Convolutional Network Feature Hierarchy for Hyper Spectral Image Classification

A DISSERATATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE AWARD OF THE DEGREE

OF

MASTER OF TECHNOLOGY

IN

Signal Processing and Digital Design

Submitted by

Ankita Thapliyal

(2K19/SPD/03)

Under the supervision of

Prof Dinesh Kumar (Prof, ECE Dept.)

DEPARTMENT OF ELECTRONICS AND COMMUNICATION

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi -110042

JULY 2021

CONTENTS

CANDIDATE'S DECLARATION

I **Ankita Thapliyal** student of M.Tech (Signal Processing and Digital Design), hereby declare that the project Dissertation titled **"Convolutional Network Feature Hierarchy for Hyper Spectral Image Classification"** which is submitted by me to the Department of Electronics and Communication Engineering, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

CERTIFICATE

I hereby certify that the Project Report titled "**Convolutional Network Feature Hierarchy for Hyper Spectral Image Classification**" which is submitted by **Ankita Thapliyal**, **2K19/SPD/03** of Electronics and Communication Department, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology, is a record of the project work carried out by the students under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

Place: Delhi
Date: 30th July 2021

Supervisor:

ACKNOWLEDGEMENT

A successful project can never be prepared by the efforts of the person to whom the project is assigned, but it also demands the help and guardianship of people who helped in completion of the project. I would like to thank all those people who have helped me in this research and inspired me during my study. With profound sense of gratitude, I thank Prof Dinesh Kumar, my Research Guide, for his encouragement, support, patience and his guidance in this research work. Furthermore, I would also like to thank the Head of the Department, Electronics and Communication, Prof N. S. Raghava and Prof. S. Indu, who gave me the permission to use all required equipment and the necessary format to complete the report. I take immense delight in extending my acknowledgement to my family and friends who have helped me throughout this research work.

Ankita Thapliyal

ABSTRACT

Hyper spectral image classification is the recent technology that is famous among the researchers nowadays. It is simply an application of remote sensing methodology. The results of remote sensing are basically needs to get studied by the scientists properly so that they can analyze the surface area and the target information accordingly. Then next to the study of target information further conclusions can be made about that area successfully. Remote sensing is the first step to analyze the whole process in which the satellite helps to provide several images of a particular land area or vegetation portion. These images can be obtained by using active or passive remote sensing depending on the choice of user. As soon as the images are received by the sensors we not only analyse them in visible spectrum, but we do recognize them in ultra violet and infrared region of the electromagnetic spectrum. This type of technique is known as hyper spectral imaging. We use hyper spectral sensors to perform this type of imaging. This method has so many advantages over multispectral imaging in which number of spectral band information is comparatively less. Since the number of bands in hyper spectral imaging is greater than the band information in multi spectral imaging, the recognition of images and target is more specified and accurate for hyper spectral data. More significant information is obtained through hyper spectral imaging. Since we receive the data through hyper spectral imaging we need to apply the upcoming tasks to know the target area in deeper way. As soon as the input is received in the form of images that are now three dimensional due to the hyper spectral view, we need to classify these images into the categories they are having. For instance, we get the information of a vegetation area we need to classify this three dimensional image data into the different categories of vegetation in that particular portion of land. This whole process is known as image classification which is the latest topic for machine learning methods. The use of deep neural networks at present helps in doing the classification of large number of images at a time with much more accuracy and reduced complexity. In past few decades many researchers have provided their own supervised models to implement the image classification over a huge dataset to classify the images successfully. But due to the drawbacks like less accuracy and higher complexity, these models have been over take by convolutional neural networks. Supervised technology is a type of machine learning task where the model learn itself on the basis of input and the outputs provided at the time of training. The methods like SVM and CNN are supervised methods that we us for the classification purpose. Hence in this project instead of using multi spectral data, we have discussed the use of hyper spectral data. This chapter consist of six chapters. In chapter 1 we are discussing the basic of remote sensing and its types. Chapter 2 will tell us about the type of imaging method and their advantages and disadvantages so that we can prefer the suitable one to perform the objective. In chapter 3 we are looking over the multiple supervised methods like SVM, CNN and ANN that helps in the classification of hyper spectral image data. Chapter 4 and 5 are the discussion of latest model with increased accuracy and reduced complexity.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Full Form
EEG	Electroencephalogram
ECG	Electrocardiography
MEG	Magneto encephalographic
CNN	Convolution Neural Network
M3D CNN	Multi-Scale Convolution Neural Network
ANN	Artificial Neural Network
SVM	Support Vector Machine
PCA	Principle Component Analysis
HSI	Hyper Spectral Image
ML	Machine Learning

CHAPTER 1

REMOTE SENSING

1.1 Introduction to Remote Sensing

Now a days we can generally observe the process known as remote sensing that is more common for utilizing the physical characterstics of a given area through observing the radiated and emitted waves by that area. In this technique we monitor and recognise the general features of the target zone. Typically there are some special devices to gather these remotely sensed images which help the scientist to have an idea about the earth. In other words it is basically the capturing of information without having any physical touch with the target or can say it is an on-site experiment. The application of remote sensing can be viewed in various fields like geography, examination of land and many of science departments. For instance hydrology, meteorology, glaciology and ecology like fields. We can also see its demand in other areas like military, planning, commercial, economy and many more. At present situation we can just simply refer remote sensing as the utilization of satellite or technologies that are based on aircrafts, which helps in detecting and classifying the targets on earth. Today the place is totally covered by the use of radars and lasers that proved out to be a breath taking technologies with significant features. Talking about the target, it can be the surface and the ocean region depending on the travelled electromagnetic signals. This phenomenon categorises remote sensing as active and passive.

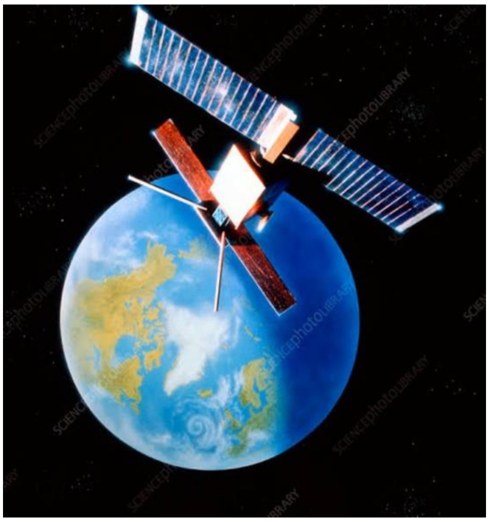


Figure 1.1: Satellite observation of Earth

Over a larger period of time remote sensing has become one of the widest technologies that have seen many applications in different spheres of human life. One of the majority groups that are utilizing the advantages of remote sensing is farmers. They can see the effect on their daily basis. The important decisions that are taken are because of the data received from Terra SAR-X, RADARSAT and many others.

1.2: Types of Remote Sensing

1.2.1. Active Remote Sensing: To perform illumination unlike passive sensors, active sensors provide their own energy source. The sensor generates their own energetic waves towards the object that needs to be observed or investigated. As soon as the target receives this radiation, it starts generating some radiations. These radiations help in the continuous monitoring of the target zone. Active sensor keeps a record of these radiations. The biggest advantage of active sensors over the passive one is that these sensors can obtain the track of the target zone any time they want, the measurements can be noticed by sending the energy waves and then keeping the record of the reflected one's by the object. It does not matter at what time or at which weather or season the observation is made. There are some of the wavelengths that cannot be generated by the sun, in such cases active sensors plays a significant role. The example of such wavelengths is microwave, which cannot be radiated by sun. But at the same time active sensor deals with one of the major drawback which says that they require huge amount of energy so that they can fairly and adequately illuminates the objects under observation. Some most common example of the active sensors is laser fluoro-sensor and the other one is synthetic aperture radar, indicated as SAR. This remote sensing needs to have a transmitting electromagnetic wave which is coherent in nature to strike the target. The target on the other hand can be directed toward the earth or it can be a celestial body.

There are two major features shown by active sensors which are not present in passive sensors

a) The amount of time taken by the transmitted electromagnetic waves to reach to the receiver. It mostly matter in the case when the transmitter and the receiver are at different locations.

b) In addition to the time, the second factor that is important is the phase information of the received electromagnetic wave.

Polarization phenomenon is shown by both these sensors in many of the applications, but due to the Doppler's shift there is extra information that is gained by the radar. One can easily calculate the distance between the transmitter and receiver emitting and receiving these electromagnetic waves respectively. This is because the speed of these waves is same as the speed of the light which helps in determining the range easily, which is measured by the radar. Another device that is commonly known as Scatterometer helps in the measurements of strength and polarization of the electromagnetic waves at the end of the receiver. These waves interact with the object by means of scattering. It is to be noted that scattering is neither absorption nor emission. The last instance is known as Imaging Radar that are basically constructed to anchorage the shift provided by Doppler's effect to increase the resolution if ground at higher level.

1.2.1.1: Active Remote Sensing Instruments

a) Radar: This is a kind of sensor that basically works with the radio waves and commonly known as Radio Detection and Ranging. It is a system that works with radio waves to find out the distance, acceleration, angle, phase and velocity of the targets. The detection plays a useful role in surveillance as it helps to detect aircrafts, ships, missiles and spacecraft etc. Basically radar was invented many years ago to play a role for military purpose. Radar is suitable in providing the data that includes the distance or range or position of the target from its scanner. Hence we can say that it plays a vital role where the positioning of the objects is very crucial. It has shown foremost applications in detecting the atmospheric, ground and sea targets.

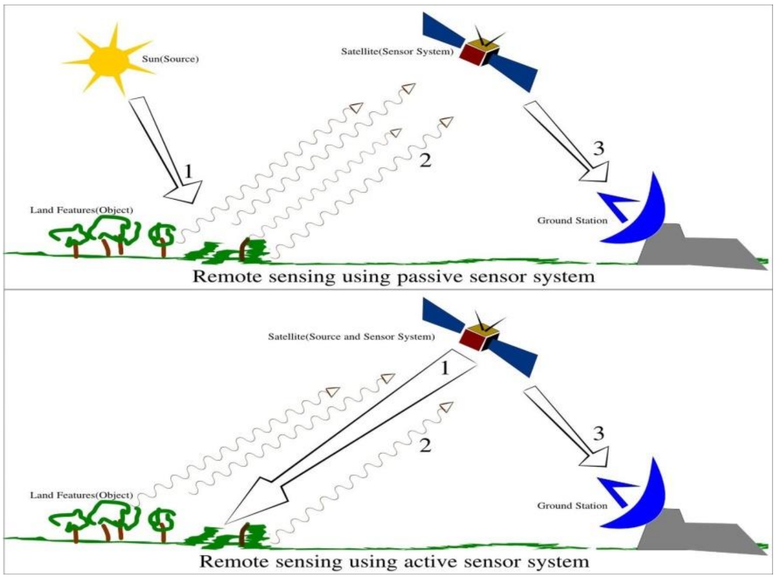


Figure 1.2: Depiction of Active and Passive sensors

Thus the aircrafts carrying the radar with them not only warn the other spacecraft in their way but also gives the weather information and correct readings for altitudes. The radar systems that are built in the military fighter aircrafts are in general equipped with air to air focusing radar system so that they can capture the enemy devices.

At the same time, bigger specialized aircrafts of military forces takes powerful and strong airborne radars with themselves to figure out the traffic present in air over a broader range and so can point the fighter aircrafts to them easily. There are marine radars that can be used to detect the range of ships to avoid the accident with other ship present inside the sea. It helps to navigate and to fix the coordinates of the ships inside the sea within a shore range. Radars are also used by the meteorologists to capture and monitor wind and precipitation. These are a type of secondary tools that serves as a platform to capture weather data like thunderstorms, winter storms and tornadoes. This method has generally been observed in the vital detection of sign monitoring and human activity detection.

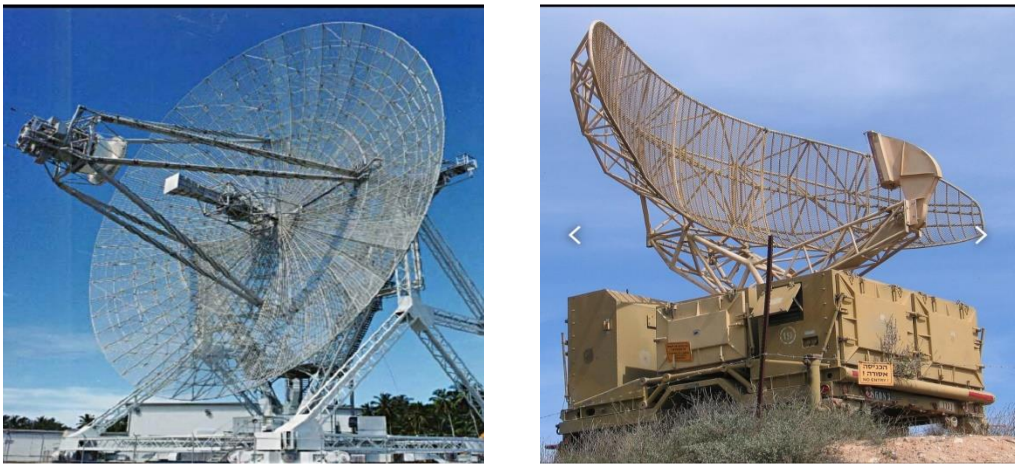


Figure 1.3: a) Radar to detect space objects b) Radar to detect aircrafts as it rotates steadily using a narrow beam

It plays a role in medical also to capture and record release of blood into the vessels of human body. It also measures the human body movement and the inhalation and exhalation of air into the lungs. Further the records are analysed and the patterns can be learnt by machine learning algorithms.

b) Lidar: Lidar stands for laser detection and ranging. This is a technique that is used to measure variable ranges focusing on an object by means of a laser. The time is also recorded at which the propagated wave is received after the transmission. The application of lidar lies in digital three dimensional representations of various areas of the earth surface, bottom areas of oceans and many land cover vegetations. The laser return of every surface is different in terms of time and wavelengths. The application also lies in terrestrial and mobile areas. Lidar is a beneficial method that is used to generate high- resolution maps and surveys of various geological and geographical areas. Many fields like forestry, seismology and archaeology is also served by Lidar effects. Lidar is generally defined into two categories which are direct detection of energy or an incoherent scheme. The other one is coherent scheme. Coherent scheme serves excellent in case of measuring the shift generated by Doppler's effect or phase change in the received light particles.

This system generally uses the optical heterodyne capturing which is much more sensitive and accurate as it allows operating at lower power. But the drawback is that it requires complex transrecievers. The incoherent detection measures the change in amplitude in the sensed reflected waves. The intermittent burst of energy is evaluated by the micro pulse devices while. The base of these systems is ever increasing computer technologies with modification in laser methods. The energy and power consumed by laser is near about one micro jule which is considered to be very low. Hence these can be used without safety precaution to human eyes as they are significantly very smaller. Lidar operates on a limited power because it contains an automatic shut off device which turns off the laser at certain altitudes to preserve the human vision. The most general lidar is 1500nm laser based which is safer for human eyes too at high power levels but their accuracy are low due to lack of advanced technologies. These lasers are not visible at night time, thus can be used for military purpose even with night vision goggles.

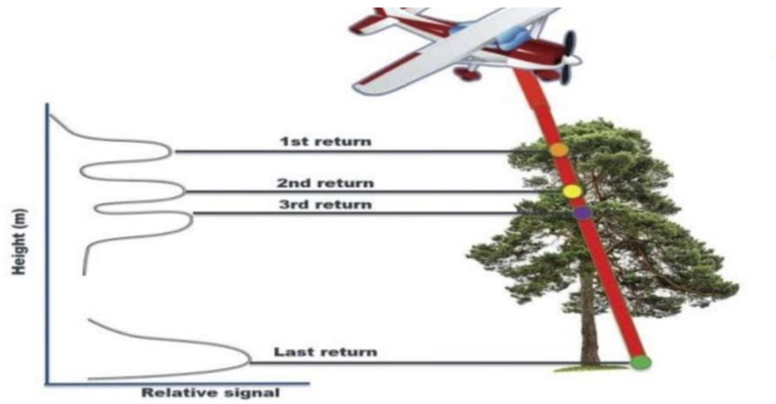


Figure 1.4: Functioning of Lidar Systems using laser beams through multiple returns

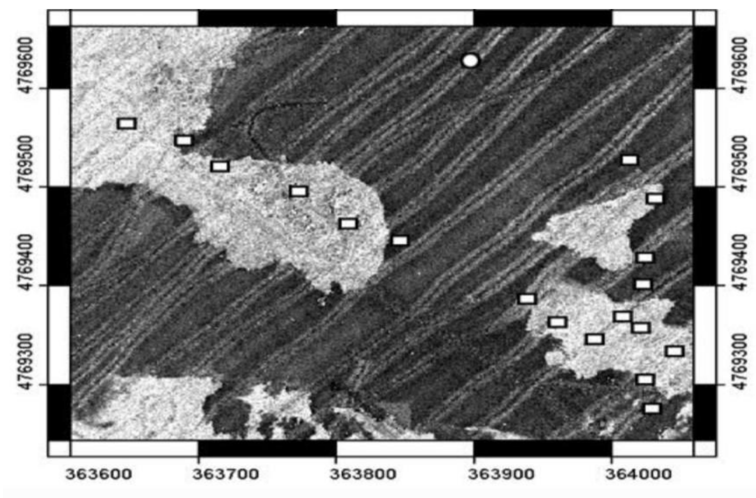


Figure1.5: Coastal Areal Information gathered after Laser beam Detection

Both types of lidar ways works using pulse framework. It can be either micro pulse or higher energy pulse.

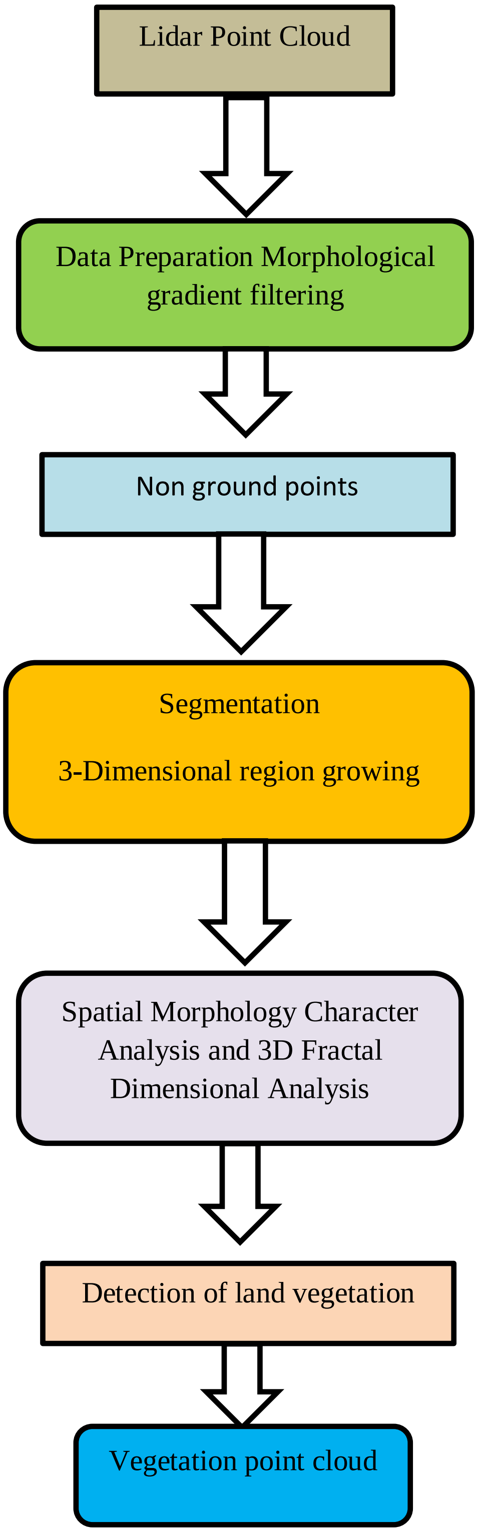
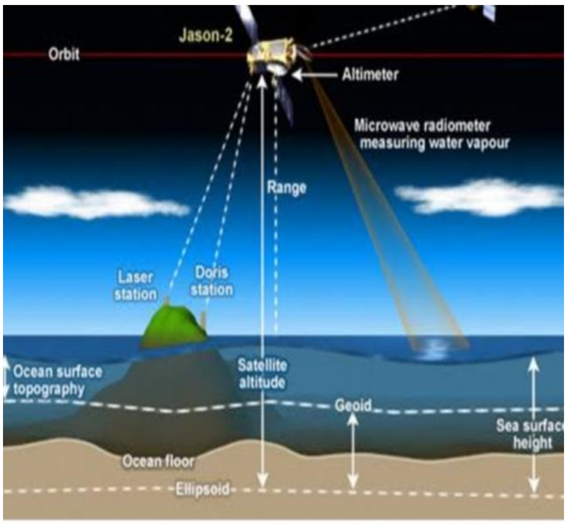


Figure 1.6: Flowchart depicting the procedure of Lidar functioning

c) Altimeter: It is also known as altitude measuring device. Basically it is used to measure the height of the target above certain level. This whole process of height measurement is also known as altimetry. Altimetry is somehow related to the term bathymetry as it is the measurement of water depth. The altimeter that is present in the satellites helps in the remote sensing of the topography of the surface above oceans and seas with precision value in few centimetres. As the orbital shape can be viewed independently, an accurate record is there to measure the ocean surface and its shape along with a line given by the satellite.

