yolo v3 is used for creating the grid of the cells. The anchor box is used and created at each pixel location for the predication of the object location. The dataset is divided into two sets, one for training on seen object classes and other set for unseen object classes. Experiment is performed on different methods to evaluate the performance. Two public benchmark datasets are used and better results are achieved but the method needs more improvement.

Random forest [71] which is machine learning model is used for cloud detection from remote sensing images using multiple channels. The model is used for classification. It is strong learning model and it works as a single classifier. It is used because of its small learning samples and fast computation and accurate classification. It uses the multiple classification features and radiation correction is performed at training phase. In this phase, the texture and spatial features are extracted used to train the data for cloud detection. In testing phase, the trained images are passed as an input to the RF model and results are obtained. Filters are used for accurate cloud mapping and to compute the cloud covers. Radiometric calibration is done to remove the error and for this normalization is done between the parts of single or multiple images. Feature extraction helps to increase the accuracy. Single feature gives the results but the multiple features gives better accuracy. It improves the cloud detection by capturing the texture details of the image and the channel band including RGB and near infrared band. Channel bands helps to detect the cloud by discriminating the clouds from different ground surface types. It is more challenging to detect thin clouds for this RGB color converted into IHS which is pixel value of the cloud in intensity channel. The intensity and saturation of the image is calculated. Dark channel features are also extracted using some formula. Whiteness index feature is proposed to remove the non-cloud pixels. To distinguish clouds from ground surface, all the features of the images are mapped with their probability value. Gabor transform feature is used to obtain the scale properties. The method is applied on different datasets and high accuracy is achieved. It is also compared with deep learning methods which conclude that it takes less training time and less dependent on computing device performance. Ship detection from remote sensing images is another challenging problem. For this, the scheme [72] is proposed in which deep learning approach is used containing Yolo v3 algorithm. Object detection is done by passing some images as an input to the model. Boundary box function is used and then normalization is done on the images. The encoded pixels are dropped using some criteria. The hash algorithm is applied to the images to count the objects. Two modules are presented and the other is used to train the dataset. Two algorithms one used but the Yolo v3 algorithm is used for ship detection. The images are selected by using function and then combined. Function is used to define the boundary box. Then the dataset is normalized by removing some extra pixels from the boundary box. Hash function is used for the combination of counting and boundary box location. The other algorithm is used to count the number of objects which is the number of ships. The trained images are passed to this count algorithm and anchor boundary boxes are used in each layer. Incorrect classification is avoided by the classifier. Each image contains multiple ships and each ship is represented by these anchor boundary boxes. If the objects are closer to each other the boundary box overlaps them. The third algorithm is used for the calculation of hash value. The proposed model is accurate for real time dataset. It is compared with other models and it gives the better performance. It reduces the processing time of the method.

Automatic Yolo based ship detection method [73] applied on the remote sensing images and used in practical remote sensing applications. The improved version of Yolo algorithm is proposed for the object detection. The method is based on CNN deep neural network. The method consists of two stages which are one stage and two-stage. The one stage helps to increase the speed and the two-stage gives the high accuracy to detect the position of the object. For real time detection, Yolo model is improved which contains Dilated Conv, Depthwise Convolution (DW Conv) and SE layer modules. To improve the robustness, anchor computation and image scaling is done for the dataset. Dilated Conv and SE layer modules are used for the feature extraction. The features of the images are obtained which contains multiple objects with different sizes. The image is divided into small parts and passed to the convolutional layer. The output is in the form of feature representation. The modules are used for the multi-scale feature extraction at different levels. The feature fusion is linked and the multi-scale information of the features is extracted using Dilated Conv module in Yolo network. Multiple parameters are used for more feature extraction in this module and it takes less computer resources. The Yolo algorithm is modified and it resolves the issue of failed detection by using standard convolutional layer. Semantic feature helps to extract the object features which are of different sizes. DW Conv module is used to reduce the number of parameters using down-sampling and SE layer is used to filter the features. The method is applied on the dataset and high accuracy is achieved. There are still some issues like false positive detection, noisy data and to locate the objects of shapes. SR-Yolo model [74] is proposed for the detection of the plane. Low resolution images give less accuracy on complex scenes. For this purpose, convolution neural network is used; it firstly increases the size of low resolution image to its object size. Then the feature mapping is done using convolutional layer. Residual network is also used for the enhancement of the image. Texture features of high resolution image are obtained and loss function is used to recover the loss details of texture feature. Object detection is based on two stages. One stage classifies the object area and calculates the probability and coordinates the value of the object. The accuracy of one-stage algorithm is low and two-stage algorithm is high while the speed of one-stage algorithm is fast and two-stage is slow. For this boundary box algorithm is used to detect the objects. Weakly supervised learning algorithm is proposed for plane detection which involves end to end training of dataset. Quality related features are extracted and the trained features used more information for object prediction. SRGAN module is used for unstable training process and to enhance the image quality. The small objects are detected by the Yolo algorithm by storing the positioning signals in low level features. This is done to use the short path detail of the object. Noise is caused by pooling layer so it is avoided. The experiment is performed on the datasets. The method is proposed to solve the issue of false object detection. The method improves the performance and reduces the complex computation.

2.4 Comparison of Existing Cloud Detection Schemes

Different cloud detection methods with their dominance and limitations are given in the table.

Table 2.1: Existing Cloud Detection Schemes Comparison

Methods	Datasets	Dominance	Limitations	Results
Algorithm using Radiative Transfer Equation [23]	LandSat-8 Images	Thin clouds are detected using spatial details.	Effective results are not obtained because the number of thin clouds is less.	96%
Spectral Unmixing Technique [24]	LandSat-8 OLI MS Images	Thin clouds are detected and removed.	Unable to identify the thick clouds.	22% improvement from previous studies.
Multi-scale Decomposition [25]	LandSat-8 Images	High accuracy achieved using texture features.	Spectral data is required for misidentified cloud area.	90%
Hybrid Multi-scale Features [27]	GF-1 Images	The method reduces the speed.	Differences in parameters are needed for more accurate detection.	90.8%
Fuzzy Logic and	MSG SEVIRI	Both networks are	Thin clouds are	99.6% and

MLP [28]	Images	efficient as they train without meteorological information.	not detected due to its spectral detail similarity with clear sky pixels.	88.4%
Modified Differential Evolution (MDE) Algorithm [30]	Satellite Images	It enhances the image contrast and brightness.	The method is computationally complex and requires high computing power for cloud detection.	99%
Cuckoo Search Algorithm [31]	Pl´eiades Satellite Imagery	It reduces the execution time complexity.	More threshold values are required for segmentation.	98%
Angular Difference Method [32]	ZY-3 Images	High level classification is done due to the angular properties of the image.	Optimal performance is not achieved due to multi-level integration of the network.	90%
Extreme Learning [33]	Cross-track Infrared Sounder (CrIS)	Infrared data provide the good classification results.	More accurate dataset is needed to improve the performance.	80%
Multi-layer Perceptron (MLP) [34]	Cross-track Infrared Sounder (CrIS)	Channel pairing of the features of images improve the results.	Ensemble techniques provide better pairing.	84.5%
Domain	LandSat-8 and	It reduces the	During result	92%

Adaption [35]	Proba-V	statistical differences between the images.	generation some blue and green channel bands are missing.	
Random Fractal Model [36]	NWPU-RESISC45	It reduces the requirement of input data.	The method is time consuming and the speed is low.	88%
Multi-Layer Algorithm [37]	CPR, CALIOP, MODIS, AHI	The accurate detection due to radiative reflectance and temperature.	Thin clouds are not detected due to misclassification.	85% Day Time and 79% Night Time
Multi Feature Fusion [38]	Panchromatic Images	Point and block features of images gives high accuracy.	Small samples of datasets are used.	96%
Random Forest Model [39]	CALIOP, VIIRS, MODIS, ABI, AHI, SEVIRI	Applying on other datasets gives the efficient classification.	The training quality is poor because of spectral structure of the image.	98%
RFMask Algorithm [40]	L7 Irish	Thin and broken clouds are identified due to spatial and spectral information.	Low resolution spatial details lack to differentiate snow and clouds.	93.8%
Deep Learning Algorithm [21]	LandSat-8 OLI, NPP VIIRS,	It takes the minimal trained samples.	Due to limited spatial details training dataset	94%

	MODIS		exclude the surface types and cloud shapes.	
Recurrent Neural Network – Random Forest Algorithm (RNN-RF) [42]	UC Merced (UCM)	Feature extraction and accurate classification is done produces the better performance.	There is a need to enhance hyper parameters and correct classes prediction rate is too low.	87%
Fully Convolutional Network (FCN) [43]	LandSat-8 Infrared	High level features and multi-scale information helps to extract the features.	The details are loss due to compression and cropping of the images.	95.3%
Spatial folding-Unfolding Remote Sensing Network [44]	Gaofen-1 (GF-1)	It enhances the images using cloud area segmentation network.	It losses the ground information and hence there is a reduction in accuracy of the method.	96.8%
Convolutional Neural Network (CNN) [45]	WV-2 and Sentinel-2 Imagery	It eliminates the complicated algorithm which contains threshold values.	The limited Infrared bands can improve the performance.	88%
Convolutional Neural Network (CNN) [46]	ZY-3 Satellite Images	Multi-layer non-linear transformation is done which helps in cloud removal.	It cannot use the multi-source data and cannot apply on land cover dataset.	90%

