Effective Preprocessing of Satellite Images

S.Asha[1] ,M.Parvathy[2] ,Suriya.D[3]

[1]Assistant Professor, [2]Prof & Head, [3]Assistant Professor,

Department of Computer Science and Engineering,

Sethu Institute of Technology

Abstract— This study provides a theoretical and practical overview of the topic of preprocessing and production of satellite images. Specifically, the entire lifetime of the process of creating satellite images is covered. The method of remote sensing, as well as additional mapping and fusing of satellite pictures within captured regions of the electromagnetic spectrum, are given particular explanatory attention. Furthermore, retrieval techniques for satellite pictures have been demonstrated in practice. The Sentinel 2 dataset was used to extract the satellite imagery data, which was then examined using Python image processing packages.

Keywords — Computer vision, satellite image, image formation, image preprocessing, NDVI, Sentinel 2, satellite, imagery, python, sentinel, image preprocessing, remote sensing

I. INTRODUCTION

Today is a great day to be alive, as anyone may get the most recent free satellite photos of our planet. Thanks to satellite imagery sources like Sentinel, Landsat, USGS, NASA Earthdata, and DigitalGlobe, we can see back up to 40 years at nearly every location on Earth.The topography, vegetation, and resources of the planet may all be seen in amazing detail in photographs and other images of Earth shot from space and the air. Remotely sensed imagery, which includes satellite and aerial photos, enables precise mapping of land cover and clarifies landscape features at the regional, continental, and even global levels. For individuals who work closely with satellite imagery, it is crucial to comprehend the process of taking and preprocessing satellite images before they are indexed and utilized for analytical reasons.

This paper aims to provide an overview of the basic and practical aspects of satellite image formation and preprocessing. It begins by dispelling some common misconceptions and progresses to a description of the fundamental physical concepts that satellite imagery relies on. Additionally, it covers some common techniques and approaches used to address issues that arise in this field of study.

2.IMAGE FORMATION USING REMOTE SENSING

The idea that satellite photos are photographs is a frequent one. Actually, they are very dissimilar. Satellites gather digital data through remote sensing. Computers are used by people to transform this data into pictures. Information is gathered on using remote sensing to something without making contact with it. The science (and, to some extent, art) of learning about the Earth's surface without physically touching it is known as remote sensing.[1] This is accomplished by detecting, logging, and processing energy that is reflected or emitted, then using the data for analysis and application.

It is possible to detect electromagnetic energy electrically or by photography. Chemical processes on the surface of light-sensitive film are used in the photographic process to identify and capture energy variations. In remote sensing, it's critical to understand the difference between images and photographs.[1] Any visual depiction is referred to as an image, regardless of the wavelengths or remote sensing equipment used to find and capture the electromagnetic radiation. Images that have been identified and captured on photographic film are explicitly referred to as photographs. The visible portion of the spectrum was used to capture the black and white image on the left, which shows a portion of Ottawa, Canada. Typically, photos are taken between 0.3 and 0.9 mm in wavelength, which corresponds to the visible and reflected light infrared. We can conclude from these definitions that all pictures are photographs, but not all images are photographs.

2.1 Electromagnetic Spectrum:

Radiation is released by everything. Furthermore, all that receives light from the sun reflects radiation. Our lives are coloured by this radiation because it allows us to see and perceive different visible spectrum wavelengths, which our brains process to perceive and interpret colour.[2] On the other hand, other sections of the spectrum are invisible to our sensors, or eyes. We can see across the visible spectrum thanks to remote sensors. In other

words, humans are unable to view radiation such as gamma, x, or radio waves that are not visible to the human eye. Consider bee guides that are invisible to the human eye but visible through the use of infrared photographs. Crucial point Different wavelengths are absorbed and reflected by different things. Greenery, for instance, reflects in the infrared is fairly good. For this reason, remote sensing technology allows us to see the world from fresh perspectives. Scientists and engineers work together to create sensors that monitor particular elements of the EMS, such as the difference between the atmosphere and the ocean or the land.
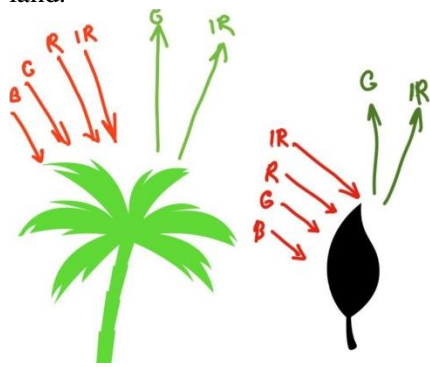