Altimetry evaluates the time taken by a radar pulse to travel both side between the satellite antenna and surface of earth. The pulse waves that are bounced back due to the reflecting facets are recorded at the footprints of an altimeter. In case of water surface these pulses are not exactly penetrates the water surface, but it is reflected back which is sensed by the on-board sensor at the receiver end. The travel time due to two ways propagation and the received echoes are traced by the sensor, which is helpful in keeping the window at accurate positional angle that is appropriate to gather complete waveform.

The correct evaluation of the distance between the satellite and the object is gathered by the retrace that functions a fair correction to indicate the coordinates of the travelling echo waveform. Altimetry techniques were developed to detect the ice sheets within the oceans and for their complete monitoring through successful mapping. Up till now there is no altimetry methods designed for monitoring the inland of water, even the satellite keeps on tracing the continental parts.



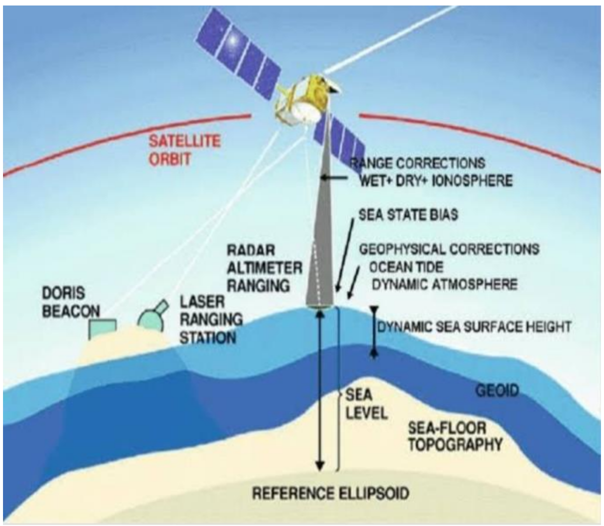


Figure 1.7: Measurements for sea levels by using Altimeters via remote sensing

d) Scatterometer: In remote sensing it is a type of diffusion meter instrument that predicts the return of a light wave or radar beam which are scattered because of diffusion inside a medium like air. These meters basically work using visible lights to detect the visibility horizontally. The scatterometers that uses radio or microwaves are known as radar scatterometers. They help in determining the normalized radar cross section for a given surface. More oftenly these can be seen as a mounting device on the weather satellites to find the speed of wind and its direction. There are some optical diffusion meters that can be used in the field of meteorology so that the optical range can be visualized horizontally.

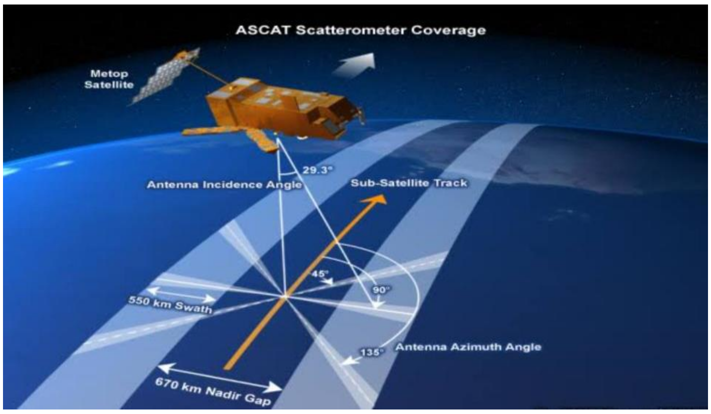


Figure 1.8: Illustration of Spectrometer Coverage over the surface of earth to target the object

They comprises of a light source that is commonly known as laser and a receiver. The placement of both these devices is made at an angle of 35 degree downward, so that focus can be made at a common area. In the direction of light beam there is a lateral scattering through air that is further quantified as attenuation coefficient. If there is a departure observed from the clear air, it is measured as extinction coefficient and is inversely proportional to the visibility. At automatic weather station we can find these types of instruments to look for general visibility, in the airport runways to determine runway visual range and also across the roads for visual situations. With these applications there is one major drawback which says that the measurement can be performed for a very smaller volume of air within the transmitter and the receiver. Therefore the traced visibility only represents the simpler situations near around the device in generalized manner. For instance, synoptic fog is the suitable way to explain the above case.

e) Sounder and Range measuring Devices: These are the instruments that figure outs the estimation of range between one or more than one same devices present at different platforms that are sending and receiving signals from each other. They evaluate the weather conditions in vertical direction by generating impulses if it comes under the active category.

1.2.1.2: Applications and Advantages of Active Remote Sensing

Since there are many kinds of practical implementations regarding active remote sensing this technique have no limitations in research fields. As already discussed that active remote sensing can function all day and for all weathers, it seems out be an independents functioning. There are many kinds of active sensors that find their way both in scientific industries and practical branches. To gather earth's elevation data a mission was discovered known as Shuttle Radar Topography. On the other hand the lidar active remote sensing happens in the sky which keeps on assisting the elaboration of some digital framework of our planet surface. Data that is acquired through remote sensing devices serves very well to agriculturalists and foresters. It is very difficult for them to reach areas in marine sciences and then rescue the objective. Sounders support the development of weather forecast having a vertical profile for humidity, temperature, precipitations and without or with the clouds.

1.2.2. Passive Remote Sensing: The first type of remote sensing is passive remote sensing in which the sensing devices captures the energy that is available naturally. If and only if the energy is naturally present, the passive sensors play their role significantly. One can say that this majorly happens only when the earth is being illuminated by the sun. It can only happen in day time as the reflected energies are not available in the night time. On the other hand there are some of the energies that are natural and available at both day and night times like infrared energy. The only condition is that amount of energy should be large enough. These sensors never streamline energy at their own to pass on at the specific target or surface. Being opposite to the active sensors, passive one usually depends on the natural sources like sunrays.

Therefore passive sensors can only work with proper sunlight environment so that there can be something to reflect. This technique possess multispectral and hyperspectral sensors which captures the earned quantity having multiple combinations of bands. All these combinations vary in the number of channels or the number of spectral bands they might contain. There can be one or more than one wavelength. These bands include infrared, ultra waves that are beyond the vision of humans' eye and also they have visible rays that can be viewed by human eyes.

1.2.2.1: Passive Remote Sensing Devices

a) Spectrometer: It helps in differentiating between several spectral bands and in their observation. This is one of the latest methodologies in the field of remote sensing which is technically feasible through the airplanes and space shuttles. In the beginning the results indicates that the direct remote identification and observation of target objects present over the surface of earth in a picture element basis must be executed by perfect sampling of the attributes undergoing absorption present in the reflectance spectrum. Both the airborne and space borne sensors have a capability of attaining several images simultaneously within 100 to 200 contiguous spectral bands. This property to attain thousands of laboratory kind of spectra via remote sensing is a biggest advance in the field of remote sensing.

All the concomitant research in computer and electronics technology to decrease these potentially massive information inputs at hand and to provide their storage, several latest analytical methods has been discovered to extract the complete information content of the given input. This emphasis over the deterministic method to tackle the multispectral input and its analysis unlike the statistical manner had been used in the past needs to be stimulates along the development in latest digital image processing technologies.

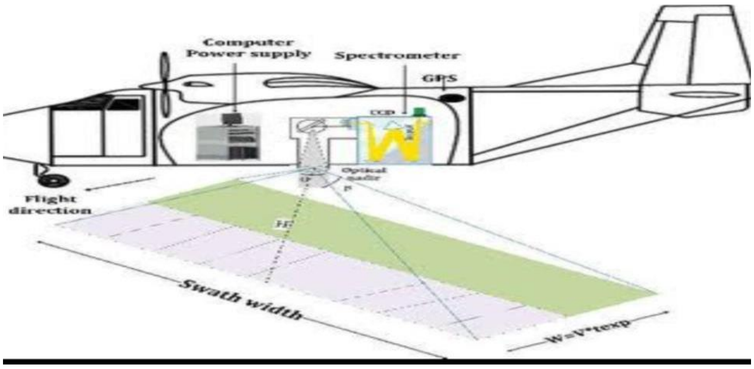


Figure 1.9: Demonstration of Spectrometer over a bounded surface area

b) Radiometers: This device is generally very sensitive and difficult to construct to capture thermal radiations from the electromagnetic spectrum which is generated by the atmospheric gases. Microwave radiometers are used in several applications like environmental and engineering areas. It also includes weather forecasting, astronomy of radio waves, propagation studies and proper monitoring of climate. The functioning of radiometers is basically dependent on the size of target, various kinds of obstacles lies in the path of travelling waves and on the temperature of atmosphere. A microwave radiometer contains an antenna system along with the components having microwave-radio frequency in the front end, they have a back end to process the signals at the intermediate frequencies. The signals in the atmosphere are sometimes very weak in nature and therefore they needs to get amplified near around 80db. More oftenly there are some heterodyne methods that are used to convert the signals from higher frequencies to lower ones and further provides the use of signal processing and commercial amplification. At lower noises the amplifiers increasingly becomes sufficient at the higher frequencies. For instance near around 100GHz, which makes this heterodyne method more accurate and significant. To reduce the drift present in receiver thermal stabilization is very necessary. The applications of radiometer are as follows:

- To measure the rate of rainfall
- To observe the temperature of Sea surface
- To detect wind speed
- To measure sea/ice contents and their concentration with their types
- To monitor cloud liquid water
- Soil moisture
- To detect atmospheric water vapours

c) Spectroradiometer: This instrument helps in finding the power of radiated waves in different spectral bands.

d) Hyperspectral radiometer: This is the most accurate device in case of passive sensing due to its very high resolution quality. Hence it is able to differentiate thousands of narrow spectral bands that are found ultimately in infrared, ultra and visible regions. Unusually big radiometers antennas have been placed on the top of satellites to measure the salinity. Width of these radiometer beams changes inversely along to both radio frequency and the aperture of antenna. The conventional radiometer antenna that uses microwaves has frequency different than 1.413 GHz as they are quite lower than this. The antenna having the larger aperture and diameter need to avoid excessive waves width and thus the size of bigger footprint. For instance, a footprint of 60 km needs near around an aperture antenna whose radius is 6m while on the other hand the conventional microwave radiometer antenna has the same radius of 1 to 2m only. So, to reduce the footprint by a factor of two it requires to double the size and dimensions of the antenna.

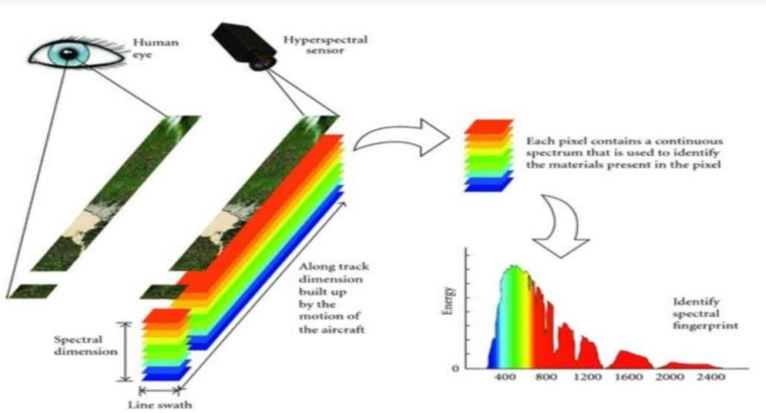


Figure 1.10: Hyper Spectral Radiometers demonstration

There are so many thinned and filled methods for the antennas having larger dimensions which have also reached to a greatest development stage where the functioning of saline remote sensing is excellent in its applications. The radiometers can be based on the use of radio waves or micro waves of the infrared waves. They basically functions to capture the noise power radiated because of the atmosphere with a significant accuracy. They are very sensitive and well calibrated with the radio tracers. Considering all the passive methods, we see two ways to measure the input out of vertical profiles. These are determined as multispectral and multiangular or sometimes both. For a given frequency and noticeable angle the thermal emission in the atmosphere can be expressed in terms of brightness temperature Tx which is also known as radiance.In other terms it is the weighted mean of mixture of water vapours, the absorption path between the layer and the tracer, a local absorption variable, and temperature at that time and the profile of liquid clouds. Every layer in the atmosphere contributes in giving a Tx value along with a weighting function that is not independent of the given temperature. If we change the contribution of every layer by a small change in the frequency or the phase at which the generation is noticed, we can easily get the data about the individual layers of the atmosphere. Further, in opposite to the ranging instruments like radar and lidar, the layers are not well defined and hence some mathematical functional techniques comes in role to derive the parameters or variables.to retrieve the representation of the given medium. However the present retrieval methods are applicable to well calibrated inputs that have radiometric production of excellent information to further process the boundary of layered temperatures. There are two most common gases in the troposphere that radiates the power in the microwave region. Energy emission is due to the molecular water vapours because of their rotational line in the spectral area at 22.235Ghz.The rotational bands helps in generating the molecular water vapour near about 60 GHz that extends to 40 and 70 GHz. As there is comparatively very weak absorption of the water vapour line, the radiation nearly generates from the troposphere entirely within 22.235 GHz. Therefore only the integrated variables like water vapours that can be precipitated are deterministic. For all such spectral window areas like 30 to 90 GHz, the radiations through the clouds are detected to generate some other integrated measurements like liquid form of water paths. But about the center of the oxygen band, the radiated measurements possess information of the vertical temperature profile around 500m at the beginning from above the ground. Here, we can see the concentrated oxygen band emission.

e) Imaging radiometer: Sense the target or a particular surface area to reproduce the image. This device is generally considered as the backbone of the visible remote sensing technique in order to observe the target and objects on the surface of earth from the instruments like aircrafts, drones, underwater submarines, spacecraft and satellites for both night and day time observation. The typical range covered by these radiometers is near around 0,3 to 2.5 micro meters which is significantly very small in the visible and infrared region of the electromagnetic spectrum. On the other hand it covers around 3 to 13 micrometres in the thermal infrared region which is slightly larger than the previous ones. The range of spatial resolution starts from very few centimetres to few kilometres.

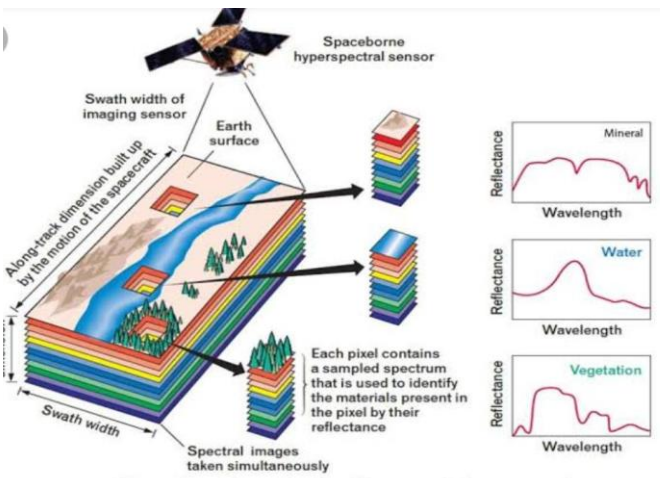


Figure 1.11: Reflectance curve variation in terms of mineral, water, and vegetation

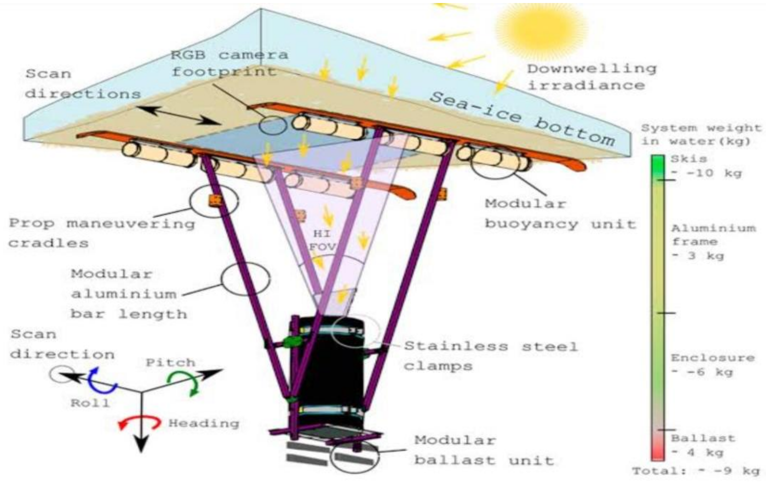


Figure 1.12: Under water Spectrometry

The typical range covered by these radiometers is near around 0,3 to 2.5 micro meters which is significantly very small in the visible and infrared region of the electromagnetic spectrum. On the other hand it covers around 3 to 13 micrometres in the thermal infrared region which is slightly larger than the previous ones. The range of spatial resolution starts from very few centimetres to few kilometres.

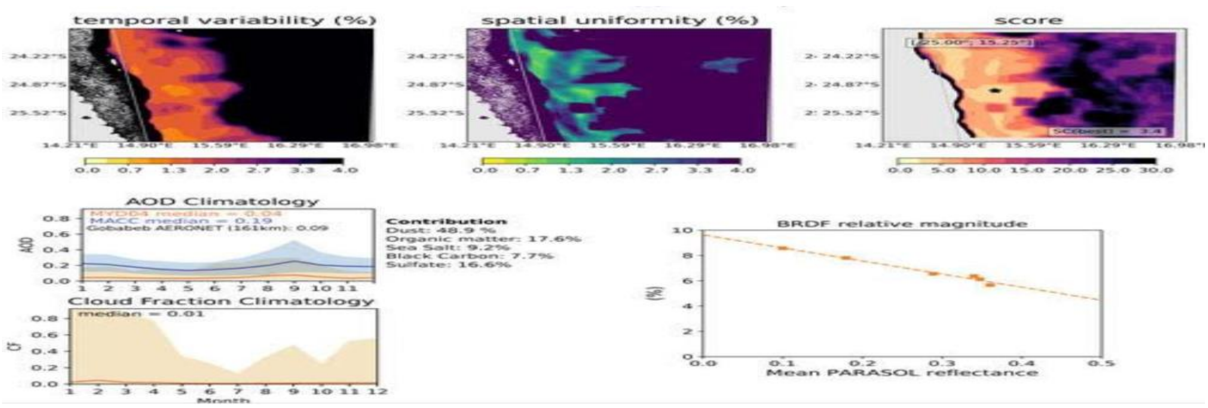


Figure 1.13: Identification of a site area by means of imaging radiometers

The experiments states that they can be used for a larger number of applications like:

- Taking several images from the earth surface also called as imagery,
- Changes in the land cover,
- Development in the urban and rural areas,
- The vegetation survey and its quality for agriculture purpose,

- Production of food,
- Types of land cover vegetation,
- Measurements of temperature,
- Characterization of sea and ice areas,
- To detect the colour of ocean ,
- To measure the water quality and lunar reflection,
- Detection of change in climate and
- Monitoring the endangered species.
- To recover the obtained images from sensors
- Forestry and Vegetation observation

For different quantitative uses the imaging radiometers have been calibrated and the data results are validated respectively. With the proliferation of the minute development in the applications, in the drones in latest years, we have expected a successful growth in the visible and infrared imaging extraction for remote sensing in the upcoming decades. The application can further be expanded because of the increased volume of input data generated by the remote sensing systems. The exponential growth will definitely increase the demand and stimulates the data analytics technologies with new open opportunities in future.