Multi-scale Features-Convolutional Neural Network (MF-CNN) [47]	LandSat-8	Complex clouds are identified by using low-level spatial and high-level semantic information.	Thin cloud detection is not done.	93%
CloudSeg Net [48]	SWIMSEG dataset	Day and night cloud detection is done using image segmentation.	Large dataset is needed and there are problems related to images which improve the performance.	92%
Spatio-temporal Patch Group [49]	Sentinel-2 MSI and LandSat-8 OLI	The loss function gives the optimal training model.	Cloudy areas are not constructed without temporal information.	81% and 79%
Cloud-AttU Network [50]	Landsat-8	Multi-scale information is achieved using low-level and high-level features.	Noise clearance issue decreases the performance.	97%
Multi-scale Content-aware (MCNet) [51]	38-Cloud	Deep Semantic information helps the accuracy of cloud masks.	There is a need to enhance the model for optimal results.	96.4%
UDNet [52]	Goafen-1 (GF-1), WFV images	Thin clouds are detected using dark channel prior module.	More color channels can improve the model.	98.5%
DABNet [57]	GF-1, GF-2	It reduces the	The model is not	97.9%

	and WFV	number of parameters and computations.	performing well in order to detect the thin clouds.	
CD-SLCNN Algorithm [58]	Landsat-8 OLI, MODIS and Sentinel-2	Clouds of different sizes and types are detected.	There is a misclassification as it unable to distinguish clouds from ground objects.	95.6%, 95.3% and 94.2%
Cloudformer [59]	AIR-CD and 38-Cloud	Rich feature information is extracted for correct position information.	The speed of the model decreases.	98.2% and 96.3%
Semi-Supervised Learning (SSL) Network [60]	Landsat-8 OLI and WFV GF-1	Limited number of labeled samples gives the better accuracy.	Auxiliary information is needed for detection on different surface types.	97.1%
Light-Weighted Cloud Detection Network (LWCDNet) [61]	GF1_WHU and L8 Biome	It reduces the feature redundancy and false detection rate.	Low cost and cloud segmentation is needed.	89% and 94.6%
Remote Sensing Network (RS-Net) [63]	Landsat-8 Biome and SPARCS	Accurate clouds are detected using spatial and spectral features.	Hyperparameter optimization is needed for global features.	93.8% and 92.4%
Cloud Detection Network	Landsat-8 and WFV GF-1	Multi-scale texture features are	Satisfactory results are	97.1% and 96.7%

(CDNet) [64]		extracted without the loss of information.	obtained due to omission pixels of thin clouds.	
MLCG-OCNN Method [66]	Vaihingen and Potsdam trained dataset	Deep features are extracted which refine the images.	Small objects are not detected because of their irregular shape.	80% and 81%
Yolo v2 Algorithm [67]	LiDAR 3D Images	Automatic object detection using multi-spectral images.	Accuracy of diameter estimation needs to be improved.	98%
Yolo v3 Algorithm [68]	RSOD and USC-AOD	Feature extraction network is improved to enhance the small object detection.	It reduces the speed of the system.	79% and 86%
Dense Attention Fluid (DAF) Network [69]	ORSSD dataset	Low features are mapped to high features avoiding the global details.	It is inefficient to deal the tiny objects.	85%
Few-Shot Learning [70]	DIOR dataset	Correct object location is detected using boundary boxes.	Objects of all classes are not identified.	76%
Random Forest Cloud Detection (RFCD) [71]	Sentinel-2 and GF-1	It reduces the dependence of channel band information.	The features containing RGB and NIR cannot distinguish the cloud from ground objects.	95%
Yolo v3	Kaggle ships	It reduces the	The method needs	93.4%

Algorithm [72]	dataset	processing time of the method.	to be performed with video input.	
CYSDM Method [73]	Ship datasets	It reduces the number of parameters.	There are some issues like false positive detection and noisy data.	98.6%
SR-Yolo Method [74]	UCAS-High Resolution dataset	It reduces the path detail of the object.	Noisy data obtained as an output.	96.1%

2.5 Research Gaps and Directions

Different classical deep learning and machine learning methods are proposed for cloud or object detection. In the above sections, many methods are discussed in literature. These methods are efficient but there are still some challenges that are not focused. These challenges are as follows:

- For accurate classification of the images, more classification details are required. The methods used take less time and requires less computation cost and computer resources. Complex and thin cloud detection is still challenging.
- Many methods are proposed for semantic segmentation of the images. The feature details are extracted for the object identification but some information loss which in return gives the low accuracy rate.
- Limited channel bands are used and texture details of the features are not used for the extraction. It also causes the loss of information and the proposed model misclassifies the thin or small objects.
- Some models are trained on specific datasets. If they are applied on other datasets then they perform inefficiently due to different spectral and spatial details of the image features.

- Mostly cloud detection techniques are applied on day time dataset. Some techniques are applied on night time dataset with better accuracy. But with day time dataset there are still some deficiencies as the network misclassifies the thin clouds.
- There is a need to find the exact boundaries of the clouds and the better image quality for the correct object detection.
- Limited datasets are used for testing. Large dataset can evaluate the performance of the model and measure the correct accuracy.
- Generalize network is needed for semantic segmentation which protect the features information from loss and exact details are extracted.

Machine learning applications are used for the cloud detection to train the dataset and to eliminate the earth surface. AHI model [75] is used to detect the clouds in daytime and also in nighttime. The model is useful for satellite and climate prediction applications but there is a need to accurately detect the clouds. Semantic segmentation methods are proposed for the pixel wise detection of clouds. Three dimensional U-Net [76] which is deep learning method is used to detect the clouds taking color channel information of the images. The method is useful for remote sensing applications but having small dataset and complex training. Many schemes are proposed for cloud detection at pixel and object level. The pixel level approach is fast but time consuming as compared to object level. Different methods including classical, machine and deep learning based approaches are proposed in these schemes. All the methods are efficient in terms of speed, performance, accuracy, memory, computation cost etc. based on which network is used. But these methods still have some deficiencies. The challenges are also discussed in their schemes and the future work is also proposed. The semantic segmentation of clouds is still challenging which is pixel level cloud detection and there is a need to improve its accuracy.

Based on the above findings, there is a need for the accurate cloud detection from remote sensing images. The proposed model contains the semantic segmentation of clouds for better accuracy and applies the algorithm on the high resolution image dataset.

2.6 Summary

In this chapter, different cloud and object detection techniques are discussed. Their methodologies are also discussed in detail. Procedures are explained which are used to detect clouds and other objects at pixel level and object level. Comparative analysis of the schemes is also discussed in terms of methods, datasets with their limitations. Research gaps are also identified and discussed in detail. Pros and Cons of these methods are also discussed in this literature.

CHAPTER 3

DEEP LEARNING METHODS

3.1 Overview

Many algorithms are proposed for cloud detection and removal. Semantic segmentation is done which is pixel wise or patches wise detection. Many methods are presented for object detection or scene detection through semantic segmentation. Some of the deep learning pixel wise segmentation methods are discussed in this chapter.

3.2 Yolo Method

Many fast and accurate algorithms are proposed for fast and accurate object detection. Classifiers are used to detect the objects, some models used the R-CNN architecture which classifies the object by passing the classifier to the boundary of the object then other methods are used for the refinement of the object and to remove the repeated objects. These methods give the high accuracy but their performance is slow. Yolo which is object detection method [20] is proposed which not only detect the object but also indicate the location of the object. It is fast object detection method, uses the boundary boxes for the detection of the objects. The model contains the convolutional layers. The architecture diagram of the algorithm is shown in Figure 3.1.

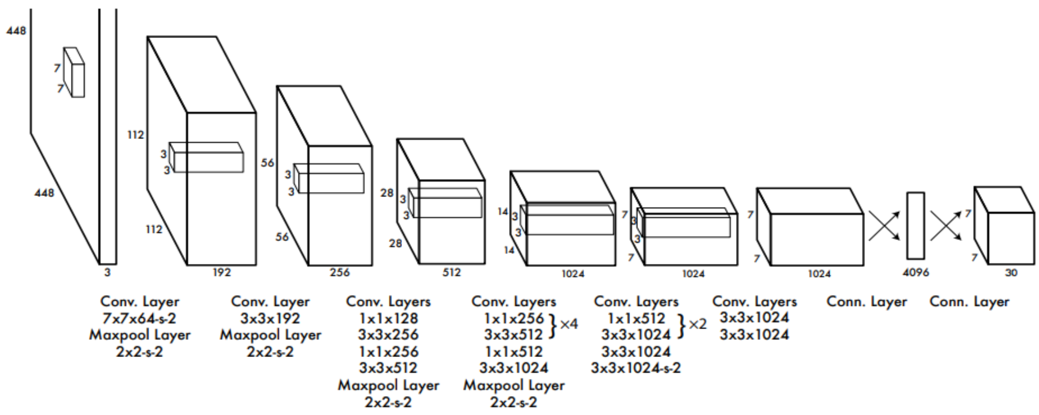


Figure 3.1: Architecture of Yolo algorithm [20]

The architecture of the Yolo shows that the convolutional layers are used for the extraction of the features of the image with probabilities. To classify the image GoogLeNet model is used and the model contain multiple convolutional layers with two fully connected layers. To detect the object using boundaries few convolutional layers are used and filters are also used for their refinement. For fast detection training and testing parameters are same. The model is trained on the ImageNet dataset for the classification and only few convolutional layers including pooling and fully connected layer are used for training. Features of the image are used for the detection of all boundary boxes of all objects.

The proposed model needs correct spatial information to detect the boundary boxes. It predicts and detects the small objects and generalizes them. Errors are also detected using the boundary boxes. The method is also compared with other detection methods in which sliding window, region based and other different detection methods are included which are classifier-based approaches, but Yolo gives the better performance. It is proposed for general-purpose and fast object detection.