Figure 1.Wave reflection on different objects

Chlorophyll, a substance found in leaves, is a strong absorber of red and blue light while reflecting green light. We perceive leaves to be greenest in the summer, when the amount of chlorophyll in them reaches its peak. Autumn leaves appear red or yellow because they have less chlorophyll, which causes decreased absorption of red wavelengths and a corresponding increase in reflection of those wavelengths (yellow being a mixture of red and green wavelengths). Healthy leaves' interior structure is a great diffuse reflector of near-infrared (IR) light. Trees at these wavelengths would appear incredibly bright to us if our eyes were capable of detecting near-infrared light. In actuality, one method by which scientists can assess the health (or lack thereof) of vegetation is by measuring and tracking the near-IR reflectance.
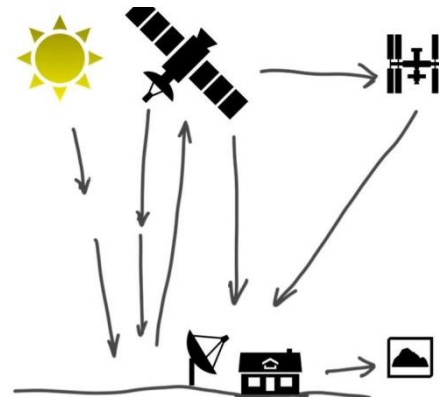

Figure 2. The full-travel process of remote sensing

Satellite sensors measure the quantity of reflected energy that is communicated to them. These sensors are detecting energy from objects below remotely. This data is saved and transformed into an image format. The complete remote sensing method is depicted in the figure 2. In order to help us grasp what might or might not be possible to do using remote sensing and to help you start to learn more about the meaning of images, we'll be walking you through a portion of this process here.

Different portions of the electromagnetic spectrum are reflected by different types of objects. For instance, plants reflect light in a different way than structures or rocks do.[3] This holds true for both the visible and invisible halves of the spectrum. reflected energy, such as infrared or UV light. Scientists use the colour of reflected light to examine Earth from space. An object's characteristic is described by the amount of light reflected. Ask the students to graph the quantity of light reflected by the items they used in worksheet 2's activity using the chart they just made. The red, green, and blue light will be on the x-axis, while the amount of reflected light will be on the y-axis. Scientists use the colour of reflected light to examine Earth from space.

An object's characteristic is described by the amount of light reflected. Ask the students to graph the quantity of light reflected by the items they used in worksheet 2's activity using the chart they just made. The red, green, and blue light will be on the x-axis, while the amount of reflected light will be on the y-axis. The amount

of light reflected by these sensors is measured numerically, and the data is then graphed. The spectral signatures, or graphs, of many substances, such plants, rocks, and water, are distinct from one another. A scientist can identify similar objects across a wide geographic area by using the signature of a specific object. We can create an image with just the numbers since the amount of light reflected can be expressed as a number. Sensors Capture Reflected Energy Intensity Numerically Satellite imagery consists of tiny squares, each with a distinct grey tone or colour, much like the images on your screen. These squares, known as pixels (short for picture components), show the proportion of reflected light energy that was captured for that area of the picture.
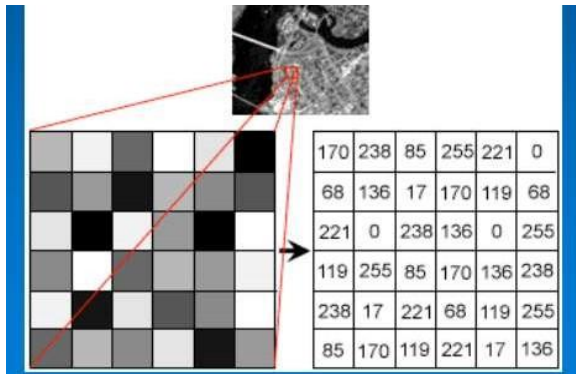