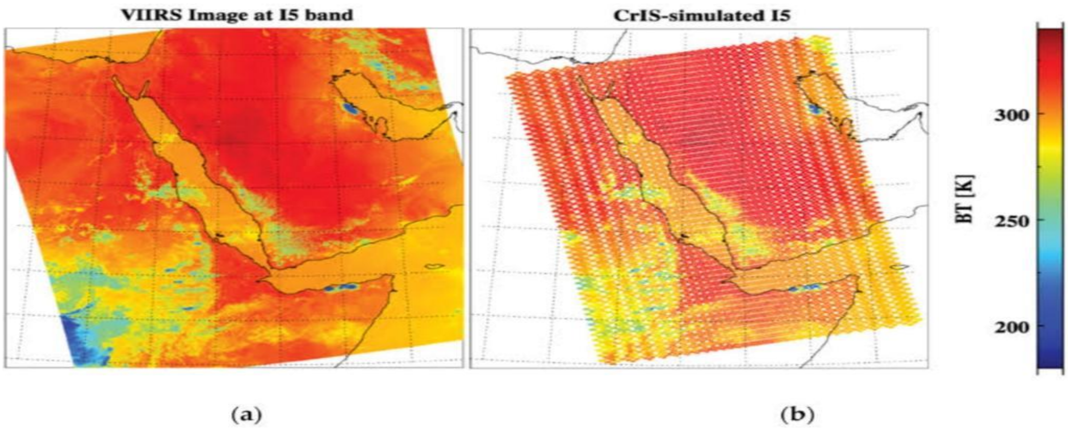


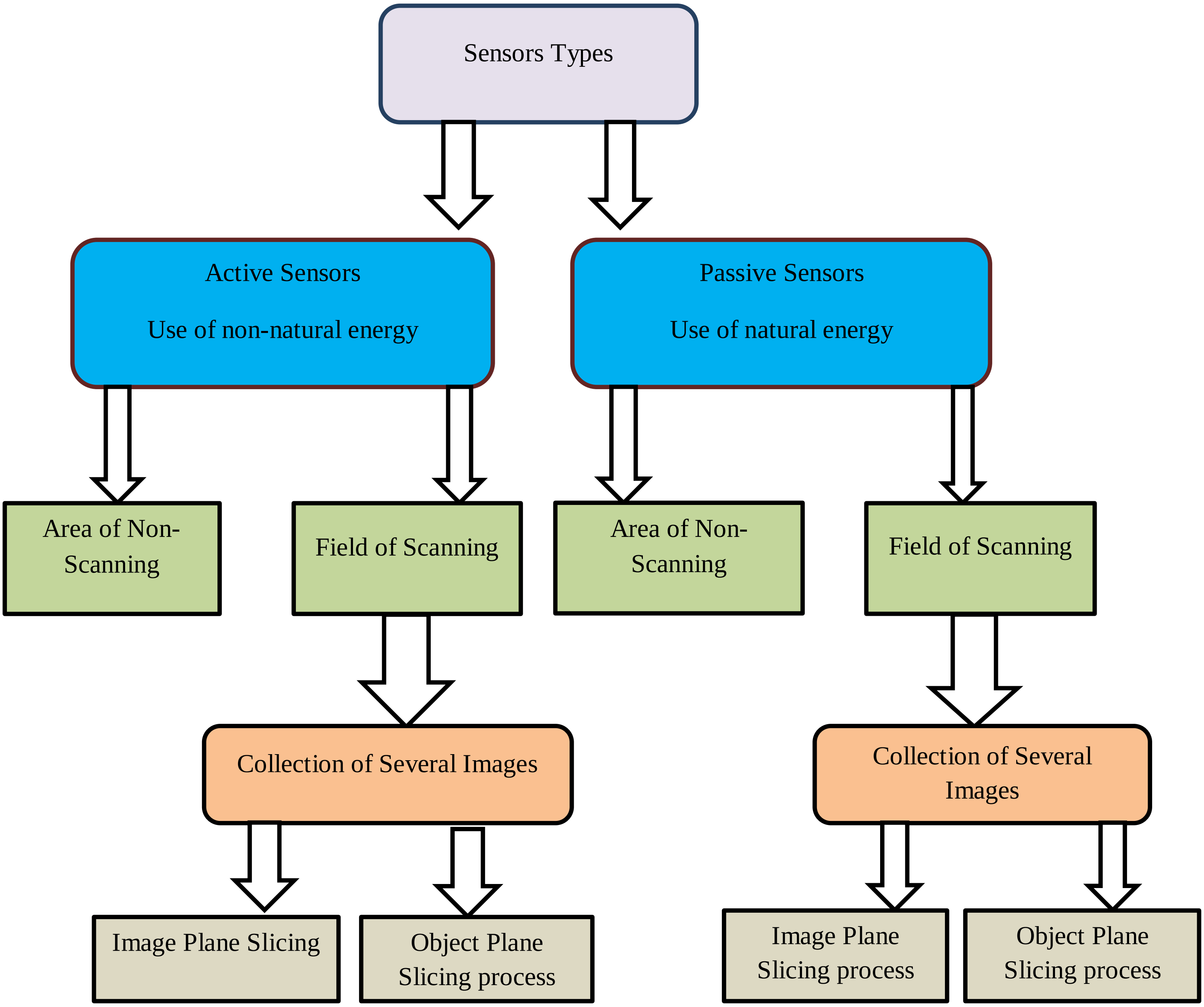
Figure 1.14: Fast and Accurate collection of Visible and Infrared Remote Sensing

f) Sounder and Accelerometer:

➢ These devices help in sensing the conditions within the environment vertically and they also detect the change in the speed with respect to time which is a physical quantity called acceleration.

- Along with several devices mentioned so far sounder and accelerometers also plays a significant role in the field of remote sensing and thus they are also in huge demand as passive sensors.
- Nowadays, remote sensing is incredibly enhancing due to the lasts advancements in these devices. The future research can be build based on the use of the instruments.
- Below is the flow chart explaining the active and passive sensor that we have discussed so far in the field of remote sensing.

Figure 1.15: Flowchart: Depiction of types of sensors with different categories



1.3: Discussions made so far

Till now in this chapter we have discussed the concepts of remote sensing and the applications across several fields. We have seen the types of sensors that help in monitoring the surface of earth so that one can get the idea about the target material. One of the most common applications is imaging. It is done by means of the passive sensor which is basically the sunrays. As soon as the sun rays falls on the surface of earth, the reflected radiations are tracked down by the satellite. After this based on the spatial and spectral attributes one can study these data inputs more accurately. In next chapter we are going to see about imaging and it's types with the applications so far. Researchers have been going on these fields with different methods by the researchers nowadays.

In next chapter we will see the study of multi spectral and hyper spectral images in the field of remote sensing. It will help to understand the further need of classification and detection for images successfully. The surveys tell us that there are so many advantages of imaging in common life. Basically it plays a role in medical and also in the surveillance. The advantages can be seen in:

a) Medical

b) Forestry and agriculture

c) Weather forecasting

d) Under water object detection

e) Measurement of atmospheric particles

f) Detection of targets under dense covers

h) Classification of Hyper Spectral Data

Remote sensing is the latest technology that commonly helps in the surveillance purpose to many countries. We are able to detect the targets under forest covers and under the dense areas. This brings us to the fact that observation and detection of these targets is the most significant challenge to our day to day data.

One of the common studies that we will do is the imaging of land cover vegetation over different spectral bands within the entire electromagnetic spectrum to obtain complete idea about the forestry. Latest research has been done using machine learning tools for higher accuracy and precision. These supervised tools help to get a full view of land cover areas that we will discuss in the upcoming chapters. Many devices have been developed like computer topographic spectrometer, fibre reformatting tool, hyper spectral array, lenset array spectrometers, integral filed spectrometry and multiple spectrometers.

CHAPTER 2

IMAGING TECHNOLOGIES

2.1: Introduction

It is a recent technique which explains the representation or production of a form for a given object. Basically it is a visual representation. It provides a platform to generate, capture, preserve and duplicate the original images any number of times we want. Science behind the imaging process is a multidisciplinary area related to the creation, gathering, duplication, observation, enhancement and proper visualization of the given sets of images. It also includes those images which are very hard for the human eyes to detect.

Imaging technologies are emerging very fastly and with time they have covered research and researchers from various departments like maths, physics, electronics, computer vision, electrical engineering, and psychology and computer science.

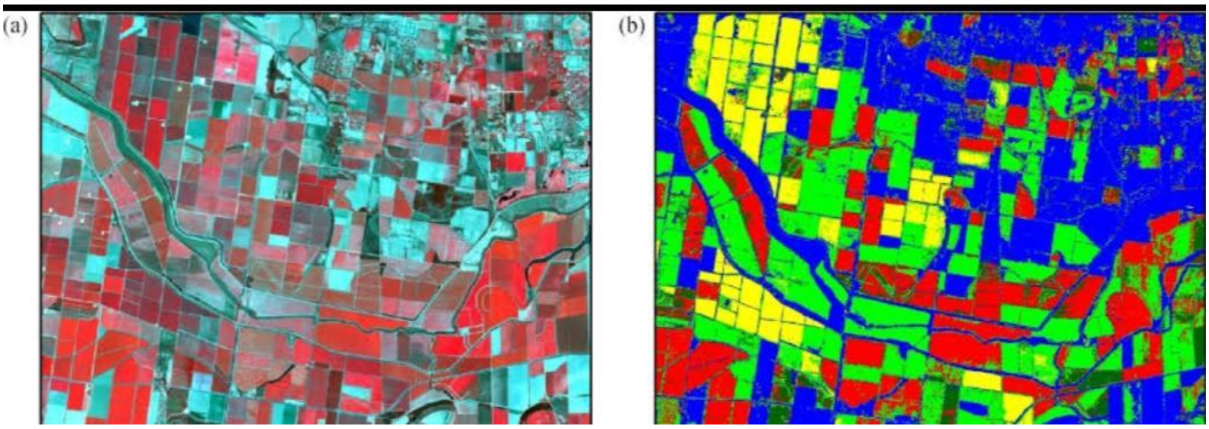


Figure 2.1: Data Set obtained from Satellite view as a result of Imaging

2.2 Components of Imaging

a) The visual system: The human visual system has been considered by the developers for the psychological methods that take place inside the human bodies because they have s sense in receiving the data sets as input through various visual systems.

b) Topic of Image: For creating an imaging system the focus of the designers must have the consideration regarding the observables related to the topics which are going to be images by the systems. These experiments generally take the form of radiated energy or can say a reflected energy like electromagnetic and mechanical energy respectively.

c) Receiver Device: As soon as the observation linked with the target starts characterizing themselves, the designer or the developer must get an idea to identify and integrate the methodologies that are require to detect those observations. For instance, in the case of buyer camera having digital properties, the method not only includes optics for gathering the power but also the energy lies in the visible range of the electromagnetic spectrum. They might as well have some electrical detectors to convert the electromagnetic energy to an electrical signal form.

d) Processor Device: In case of every digital imaging device, the electrical signals are generated by the captured device. These signals when generated must be manipulated by the mathematical algorithms that will help in formatting the signals. In this way they can be projected as an image. In general, there are various multiple processing tools that are more often involved in the generation of a digital image.

e) Display Tool: The display of any image will take the given electrical signal that has been changed and modified by the processing tool and render itself to some medium that shows visual behaviour. For instance, printed papers for hard copies, images, television, projector and monitors. Basically we can explain that imaging is the manipulation, generation and storage of an image and it basically emphasis on retrieving the information from a given image. It sometimes plays a vital role in diagnosis and medical fields, radiologies and fluoroscopy. Imaging gives us an idea to notice every minute detail regarding that target. It provides the analysis of molecular portion and their operation at very lower energy levels significantly. It is basically done by taking the image for a particular area over the surface. The other names for imaging are computer imaging or digital imaging which is further processed by the imaging software to generate, edit, and duplicate the given image at a time.

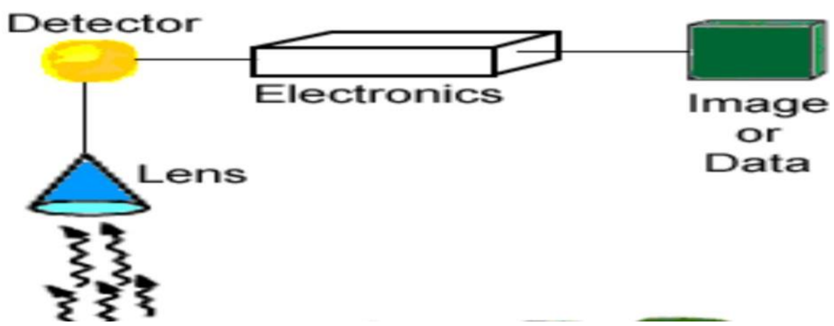


Figure 2.2: Working building blocks required to function Imaging

2.3: Basic Application Fields

There are numerous of subfields where the process of imaging plays a significant role nowadays. We can see that various subfields where it is applicable includes :

a)Image Processing: It is the process of analysing the digital image through an mathematical algorithm using digital computers. It has so many advantages over the analog image processing because of the fact that it allows the larger range of mathematical functions to work on the input information such that the problems like noise and distortion in the signals at the time of processing can be avoided. All the images are two dimensional data and thus the digital image processor may be designed to function as multi-dimensional systems.

The creation and modification of all such digital image processor may be effected by primary factors like development in the computer technology, enhancements in the mathematical tools and further the changes in the discrete mathematical solutions. The demand of larger range of use lies in the forestry, military, industrial zone and medical sciences has increased significantly in the past decades.

b) Computer Vision: This is science field that is interdisciplinary to handle the system to gain a high level knowledge from digital images or the digital videos.

As looking from the engineering point of view, the perspective is to understand and create some tasks that the human naked eyes can perform. The areas of computer vision include the techniques of acquiring, processing and viewing and getting knowledge about the images and their extraction that are higher dimensional data.

Objective: The main objective is to pick the data from the real world and thus generate various symbolic data that are numerical in nature. This can also be termed as the transformation of visual images to the data into which the people can apply processes and can edit sufficient actions. This detangling data of numerical input from the images has been modelled as constructed version for geometry.

If we say scientifically, then computer vision is related to the theory of any artificial device that gains the data from a group of images. Then after this extraction, the data can have multiple forms. For instance, video sequence, 3D scanner generated multi-dimensional image, camera generated views and medical images. These technological information about the set of images after the transformation helps in modelling the construction of newly set of computer vision devices.

Fields Applicable: The area to which this technology can be applied includes reconstruction of particular scene, detection of object from any scene, estimation of three dimensional data, to analyze the motion, indexing and restoration or classification of images.

c) Astronomical Imaging: This is the photography in the domain of astronomical images and objects, it is basically related to the celestial bodies observation, and all the views at the night times. The very first data that was taken astronomically was of the moon, but now there are so many advancements in this field regarding the observation of the images. By means of this technology one is able to view the galaxy in very deep and detailed manner, thus is very common to consider this as an emerging branch of imaging. As imaging has so many advantages this is served as the best possible one for the researchers in past few decades. There are many researches going in this area to visualize the latest enhancements. Astronomy keeps on viewing the celestial targets and thus they can generate the vision of the object at the day or night time depending on the users. We can hence prove that imaging is related to real and practical analysing of datasets.

Figure 2.3: Astronomical View of Celestial bodies.

The extended targets are being recorded such as sun, moon and stars but apart from that astronomical imaging allow the observation at minute level and it also has the capability to analyse the invisible object to the naked human eyes. For instance, dim stars and so many galaxies present in the sky. The long term observation may be able to accumulate large number of photons over these targets. At present astronomy is serving the wide range of application and methods. With very less hope, astronomy science is employing greater exposures because both digital and optical instruments can handle the light particles known as photons for a longer period of time.

d) Imaging in Medical: This simply means a type of imaging method used in biological field. For instance:

- There is a method that is used for determining the animals present inside a laboratory using some kind of luminous protein. This is known as bioluminescence imaging process.
- Another field is calcium imaging that can be used to calculate the status of calcium for a tissue with the help of fluorescent.
- The next one is about diffuse imaging using optical tools to produce various numbers of images of human bodies using infrared light.

- Similar to this term is diffused weighted imaging process that is a simple MRI which basically works over water diffusion.
- Fluorescent imaging is a lifetime key that is used to slow or reduce the rate of a given fluorescent sample.
- The other filed is known as gallium imaging in which a nuclear process happens to detect the infections and cancer inside the human body.
- There are some chemicals designed so that clinical are allowed to check if there is a mass benign or malignancy present.
- There are several medical technologies that include imaging study.
- In medical imaging we can see the creation of images of a specific body part of human body so that further diagnosis of the disease can be made on the basis of which the patients can be treated nicely.
- Microscopy is the way by which objects and attributes that are created can be noticed by the naked human eyes. This method helps to simplify the visualization process in an easy manner everywhere.
- In order to get a chance to study the molecular object we have a method known as molecular imaging by which the inside organs and the study of their functioning can be made easily in a specific manner.
- There is an area of thermography which serves as a platform of thermography through which one can derive the diagnosis indication with the help of infrared rays for the human body itself.
- Nuclear medicine is the branch of science technology that generates images to a study the internal organs using the administrated radioactive materials. Hence the nuclear advancements are very advantageous nowadays.
- If there is use of light particles for all the activities that we have discussed so far, the method is known as optical imaging. There is an investigation related to biological research processes and the diagnosis for medical uses.
- Opto acoustic imaging is new keyword in this area which is based on the photo thermal effect to know the accuracy in the ultrasound with the spectroscopy having a deep depth. The precision is also known in this case.
- Another word is photo acoustic phenomenon to find out the vascular disease and disorders like cancer and non-ionizing cells and large pulses.
- The last one is ultra sound imaging in the medical sciences that has tremendously changed the world. A very high frequency sound is used to analyze the inside view of internal organs perfectly to detect the disorders.

e) Digital restoration for images: It is method to restore or save the digital copy of the appeared physical image that has been distorted in its appearance by some natural activities or some man -made processes or can say some kind of environmental operations. Another way to say this is neglecting the real image due to some cause.

Figure 2.4: a) Image before restoration b) Image after restoration

The restoration of a digital photograph makes use of several editing in the photograph so that there can be a removal of the visible distortion and the aging effect from the copied version of the real images. For repairing the distorted image we uses raster graphics methodology such that the appearance of any kind of digital image can be added to the digital version in order to replace torn and the missing portions of the real picture. The evidences of the distortions and damage like scratches and any other unwanted signs of the photographic versions are differentiated from the original image without using any software manually or by painting over the pieces meticulously. The unwanted colours are also removed or altered by changing the brightness and contrast of the photograph by some image processing devise like image enhancement or the restoration of any physical image. This is also true for digital image photography.

Photographic particles are very susceptible to any change in the physical and chemical damage or if there is any change in the biological area they can also be determined easily. There are some processes to which it is sensitive and they are fire, water, vandals, thieves, some external forces, light energy, unwanted incorrect temperature, humidity level and some dislocation in the position of the objects in the real image.

Conventionally, there are some efforts made in the past to focus on the physical appearance of the real photograph but the restoration of the digital surrogates also become very important at the same time. The original of the real images that are very sensitive and fragile or can say very valuable can be easily protected as soon as the digital surrogate tries to protect them. All the images that are digitally generated are saved more oftenly and they reflect the intension of the person who takes them is an exclusive copyright owner for that image. It is not suggested that the changes or the conversion regarding the original images can be personal or opinions based. Further, with the absence of copyright permission, the museum can create and restore the images so that the informational purpose can be fulfilled.

f) Medical Imaging: This is a technique which is related to functions the internal part and processes of the human body. A body full of clinical and internal biological processes needs an analysis and an intervention for the successful visual representation to function as the analyser of some of the body parts or cells and tissues that is also known as physiology sometimes. Medical imaging is a scientific method that is seeking to generate the internal functioning which may be sometimes hidden under the skin and the bones of human body. To well diagnose the human body and the disease present inside the organs medical treatment is very important. It also serves as a base to database of much normal anatomy such that it would be possible for the doctors to find out the abnormalities inside the organs which are also called as physiology. Further the organs that arte removed from the body can also be imaged to serve as the best medical treatment and these procedures are basically known as the part of pathology, they are not under the medical imaging.

The biological imaging sometimes incorporates the widest portion of the medical filed as it also includes radiology in which there is the imaging with the help of X-rays radiography, also the medical imaging find the application in magnetic resonance, ultrasound, thermo logy and nuclear medicines branch. They also serve their best in tactile imaging, endoscopy, and PET scan process. The PET scan is generally the whole information imaging of human body which is known as positron emission tomography. The other application is SPECT that is single photon emission computed tomography.

Measurement has shown that they are primarily constructed to develop images like EEG, MEG, ECG and several other:

- EEG: Electroencephalogy
- ECG: Electrocardiography
- MEG: Magnetoencephalogy
- Some methods are data susceptible in order to show the graphical representations as a parameter graph with time or information which might have inputs about the observation of a given location.
- Seeing a limited comparison there are some methodology that are considered as medical imaging.
- Application also lies in CMOS industry.
- Digital signal processors and microprocessors.
- Micro-controllers and biosensors
- The semiconductor based power device like image sensor
- It can be considered as designation of various set of technologies that generates images non-invasively at the internal parts of the human body.