3.3 DNLNet Method

The non-local model based method is proposed which uses the self-attention technique. For pixel wise modeling, features of the pair of the pixels are extracted and their relationship is measured. The detail information of single pixel is also computed. The need of the information for the pair of pixel is to identify their dependency and influence on each other. The visual characteristics of each pixel are observed and the dataset is trained on two networks individually and both models are present in non-local block. One model learns the position of the object and the other model detects the boundaries of the object. When both models are combined in the non-local network the clear results are not obtained and all the feature detail is not extracted clearly.

The new network DNL [77] is proposed which contains the softmax function as shown in Figure 3.2. In this scheme, non-local block is discussed in detail due to which this network is proposed. Both models which are discussed above are used in this network but with new function and proposed matrix. The network is the combination of non-local block, pairwise and unary non-local. After passing through the convolution layer, the dataset is passed to the generalized block of attention technique for feature extraction. The purpose of the network is to eliminate the influence of both models and to achieve the better results. The pixels of the image are then trained on non-local block for pairwise pixel relationship. The pairwise model determines the relationship between two pixels and their impact on each other. The computed output tells the region of the object through these pixels. Then the other model is used for the detection of clear boundaries of the object. Softmax function helps the models to separately train the dataset. The network learns the region and boundaries clearly. This semantic segmentation is applied on benchmark datasets and efficient results are produced.

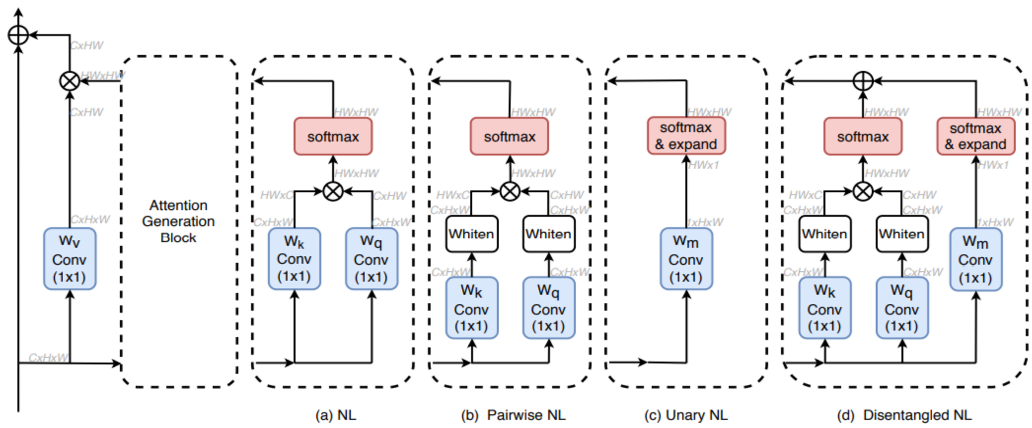


Figure 3.2: Architecture of DNL Network [77]

3.4 DDRNet Method

For road scenes dataset, deep neural model named deep dual-resolution network [78] is proposed. Almost all the methods contain encoder decoder architecture. In this technique, encoder is used to decrease the spatial details of the pixels and to get the context based information. It works as a backbone and reduces the computation cost. Decoder is used to recover the resolution of the pixels. Some details loss in encoder decoder architecture which is recovered by two path-way architecture. It consists of two modules; one for the semantic information and other for the multiple feature extraction. Pyramid pooling module is used to extract the contextual information of the pixels.

The proposed model is the combination of two modules. One is deep dual-resolution whose backbone is ResNet deep learning model. It is used for feature representation of the pixels to extract the spatial and semantic detail of the image. Many convolutional layers are also used to map low resolution features to high resolution. The other module is pyramid pooling for the extraction of contextual detail. The method is applied on benchmark datasets and it gives the accuracy. The method increases the speed and requires less computation.

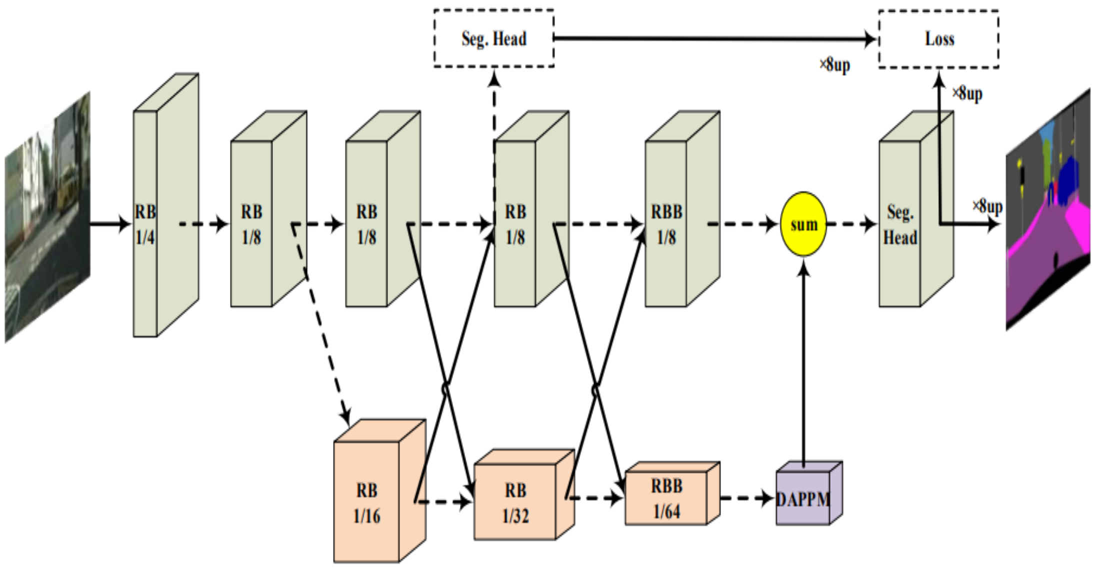


Figure 3.3: Architecture of DDRNet Method [78]

3.5 PP-HumanSeg_Lite Method

Another scheme is proposed for image and video dataset named as portrait segmentation [79]. Portrait segmentation is a sub type of semantic segmentation. Different real time semantic segmentation approaches are proposed but optimum results are not achieved. Many methods are proposed and general semantic segmentation is done which is pixel-wise classification like if it detects the bird it ignores the completeness of the bird object. For this completeness of semantic results, the new model is proposed. A new framework is proposed in which labeling is done at pixel and video level. The method is proposed to improve the quality of semantic results. It is an open source light weight model for portrait segmentation in which video portrait dataset is used.

In the proposed method, the ground truths of the images are obtained and the connected components are extracted from these ground truths for image processing. The connected components are the ground truth and the predictions. Encoder decoder network is used to extract the features but it loss the spatial detail and lower the resolution of the feature map. So, the encoder decoder network is used to connect the texture features of the images to get more enhanced ground truth. Semantic connectivity model is used to check the consistency to get the ground truth labels and different convolutional layers are also used. The proposed method is compared with other algorithms and gives the better results and requires less computation cost. The Figure 3.4 shows the lightweight model of the method.

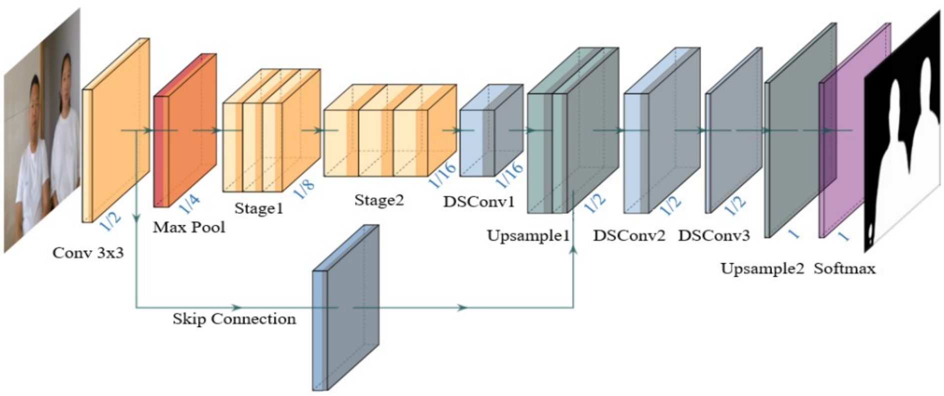


Figure 3.4: Lightweight model for Semantic Segmentation [79]

3.6 Summary

This chapter discussed the different deep learning methods for the semantic segmentation. These methods are used for object detection and for action recognition tasks. Different popular benchmark datasets are used which contains the natural images. The objects are detected with their high resolution ground truths. These methods are used for cloud detection which is discussed in chapter 4.

CHAPTER 4

METHODOLOGY

4.1 Overview

Different datasets are used for cloud detection. 38-cloud and HRCD datasets are publicly available and used in the proposed solution. The datasets and the proposed solution both are discussed in this chapter. How the performance of the method is measured and the comparison of algorithm is also discussed.