Figure 3. Digital representation of satellite imagery

Electromagnetic energy is electronically recorded by sensors and stored as a digital array of integers. The information is recorded by satellite sensors as a single numerical number that represents the energy intensity they observe in a specific area. For each 30 m x 30 m rectangle, the Landsat sensor calculates the quantity of reflected energy. It performs this function for every one of the electromagnetic spectrum's seven parts. It keeps the values for every spectrum segment apart. These numbers are noted and can be transformed into a greyscale to create a pixel-based black and white image. How does the computer determine which areas of the picture ought to be bright and which should be dark? Binary numbers, which are sets of numbers made up of 0s and 1s that function

as an on-off switch, are a numerical language that computers can understand. Our decimal system is converted to binary numerals as follows: 00 = 0, 01 = 1, 10 = 2, 11 = 3. Since computers are all finicky and prefer to turn on and off only, we are unable to use decimal numbers. Think of an image, for instance, that consists of 8 columns by 5 rows of pixels.

There are four tones in this figure: light grey, dark grey, black, and white. The binary numbers 00, 01, 10, and 11 correspond to the darkest point, light grey, and brightest portion, respectively. Consequently, the spacecraft indicates that four of our pixels (B5, C4, D7, and E2) are 00. Three light grey pixels (D3, D6) and three dark grey pixels (B3, C2, C6, and E6) have been assigned the binary number 01 respectively.

There are four tones in this figure: light grey, dark grey, black, and white. The binary numbers 00, 01, 10, and 11 correspond to the darkest point, light grey, and brightest portion, respectively. Consequently, the spacecraft indicates that four of our pixels (B5, C4, D7, and E2) are 00. Three light grey pixels (D3, D6) and three dark grey pixels (B3, C2, C6, and E6) have been assigned the binary number 01 respectively.

This yields 256 grayscale images ranging from black to white. Every pixel in every row and column receives these binary values from the spaceship, which range from 0 to 255. To keep track of every number for every pixel, a computer is required.[4]

2.2. Assigning Colors to Bands:

For humans to comprehend the information more fully, visual representations of these data are also necessary. We convert the numerical scale to colour to make an image easier for humans to read. We can choose a colour for each band in the data such that, for example, gradations of red would appear in place of black and white. Light that is infrared or ultraviolet does not affect human vision. We need to employ the colours red, green, and blue from the visible region of the electromagnetic spectrum to create a composite image from remote sensing data that makes sense to human eyes.

TABLE I.ELECTROMAGNETIC SPECTRUM WITHIN WAVELENGTH

ETM+ BAND	Wavelength(um)	Colors
1	0.450-0.515	Blue
2	0.525-0.605	Green
3	0.630-0.690	Red
4	0.750-0.900	Near infrared
5	1.55-1.75	Infrared
6	2.09-2.35	Shortwave infrared
7	10.40-12.50	Thermal infrared
8	0.52-0.90	Panchromatic

Figure 4.A Landsat scene across 7 bands

A portion of a Landsat scene is shown here. Until we give a colour to the data, all seven

Figure 5. Landsat scene across 7 bands with false colors

bands of data—including those that represent visible regions of the spectrum (blue, green, or red reflected light—are in grey scale. In this instance, we designated the colour red to signal red light, the colour green to represent green light, and the colour blue to indicate blue light inside the data band. This creates an image that's similar to what it would seem like to be atop one of these satellites and is referred to as a true colour image.[5]