Many mathematical algorithms can be seen as the solution to all these problems. This indicates that the cause of problems in the living cells is derived from the observed wave effect. For the case of ultrasound the ultrasonic pressure is carried out by the probe that have waves and signals creating echoes to enter inside the tissues or cells of the human body to get an idea of the internal structure. For projection radiography processes the X-ray radiations are carried by the probes that can absorb the waves at different rates and time intervals by so many cells like bone, fat and the muscular tissues. The definition of non-invasive is basically used to represent the donation of a method to work without the instrument to introduce waves inside the human body which is considered to be the most common case for any patient. Below is the imaging of CT scan showing abnormalities and successive sections having transverse behaviour about the patient.

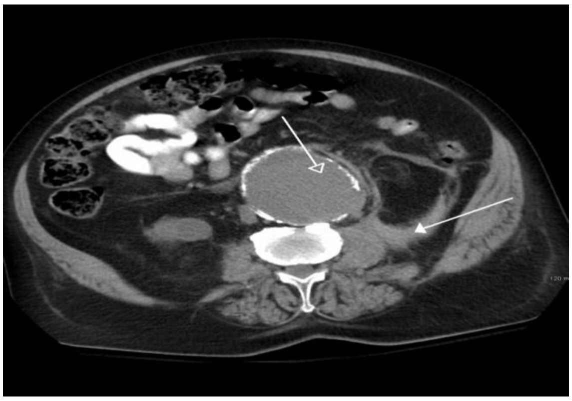


Figure 2.5: Image of CT scan indicating abdominal aortic

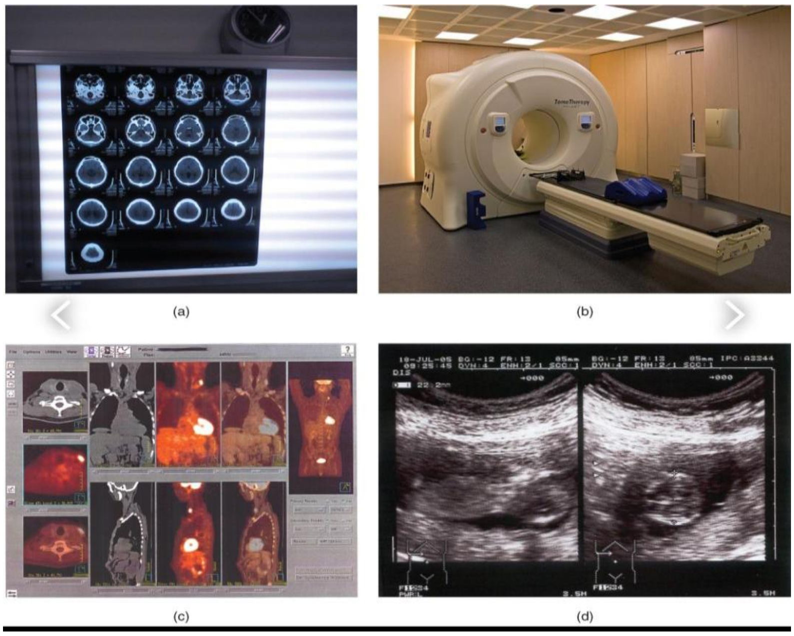


Figure 2.6: a) Results of CT scan for successive sections b) An MRI generated for any patient randomly c) PET scan of a patient that is based on the use of radiopharmaceuticals

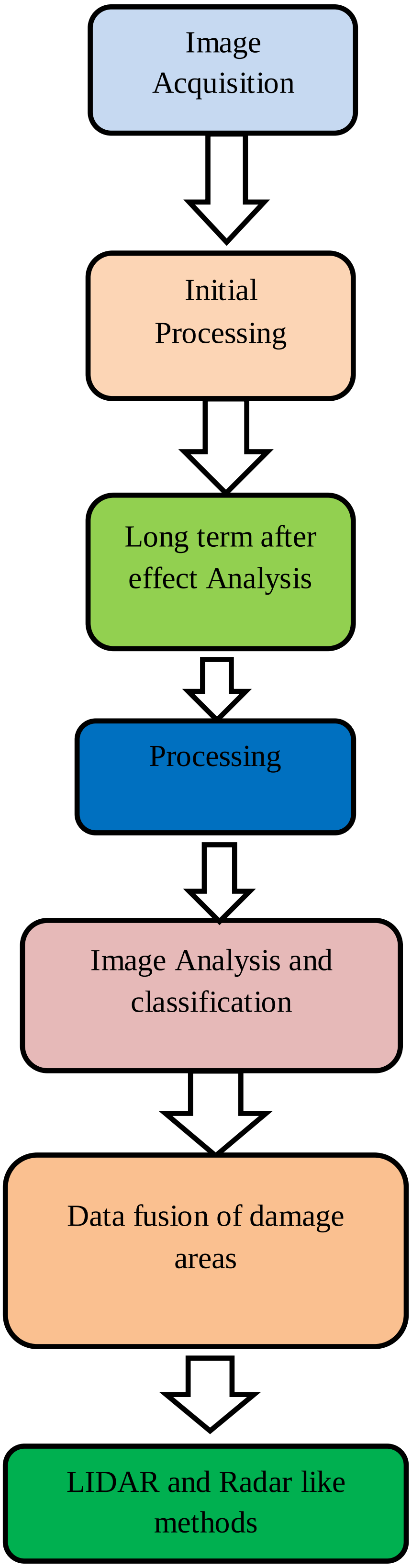
Sometimes magnetic resonance imaging device scanner or the nuclear magnetic resonance imaging scanner uses the powerful magnets so that the polarization of the excited hydrogen can take place. So far we have discussed the various types of imaging. There are two basic types of imaging that we are focusing in this project work. These are Hyper Spectral and multispectral imaging. Then we will study the classification of hyper spectral images by the use of passive sensor to get the detailed information regarding any sub area in the earth surface. First we will quickly have a look over the multispectral imaging and then we will study deeply about the hyper spectral imaging.

2.4: Multi-spectral Imaging

In multi-spectral imaging there is a way to record the image given as input in certain specific wavelengths that are widely spread across the electromagnetic spectrum. These wavelengths can be differentiated by the filters or can say is observed by the use of several devices that are proved out to be the most sensitive instruments for a given particular wavelength. This ranges from frequencies within the visible light waves to infrared or ultra violet rays. This kind of spectral imaging provides the observation of some extra data inputs that the normal human eyes is not able to capture properly with the presence of visible receptors for different colours like red, green and blue which are considered to be the primary colours. Initially, it was discovered for only the military and surveillance purpose so that objects can be identified perfectly for reconnaissance. Very initial stage of imaging technologies has included multi spectral methods so that they would be able to map the areas of earth that are related to details like coastal edges, forestry and agriculture, and different landform. In multi spectral imaging process we have find that it calculates the light in a significant narrow and small number of band. For instance typically near around 3 to 15 bands. We cannot have hundreds or thousands of contiguous number of spectral bands availability. We can have a look on the applications on which we are able to use multi-spectral imaging successfully. It is given as follows:

a) Target Tracing for Military Purpose: In multi spectral imaging the measurement of light emission is done with the help of multi-spectral imaging and this method if more oftenly used in the detection and observation of tracing the military objects successfully. Past in 203, many scientists have reported a combined research sitting at the United States Army Research Lab and another collaboration of Federal Technological Alliance, that the dual spectral bands for multi- spiral imaging contains a focal array of different planes. In this focal plane so many scientist have successfully created a view at two different planes in the infrared spectrum with no time difference. As the mid-wave and the long waves in the infrared region are used to detect the wave energy that are with same phase to the targets and hence there is no requirement of some other external light waves as energy, therefore there is also a thermal imaging method based on this technology.

Figure 2.7: Flowchart representing the acquisition of image in given order



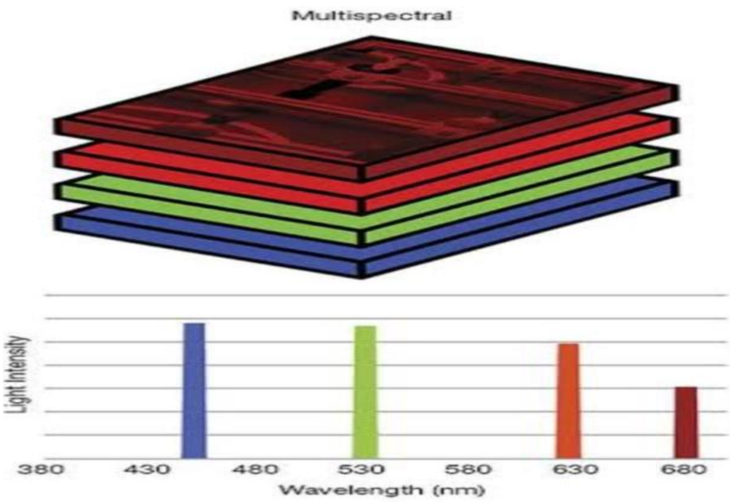


Figure 2.8 Representation of Multi-Spectral Imaging over several wavelengths

We can see some key points regarding this imaging:

- The contrast property of the input image set generated through imager is totally dependent on the targets creating the emissivity and the thermal temperature.
- Each and every target has their own infrared spectral tendency which can aid the identification of various targets easily.
- These signatures possess by every target are very less in terms of the pronunciation for the hyper-spectral devices in which the image dataset have more number of spectral bands as compared to this multi-spectral images.
- If so whenever there is an exposure to the wind, rain and more commonly say the dramatic reasons, the surface of the objects can be able to reflect the infrared zone energy. Due to this reflection there can be some mis-construe in the true and correct readings for the target inherent radiation in a given spectrum.
- Very few imaging devise which are based on the use of MWIR methodology plays excellent with the reflection caused by the solar energy.
- Hence as a reason more number of definitive images can be generated for the hot target at that time. For instance, we have engines that are much better than the LWIR based techniques as they are based on MWIR methods.
- Although the LWIR methods can have some of the advantage as they serves better in case of hazy environments and smoke or fog as there is very less scattering present for the longer wavelengths.
- So LWIR are beneficial in every aspect. In terms of environmental condition use of LWIR method suits the best result. At the end there so many researchers that are claiming the dual band methodologies together to the advantages so that there can be more information extracted from a given image, especially in case of realm object

tracking. In case of detection of target in the night time, there is thermal imaging serving best for the single band operation under the multi-spectral imaging.

- MWIR and LWIR are dual bands methods that have shown foremost results in very special visualization at the night duration as compared to use of MWIR or LWIR alone. In USA the army has significantly reported that there is dual band LWIR and MWIR plane array platforms that are demonstrated better as compared to visualization of vehicles that are tactical over the MWIR only after the detection with the help of night and day time devices.

b) Land Mine Extraction: To find out the emissivity of the ground areas, this technology called multi-spectral imaging is able to extract the presence of the missiles that are under the ground. Surfaces and the lower areas under the soil possess several kinds of physical and chemical features that may be appearing in different spectral bands. The soil that has the disturbed compositions has increased emissivity in the wider range between 8.5 to near about 9.5 microns. There is no further change in the demonstration of the wavelengths that are greater than 10 microns.

- In the US Army research labs there is a dual MWIR and LWIR based plane arrays that is based on the use of three primary colours like red, blue, and green
- This process simply tries to enhance the emissivity. We can see that the red detector behaves like a backdrop in order to verify the trams within the undisturbed soil compositions.
- The reason is that they are very sensitive to the 10.4 micron meter frequency. The blue detectors are used to find out the search locations and they are also sensitive to 9.3 micron frequencies.
- In order to see the intensity of the blue images there can be some changes made while scanning, and that area is more likely to get disturbed. The researchers are submitting their fusing image based detection qualities.
- In recent works there is unmanned vehicles that are using the remote sensing to work through the light weights in the multi-spectral infrared region of the electromagnetic spectrum.
- These waves are allowed by the thermal sensors to work as a rapid broader region to find out the contamination and the mapping surveys. There are some of the techniques that are based on the development and testing of an automated method of landmine recognition and detection for the scatterable particles with auto-personable landmine broader region surveys.
- These methods which calibrates the recognition utilizes a plastic body which is encapsulated in liquid form with an explosive polyethylene in the design.
- The first research base on the analysis of multispectral and data set that are designed thermally have been collected by the UAV devices under the surveys featuring a PFM-1 in which there is a scattered landmine as a target particle and the current outputs rely

on a supervised learning to automate the detection with an effort to use the convolutional neural networks.

2.5: Hyper Spectral Imaging

So far, we have discussed about the concepts of multi-spectral imaging. But there is another type of imaging that is commonly known as Hyper Spectral Imaging. It is similar to the fact that this technique also collects and processes the data as an input through across the data received from the electromagnetic spectrum. The main objective of this imaging is to receive the spectrum corresponding to each pixel present in the entire image at a time. Then the second task is to find the objects and then to recognize them which is followed by their detection. Spectral imagers are basically categorizes into three different branches as follows:

- These are known as push scanners. The task of a push scanner is to read a given image at an instant of time.
- The second one is band sequential scanner whose task is to capture the spectral information to obtain the images carrying several portions of areas that have characteristic of different wavelengths.
- At last the third one is snapshot imager that is very helpful in the generation of image at any instant using the string arrays.

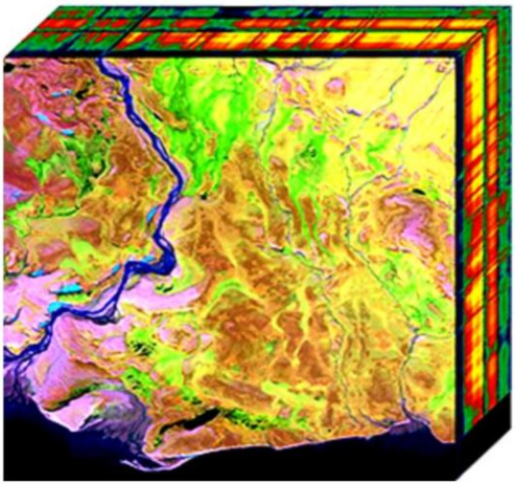


Figure 2.9: Hyper Spectral Cube in two dimensional view

When we talk out the human eyes we can see that our eyes can only perceive the rays reflected from the visible portion of the electromagnetic spectrum. This and is also captured basically in three different band colours which are green, blue and red. In general the green is the representation of medium wavelengths while the red is longest and blue is the shortest in the visible portion. In the Hyper Spectral Imaging concept we are using this visible band information along with the information from others band too. These bands are infrared and ultra-violet. The resolution parameter covers a broader range of the spectra. This method is very useful in the measurement of spectral bands that are continuous in nature unlike the multi-spectral imaging because it only captures the discrete spectral bands that are spaced. Our researchers have developed some sensors and the devices for processing the task so that

we are able to apply it in the fields like astronomy. Forestry, agriculture, biomedical science technology, physics and security and surveillance purposes in the military. For an instance we can see the condition where oil fields are found out because of the spectral signatures that provided the geologists some new information. We can have a look over the methods that are helpful in the scanning process in Hyper Spectral Imaging. These are spatial scanning, spectral scanning and non-scanning discussed below.

2.5.1 Spatial Scanning

This is the platform in which we take individual two dimensional sensors to generate an output having a full slit spectrum. In the Hyper Spectra devices the spatial scanning generates the slit spectra by making a projection of a strip for a scene in the slit and after that there is a dispersion which is granted by the prism. The devices have some disadvantages that there is presence of images per lines through the boom scanners and at the same time the mechanical portions are added with the train of line scan systems. Due to these situations the dimensions across the spatial boundaries is received by the line scanners. To reconstruct back the original image we need an accurate and precise platform to stabilize the pointing.

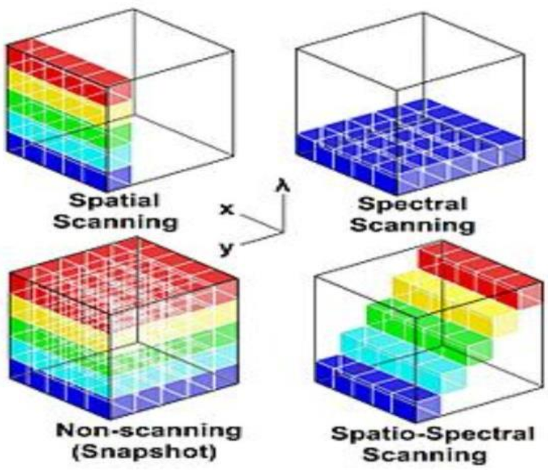


Figure 2.10: Visualization of sections in spatial dimensions in HSI image cross-section

None other than this for remote sensing application the line scanner plays a significant role. They are also very sensitive to the mobile platforms. Point scanning the great example that comes under the line scanning in which slit is replaced by the use of point aperture and the dimension of the sensor is only one dimensional unlike in the line scanning.

2.5.2 Spectral Scanning

This is the second important parameter in the hyper spectral imaging where every two dimensional sensor has a single colour or commonly called monochromatic nature. The spatial mapping is represented by (a,b) map in the scene. HSI tools that we need to the spectral scanning generally uses optical band pass filter which can be either tunable or fix.

The image area is scanned via interchanging two filters with each other spectrally where the platform does not change. The phenomenon of spectral smearing can happen in scanning devices which are stared or stationary, this is due to a movement or a motion event within the scene. This negates the correlation and detection

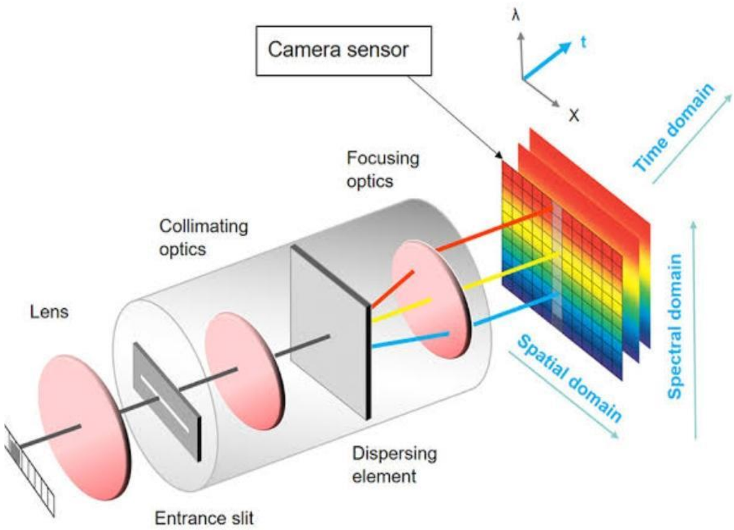


Figure 2.11: Spectral Scanning depiction

among the spectral data. This also provokes a drawback at the same time for providing the picking and selection of spectral bands, and to describe a direct representation for the spatial figures that is present in the image. If we are using any imaging device over a moving platform, like we can say the spacecraft, airplanes and many others, there is observation of different images at multiple wavelengths in respective to multiple areas within that scene. The spatial attribute helps to talk about the respective pixel.

2.5.3: Non-Scanning

The third case is non-scanning where a single output valued sensor is used to represent all the spatial (a,b) and spectral attributes (μ). This sensor is two dimensional and it can obtain the complete data cube at a time with the help of non-scanning without the use of any other scanner. If we talk figuratively we can say that the use of single screenshot helps to develop a perspective projection that finally helps in generating the reconstruction of the original data cube.

2.6: Results and Discussions

In this chapter we have seen that remote sensing is applicable in various applications and it make the use of imaging systems. We have discussed the different imaging techniques that have been used for the remote sensing purpose. Two of them are multispectral imaging and hyper spectral imaging. The basic difference between them is given below:

- The number of bands in multi-spectral imaging is less in numbers while there is a lot of spectral band information in hyper spectral imaging. For instance, there are 3 to 20 bands present to cover the spectral attributes while in hyper spectral system there are hundreds of such bands from the visible, infrared and ultraviolet region of an electromagnetic spectrum.
- The bands are much broader in the multi-spectral imaging as compared to hyperspectral imaging.
- The hyper spectral sensors are more sensitive to changes in the reflected energy waves.
- As the number of bands represents the data information, it is clear that the information is very well represented by hyper spectral sensors.
- We can have a look on one example of forestry that multi spectral sensor can only tell the areas of forests while hyperspectral sensors can view the species of tree present.