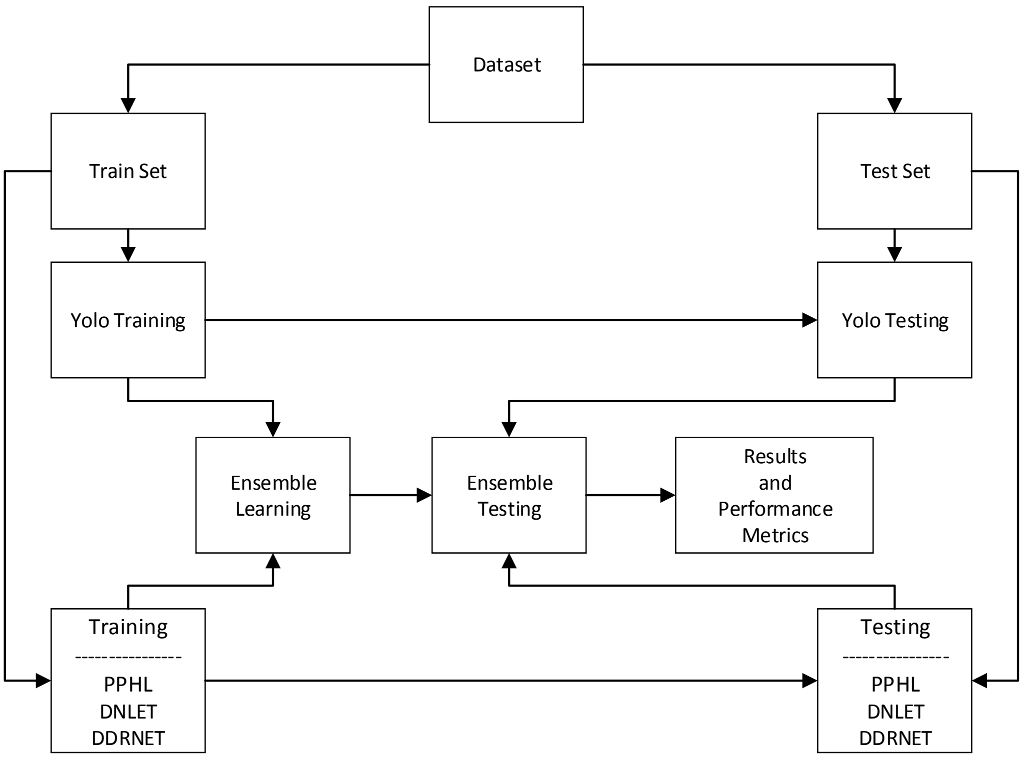
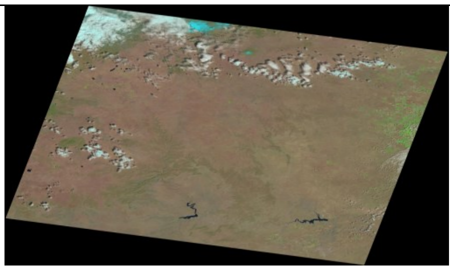


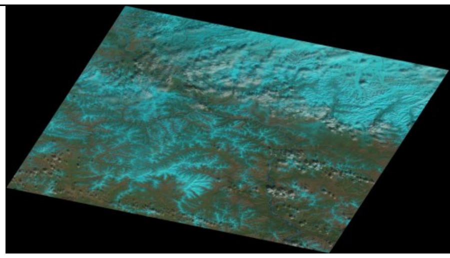
Figure 4.1: Block diagram of the proposed method

4.2 38-Cloud Dataset

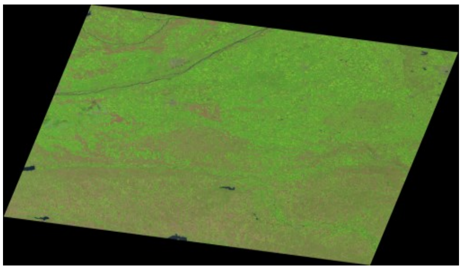
The dataset is available at different websites. The images of the dataset are taken from Google Earth. The dataset is used for cloud detection. The cloud masks or the ground truth of the images are manually extracted. The datasets contain 38 scenes or the images. The images are further divided into patches as it is not possible to process the whole image. The RGB channel of the image is used. From these patched, 80% is used for training and 20% data is used for testing. The images of the dataset have low resolution and contain the pixel resolution. The dataset is also used in these schemes [4], [80] and gives the better accuracy.



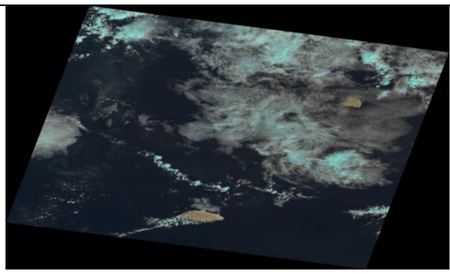
(a)



(b)



(c)

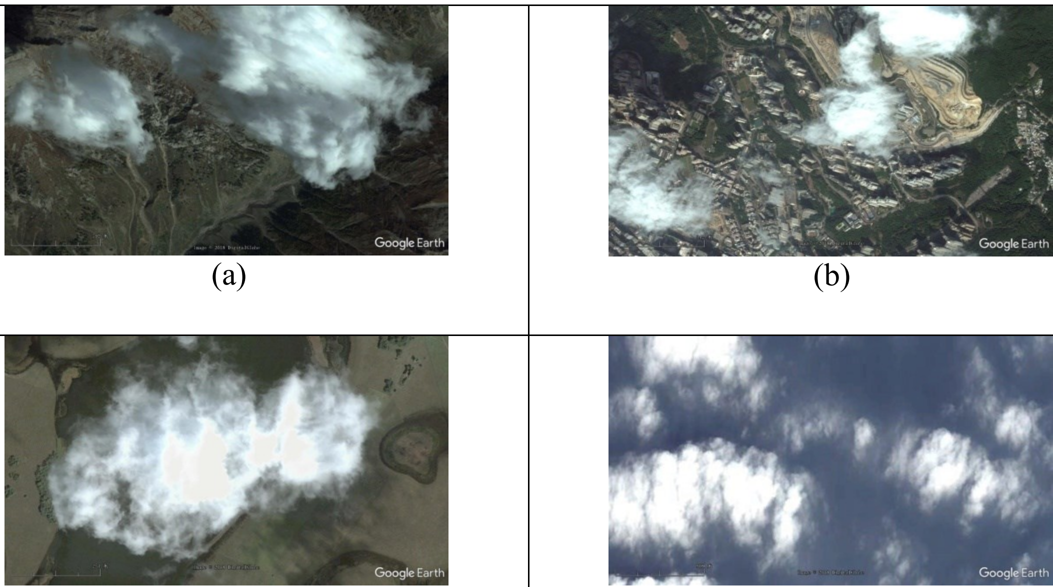


(d)

Figure 4.2: 38-Cloud dataset [4], [80]

4.3 High Resolution Cloud Detection Dataset

High Resolution Cloud Detection (HRCD) dataset is taken from the scheme [3], named as HRC_WHU. Dataset is used for training and testing purpose applying deep learning models. Total 150 images are taken and the pixel resolution of all images is not same, it varies from 0.5 to 15 m. The images are taken from Google earth and these images are from different satellites containing aerial images. The images with their RGB channels are obtained. The cloud masks of these images are computed. The images are manually labeled using Adobe Photoshop and its tools. The cloud and non-cloud pixels are labeled as 255 and 0. The dataset is publically available. Some images and their masks are removed due to cloud and cloud shadow errors. The 80% of the data is taken for training and rest is used for testing. The whole image cannot be processed, so the patches of each image is taken and used for training purpose.



(a) (b) (c) (d)

Figure 4.3: HRCD dataset [3]

4.4 Ensemble Learning

Ensemble learning approach is applied in the model in which multiple models are combined to give the better performance. In our proposed solution, above mentioned datasets are used. Images of 38-cloud dataset are firstly divided into patches. The patch size of each image is 1024x1024. HRCD dataset is used whose cloud masks are already with the dataset. Hybridization model is used in which different algorithms are combined with Yolo algorithm for better results. DNLNet, DDRNet and PP-HumanSeg_Lite are the methods used in which each model is combined with Yolo algorithm. Seven vectors are used in which the RGB values of datasets, probability values of Yolo and the probability value of one of the algorithm is computed. Then the ground truth of all the three models is calculated.

4.5 Random Forest

Random forest which is a machine learning model is a tree like structure is efficient for training purpose. The model is used for the classification of the dataset. It works as an ensemble. The features of the dataset are calculated by the model and after training it construct the classification model and also done the cross and separate validation.

Both datasets 38-cloud and HRCD are used by individual methods DNLNet, DDRNet, Yolo and PP-HumanSeg_Lite. As these methods are applied in previous studies for semantic segmentation, object detection and video action recognition. The methods give the better performance with less computation and memory applied on previous studies and datasets. Random forest is used to train the dataset. In 38-cloud dataset, total test images are 165 with 1024 pixel resolution and HRCD dataset contains 30 test images.

4.6 Performance Metric

Performance metric is basically used to check the measurement of the technique applied. It tells the working and performance of the method. Small errors can be identified and it helps for the improvement of the method. It is used to evaluate the performance of the proposed model. It is the calculation of all the correct predicted values. The computed correct results from all the total values where the true positive and true negative values both are considered as correct.

In our proposed method, accuracy is used to evaluate the performance of the models which is the calculation of all the correct predicted values. The computed correct results from all the total values where the true positive and true negative values both are considered as correct. All the four methods are individually applied on both datasets and then the three hybrid models are proposed in which these three models are combined with Yolo algorithm to achieve the better accuracy and the better results are obtained.