As a conclusion it is better to use hyper spectral imaging to generate the view of target in more depth. The computational costs that are required to design these tools are very high. So the effort to decrease the expensive cost demands a perfectly non costly spectrometer and a non-scanning instrument that is a prototype of optical computing tool with multivariate features. The demonstration of these tools is based on the multivariate optical measurement for the spectral calculations and the light modulators. These platforms provides the chemical data as input for calculation that is prior to imaging methodologies because they rely on the old fashioned cameras devices that have no computing for them. The drawback says that there is absence of spectral data in these systems due to the chemical data processing and further the regeneration is not possible in any terms. Hyper Spectral Imaging is applicable in one of the basic task called as Image Classification. To view the vegetation area in different species and more attributes the images capture as a process of remote sensing is then served by the hyper spectral sensor for their classification. Many works in the past have been done to perform the classification using deep neural models and unsupervised methods. In next chapter we will see the method in past for the classification for hyper spectral images just to understand the nature of imaging. So far we have only understood that there are many images that we receive on daily basis from the remotely sensed area and we need to do operation on them in order to recognise them. It would be really helpful if these images are from hyper spectral choice as they contain a significant amount of information and hence helps to detect the target successfully. Many field use this principle as their daily operation just to classify the images and we are discussing this principle in the upcoming chapters. Many field use this principle as their daily operation just to classify the images and we are discussing this principle in the upcoming chapters.

CHAPTER 3

IMAGE CLASSIFICATION

3.1: Introduction to Image Classification

In the beginning of this work we have discussed about the task of remote sensing. Image Classification plays a vital role in field of remote sensing. The remotely sensed images are passed to various methods so that further observation can be made to predict the nature and target properly. In past vey decades hyper spectral images enjoys a lot more models and methods have seen many advancements with time. In this chapter we will discuss about the image classification and the previous methods that have played a role. Basically image classification indicates about the action of getting so many information or the extraction of various attributes from the given hyper spectral or multi spectral scene. Thematic mapping is then done as soon as the information is extracted from the images successfully. This is the task of an analyst to perform such classification and then do the computing further. Classification can be of two types.

- Supervised Classification
- Unsupervised Classification

In this chapter we will only discuss the unsupervised methods more deeply while at the same time we will have a look over the supervised methods. There is full version of multi variated devices to analyze the supervised and unsupervised techniques. The classification method is not a single step task, but it requires multiple steps and thus the toolbox has been generated to provide integrated environment. The toolbar provides a support in the functioning of these methods for classification and also generated some extra functions to recognise the data properly, to generate training images and sample documents, to create a good quality recommendation to attain classification and multivariate processes.

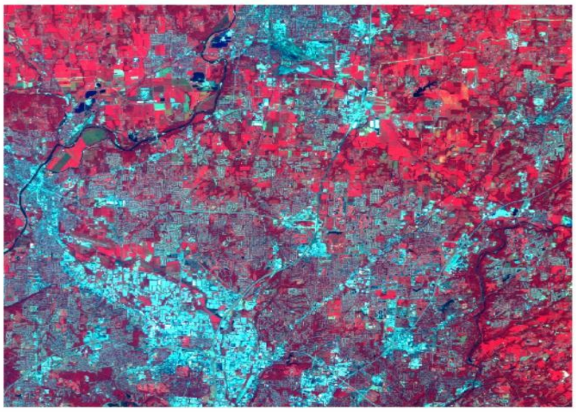


Figure 3.1: Landsat Image Classification

Above is the depiction of Landsat image which is classified properly and accurately using one of the classification techniques. The figure shows that the accuracy states that image classification is very helpful in case of hyper spectral imaging just to know the target more deeply and thus the applications in so many fields is helping the world to resolve so many problems. Now we will see the block chart representation just to understand the concept of unsupervised and supervised classification. The block chart of unsupervised classification is as follows:

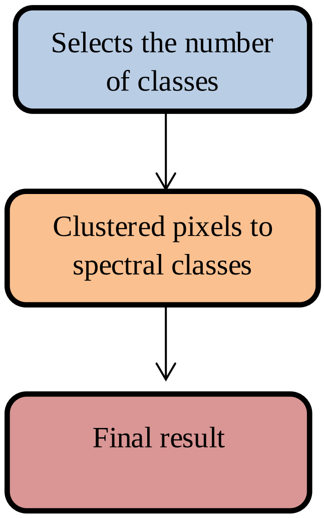


Figure 3.2: Block Chart to understand unsupervised classification

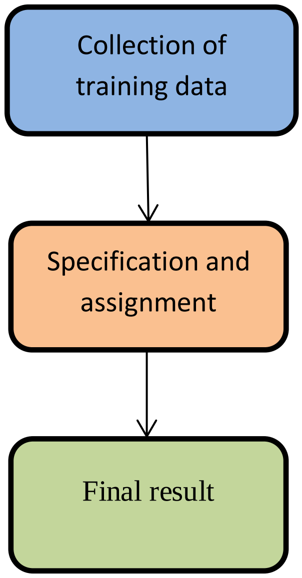


Figure 3.3: Block Chart to understand supervised classification

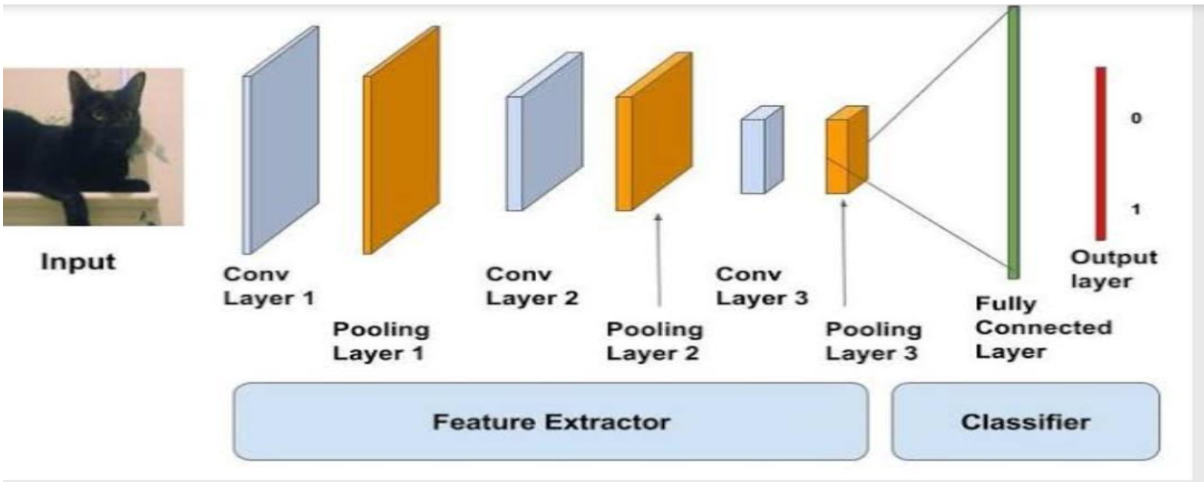


Figure 3.4: Supervised way to extract the image attributes successfully

The previous works for image classification that uses supervised methods are discussed below one by one:

3.2 Support Vector Machine

This method is well known supervised way of classifying the images successfully. The learningtakes places where the models are associatedwith the algorithms that uses the information for recognising and firther the classification. We can also use SVM in the problems associated wuth regression. But this is not our concern in this work. We are only associting our work with SVM for classification of hyper spectral images to draw the attirbutes information correctly. Machine learning is a wonderful phenomenon that is very benificial for us regarding artificail intelligence and also it provides the system to function itself which is also called a slef learning. It meant there is absence of explicit programming.

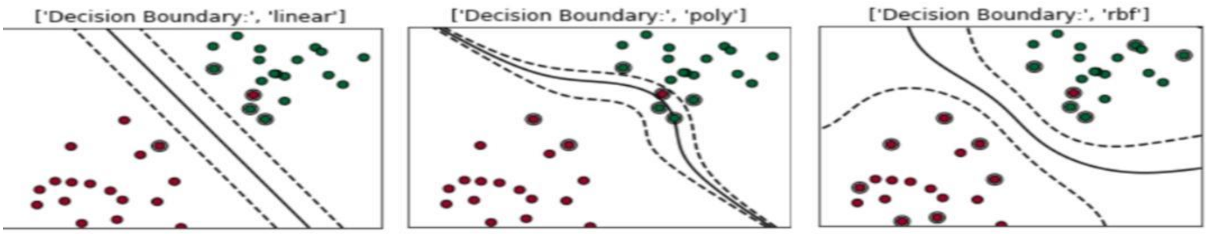


Figure 3.5: SVM kernal representation

Scifit is a software that is free to use in case of machine learning and the platform of programming is Pyhton based libraries. Support Vector Machine is a subset under Scifit. We can say that a supervised set of mathematical operations that learns itself helps us to solve classification challenges. In SVM tricks we map every input as a combination of points in n dimensional, where n is the number of attributes we are having. Every value of feature is represented by a coordinate value. After that classification is done by recognising a hyper plane in order to find a difference between every two individual class cetegory.

There are important parameters in SVM that are discussed below:

- Gamma: The parameter states the influence of each training instance and its output values leading to a output having a biased nature.
- C: The mis leading in the calculations is controlled and observed by C parameter to make smaller misclassification. Hence smaller C is always good and larger C is poor.
- Kernal: The algorithm on which SVM relies uses a combination of functions whose mathematical nature are described as a kernal. There are several types of kernal from linear, RBF(radial bias) to polynomial one.

The important job in this case is that the computer is able to learn and recognise the input data as the deive depicts the input in different manner. It observes the input set of points as an array of three dimensions. For example we have in image of size 300x300 then the devie gets it as 300x300x3 where the third dimension simply indicates the RGB channel values in it. Intensity of every pixel is determined by the range 0 to 255. The svm model needs to get train properly and the stepsto be followed are:

- Takeinput values
- Draw the model
- Train it successfully
- After training apply testing
- Final evaluation of the model

In taking the input classes, multiple different categories are taken amd they are labeled as 0, 1, and 2 respectivley. As the SVM gets the data of same dimension, the data needs to get resize to a fix dimension. In pandas there is a dataframe called df which isused to perform the required task. To talk about the model construction we simply use svm classifier in python platform using GridSearchCV parameter grid. Hance the model is generated using the abouve method after which weneed to train it significantlyfor which the input needs to split into different classes called testing and training. Training perform the mdel to learn something about the data itself while in testing we generally the model performance using a particular set of inputs. If we need to split the input we must have use train_test_split() just to perform the task from the provided library called sklearn. The model ahoukd be tested whether it is accurate or not. This is done with the help of accuracy_score() that we get from sklearn metrics. The last step in the SVM functioning is to finally evaluate the designed model over the new set of inputs and hance the final output is obtained successfully. The advantages that makes SVM better from other regression method are the formula of higher accuracy over logistic regression and decision trees. The kernal present in the SVM manges the input spaces that shows the non-linear nature. Various applications uses the SVM classifier like detecting the face, to classify emails, web pages and articles, to identify the hand writing and also to classifythe remotely sensed images into multiple land categories. The concepts behind the implementation is very simple. The svm generation just separates the input data information with the help of provided hyper plane having an extra large margin. The selection ofan optimal hyper plane is done by the SVM classifier itself. This iterative procedure helps in the reduction of error and mis classification. The maximum marginal hyperplane is supposed to suitable to classify the classes into multiple categories and these data point that have been selected is closest to the hyper plane. The term margin explain the fact that there is a gap between two different lines over the closest class point set. The calculation is performed by knowing the perpendicular distance among the line and the data points. The calculated margin is bigger in case of worst performed classification as the major goal is to differentiate the input in excellent manner in the discussion given below:

- To generate the hyper plane which helps to separate the classes in the optimal manner. The figure shows that in left side we have three best possible hyper planes out of which one has maximum classification error and one has theminimum classification error. But one with minimum classification error is seperating the classes more finely and accurately.
- Now after the selection of correct hyper plane that helps in maximum segregation from the nearest point of input set is shown in the right hand side of the below figure correctly.

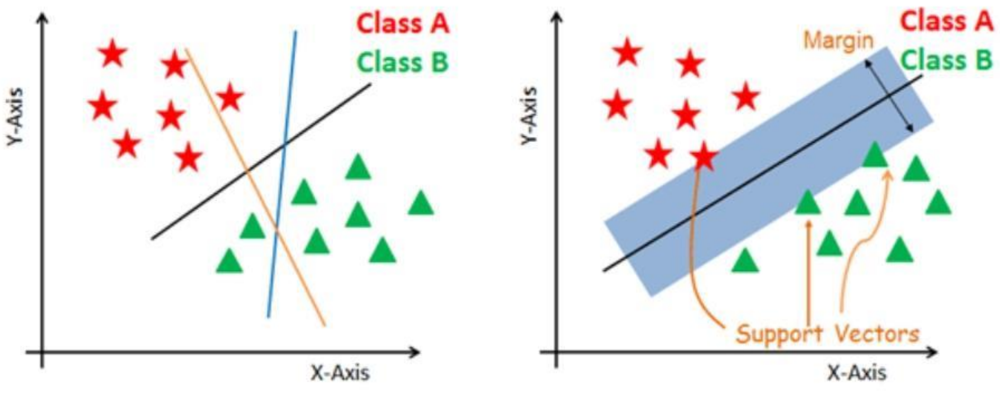


Figure 3.6: The arrangement of selecting right hyper plane

But there is a limitation that some problems can not be solved using the linear hyper plane as we have indicated in the left half of above image. So the solution to this problem is to trick a transform to convert the input space vector into higher dimensionality as we have showm in the right half of the above figure. In these cases the inputs are projected over the x axis and y aise as sum of squares and hence the segregation is quite simple and easy as it was very complex and difficult in previous case.

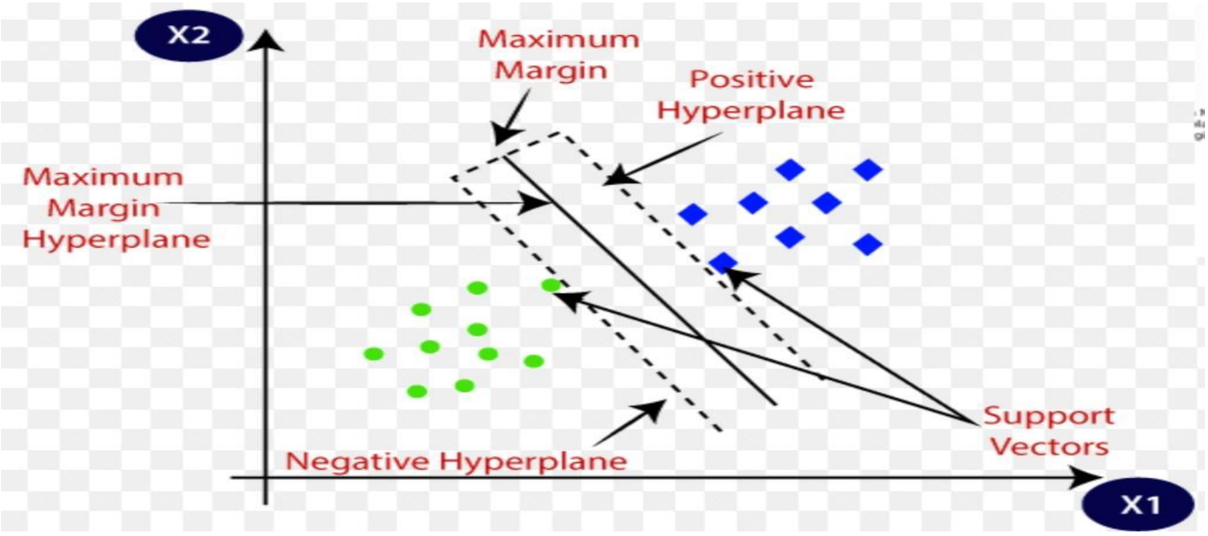


Figure 3.7: Representation of SVM classifier

As we have already mentioned the types of kernel that are used in the svm theory. Now we will see the expressions that we use to indicate each kernel. The expressions are as follows:

- Linear kernel: The use of this kernel is to perform the normal dot product between any two random observations. The product among the two different vectors is mathematically equals to sum of their multiplication for every pair.

  $$M(y,y_i) = \mathrm{sum}(y \times y_i) \qquad 3.1)$$

- Polynomial kernel: it states that a generalized form to enhance the linear kernel can be done when there is a difference among the curved or non-linear vector space. Here d is the representation of the degree and if we consider the case of linear transformation we can take the value for d as unity. We need to indicate and specify the degree manually before going into further calculations.

  $$M(y,y_i) = 1 + \mathrm{sum}(y \times y_i)\text{^}h \qquad 3.2)$$

- Radial Basis Kernel: It is the most popular kernel among all the suggested kernel for svm classifier.as it is able to map the input space vector into the infinite dimensional space. The following equation indicates the radial basis where gamma is the parameter whose values lies in between 0 and 1 and the higher values near 1 are suggested to be perfect. This perfect fitting for the training dataset can some time cause the over fitting and hence taken as a good default. The value of gamma is also manually suggested in the procedure.

$$M(y,yi) = 1 + exp(-gamma\ x\ sum(y\ x\ yi^2) \qquad 3.3)$$

3.3: Neural Networks

Nowadays neural networks are taking lead over the previous defined methods for applications like forestry, medical, military and surveillance. Neural networks have so many advantages that the researchers have found to implement more usually as compared to other techniques. The concepts behind the neural networks are inspired by the biological fact behind the human brain. The important task that we always perform in this supervised methodology is to train the model designed successfully as we know the values of given set of input data points and the corresponding output values. There is weight association between the corresponding input values that helps in the forward propagation. The training is the initial step which is totally determined by the difference among the obtained output via the model and the true value of output. Sometimes we call this value as error and based on this error value the training proceeds. The weight values are adjusted as per the error values at the time of back propagation by taking the gradient value. The purpose behind these adjustments is to make the desired output value similar to the true value of output and hence minimization in the error value. After so many changes in the weights and the output values the training is stopped using some criteria and this whole process that we discussed is simply the technique behind a supervised learning.

These devices just learn their self through the input1 an11d out1p1ut based on some certain rules. For instance, we can have a look over the recognition of image, where the task is learn some specific kinds of images like cat or dog and so we use to label them as ‘dog’ and ‘no-dog’. The network performs this recognition without having any predefined knowledge about them. It simply learns it from the dataset we are providing with their respective output values. And hence in the end they generate the real name for every image based on the training that happens for certain number of iterations. Talking about the neural networks we have several classes in it like ANN, CNN, and RNN. Every class is different in their own way. Each of these sub neural networks has different applications corresponding to their structure and the technique they use.

3.3.1: Artificial Neural Network

First we will see the ANN networking performance that are said to be a collection of so many number of smaller units which we usually call nodes. These nodes have a similar resemblance as that of the neurons in the human brain. The characterstics of the nodes is that they have to pass the biological signal that they receive immediately to the next node so that the specific action can be performed. The signal that is received at the input of a node is a real number but the processed output at the last node is the non-linear function of the respective input value’s sum. This joint is known as the edges. Nodes and the edges have a feature called weights that is the most significant parameter in the adjustments of the values. The value of the weight either increases or decreases and hence the value of the signal at the output. Every node value has a threshold value in such a way that the input values will only propagate in forward direction if and only if it crosses this threshold value limit. So in the entire process the input value is transmitted from the input node to the last node and then depending on the error value the weights are adjusted at the time back propagation up to some fix value of the iteration while training. As the iterations are over the training is terminated and the final value at the output node are obtained that are similar to the true output value with the minimum error value. In the architecture of artificial neural network there is a set of simulated neurons that shows the resemblance to the neuron present inside the human brain. There architecture is made up of so many layers and in each layer there are multiple neurons which are also termed as nodes. They are simultaneously lined side by side to each other pass the signal ahead. The very first layer is called as input layer and the last layer is called as the output layer. And among them there are so many hidden layers. Sometimes the network containing a single layer is also used in many of the applications. But we are discussing the multiple layer architecture here. If every layer is connected by means of every node in that particular layer to the adjacent node then that kind of network is known as fully connected network. The second kind of architecture defines that all the nodes in one layer is connected to only one node in the adjacent layer then this is termed as pooling. Effect of pooling reduces the number of neurons in the layer. The direct acyclic graph representation is formed only if pooling is present. The type of architecture that we are discussing is known as feed forward network. On the other hand if the architecture provides a link among same layer and the adjacent layer neuron then the networks are simply known as recurrent network. The important key features of ANN networks are given below:

3.3.1.1: Hyper parameter

It is a constant value that needs to be set before we start the learning method. We can derive this value with the help of learning. The example of any hyper parameter in the ANN networks are learning rate, batch size and the amount of layers that are hidden between the input and the output layer. The values of the hyper parameters are not independent, but they are dependent over the other values. For instance the layer dimension is totally dependent on the hyper parameter referring to the total number of layers present.

3.3.1.2: Learning

Learning allows the model to adapt the features in such a way that it is able to handle the task in much better manner by taking some sample values. It basically includes the adjustments in the weights values as discussed earlier just to increase the accuracy of the output result. And the process to increase the accuracy can be done by reducing the error value. As soon as there is no significant change in the error value or the error rate become constant we say that the training is terminated as the learning of the model is completed. But after the successful completion of learning it is not defined that the error rate is set to zero, but is has some non-zero value. Now we take a look on the value of error rate if it is high or low. The high error rate is there automatically one should understand that we need to redesign the model. And we are able to this task with the help of cost function which is examined at the time of training periodically. If the output received is continuously declining then only the learning is said to continue. The cost function is a statistical approach which is said to be approximated always. Hence the output values decided that the error received is low because of the real value, where error is the difference between the real and the expected value. Learning that we are interested in simply decrease this error value in every iteration and hence most of our frameworks are said to be optimized and can be applicable in statistical estimation.

3.3.1.3: Cost Function

The cost function is a choice which we decide by the features of the function like convexity and just the fact that it is raised from the framework. For instance, in a given model that shows the probabilistic features, in that case posterior probability for that framework can be used. And this is termed as inverse cost.

3.3.1.4 Rate of Learning Process

The dimension of the correct steps for a designed framework in order to arrange itself corresponding to the error values in every measurement is basically termed as learning rate. If the value of learning rate is high then it will reduce the time required for the training. Whereas if the learning rate is short then obviously the time for training is very high that is worse. But in this case the accuracy is significantly good. Quick prop is an optimization method which basically targets the reduction of error in short amount of time. While the other improvements help in adjusting the requirement of reliability. To avoid the situation of oscillation in the internal section of a neural network like the connection among the adjacent weights and the improvement in the convergence rate or the refinements that make use of the adaptive learning process and hence there is either increase or decrease in the process. Momentum is a physical term that helps to create a balance between gradient value and the change in the weights value that were made earlier. The change need to be adjusted such that the weighted adjustment will depend on the degree in the change occurred in previous values. If the momentum value is closer to zero it simple means that there it will emphasize the last change in the gradient.

3.3.1.5: Back propagation

This is a phenomenon that occurs in every kind of neural network including RNN and ANN. This method simply makes the use of adjustments in the weight value just to make the error less and compensate them to increase the accuracy. For the links the value of error is effectively divided. The gradient at the time of back propagation is calculated with the help of back-prop for the cost function that is linked with a defined state for a given input value. This updating in the weight value is performed due to stochastic gradient descent and many other techniques like no-prop framework in machine learning, or performing training in the absence of backtracking, the network having no weights and at last the neural network that have characterstics of non-connections.

3.3.1.6: Learning Parameters

These are classified as supervised, unsupervised and reinforcement. In supervised we use a combination of input values and the outputs. The incorrect deduction is related to the fact that each input has to produce the corresponding output value. In the supervised learning the input and the outputs are defined at the same time and given into the designed framework and so it is known as supervised learning in the presence of the corresponding output values. This technique helps the model to learn itself.

3.3.2 Convolutional Neural Networks

If we talk about the convolution neural network we can see that it is a branch or subset of deep neural networks. They basically help in the determination of various sections across the image dataset. The other name for CNN is shift invariant or it is also pronounced as space invariant artificial network as it is dependent on the architecture where the weight is shared using the convolutional kernel over different regions in the image. Kernel is also known as filtering that transverse along the image to capture the input attributes and create a translation across the responses and generates the feature maps. Counter to these maximum cases involves only equi-variant convolutional neural networks. These networks do not follow the same behaviour with translation. The application of CNN lies in the following fields at present research works:

- Recognition of video and image datasets
- Recommender devices
- Classification of images and their detection
- Segmentation
- Analyzing medical images to detect the disease in the patients
- Natural processing of languages
- Interfacing of brain-computer
- Financial analysis of time series functions

CNN are those versions which are a regularized method are containing multiple perceptrons that are fully connected to each other. As the CNN shows a fully connected feature, each node is connected to every node in the adjacent layer. But the disadvantage of fully connected feature is that it increases the number of features in the model which leads to overfitting. The ways to reduce the overfitting are also there, of which one is a regularization method called dropout. The insertion of dropout after the convolutional layers will help in the reduction of overfitting. This method will penalize the parameters while the model is in its training period. For instance decay in the weight value and trimming in the connection. CNN are very helpful as they take the advantage of the pattern in a hierarchical manner in the input and arrange the sets in order of increased complexity through the least valued patterns that can be embossed in the data values. Hence CNN shoes a drastic advantage in case of complexity and connectivity.

CNN frameworks are a inspired version from several biological methods. We are saying this because CNN resembles the connectivity behaviour of biological nodes with that of nodes present inside them. Hence this creates an optimization among the kernels so that they automatically learn the conventional algorithms that are hand-engineered. It is independent from two factors like human intervention and knowledge and hence it is a major advantage in the feature extraction. We can see from the term that CNN is based on the mathematical operation called convolution. CNN generally uses convolution rather than using simple multiplication which is used in other techniques. The convolution is the sum of simultaneous multiplication among the pixel and the kernel values of same window dimensions. The CNN consists of the following features:

- Filters or kernels that are specifically determined by their width and height (i.e. hyper parameter).
- The number of input and output nodes in the respective input and output layers. Also the input values to one of the layer must be equal to the depth of the given input.
- Many other hyper parameters are there on which the CNN model is dependent.
- These are stride value, padding and dilation process.
- Convolutional mathematical operation.
- The convolution among the image and the kernel can be of different dimension and based on that CNN is classified also.
- The different types of CNN are 1-D, 2-D and 3-D.
- In every convolution the process and feature remains the same but the only thing that differentiate them form each other is that the kernel dimension is different in each case and so the side of the input data image.
- Sometimes we can use the combination of 2-D and 3-D CNN just to have a good accuracy.

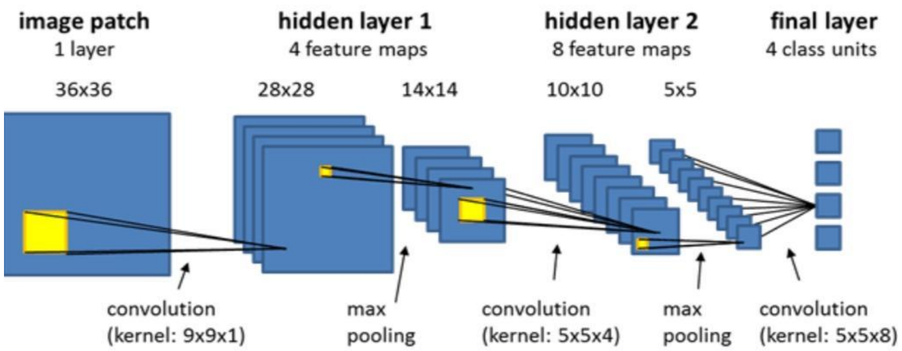


Figure 3.8: A typical representation of CNN framework in machine learning

Convolution layers can include pooling layers sometimes along with the normal layers that they always contain. This introduction of pooling generally helps in the reduction of the number of features, which in return helps to reduce the problem of overfitting in the CNN framework.

3.4: Results and Conclusion

In this chapter we have discussed that image classification is one of the major task that needs to be performed as soon as we receive the images obtained from the satellites. The main objective of our report is that we are keenly observing the remotely sensed satellite images and hence in the beginning we discussed about the concepts and types of remote sensing. The type of images that one can obtain via remote sensing is also categorizes by us chapter 2. One is multispectral images and the other one is hyper spectral images. But we are interested in the study and classification of hyper spectral images. Now we need the models designs with the help of supervised machine learning to classify the image into categories. So after the imaging techniques we discussed about the older methods like SVM, Neural Networks. Nowadays neural network have taken the place over svm due to good accuracy and less complexity. Neural networks are also divided into so many parts out of which we have discussed the introduction for 2D CNN in this chapter including artificial neural network. Not only 2D CNN helps in the classification of hyper spectral images but also 3D CNN, multihybrid CNN and a combination of 2D and 3D CNN also do the same. The only difference in these frame works is in terms of accuracy.

This chapter has only told us about the old methods that we can use to classify the images. The accuracy provided by these models is very high as compared to other methods which are always the main target. This is the fastest way of approaching any problem these days. Researchers generally work on different kind of models by introducing different layers like pooling, normalization or adding extra dropouts to reduce the overfitting. Sometimes the number of iterations also affects the training of model and the batch size too. Hyper spectral images contains a lot of information in the three dimensional system, and hence they require a model which is less complex and easy to implement so that all the information can be gathered in more abstract form. This so is due to the fact that none of the spectral band is left behind. Every spectral band represents some sort of information that needs to pick to classify every portion successfully. In the upcoming chapter we will go through these older techniques a little bit and then discus about the new models that we have designed at present to classify the images with almost 100 percent accuracy and less complexity with minimal cost of implementation and training. The models that we have designed are typically of two types that we will see ahead. One of them include maxpooling layers after the hybrid 3D and 2D CNN layers that helps in reducing the complexity while the other model also contains batch-normalization layer to overcome the time issues for training the model. Both models serves good in almost every aspect of classification. We will discuss that how these are better in terms of accuracy, complexity, overfitting, time requirement and cost issues. In every fact our models outperform the older methods like SVM, Hybrid 3D-2D CNN, 2D CNN, and 3D CNN models that have been designed in the past few decades. We will see the requirements to implement these models with the available datasets and the basic factor on which the functioning of the framework is dependent. All these observations proved that model is good to implement and classification can be done easily. The models that we are generating hence serve best in every way as compared to the older ones.

CHAPTER 4
DEEP INTENSIFIED ARCHETYPICAL CNN

4.1 Introduction

In the field of remote sensing to analyze remotely sensed images, Hyper Spectral Image Classification and attribute extraction is playing a significant role. Nowadays the advantage is taken by using Convolution Neural Networks in deep learning methods for processing data. CNN is more frequently used in the latest works for HSI classification. To perform HSI classification more accurately parameters like spectral and spatial information needs to get considered. The latest works are based on the use of 2D CNN or 3D CNN alone. Due to complexity issues use of 3D CNN is not suggested. To overcome this issue we are purposing a deep intensified archetypical 3D-2D neural net with the incorporation of maxpooling. The use of 3D CNN is solving the representation of joint spectral-spatial attribute information from a group of several spectral bands. On the other hand, 2D CNN is extracting spatial information on a more abstract level. Further, we have reduced the number of features by introducing maxpooling after 2D convolution. By reducing the number of features overfitting is avoided up to some level. It also helps in reducing the optimized complexity and computational cost to train the model. At the same time, maxpooling provides translational invariance to the feature maps generated by the convolutional layer. The performance of our model is tested over all three publicly available datasets namely Indian Pines, University of Pavia, and Salinas Scene in the classification of Hyper Spectral Images. The outcomes of our model are then compared to previous methods in terms of accuracy and computational complexity.

Keywords: 2D CNN, 3D-CNN, Hyper Spectral Image Classification, Deep Intensified Archetypical 3D-2D Neural Net, MaxPooling, spatial-spectral.

To identify and recognize different materials successfully Hyper Spectral Images play a significant role as they have bountiful spectral and spatial information. In the last few decades, Hyper Spectral Image Classification has shown foremost anticipation and efforts [1]. We can see the applications of Hyper Spectral Images in diverse fields like forestry, military, medical, natural language processing[2], security and forgery observation[3], and object detection[4]. The spectral and spatial correlation within multiple bands of images delivers convenient details. But the volume of data is very high in Hyper Spectral Images due to the presence of several bands which makes the evaluation more difficult. Due to complex spatial and spectral models, large dimensions, and fewer training specimens, we can observe some limitations and challenges in the classification of HSI data [5]. Hand-designed attribute extraction and learning-based attribute extraction are two different methods to deal with HSI classification. A collective methodology is based on the use of a local robust dictionary to reduce the negative effect of unwanted pixels and ameliorates the classification of Hyper Spectral Images[6]. A novel local covariance matrix has been purposed to identify the correlation within distinct spectral bands where these matrices helped in training the data and further classification is achieved using SVM[7]. One of the hand-crafted techniques states the combined sparse framework that is based on discontinuity preserving relaxation to smooth the outcomes locally by analyzing discontinuity in advance [8]. The latest joint spatial-spectral method is multiscale superpixels and guided filter which makes use of local information from various regions to gather spatial information [9]. It has been noticed that techniques that utilize the combination of spatial-spectral information together provide more accuracy in terms of classification. In recent years deep learning has seen many advantages in research fields like speech recognition[10], computer vision, image processing, and classification[11]. A common survey can explain that deep learning models are the best way to classify Hyper Spectral Images rather than going for old hand-designed methods of attribute extraction. A deep fusion framework had been suggested that combines the outcomes of various layers to increase the accuracy[12]. An off-shelf CNN model is purposed as an effective way to gain deep spatial attributes in the absence of extra retraining[13]. Instead of using SVM to train data, we can see several deep learning-based methods that consume very low power. Many of these methods are deep feature extraction of HSI data using 3D CNN[14],HSI data classification using 2D CNN[15],spectral-spatial residual framework[16],multi-scale deep 3D CNN[17],HybridSN using 3D-2D CNN[18].

All these methods have some amount of shortcomings. Individually 2D and 3D CNN do not show prominent results in classification. 2D CNN alone is unable to pick strong attributes from spectral bands of HSI data due to its volumetric nature. On the other hand, 3D CNN alone shows a very complex nature and performs very poorly in the case of spectral redundancy. SSRN is based on a residual block that takes the advantage of identity mapping to connect every 3D layer, but a significant accuracy can be achieved further[16]. A hybrid framework can solve all these problems to gain both spatial and spectral information with less complexity and maximum accuracy. We have integrated this network by the incorporation of maxpooling layer which reduced the number of parameters. It reduced the computation cost to train the model. At the same time, maxpooling is reducing the dimension of the feature maps generated from the convolutional layers that are decreasing the chances of optimized complexity and overfitting. Our purposed model is providing an abstract-level feature representation with significant accuracy, much lesser complexity, and cost. This chapter is suggesting a deep intensified archetypical 3D-2D neural net with the incorporation of maxpooling in Section 2, experiments and results in Section 3, and conclusion in Section 4.

4.2. Proposed Method

We are applying principle component analysis across the spectral bands to reduce the data redundancy. Thereafter an extensive archetypical approach is introduced using the conglomerate 3D-2D CNN model and maxpooling. The introduction of maxpooling dispenses downsampling in the features given as input. The overall technique helps to achieve significant accuracy in the successful classification of Hyper Spectral Images.

4.2.1 Elimination of Data Redundancy

The hyperspectral image data is a three-dimensional cube represented as I. Let us assume the dimensions for this spatial-spectral image cube as AxBxC. Considering I as the original input, A as the width, B as the height, C as the number of spectral bands in the image data cube. Dissimilar land categories are characterized by each spectral band in the image cube that forms one hot label vector K= (k_1, k_2 ………k_z), where z is the number of land categories. A strong local correlation within the bands is examined due to the quasi-continuous nature of the spectrum. This band correlation causes the data to be redundant. We proposed a method to tackle this problem by imposing principal component analysis over the spectral bands that reduces them from C to D. At the same time we are preserving the spatial information so that recognition of image can be done efficiently. Talking about spatial and spectral facets we can observe the contiguous pixels in hyperspectral images manifesting some homogenous nature. So to resolve this issue, image slicing is performed after PCA with a window size of PxP to maintain the same labels at the center and surrounding pixels. Finally we have a set of overlapping 3D square patches as V= $\{Q_{x,z}\}$ where x= 1, 2 …A and z=1, 2…B. Here Q represents the square patch with (x, z) as pixel location. Also for the CNN model, there is a need for profuse training samples so that substantial weights can be adjusted in the network. These square cubes are used together to obtain appropriate weights in the neural net which is beneficial for further processing tasks.

4.2.2 Architecture of CNN Model

After the successful completion of dimensionality reduction and image slicing, a fused model is implemented over the 3D patches. This neural net incorporates 3D convolution layers to capture the spectral information, 2D convolution layers to encapsulate spatial information, and maxpooling layers to extract cardinal features. Convolution of a 3D kernel [19]with a 3D square patch is utilized in 3D convolutional layers which engender their feature maps on top of various proximate bands provided by the input layer. 3D convolution expresses the linear combination of input cubes and kernels. The output contains modified attributes that are utilized by the relu function to instigate non-linear behavior. We can determine[19] the activation value for the ith layer in jth attribute at spatial position (a, b, c) as:

$$T_{i,j}^{a,b,c} = \Psi\left(c_{i,j} + \sum_{\tau=1}^{d_{k-1}} \sum_{\mu=-\omega}^{\omega} \sum_{p=-\Omega}^{\Omega} \sum_{\pi=-\delta}^{\delta} w_{i,j,\tau}^{\pi,p,\mu} \text{ x } v_{i-1,\tau}^{a+\pi,b+p,c+\mu}\right) \tag{4.1}$$

Here, Ψ is activation function, $c_{i,j}$ is bias parameter of ith layer for jth attribute, d_{k-1} is the number of attribute in (k-1) layer, $w_{i,j,\tau}$ is weight parameter. After 3D convolution data is reshaped to perform 2D convolution that will enclose complete spatial particulars through striding kernels on top of two-dimensional inputs.

4.2.3 Pooling Layer

Immediately after convolution one maxpooling layer is used to extract sharp and smooth attributes from the data that are the best low- level representation. It also helps in turning down the variance and computational cost as it diminishes the number of variables to ascertain and also ensuring local translational invariance. CNN layers recognize the accurate location of attributes in the data which is a limitation for small movements in the location of features described for the given input. So introducing maxpooling after convolution solves this issue by downsampling as they generate pooled attributes at the same location. At the same time problem of overfitting can be reduced due to dimensionality reduction that further lays down simple optimization.

4.2.4 Model Description

In our purposed model 3D convolution will strongly capture the spectral information while 2D convolution differentiates the spatial information within several spectral bands without losing any spectral attribute. The introduction of maxpooling will reduce the computational cost by decreasing the number of parameters and it highlights the prominent attribute maps. At the same time, our model becomes less complex and overfitting can be avoided because the dimensions of feature maps are reduced due to maxpooling. Thus, significant accuracy is obtained after the implementation of the suggested framework. By taking this precedence of instinctive feature learning we have purposed a deep intensified archetypical 3D-2D neural net with the incorporation of maxpooling. The flow diagram of our model is shown in Figure 1. It consists of two 3D convolution layers (2), two 2D convolution layers (2), one maxpooling layer (1) followed by a combination of one 2D convolution and maxpooling layer. Thereafter it has three fully connected layers and Softmax in the end for class labels. We can see a comprehensive outline of the designed model regarding types of layers, the number of parameters involved and aspects of output maps in Table 1, Table 2, and Table 3 for Indian Pines, University of Pavia and Salinas Scene datasets respectively. It is clearly visible that the first 2D convolutional layer contains the maximum number of parameters. There are 16 distinct classes present in the Indian Pines and Salinas Scene dataset and 9 distinct classes in the University of Pavia dataset that are illustrated by the last dense layer in Table 1, Table 2, and Table 3 as it contains the same number of nodes respectively. The total number of classes that are present in a dataset is determined by the total number of parameters defined for the neural net. The training is done by accounting Adam optimizer with Softmax loss. At the same time backpropagation methodology randomly initializes the weights. The final training of the neural net has been implemented using 50 epochs with a minimum batch size of 256 in the absence of batch normalization and data augmentation. But translation invariance is considered through maxpooling. The total numbers of weighted parameters that are trained by means of our model are 261,097 for the University of Pavia Dataset, 354,160 for the Indian Pines dataset, and

274,224 for the Salinas Scene dataset. Further to make this framework more vigorous to oppose overfitting, a dropout of 40% is taken using fully connected layers.