4.7 Google Co-Lab

All the implementation is done in python using Google Co-Lab; its GPU specification is as follows:

Table 4.1: GPU Specs

Parameter Descriptions	Google Colab
GPU	Nvidia K80/T4
GPU Memory	12GB / 16GB
GPU Memory Clock	0.82GHz / 1.59GHz
Performance	4.1 TFLOPS / 8.1 TFLOPS
Support Mixed Precision	No / Yes
GPU Release Year	2014 / 2018
No. CPU Cores	2
Available RAM	12GB (upgradable to 26.75GB)
Disk Space	358GB

4.8 Summary

This chapter includes the explanation of the datasets which are used for the cloud detection. The technique is discussed in detail which is used for the proposed model. The experiment is performed on which platform that is also discussed with its specification. The model which is used to train the datasets is also discussed in this chapter.

CHAPTER 5

PERFORMANCE EVALUATION

5.1 Overview

In this chapter, the result of cloud detection methods on both datasets in terms of performance metrics is given. The graphs show the comparison between the methods on both datasets.

5.2 Performance Metrics in terms of Accuracy

It is the calculation of all the correct predicted values. The computed correct results from all the total values where the true positive and true negative values both are considered as correct. Below equation [81] shows the accuracy equation which is the combination of true positive and negative values by all the values.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{5.1}$$

Above equation 5.1, contains four values which are True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). True positive gives the correct positive class in result, true negative tells the correct negative class, false positive wrongly predict the positive class and false negative wrongly predicts the negative class. Confusion matrix is shown below in Figure 5.1. Cloud detection methods are applied on both 38-cloud and HRCD datasets. It is shown in the table 5.1, when the methods are applied individually on both datasets the Yolo algorithm has given the low accuracy. When Yolo is combined with other methods individually, it gives the high accuracy. Yolo + PPHL give the highest accuracy which is 96% on 38-cloud and 93% on HRCD dataset as compared to other hybrid models. Figure 5.2 shows the image of dataset HRCD with its ground truth image and other methods result. Figure 5.3 shows the image of dataset 38-cloud with its ground truth image and other methods result.

	Actually Positive	Actually Negative
Predicted Positive	True Positive (TP)	False Positive (FP)
Predicted Negative	False Negative (FN)	True Negative (TN)

Figure 5.1: Confusion Matrix

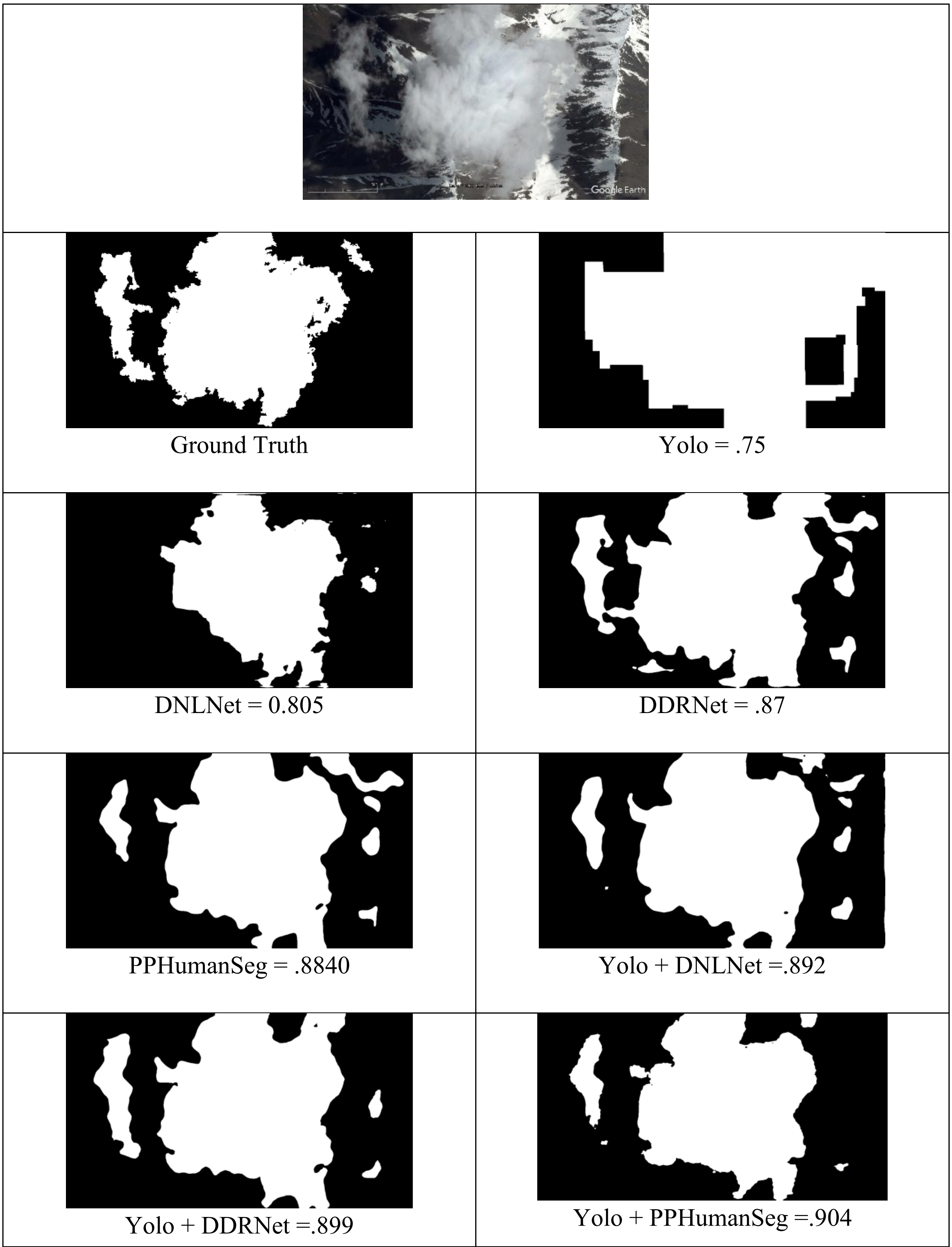


Figure 5.2: Comparison of different algorithm based on accuracy on the image of HRCD dataset.

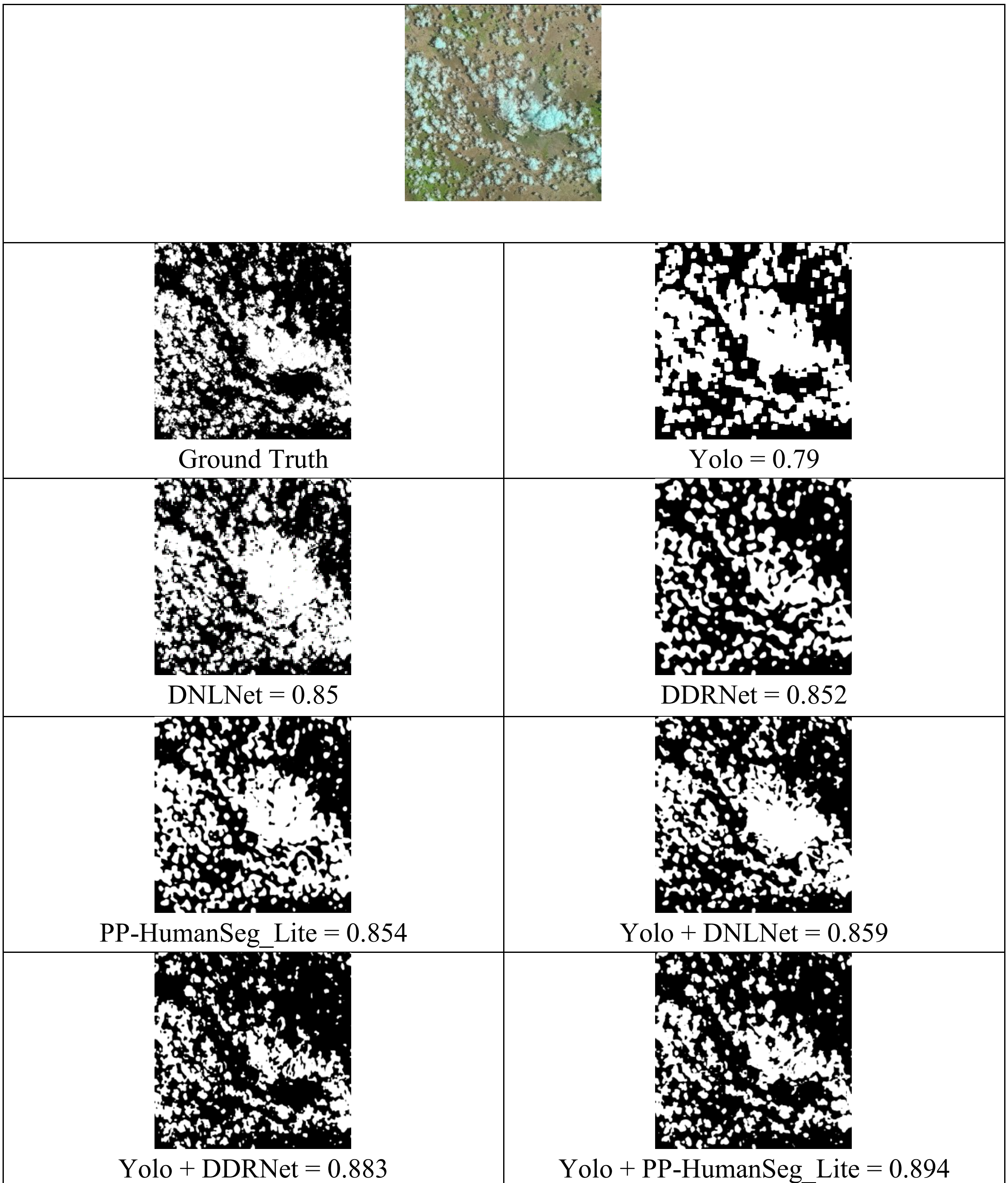


Figure 5.3: Comaprsion of Hybrid models based on accuarcy

Table 5.1: Accuracy based comparison of deep learning algorithms for cloud detection

Method	HRCD	38-Clouds
Yolo + PP-HumanSeg_Lite	0.960	0.933
Yolo + DDRNet	0.955	0.923
Yolo + DNLNet	0.949	0.920
PP-HumanSeg_Lite	0.942	0.903
DDRNet	0.951	0.914
DNLNet	0.948	0.883
Yolo	0.912	0.815

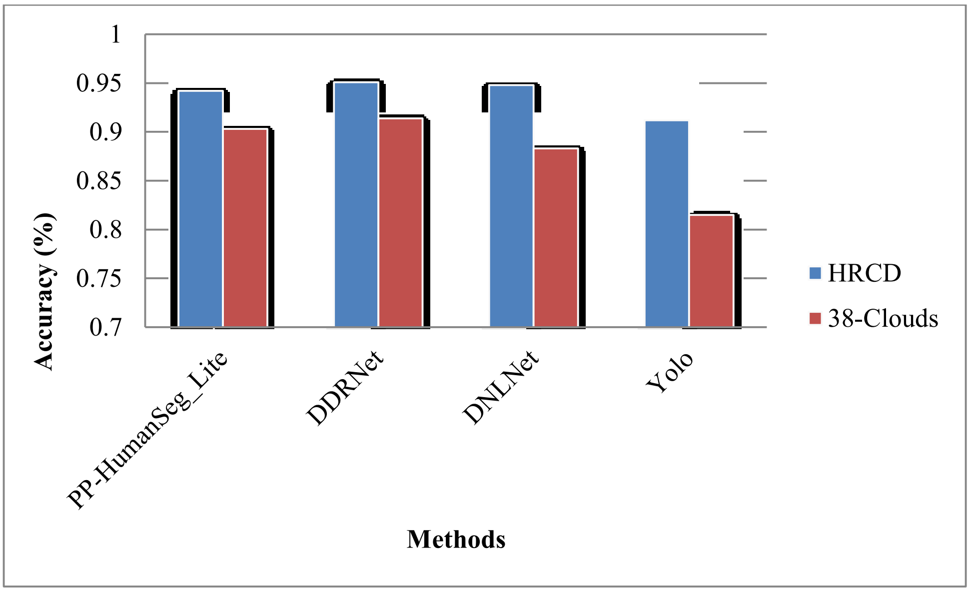


Figure 5.4: Comparison of both datasets applying individual models

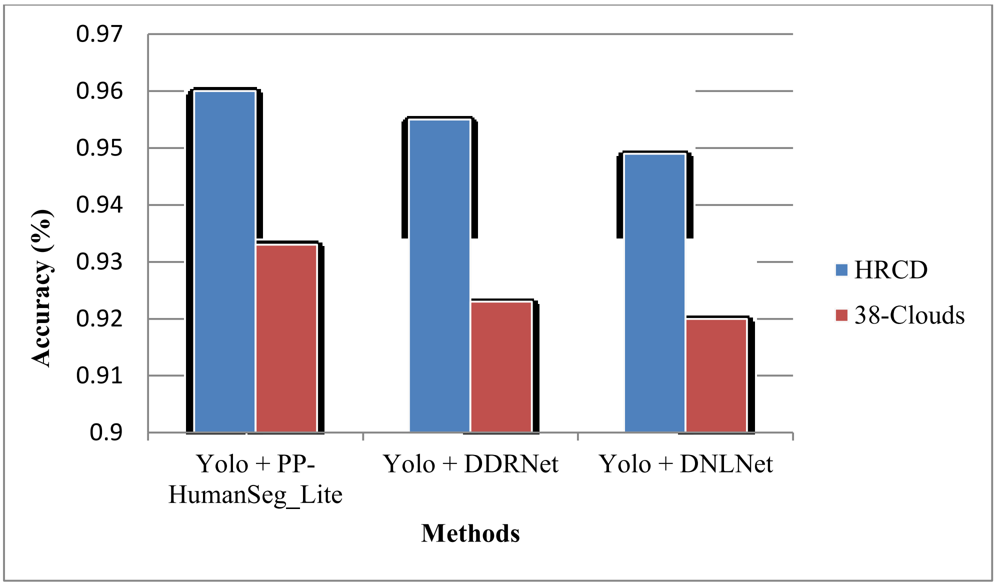


Figure 5.5: Comparison of both datasets applying Hybrid models

5.3 Summary

This chapter shows the result and analysis. The graphs show the comparisons between the methods are given. The output images of cloud detection of both datasets are also shown in figures. Performance metrics is measured in terms of accuracy is also explained in detail.

CHAPTER

CONCLUSION AND FUTURE WORK

6.1 Overview

Performance of the proposed model is measured in terms of accuracy. Different methods DNLNet, DDRNet, PP-HumanSeg_Lite and Yolo are applied on the datasets. Methods are also compared with each other and their combined model is also compared. The model gives the better performance for cloud detection of high resolution datasets.

6.2 Conclusion

Cloud detection is challenging problem now a days, as many methods are proposed for remote sensing applications to monitor the land, for disaster management and for the agriculture field. Different classical, machine and deep learning methods are used for pixel level cloud detection but still there are some challenges related to the performance, speed and accuracy. Deep learning algorithms are fast and they provide more accurate results as compared to other classical and machine learning algorithms. Yolo (You Only Look Once) algorithm is used for object detection of natural indoor and outdoor images. It gives the better performance on different datasets. The algorithm is combined with the state-of-the-art methods Disentangled Non-Local Neural Network (DNLNet), Deep Dual-Resolution Network (DDRNet) and Practical Portrait Human Segmentation (PP-HumanSeg_Lite). These algorithms are used for semantic segmentation in which the network is based on encoder decoder. The features are extracted for the enhancement and the classification of the images. Semantic segmentation is done on real life objects. These methods are combined with Yolo algorithm to improve the accuracy. The methods are combined using ensemble learning technique for accurate cloud detection. Random forest is a machine learning model used in training process for preprocessing the datasets. Experiment is performed on both dataset 38-cloud and high resolution dataset applying each algorithm individually. The results show that the Yolo gives low accuracy 81% but when these deep learning methods are combined with Yolo algorithm. Yolo + PPHSL method gives the high performance on both datasets which is 96% on HRCD and 93% on 38-cloud.

6.3 Future Work

In future, methods will be applied on the other datasets and the Yolo algorithm will be combined with different algorithms to achieve more accuracy and performance.

REFERENCES

[1] S. Metsämäki, M. Luotamo, and A. Klami, "Multi-scale Cloud Detection in Remote Sensing Images using a Dual Convolutional Neural Network," *arXiv*, pp. 1–12, 2020, doi: 10.1109/tgrs.2020.3015272.

[2] S. Mahajan and B. Fataniya, "Cloud detection methodologies: variants and development—a review," *Complex Intell. Syst.*, vol. 6, no. 2, pp. 251–261, 2020, doi: 10.1007/s40747-019-00128-0.

[3] Z. Li, H. Shen, Q. Cheng, Y. Liu, S. You, and Z. He, "Deep learning based cloud detection for medium and high resolution remote sensing images of different sensors," *ISPRS J. Photogramm. Remote Sens.*, vol. 150, pp. 197–212, 2019, doi: 10.1016/j.isprsjprs.2019.02.017.

[4] S. Mohajerani, T. A. Krammer, and P. Saeedi, "A cloud detection algorithm for remote sensing images using fully convolutional neural networks," *2018 IEEE 20th Int. Work. Multimed. Signal Process. MMSP 2018*, pp. 1–5, 2018, doi: 10.1109/MMSP.2018.8547095.

[5] Q. Zhang and C. Xiao, "Cloud detection of RGB color aerial photographs by progressive refinement scheme," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 7264–7275, 2014, doi: 10.1109/TGRS.2014.2310240.

[6] J. Yang *et al.*, "A total sky cloud detection method using real clear sky background," *Atmos. Meas. Tech.*, vol. 9, no. 2, pp. 587–597, 2016, doi: 10.5194/amt-9-587-2016.

[7] P. Shah, "Development and Hardware Implementation of an Efficient Algorithm for Cloud Detection From Satellite Images," *Signal Image Process. An Int. J.*, vol. 7, no. 2, pp. 73–80, 2016, doi: 10.5121/sipij.2016.7205.

[8] M. Mauniga and S. I. Padma, "Automatic Cloud Detection and Weather Forecasting Using Gradient Based SPS," no. 7, pp. 541–545, 2016.

[9] W. Wu, J. Luo, X. Hu, H. Yang, and Y. Yang, "A thin-cloud mask method for remote sensing images based on sparse dark pixel region detection," *Remote Sens.*, vol. 10, no. 4, 2018, doi: 10.3390/rs10040617.

[10] W. Huang, Y. Wang, and X. Chen, "Cloud detection for high-resolution remote-sensing images of urban areas using colour and edge features based on dual-colour

models," *Int. J. Remote Sens.*, vol. 39, no. 20, pp. 6657–6675, 2018, doi: 10.1080/01431161.2018.1466069.

[11] M. Le Goff, J.-Y. Tourneret, H. Wendt, M. Ortner, and M. Spigai, "Deep Learning for Cloud Detection," pp. 10 (6 .)-10 (6 .), 2017, doi: 10.1049/cp.2017.0139.