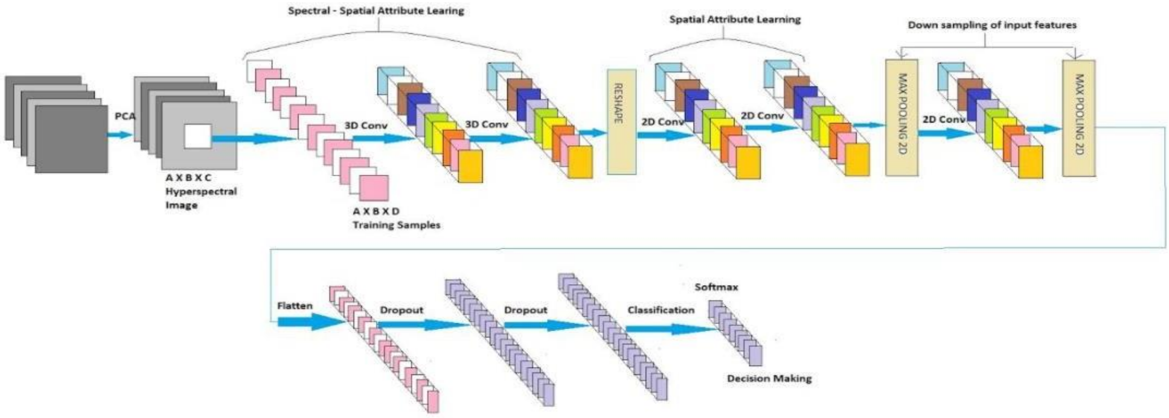


Figure 4.1.Deep Intensified Archetypical 3D-2D Neural Net followed by maxpooling for Hyper Spectral Image Classification.

TABLE 4.1: Outlining of the Proposed Deep Intensified Archetypical 3D-2D Neural Net with Incorporation of MaxPooling showing every layer involved for the Indian Pines dataset having a window size 19x19

Layers	Output Dimensions	Parameters
Input1	(19,19,20,1)	0
Conv3d	(17,17,14,8)	512
Conv3d1	(15,15,12,32)	6944
Reshape	(13,13,384)	0
Conv2d	(13,13,64)	221,248
Conv2d1	(11,11,64)	36928
Maxpooling2d	(5,5,64)	0
Conv2d2	(3,3,64)	36928
Maxpooling2d1	(1,1,64)	0
Flatten	(64)	0
Dense	(256)	16640
Dropout	(256)	0
Dense1	(128)	32896
Dropout1	(128)	0
Dense2	(16)	2064
Total parameters		354,160

TABLE 4.2: Outlining of the Proposed Deep Intensified Archetypical 3D-2D Neural Net with Incorporation of MaxPooling showing every layer involved for the University of Pavia dataset having a window size 19x19

Layers	Output Dimensions	Parameters
Input1	(19,19,15,1)	0
Conv3d	(17,17,9,8)	512
Conv3d1	(15,15,7,32)	6944
Reshape	(15,15,224)	0
Conv2d	(13,13,64)	129,088
Conv2d1	(11,11,64)	36928
Maxpooling2d	(5,5,64)	0
Conv2d2	(3,3,64)	36928
Maxpooling2d1	(1,1,64)	0
Flatten	(64)	0
Dense	(256)	16640
Dropout	(256)	0
Dense1	(128)	32896
Dropout1	(128)	0
Dense2	(9)	1161
Total parameters		261,097

TABLE 4.3: Outlining of the Proposed Deep Intensified Archetypical 3D-2D Neural Net with Incorporation of MaxPooling showing every layer involved for the Salinas Scene dataset having a window size 19x19

Layers	Output Dimensions	Parameters
Input1	(19,19,15,1)	0
Conv3d	(17,17,9,8)	512
Conv3d1	(15,15,7,32)	6944
Reshape	(15,15,224)	0
Conv2d	(13,13,64)	129,088
Conv2d1	(11,11,64)	36928
Maxpooling2d	(5,5,64)	0
Conv2d2	(3,3,64)	36928

Maxpooling2d1	(1,1,64)	0
Flatten	(64)	0
Dense	(256)	16640
Dropout	(256)	0
Dense1	(128)	32896
Dropout1	(128)	0
Dense2	(9)	2064
Total parameters		262,000

4.3. Experiment with Result

4.3.1 Elucidation of Dataset

We have used three recognized Hyper Spectral Image datasets that are publicly accessible for substantiating the execution. These datasets refer to Indian Pines (IP), the University of Pavia (UP), and the Salinas Scene (SA). The Indian Pines dataset has a geometric resolution of 20 meters and it incorporates 145x145 spatial dimensions. AVIRIS sensor captures the Indian Pines dataset located in the northwestern part of Indiana. The 224 spectral bands covering the wavelength from 400 to 2500nm out of which 24 bands are rejected as they are water absorption bands. The ground truth has 16 distinct land cover vegetation classes. The University of Pavia dataset is obtained by ROSIS sensor with a geometric resolution of 1.3 meters and it contains 610x340 spatial dimensions. The spectral information is represented by 103 spectral bands covering the wavelength from 430 to 860nm. The ground truth comprises 9 distinct urban land cover classes. In the Salinas Scene dataset, the spatial dimensions are 512 x217 and the numbers of spectral bands are 224 that covers the wavelength range from 360-2500nm. We have performed the experiments with an optimal learning rate of 0.001 for the best possible classification results. A fair comparison is made by taking the same spatial dimension of 3D cubes for the Indian Pines dataset as 19x19 x20, 19 x 19 x 15 for the University of Pavia dataset, and 19 x 19 x 15 for the Salinas Scene dataset.

4.3.2 Training and Testing Parameters

We have randomly selected 30% of the labeled data for training and 70% for testing from all three datasets to successfully estimate our model. The training is done by accounting Adam optimizer with Softmax loss. At the same time, the backpropagation methodology randomly initializes the weights. The final training of the neural net has been implemented using 50 epochs with a minimum batch size of 256 in the absence of batch normalization and data augmentation. But translational scale invariance is considered through maxpooling. The total numbers of weighted parameters that are trained using our model are 317,232 for the Indian Pines dataset, 261,097 for the University of Pavia Dataset, and 274,224 for the Salinas Scene dataset. Further to make this framework more vigorous to oppose overfitting, a dropout of 40%

is taken employing fully connected layers. With all these defined parameters for our purposed model training is done for all three datasets and our main target is to apply a window size of 19x19 with a combination of 3D-2D-MaxPooling followed by dropout. As soon as training is completed for all three datasets, we saved the model, and the test data is estimated depending on the saved model.

4.3.3 Results and Comparison

To evaluate our model performance in the classification of Hyper Spectral Images we focused on three factors namely, Overall accuracy (OA), Average accuracy (AA), and Kappa coefficient (KC). Overall accuracy states the number of specimens that are classified correctly from the total number of test specimens. It indicates the proportions that are mapped correctly. Average accuracy indicates the average of individual accuracy per class. It provides the average class-wise classification and Kappa is a statistical metric that helps in comparing the classification maps and ground-truth maps. In other words, it provides an estimation of a single classifier and the classifiers among themselves. We have compared the outputs of our suggested Deep Intensified Archetypical model with previous methods that are based on supervised learning. These are SVM[20], 2D CNN[15], 3D CNN[21], CNN-PPF[22] and Multi-Scale 3D Deep Convolution Neural Network (M3D-CNN)[17]. Based on publicly available codes for these techniques, we have made our comparison in terms of classification accuracy. In Table 4.4 we can observe the outputs of distinct models in terms of Overall accuracy, Average accuracy, and Kappa coefficient. It can be seen that the model presented by us is outperforming previous models for all three datasets. From Table 4.4 we can observe that 3D CNN is showing poor results for the Salinas Scene dataset as compared to 2D CNN. With our best observation, we can say that this can be due to the similarity between two classes (Grapes untrained and Vinyard untrained) among several bands in the Salinas Scene dataset. In our model, we have eliminated this issue of spectral redundancy by using PCA over spectral bands. Also, 2D CNN and 3D CNN cannot individually represent the distinct attributes, so we need a combination of 3D CNN and 2D CNN which is utilized by an archetypical 3D-2D neural net. In our model, we are incorporating maxpooling that has reduced computation cost by reducing the number of parameters. Only prominent features are highlighted after maxpooling. At the same time, our model supports translation invariance because of maxpooling. Optimization complexity is reduced due to a reduction in the dimensionality of the features map generated through convolutional layers. Thus the introduction of maxpooling makes our model better in terms of complexity, overfitting issue, and computational cost to train the model. Also, the results of our framework are way better than M3DCNN[17] and SVM. Different results are shown in Table 5 regarding different spatial dimensions. We noticed that a window size of 19x19 is suitable for our framework in the case of all three datasets. We can say that our suggested model is outperforming all previous methods in terms of accuracy and computational complexity desires. Results can be viewed in Figure 2 explaining the ground truth and the predicted classification maps for Indian Pines, University of Pavia, and Salinas Scene datasets.

TABLE 4.4: Representation of comparison among various methods over IP, SA and UP datasets in terms of accuracies for Hyper Spectral Image Classification

Methods	IP			UP			SA		
	OA	KA	AA	OA	KA	AA	OA	KA	AA
SVM	83.30	83.10	79.03	94.34	92.50	92.98	92.95	92.11	94.60
2DCNN	89.48	87.96	86.14	97.86	97.16	96.55	97.38	97.08	98.84
3DCNN	91.10	89.98	91.58	96.53	95.51	97.57	93.96	93.32	97.01
M3DCNN	95.32	94.70	96.41	95.76	94.50	95.08	94.79	94.20	96.25
SSRN	99.19	99.07	98.93	99.90	99.80	99.84	99.98	99.97	99.97
Our model	99.84	99.82	99.46	99.91	99.82	99.82	100	100	100

4.4. Conclusion

In this work, we have introduced a deep intensified archetypical 3D-2D neural net with the incorporation of maxpooling to classify Hyper Spectral Images. This purposed model joins spatial-spectral information and spectral information using 3D CNN and 2D CNN respectively. In addition to this, the computational cost and complexity are also reduced by incorporating maxpooling after 2D CNN. This overall framework provides a significant accuracy with a chance to avoid overfitting to some extent that confirms its superiority over previous methods. This model is more efficient and accurate as compared to 3D CNN and 2D CNN models, also less costly and less complex as compared to Hybrid 3D-2D CNN.

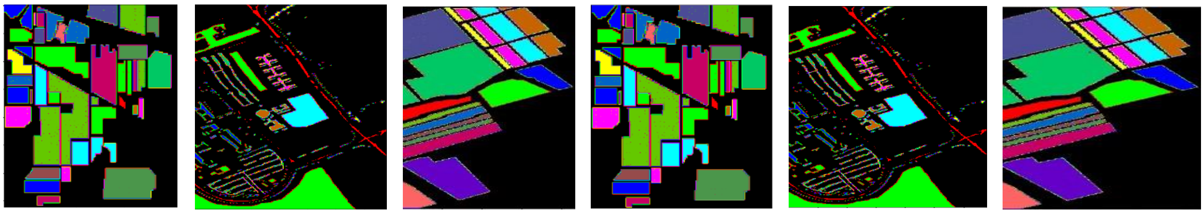


Figure 4.2: a) to c) Ground truth and d) to f) Predicted feature maps for IP, UP and SA datasets respectively.

TABLE 4.5: Comparison of overall accuracies with different spatial window dimensions over IP, UP, and SA datasets.

Window	IP(%)	UP(%)	SA(%)
17 x 17	99.69	99.95	99.98
19x19	99.84	99.91	100

CHAPTER 5

A DENSE CNN NETWORK FOR HSI CLASSIFICATION

5.1 Introduction

Nowadays, Hyper Spectral Image Classification finds its application in the recognition of assorted remotely sensed images. Looking on the previous recent works we can observe the HSI classification utilizing 2D and 3D CNN individually. We are suggesting a new deep learning based 3D-2D convolutional approach followed by Batch Normalization and maxpooling to solve the overfitting and complexity issues in previous works. The overall technique is not only solving the representation of joint spectral-spatial attribute information but also extracting the spatial information on higher abstract level. We have tested the performance of our model over three publicly available datasets namely Indian Pines, University of Pavia, and Salinas Scene to demonstrate the comparison with previous works.

Keywords: Batch-Normalization, 2DCNN, 3DCNN, Hyper Spectral Image Classification.

Hyperspectral Images consists of several spectral bands covering a wide range of electromagnetic spectrum from visible to short waves. These images have plentiful spatial and spectral information which plays a significant role in object detection and classification. We can find the applications of Hyper Spectral Images in various fields like medical, forestry, military, object detection[23], natural language processing[24], security and forgery information[25] and noise elimination. Two important aspects are taken into consideration to procure selective attributes in classification of Hyper Spectral Images. First is the bountiful spectral information that helps in precise recognition of various ground materials. Second is maximum correlation among spatial attributes to provide accurate mapping. Due to complex three dimensional spatial and spectral cube, we can observe some challenges in the HSI processing and classification. To solve these restrictions two major techniques are suggested namely, attribute engineering and learning based attribute extraction. Attribute engineering provides selection and extraction of several distinguishing characterstics by reducing the higher dimensions in HSI pixels[26]. Further, the training of general purpose classifiers is performed utilizing distinguishing attributes from attribute engineering. As compared to attributes extraction, attributes selection helps to detect significant characterstics from HSI data without any transformation. For training the classifier, adopted manifold ranking which is an unsupervised attribute selection technique, selects the most significant band[27]. Also, Markov random field has implemented a multi-task joint representation method[28]. Both these semi-supervised techniques utilize the spectral bands from every pixel for attributes selection. Spatialized input and post-processing are two different approaches to include spatial data in HSI classification. In this chapter, we have discussed literature survey in Section 2, we have proposed a deep archetypical 3D-2D framework followed by maxpooling and batch normalization layers in Section 3, experiments and results in Section 4, and conclusion in Section 5.

Several works stated that enhancing the spatial attributes in input data provides an improved classification performance[29]. For instance, a SVM classifier learns the categories in hyperspectral data by utilizing region based kernel that extracts both spatial and spectral attributes[30]. On the other hand, an initial knowledge of smoothness is used in post-processing techniques which states that adjacent pixels with indistinguishable features have the maximum probability to be associated with same land cover vegetation. For example, the classification outputs of SVM kernel can be enhanced by imposing a probabilistic graphical method in the post-processing[31]. Although several experiments have been carried out using combination of attribute extractor and trainable classifier, these methods face some common challenges. One of them is the restricted capacity to represent for a complete use of profuse spectral and spatial attribute. This drawback in attribute engineering has shifted the focus of researchers towards supervised deep learning framework where the principle goal is to allow the framework to get train and choose the significant attributes itself with some constraints. Therefore, a sufficient training of deep learning frameworks helps in learning the attributes of Hyper Spectral Images that is far better than SVM[32]. Several works have implemented the convolutional neural networks to provide the state of art outcomes for classification purpose using spatialized data[33].

For instance, in [34]we can see the CNN model extracting the spatial and spectral attributes that are consolidated together for classification. The model has learned through the balanced local discriminator. However the drawback is the loss of spatial-spectral information from spatial attribute extractor. To solve these issues 3D CNN models were proposed to gain deep spatial-spectral characterstics out of HSI data to deliver excellent results [33]. Further, in [35] the 3D CNN is implemented utilizing input cubes that have smaller dimensions. This approach creates a mapping that helps in generating raw Hyper Spectral Images with a constraint of decreased accuracy with deeper network. This drawback is solved by another method known as supervised spectral-spatial residual network[36] in which successive elements extracts the features of input data and identity mapping is utilized to join every 3D layer. The spatial-spectral attributes are obtained using spectral-spatial residual learning blocks. Therefore, due to shortcut links between every convolutional layer, SSRN is able to learn significant spatial and spectral characterstics from the raw input. However, better accuracy can be achieved further and also the framework is unable to differentiate between spectral and spatial characterstics.

Many other works that have been proposed to extract deep attributes of HSI data are 2D CNN[32], multi-scale deep 3D CNN[37], HybridSN using 3D-2D CNN[38]. Individually 3D and 2D CNN don't provide significant classification accuracy as 2D CNN cannot grab strong characterstics from the spectral bands present in input data and 3D CNN lags in case of spectral redundancy. Also, all these methods are much complex in nature which is an issue regarding the model implementation. Hence in this paper we are suggesting a Hybrid 3D-2D CNN framework with incorporation of maxpooling and batch normalization layers to overcome the complexity and overfitting issues with significant classification accuracies. The arrangement of 3D and 2D layer in the model will help in the extraction of both spatial and spectral attribute maps up to full extent. Also, our model is best in terms of computational cost and learning rates defined for training purpose.

5.2 Proposed Method

The first step of our work is reduction of spectral redundancy using unsupervised technique called PCA followed by Image Slicing. Further, the hybrid 3D-2D CNN model is implemented which incorporates both maxpooling and batch normalization layers.

5.2.1 Suppression of redundant information

The original raw input data is a three dimensional HSI cube that can be represented as I with dimensions as XxYxZ. Here, X is the width, Y is the height and Z is the number of spectral bands present in the input cube. Every spectral band in the original data helps in characterizing the dissimilar land cover categories which is encoded by one hot label vector V= $(v_1, v_2 \ldots\ldots\ldots v_m)$, where m is the number of land cover categories. We are utilizing principal component analysis to tackle the data redundancy issue that arises due to quasi-continuous behavior of the spectrum. This behavior causes a strong correlation within the bands which needs to be solved. The application of principal component analysis helps in reducing the number of spectral bands from Z to W. To observe the image successfully the spatial information has not been distorted. The XxYxZ image is now enhanced to XxYxW. After pca the next step is to deal with spatial and spectral aspects. We can determine a homogenous trend as demonstrated by adjoining pixels in the hyper spectral images. To handle this situation we are using slicing process after pca incorporating a window of dimension NxN. At last we have produced a set of coinciding three dimensional square speckles as S = $\{K_{cv}\}$. Here range of c is from 1 to X and v ranges from 1 to Y. K is the square speckle at pixel location {c,v}.

5.2.2: Convolutional Network Framework Architecture

After image slicing there are several number of square speckles over which we have implemented the amalgamated model. The truth labels of three dimensional square speckles are determined by the label of centered pixel. Our model consists of 3D convolutional layer that helps in observing the spectral information from the images, 2D convolutional layers to capture spatial information on an abstract level, batch normalization to stabilize the learning process, followed by maxpooling to pull out primary attributes. The three dimensional convolutional layers performs the convolution between 3D square patches and three dimensional kernels. This convolution brings out many feature maps over the top of proximate bands that are present in the input layer[39]. The relu function instigates the non-saturating and non-linear nature in the modified features after 3D convolution. This behavior of relu allows the easy backpropagation of errors which activates multiple layers of neurons. The activation value at spatial position (u,v,w) for the k_{th} attribute map of m_{th} layer is represented by $d_{m,k}^{u,v,w}$ as given below[39]:

$$d_{m,k}^{u,v,w} = \beta \left(l_{m,k} + \sum_{\mu=1}^{h_{j-1}} \sum_{\Psi=-\pi}^{\pi} \sum_{p=-y}^{y} \sum_{\omega=-\delta}^{\delta} a_{m,k\mu}^{\omega,p,\Psi} \, \mathrm{x} d_{m-1}^{u+\omega,v+p,w+\Psi}\right)$$

5.1)

Here d is the activation function, $l_{m,k}$ is the bias parameter for m_{th} layer, h_{j-1} is the number of feature map in (j-1) layer and $a_{m,k,\mu}$ is the weight parameter. After this 3D convolution we have utilized 2D convolution which will perform the same mathematical convolution among input and the 2D kernels. Convolution is basically the sum of dot product between the given data and available kernels. The kernel is strided over the provided information covering the entire spatial dimensions.

5.2.3: Batch Normalization and Pooling Effect

The next step after convolution is taken by applying batch normalization and maxpooling. The effect of batch normalization can be seen as the time taken during the training process. Further, the problem that is more often in training the neural networks is covariant shift. The distribution in every input layer changes when training happens. This is because of change in the parameters of previous layers which slows down the training process. We have reduced this shift by utilizing batch normalization after the convolution. This layer will smoothen the loss function which optimizes the model parameters and so the training speed is improved significantly. After, this we have used maxpooling layer to obtain sharp and smooth attributes. The use of maxpooling helps in the reduction of computational cost to train the model as it reduces the number of features and confirms the local translational invariance. The layers in convolutional model can easily determine the features present in the input which is a restriction in the minor displacement for them. Hence, the incorporation of maxpooling will tackle this issue during downsampling of generated pooled attributes at same position. Maxpooling also helps in reducing the complexity and overfitting issue due to dimensionality reduction. Effect of batch normalization and maxpooling will simplify the optimization.