[12] Z. Guo, C. Li, Z. Wang, E. Kwok, and X. Wei, "A cloud boundary detection scheme combined with ASLIC and CNN using ZY-3, GF-1/2 satellite imagery," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.*, vol. 42, no. 3, pp. 455–458, 2018, doi: 10.5194/isprs-archives-XLII-3-455-2018.

[13] G. Luis and J. Mu, "CLOUD DETECTION MACHINE LEARNING ALGORITHMS FOR PROBA-V Image Processing Laboratory (IPL), University of Valencia , Spain," pp. 2251–2254, 2017.

[14] C. Deng, Z. Li, W. Wang, S. Wang, L. Tang, and A. C. Bovik, "Cloud Detection in Satellite Images Based on Natural Scene Statistics and Gabor Features," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 4, pp. 608–612, 2019, doi: 10.1109/LGRS.2018.2878239.

[15] W. Li, Z. Zou, and Z. Shi, "Deep Matting for Cloud Detection in Remote Sensing Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 12, pp. 8490–8502, 2020, doi: 10.1109/TGRS.2020.2988265.

[16] Y. Li, W. Chen, Y. Zhang, C. Tao, R. Xiao, and Y. Tan, "Accurate cloud detection in high-resolution remote sensing imagery by weakly supervised deep learning," *Remote Sens. Environ.*, vol. 250, no. December 2019, p. 112045, 2020, doi: 10.1016/j.rse.2020.112045.

[17] M. Khoshboresh-Masouleh and R. Shah-Hosseini, "A Deep Learning Method for Near-Real-Time Cloud and Cloud Shadow Segmentation from Gaofen-1 Images," *Comput. Intell. Neurosci.*, vol. 2020, 2020, doi: 10.1155/2020/8811630.

[18] J. Ding *et al.*, "Object Detection in Aerial Images : A Large-Scale Benchmark and Challenges," vol. 14, no. 8, pp. 1–17, 2015.

[19] H. Fazlali, S. Shirani, M. McDonald, and T. Kirubarajan, "Cloud/haze detection in airborne videos using a convolutional neural network," *Multimed. Tools Appl.*, vol. 79, no. 39–40, pp. 28587–28601, 2020, doi: 10.1007/s11042-020-09359-7.

[20] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 779–788, 2016, doi: 10.1109/CVPR.2016.91.

[21] L. Sun *et al.*, "Satellite data cloud detection using deep learning supported by

hyperspectral data," *Int. J. Remote Sens.*, vol. 41, no. 4, pp. 1349–1371, 2020, doi: 10.1080/01431161.2019.1667548.

[22] R. Cilli, A. Monaco, N. Amoroso, A. Tateo, S. Tangaro, and R. Bellotti, "Machine learning for cloud detection of globally distributed sentinel-2 images," *Remote Sens.*, vol. 12, no. 15, pp. 1–17, 2020, doi: 10.3390/RS12152355.

[23] H. Lv, Y. Wang, and Y. Yang, "Thin Cloud Detection Using Spectral Similarity," *Ieee Igarss 2017*, pp. 3–6, 2017.

[24] M. Nagare, Hirofumi Aoki, and Eiji Kaneko, "A UNIFIED METHOD OF CLOUD DETECTION AND REMOVAL ROBUST TO SPECTRAL VARIABILITY Madhuri Nagare , Hirofumi Aoki and Eiji Kaneko Data Science Research Laboratories , NEC Corporation," *Igarss 2017*, no. 2, pp. 5418–5421, 2017.

[25] J. Zhang, Q. Zhou, X. Shen, and Y. Li, "Cloud Detection in High-Resolution Remote Sensing Images Using Multi-features of Ground Objects," 2019.

[26] V. V. Voronin, N. Gapon, E. A. Semenishchev, A. Zelensky, and S. S. Agaian, "A block-based method for the remote sensing images cloud detection and removal," no. April 2021, p. 18, 2021, doi: 10.1117/12.2588691.

[27] Q. Xiong *et al.*, "A cloud detection approach based on hybrid multispectral features with dynamic thresholds for GF-1 remote sensing images," *Remote Sens.*, vol. 12, no. 3, pp. 1–18, 2020, doi: 10.3390/rs12030450.

[28] M. Reguiegue and F. Chouireb, "Automatic day time cloud detection over land and sea from MSG SEVIRI images using three features and two artificial intelligence approaches," *Signal, Image Video Process.*, vol. 12, no. 1, pp. 189–196, 2018, doi: 10.1007/s11760-017-1145-0.

[29] A. Asokan and J. Anitha, "Machine Learning based Image Processing Techniques for Satellite Image Analysis -A Survey," *Proc. Int. Conf. Mach. Learn. Big Data, Cloud Parallel Comput. Trends, Prespectives Prospect. Com. 2019*, pp. 119–124, 2019, doi: 10.1109/COMITCon.2019.8862452.

[30] S. Suresh and S. Lal, "Modified differential evolution algorithm for contrast and brightness enhancement of satellite images," *Appl. Soft Comput. J.*, vol. 61, pp. 622–641, 2017, doi: 10.1016/j.asoc.2017.08.019.

[31] S. Suresh and S. Lal, "An efficient cuckoo search algorithm based multilevel thresholding for segmentation of satellite images using different objective functions," *Expert Syst. Appl.*, vol. 58, pp. 184–209, 2016, doi: 10.1016/j.eswa.2016.03.032.

[32] X. Huang, H. Chen, and J. Gong, "Angular difference feature extraction for urban

scene classification using ZY-3 multi-angle high-resolution satellite imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 135, pp. 127–141, 2018, doi: 10.1016/j.isprsjprs.2017.11.017.

[33] M. Tian, H. Chen, and G. Liu, "Cloud Detection and Classification for S-NPP FSR CRIS Data Using Supervised Machine Learning," *Int. Geosci. Remote Sens. Symp.*, pp. 9827–9830, 2019, doi: 10.1109/IGARSS.2019.8898876.

[34] "IMPROVED CLOUD DETECTION MODEL USING S-NPP CRIS FSR DATA VIA MACHINE LEARNING School of Information and Communication Engineering , University of Electronic Science and Technology of China , Chengdu , China," pp. 2260–2263, 2020.

[35] G. Mateo-Garcia, V. Laparra, D. Lopez-Puigdollers, and L. Gomez-Chova, "Cross-Sensor Adversarial Domain Adaptation of Landsat-8 and Proba-V Images for Cloud Detection," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 14, pp. 747–761, 2021, doi: 10.1109/JSTARS.2020.3031741.

[36] "AUTOMATIC SINGLE-IMAGE BASED CLOUD DETECTION METHOD WITHOUT PRIOR INFORMATION Yuhan Liu , Zhenming Peng University of Electronic Science and Technology of China Schoold of Information and Communication Engineering," pp. 2611–2614, 2020.

[37] Z. Tan *et al.*, "Detecting Multilayer Clouds from the Geostationary Advanced Himawari Imager Using Machine Learning Techniques," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–12, 2022, doi: 10.1109/TGRS.2021.3087714.

[38] Z. Ning, W. Wei, S. Qin, Y. Chengzong, and Z. Xinzhong, "Cloud Detection of Remote Sensing Image Based on Multi Feature Fusion," *2020 5th IEEE Int. Conf. Big Data Anal. ICBDA 2020*, pp. 298–303, 2020, doi: 10.1109/ICBDA49040.2020.9101218.

[39] C. Wang, S. Platnick, K. Meyer, Z. Zhang, and Y. Zhou, "A machine-learning-based cloud detection and thermodynamic-phase classification algorithm using passive spectral observations," *Atmos. Meas. Tech.*, vol. 13, no. 5, pp. 2257–2277, 2020, doi: 10.5194/amt-13-2257-2020.

[40] J. Wei *et al.*, "Cloud detection for Landsat imagery by combining the random forest and superpixels extracted via energy-driven sampling segmentation approaches," *Remote Sens. Environ.*, vol. 248, no. July, p. 112005, 2020, doi: 10.1016/j.rse.2020.112005.

[41] L. Li, X. Li, L. Jiang, X. Su, and F. Chen, "A review on deep learning techniques for

cloud detection methodologies and challenges," *Signal, Image Video Process.*, vol. 15, no. 7, pp. 1527–1535, 2021, doi: 10.1007/s11760-021-01885-7.

[42] 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," *Eur. J. Remote Sens.*, vol. 54, no. sup2, pp. 383–395, 2021, doi: 10.1080/22797254.2020.1790995.

[43] L. Li, X. Li, X. Liu, W. Huang, Z. Hu, and F. Chen, "Attention Mechanism Cloud Detection with Modified FCN for Infrared Remote Sensing Images," *IEEE Access*, vol. 9, pp. 150975–150983, 2021, doi: 10.1109/ACCESS.2021.3122162.

[44] X. Li *et al.*, "Sfrs-net: A cloud-detection method based on deep convolutional neural networks for gf-1 remote-sensing images," *Remote Sens.*, vol. 13, no. 15, pp. 1–20, 2021, doi: 10.3390/rs13152910.

[45] M. Segal-Rozenhaimer, A. Li, K. Das, and V. Chirayath, "Cloud detection algorithm for multi-modal satellite imagery using convolutional neural-networks (CNN)," *Remote Sens. Environ.*, vol. 237, no. September 2019, p. 111446, 2020, doi: 10.1016/j.rse.2019.111446.

[46] Y. Chen, L. Tang, X. Yang, R. Fan, M. Bilal, and Q. Li, "Thick Clouds Removal from Multitemporal ZY-3 Satellite Images Using Deep Learning," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 143–153, 2020, doi: 10.1109/JSTARS.2019.2954130.

[47] Z. Shao, Y. Pan, C. Diao, and J. Cai, "Cloud detection in remote sensing images based on multiscale features-convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 4062–4076, 2019, doi: 10.1109/TGRS.2018.2889677.

[48] S. Dev, A. Nautiyal, Y. H. Lee, and S. Winkler, "CloudSegNet: A Deep Network for Nychthemeron Cloud Image Segmentation," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 12, pp. 1814–1818, 2019, doi: 10.1109/LGRS.2019.2912140.

[49] Q. Zhang, Q. Yuan, J. Li, Z. Li, H. Shen, and L. Zhang, "Thick cloud and cloud shadow removal in multitemporal imagery using progressively spatio-temporal patch group deep learning," *ISPRS J. Photogramm. Remote Sens.*, vol. 162, no. August 2019, pp. 148–160, 2020, doi: 10.1016/j.isprsjprs.2020.02.008.

[50] Y. Guo, X. Cao, B. Liu, and M. Gao, "Cloud detection for satellite imagery using attention-based U-net convolutional neural network," *Symmetry (Basel).*, vol. 12, no. 6, 2020, doi: 10.3390/SYM12061056.

[51] Z. Yao, J. Jia, and Y. Qian, "Mcnet: Multi-scale feature extraction and content-aware

reassembly cloud detection model for remote sensing images," *Symmetry (Basel).*, vol. 13, no. 1, pp. 1–13, 2021, doi: 10.3390/sym13010028.

[52] J. Zhang, H. Wang, Y. Wang, Q. Zhou, and Y. Li, "Deep network based on up and down blocks using wavelet transform and successive multi-scale spatial attention for cloud detection," *Remote Sens. Environ.*, vol. 261, p. 112483, 2021, doi: 10.1016/j.rse.2021.112483.

[53] J. Zhuang, J. Yang, L. Gu, and N. Dvornek, "Shelfnet for fast semantic segmentation," *Proc. - 2019 Int. Conf. Comput. Vis. Work. ICCVW 2019*, pp. 847–856, 2019, doi: 10.1109/ICCVW.2019.00113.

[54] S. Wang, W. Chen, S. M. Xie, G. Azzari, and D. B. Lobell, "Weakly supervised deep learning for segmentation of remote sensing imagery," *Remote Sens.*, vol. 12, no. 2, pp. 1–25, 2020, doi: 10.3390/rs12020207.

[55] Q. Liu *et al.*, "Daytime rainy cloud detection and convective precipitation delineation based on a deep neural network method using GOES-16 ABI images," *Remote Sens.*, vol. 11, no. 21, pp. 1–18, 2019, doi: 10.3390/rs11212555.

[56] A. Francis, P. Sidiropoulos, and J. P. Muller, "CloudFCN: Accurate and robust cloud detection for satellite imagery with deep learning," *Remote Sens.*, vol. 11, no. 19, pp. 1–22, 2019, doi: 10.3390/rs11192312.

[57] Q. He, X. Sun, Z. Yan, and K. Fu, "DABNet: Deformable Contextual and Boundary-Weighted Network for Cloud Detection in Remote Sensing Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–16, 2022, doi: 10.1109/TGRS.2020.3045474.

[58] N. Ma, L. Sun, C. Zhou, and Y. He, "Cloud detection algorithm for multi-satellite remote sensing imagery based on a spectral library and 1D convolutional neural network," *Remote Sens.*, vol. 13, no. 16, pp. 1–20, 2021, doi: 10.3390/rs13163319.

[59] Z. Zhang, Z. Xu, C. Liu, Q. Tian, and Y. Wang, "Cloudformer: Supplementary Aggregation Feature and Mask-Classification Network for Cloud Detection," *Appl. Sci.*, vol. 12, no. 7, 2022, doi: 10.3390/app12073221.

[60] J. Guo, Q. Xu, Y. Zeng, Z. Liu, and X. Zhu, "Semi-Supervised Cloud Detection in Satellite Images by Considering the Domain Shift Problem," *Remote Sens.*, vol. 14, no. 11, p. 2641, 2022, doi: 10.3390/rs14112641.

[61] K. Hu, D. Zhang, M. Xia, M. Qian, and B. Chen, "LCDNet: Light-Weighted Cloud Detection Network for High-Resolution Remote Sensing Images," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 15, pp. 4809–4823, 2022, doi: 10.1109/jstars.2022.3181303.

[62] Y. Wang *et al.*, "Remote sensing image super-resolution and object detection: Benchmark and state of the art," *Expert Syst. Appl.*, vol. 197, pp. 1–39, 2022, doi: 10.1016/j.eswa.2022.116793.

[63] J. H. Jeppesen, R. H. Jacobsen, F. Inceoglu, and T. S. Toftegaard, "A cloud detection algorithm for satellite imagery based on deep learning," *Remote Sens. Environ.*, vol. 229, no. August 2018, pp. 247–259, 2019, doi: 10.1016/j.rse.2019.03.039.

[64] J. Yang, J. Guo, H. Yue, Z. Liu, H. Hu, and K. Li, "CDnet: CNN-Based Cloud Detection for Remote Sensing Imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 6195–6211, 2019, doi: 10.1109/TGRS.2019.2904868.

[65] X. Zhang, L. Han, L. Han, and L. Zhu, "How well do deep learning-based methods for land cover classification and object detection perform on high resolution remote sensing imagery?," *Remote Sens.*, vol. 12, no. 3, pp. 1–29, 2020, doi: 10.3390/rs12030417.

[66] C. Zhang, P. Yue, D. Tapete, B. Shangguan, M. Wang, and Z. Wu, "A multi-level context-guided classification method with object-based convolutional neural network for land cover classification using very high resolution remote sensing images," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 88, no. January, p. 102086, 2020, doi: 10.1016/j.jag.2020.102086.

[67] K. Itakura and F. Hosoi, "Automatic tree detection from three-dimensional images reconstructed from 360 spherical camera using YOLO v2," *Remote Sens.*, vol. 12, no. 6, 2020, doi: 10.3390/rs12060988.

[68] D. Xu and Y. Wu, "Improved YOLO-V3 with densenet for multi-scale remote sensing target detection," *Sensors (Switzerland)*, vol. 20, no. 15, pp. 1–24, 2020, doi: 10.3390/s20154276.

[69] Z. Huang, H. Chen, B. Liu, and Z. Wang, "Semantic-guided attention refinement network for salient object detection in optical remote sensing images," *Remote Sens.*, vol. 13, no. 11, pp. 1305–1317, 2021, doi: 10.3390/rs13112163.

[70] X. Li, J. Deng, and Y. Fang, "Few-Shot Object Detection on," vol. 60, pp. 1–14, 2021.

[71] X. Yao, Q. Guo, A. Li, and L. Shi, "Optical remote sensing cloud detection based on random forest only using the visible light and near-infrared image bands," *Eur. J. Remote Sens.*, vol. 55, no. 1, pp. 150–167, 2022, doi: 10.1080/22797254.2021.2025433.

[72] S. Gadamsetty, R. Ch, A. Ch, C. Iwendi, and T. R. Gadekallu, "Hash-Based Deep Learning Approach for Remote Sensing Satellite Imagery Detection," *Water*

(Switzerland), vol. 14, no. 5, pp. 1–15, 2022, doi: 10.3390/w14050707.

[73] L. Li, L. Jiang, J. Zhang, S. Wang, and F. Chen, "A Complete YOLO-Based Ship Detection Method for Thermal Infrared Remote Sensing Images under Complex Backgrounds," pp. 1–13, 2022.

[74] Y. Y. Wang, H. Wu, L. Shuai, C. Peng, and Z. Yang, "Detection of plane in remote sensing images using super-resolution," *PLoS One*, vol. 17, no. 4 April, pp. 1–19, 2022, doi: 10.1371/journal.pone.0265503.

[75] C. Liu *et al.*, "A Machine Learning-based Cloud Detection Algorithm for the Himawari-8 Spectral Image," *Adv. Atmos. Sci.*, 2021, doi: 10.1007/s00376-021-0366-x.

[76] Q. Yan *et al.*, "Cloud Detection of Remote Sensing Image Based on Multi-Scale Data and Dual-Channel Attention Mechanism," *Remote Sens.*, vol. 14, no. 15, 2022, doi: 10.3390/rs14153710.

[77] M. Yin *et al.*, "Disentangled Non-local Neural Networks," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12360 LNCS, pp. 191–207, 2020, doi: 10.1007/978-3-030-58555-6_12.

[78] Y. Hong, H. Pan, W. Sun, and Y. Jia, "Deep Dual-resolution Networks for Real-time and Accurate Semantic Segmentation of Road Scenes," vol. 14, no. 8, pp. 1–12, 2021, [Online]. Available: http://arxiv.org/abs/2101.06085

[79] L. Chu *et al.*, "PP-HumanSeg: Connectivity-Aware Portrait Segmentation with a Large-Scale Teleconferencing Video Dataset," *Proc. - 2022 IEEE/CVF Winter Conf. Appl. Comput. Vis. Work. WACVW 2022*, pp. 202–209, 2022, doi: 10.1109/WACVW54805.2022.00026.

[80] S. Mohajerani and P. Saeedi, "Cloud-Net: An End-To-End Cloud Detection Algorithm for Landsat 8 Imagery," *Int. Geosci. Remote Sens. Symp.*, pp. 1029–1032, 2019, doi: 10.1109/IGARSS.2019.8898776.

[81] S. Images, "MSLWENet : A Novel Deep Learning Network for Lake Water Body Extraction of Google Remote Sensing Images," 2020, doi: 10.3390/rs12244140.