5.2.4: Model Architecture

In figure 1 there is full description of our implemented model. It consists of three 3D convolutional layer(2), two 2D convolutional layers(2), followed by one batch normalization layer and two maxpooling layers. We have applied the 2D convolution before flatten so that it can differentiate between spatial attributes within spectral bands without loss in spectral contents. After this we have used three fully connected layer and Softmax activation function to label the classes accordingly. The extensive outline of the implemented model in terms of dimensions of output, types of layers, and the total number of parameters is descripted in Table 1 for Indian Pines dataset. We can observe that the first dense layer contains the highest number of parameters. The last dense layer in Table 1 helps in depicting the number of distinct classes for every dataset which is 16 for Indian Pines as they contain same number of nodes. The total number of parameters describing the model helps in determining the total number of classes in every dataset.

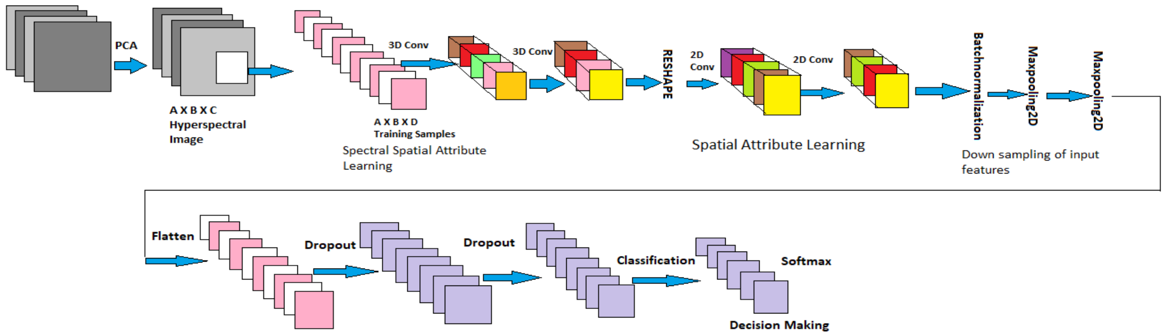


Figure5.1: The overall model description based on Hybrid 3D-2D CNN followed by Batch normalization and maxpooling.

TABLE 5.1: Structure of suggested model incorporating maxpooling and batch normalization for Indian Pines dataset having a window of 21x21.

Layer	Output Dimension	Parameter
Input	(21,21,20,1)	0
Conv3d	(19,1914,8)	512
Conv3d1	(17,17,10,16)	5776
Reshape	(17,17,64)	0
Conv2d	(15,15,160)	92228
Conv2d1	(13,13,64)	36928
Batchnormalization	(13,13,64)	256
Maxpooling2d	(6,6,64)	0
Maxpooling2d1	(3,3,64)	0
Flatter	(576)	0
Dense	(256)	147712
Dropout	(256)	0
Dense1	(128)	32896
Dropout1	(128)	0
Dense2	(16)	2064
Total parameters		318,368

5.3: Experimental Results

Here we have discussed about the data set on which we have performed the experiments. Then we have discussed regarding the parameters which we have considered for training and testing of our model followed by the results and their comparison with previous methods.

5.3.1: Dataset Elucidation

To perform our experiments we have used three publicly available datasets which are Indian Pines (IP), Salinas Scene(SA) and University of Pavia(UP). The dimension of Indian Pines dataset is around 145x145 with a geometric resolution of about 20 meters. The dataset is located in the north western regions of Indiana which is traced by AVIRIS sensor. Number of spectral bands are 224 that cover the wavelength range of 400nm to 2500nm. Out of these spectral bands 24 water absorption bands are rejected as we are not interested in that information. The original ground truth possesses 16 different vegetation categories. The spatial dimension of University of Pavia dataset is 610x340 which is traced by ROSIS sensor. The geometric resolution is 1.3 meters and 103 distinct bands represent the spectral features. The wavelength range of these bands is between 430 to 860 nm. There are 9 different land urban categories present in the ground truth image. At last the Salinas Scene dataset has a spatial dimension of 512x217. The numbers of spectral bands are 224 with a wavelength range of 360nm to 2500nm. We have made a fair comparison by extracting same spatial dimensions in 3D square speckles of the input data in Indian Pines as 21x21x20, 21x21x15 in University of Pavia and Salinas Scene datasets respectively. Every experimental observation is made by taking an optimal learning rate of 0.001 to achieve significant classification results.

5.3.2: Testing and Training Description

For each dataset we have done a random selection of testing and training dataset as 60% and 40% respectively to get successful evaluation of our designed model. We have utilized Adam optimizer for training the dataset as it provides faster results and on the other hand, the back propagation process helps in the initialization of weights. The number of iterations for which the training is done is 50 with a selected batch size of 256 in the absence of data augmentation technique. This training process helps in generating the significant results later on for the model where the window dimension is near around 21x21. In the end drop out is also provided which is a regularization method that avoids the overfitting. After the training the model needs to get saved so that testing can be performed over it. Further chances are done on the tested model.

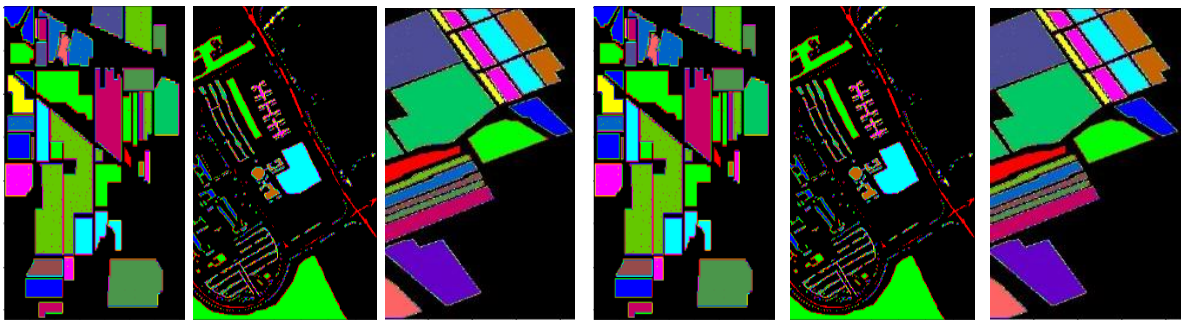
5.4: Outcomes and their Comparison

We have selected three factors to make the comparison between our model and the previous works in the classification of Hyper Spectral Images based on supervised learning. These are overall accuracy (OA), average accuracy (AA) and kappa coefficient (KC). The first factor that is overall accuracy simply indicates the total number of classes that are classified correctly out of the total number of testes specimens. The average of individual accuracy per class can be calculated by the second factor known as average accuracy. The third factor which is kappa coefficient is simply a statistical parameter that helps in the comparison between ground truth maps and classification maps. We have compared our work with previous methods such as SVM[40],2D CNN[32], 3D CNN[41],Multi Scale 3D deep convolutional neural network[37] and CNN PPF[42]. By analyzing the publicly codes we have observed the accuracies of these methods and compared them with our model that can be seen in Table 2 Our model is serving best in terms of all three factors namely overall accuracy, average accuracy and kappa coefficient. The window size that proved out to be suitable for all three datasets is 21x21 and this model resolves the issues like complexity, overfitting, computational cost of training and higher training time. We can see the results for ground truth and predicted images in Figure 2 for all three datasets. We have noticed the model's behavior and announced that it is better over all the previous techniques in terms of accuracy, complexity and time requirement.

TABLE 5.2: Representation of comparison among various methods over IP, SA and UP datasets in terms of accuracies for Hyper Spectral Image Classification

Methods	IP			UP			SA		
	OA	KA	AA	OA	KA	AA	OA	KA	AA
SVM	83.30	83.10	79.03	94.34	92.50	92.98	92.95	92.11	94.60
2DCNN	89.48	87.96	86.14	97.86	97.16	96.55	97.38	97.08	98.84
3DCNN	91.10	89.98	91.58	96.53	95.51	97.57	93.96	93.32	97.01
M3DCNN	95.32	94.70	96.41	95.76	94.50	95.08	94.79	94.20	96.25
SSRN	99.19	99.07	98.93	99.90	99.80	99.84	99.98	99.97	99.97
Our model	99.90	99.90	99.70	99.95	99.86	99.91	100	100	100

Figure 5.2: Representation of Ground Truth in a) to c) and Predicted Results in d) to f)



In this manner we have seen in chapter 4 and chapter 5 suggest the two different models to perform the hyper spectral image classification successfully.

CHAPTER 6

CONCLUSIONS

6.1: Conclusions

This work is about the implementation of classification for the image data sets that we receive through hyperspectral sensors. In this work we have presented two different models that are based on the supervised learning technique to classify the image data. Based on the past works that we have seen like SVM and CNN we have learnt that we need to suggest a model that can combine the spatial data along with the spectral data in more significant manner that is much more accurate. As the past models were only either giving the spatial information or spectral information alone, we saw that there is a need of a combined information. Hence we build a model that is able to combine the joint spectral data along with the spatial dimension with less loss in the data attributes. In both the models we have used PCA which is principle component analysis that is used to reduce the redundancy in the data input. Then instead of using two dimensional CNN or three dimensional CNN alone we have used the combination of both 3D and 2D CNN to perform the classification more accurately. This has an advantage over the past methods like 3D CNN and 2D CNN as they only gives spatial or spectral information at a time with complexity included. Now we have seen that only increasing the accuracy is not sufficient to perform the classification. There are some extra parameter that needs to get consider while performing the task.

Hence we have included the importance of time, cost, complexity and overfitting at the time of training. This change we have seen through the incorporation of an extra layer known as maxpooling in chapter 4 for model 1. This change has significantly helped in increasing the accuracy and at the same time it helped in reducing the complexity just by reducing the number of parameters referring to the attributes. This reduction in attributes has simply decreased the dimension of the data which make it less complex as compared to past models with reduced cost for training the model.

After increasing the accuracy with less complexity and we have observed that the time required for training the model can also be reduced and the accuracy can be improved further if we include another layer called batch normalization. Hence in chapter 5 we have suggested model 2 which perform the same task with more accuracy. Also the time required by this model to perform classification is very much less as compared to all the previous models. So the implementation of model 2 is better than model 1 in terms of the time requirement. Both these models are the best models that are based on the supervised technique to classify the target into multiple categories. We have used the publicly available datasets like Indian Pines, University of Pavia, and Salinas Scene that represents the land cover vegetation having 16, 9 different categories respectively.

REFERENCES

[1] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. M. Nasrabadi, and J. Chanussot, “Hyperspectral remote sensing data analysis and future challenges,” *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 2. 2013, doi: 10.1109/MGRS.2013.2244672.

[2] A. Bordes, X. Glorot, J. Weston, and Y. Bengio, “Joint learning of words and meaning representations for open-text semantic parsing,” *J. Mach. Learn. Res.*, vol. 22, pp. 127–135, 2012.

[3] C. S. Silva, M. F. Pimentel, R. S. Honorato, C. Pasquini, J. M. P. Montalbán, and A. Ferrer, “Near infrared hyperspectral imaging for forensic analysis of document forgery,” *Analyst*, vol. 139, no. 20, pp. 5176–5184, 2014, doi: 10.1039/c4an00961d.

[4] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 580–587, 2014, doi: 10.1109/CVPR.2014.81.

[5] L. Gomez-Chova, D. Tuia, G. Moser, and G. Camps-Valls, “Multimodal Classification of Remote Sensing Images: A Review and Future Directions,” *Proc. IEEE*, vol. 103, no. 9, pp. 1560–1584, 2015, doi: 10.1109/JPROC.2015.2449668.

[6] J. Yang, “Hyperspectral Image Classification via Multiscale,” vol. 15, no. 1, pp. 112–116, 2018.

[7] L. Fang, N. He, S. Li, A. J. Plaza, and J. Plaza, “A New Spatial-Spectral Feature Extraction Method for Hyperspectral Images Using Local Covariance Matrix Representation,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3534–3546, 2018, doi: 10.1109/TGRS.2018.2801387.

[8] Q. Gao, S. Lim, and X. Jia, “Hyperspectral image classification using joint sparse model and discontinuity preserving relaxation,” *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 1, pp. 78–82, 2018, doi: 10.1109/LGRS.2017.2774253.

[9] T. Dundar and T. Ince, “Sparse Representation-Based Hyperspectral Image Classification Using Multiscale Superpixels and Guided Filter,” *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 2, pp. 246–250, 2019, doi: 10.1109/LGRS.2018.2871273.

[10] A.-R. Mohamed, G. Dahl, and G. Hinton, “Deep Belief Networks for Phone Recognition,” *Scholarpedia*, vol. 4, no. 5, pp. 1–9, 2009, doi: 10.4249/scholarpedia.5947.

[11] Y. Zhang, J. Gao, and H. Zhou, “Breeds Classification with Deep Convolutional Neural

Network," *ACM Int. Conf. Proceeding Ser.*, pp. 145–151, 2020, doi: 10.1145/3383972.3383975.

[12] X. Shang, M. Song, and C. Yu, "Hyperspectral Image Classification with Background," *Int. Geosci. Remote Sens. Symp.*, vol. 56, no. 6, pp. 2714–2717, 2019, doi: 10.1109/IGARSS.2019.8898520.

[13] G. Cheng, Z. Li, J. Han, X. Yao, and L. Guo, "Exploring Hierarchical Convolutional Features for Hyperspectral Image Classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 11, pp. 6712–6722, 2018, doi: 10.1109/TGRS.2018.2841823.

[14] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks," *IEEE Trans. Geosci. Remote Sens. Deep*, vol. 54, no. 10, pp. 6232–6251, 2016.

[15] K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis, "Deep supervised learning for hyperspectral data classification through convolutional neural networks," *Int. Geosci. Remote Sens. Symp.*, vol. 2015-Novem, pp. 4959–4962, 2015, doi: 10.1109/IGARSS.2015.7326945.

[16] Z. Zhong, J. Li, Z. Luo, and M. Chapman, "Spectral-Spatial Residual Network for Hyperspectral Image Classification: A 3-D Deep Learning Framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 847–858, 2018, doi: 10.1109/TGRS.2017.2755542.

[17] H. C. Mingyi He, Bo Li, "MULTI-SCALE 3D DEEP CONVOLUTIONAL NEURAL NETWORK FOR HYPERSPECTRAL IMAGE CLASSIFICATION Mingyi He , Bo Li , Huahui Chen Northwestern Polytechnical University , School of Electronics and Information International Center for Information Acquisition & Proces," pp. 3904–3908, 2017.

[18] S. K. Roy, G. Krishna, S. R. Dubey, and B. B. Chaudhuri, "HybridSN: Exploring 3D-2D CNN Feature Hierarchy for Hyperspectral Image Classification," *arXiv*, vol. 17, no. 2, pp. 277–281, 2019.

[19] S. Ji, W. Xu, M. Yang, and K. Yu, "3D Convolutional neural networks for human action recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 221–231, 2013, doi: 10.1109/TPAMI.2012.59.

[20] F. Melgani and L. Bruzzone, "Classification of Hyperspectral Remote Sensing," vol. 42, no. August, pp. 1778–1790, 2004.

[21] A. Ben Hamida, A. Benoit, P. Lambert, and C. Ben Amar, "Three dimensional deep learning approach for remote sensing image classification," *arXiv*, vol. 56, no. 8, pp. 4420–4434, 2018.

[22] W. Li, G. Wu, F. Zhang, and Q. Du, “Hyperspectral Image Classification Using Deep Pixel-Pair Features,” *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 844–853, 2017, doi: 10.1109/TGRS.2016.2616355.

[23] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 580–587, 2014, doi: 10.1109/CVPR.2014.81.

[24] A. Bordes, X. Glorot, J. Weston, and Y. Bengio, “Joint learning of words and meaning representations for open-text semantic parsing,” *J. Mach. Learn. Res.*, vol. 22, pp. 127–135, 2012.

[25] C. S. Silva, M. F. Pimentel, R. S. Honorato, C. Pasquini, J. M. P. Montalbán, and A. Ferrer, “Near infrared hyperspectral imaging for forensic analysis of document forgery,” *Analyst*, vol. 139, no. 20, pp. 5176–5184, 2014, doi: 10.1039/c4an00961d.

[26] X. Jia, B. C. Kuo, and M. M. Crawford, “Feature mining for hyperspectral image classification,” *Proc. IEEE*, vol. 101, no. 3, pp. 676–697, 2013, doi: 10.1109/JPROC.2012.2229082.

[27] Q. Wang, J. Lin, and Y. Yuan, “Salient Band Selection for Hyperspectral Image Classification via Manifold Ranking,” *IEEE Trans. Neural Networks Learn. Syst.*, vol. 27, no. 6, pp. 1279–1289, 2016, doi: 10.1109/TNNLS.2015.2477537.

[28] Y. Yuan, J. Lin, and Q. Wang, “Hyperspectral image classification via multitask joint sparse representation and stepwise MRF optimization,” *IEEE Trans. Cybern.*, vol. 46, no. 10, pp. 2966–2977, 2015, doi: 10.1109/TCYB.2015.2484324.

[29] M. Khodadadzadeh, J. Li, A. Plaza, H. Ghassemian, J. M. Bioucas-Dias, and X. Li, “Spectral-spatial classification of hyperspectral data using local and global probabilities for mixed pixel characterization,” *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6298–6314, 2014, doi: 10.1109/TGRS.2013.2296031.

[30] J. Peng, Y. Zhou, and C. L. P. Chen, “Region-Kernel-Based Support Vector Machines for Hyperspectral Image Classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 9, pp. 4810–4824, 2015, doi: 10.1109/TGRS.2015.2410991.

[31] Y. Tarabalka, M. Fauvel, J. Chanussot, and J. A. Benediktsson, “SVM- and MRF-based method for accurate classification of hyperspectral images,” *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 736–740, 2010, doi: 10.1109/LGRS.2010.2047711.

[32] K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis, “Deep supervised learning for hyperspectral data classification through convolutional neural networks,” *Int. Geosci. Remote Sens. Symp.*, vol. 2015-Novem, pp. 4959–4962, 2015, doi: 10.1109/IGARSS.2015.7326945.

[33] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks," *IEEE Trans. Geosci. Remote Sens. Deep*, vol. 54, no. 10, pp. 6232–6251, 2016.

[34] W. Zhao and S. Du, "Spectral-Spatial Feature Extraction for Hyperspectral Image Classification: A Dimension Reduction and Deep Learning Approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4544–4554, 2016, doi: 10.1109/TGRS.2016.2543748.

[35] Y. Li, H. Zhang, and Q. Shen, "Spectral-spatial classification of hyperspectral imagery with 3D convolutional neural network," *Remote Sens.*, vol. 9, no. 1, 2017, doi: 10.3390/rs9010067.

[36] Z. Zhong, J. Li, Z. Luo, and M. Chapman, "Spectral-Spatial Residual Network for Hyperspectral Image Classification: A 3-D Deep Learning Framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 847–858, 2018, doi: 10.1109/TGRS.2017.2755542.

[37] H. C. Mingyi He, Bo Li, "MULTI-SCALE 3D DEEP CONVOLUTIONAL NEURAL NETWORK FOR HYPERSPECTRAL IMAGE CLASSIFICATION Mingyi He , Bo Li , Huahui Chen Northwestern Polytechnical University , School of Electronics and Information International Center for Information Acquisition & Proces," pp. 3904–3908, 2017.

[38] S. K. Roy, G. Krishna, S. R. Dubey, and B. B. Chaudhuri, "HybridSN: Exploring 3D-2D CNN Feature Hierarchy for Hyperspectral Image Classification," *arXiv*, vol. 17, no. 2, pp. 277–281, 2019.

[39] S. Ji, W. Xu, M. Yang, and K. Yu, "3D Convolutional neural networks for human action recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 221–231, 2013, doi: 10.1109/TPAMI.2012.59.

[40] F. Melgani and L. Bruzzone, "Classification of Hyperspectral Remote Sensing," vol. 42, no. August, pp. 1778–1790, 2004.

[41] A. Ben Hamida, A. Benoit, P. Lambert, and C. Ben Amar, "Three dimensional deep learning approach for remote sensing image classification," *arXiv*, vol. 56, no. 8, pp. 4420–4434, 2018.

[42] W. Li, G. Wu, F. Zhang, and Q. Du, "Hyperspectral Image Classification Using Deep Pixel-Pair Features," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 844–853, 2017, doi: 10.1109/TGRS.2016.2616355.

LIST OF PUBLICATIONS

Paper	Title.Author list Conference/Journal	Status
Paper 1	**Deep Intensified Archetypical approach for Hyper Spectral Image Classification.** Ankita Thapliyal, (ASIANCON)	Accepted
Paper 2	**A Dense Hybrid Neural Network approach for Hyper Spectral Image Classification.** Ankita Thapliyal, (International Conference on Advances in manufacturing technologies and application of artificial intelligence 2021)	Accepted