False colour is sometimes used to make objects easy on the eyes. 4,3,2 is a common band combination for this. Red is the colour allocated to band 4, which is the near infrared. Band 2—what you and I would see as green—is assigned the colour blue, whereas band 3, which represents light reflecting in the red region of the spectrum, is assigned the colour green. Why do this action? Remember that near infrared reflects significantly from healthy vegetation. Our eyes are particularly sensitive to red in vegetation, so we would notice that. The blue-green in 4,3,2 combinations denotes the constructed environment, which includes pavement, buildings, and roads. Water seems black or dark blue.
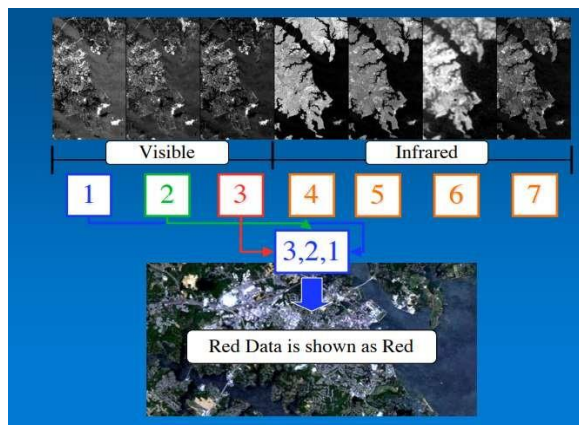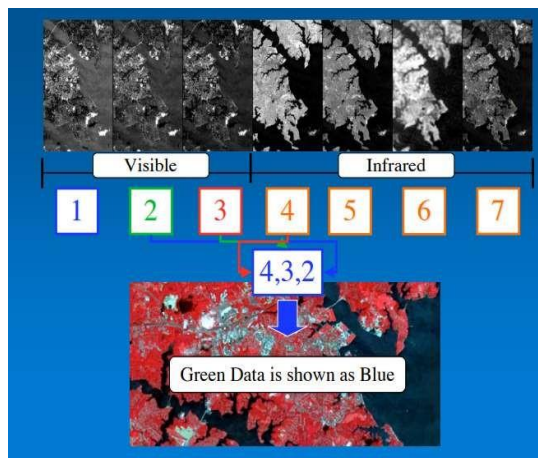
3.SATELLITE IMAGERY RETRIEVAL AND ANALYSIS IN PYTHON

Here are a few basic yet common examples of retrieving and analysing satellite imagery to provide a realistic understanding of how these photos are created and handled.

Sentinel 2 data:

Sentinel 2 image access comes in a variety of forms, the most of which necessitate website participation.Installed via pip package management, Sentinelsat is a Python module that facilitates searching, obtaining, and downloading of Sentinel satellite photos.

```
!pip install
sentinelsat import
geopandas as gpd
import folium
n_aral =
gpd.read_file('Uzb/Aral_sea.shp') m =
folium.Map([41.7023292727353,
12.34697305914639], zoom_start=12)
folium.GeoJson(nReserve).add_to(m)
```

Code 1. Data retrieval with sentinelsat Python library

By using the code above, we were able to read the Aral Sea shapefile in Geopandas and give it the name n_aral. We then used Folium to build an empty base map that was centred around points in the region, which we named m. In order to see the boundary we are interested in, we may now incorporate the Geopandas data into the foundation map we have made. We need to construct a footprint from the n_aral geometry as a final step before we can search and get sentinel 2 photos.

Since our data is in Shapefiles and has already been read as a Geopandas GeodataFrame, we will be using the Shapely Python module in this instance. We can now use the api we developed above to perform a query. Depending on your use case, there are various ways to formulate your query here. For this example, we constructed a query for Sentinel 2 pictures Level 2A that fall within or intersect the footprint (Area of study) and have cloud coverage ranging from 0 to 10. Sentinel Level 2A satellite photos captured between "20190601" and "20190626" are the only ones we are interested in for the specified time period.

```
from shapely.geometry import MultiPolygon, Polygon
footprint = None
for i in nReserve['geometry']:
    footprint = i
    show_fp(footp
    rint)
```
Code 2. Representation in GeodataFrame of footprint

NDVI with Rasterio and numpy

A Python module called Rasterio allows you to open, create, explore, and analyse georasters. The library can work with imagery created by drones, digital elevation models, and satellite photos in addition to other formats, including GeoTIFF files.

```
ndvi_tiff = rasterio.open('../tiffs/ndvImage.tiff')
fig       =       plt.figure(figsize=(18,12))
plot.show(ndvi_tiff)
```
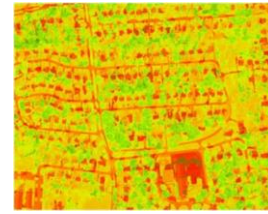

Code 3. Visualization of NDVI calculation result

CONCLUSION

There is much more to satellite imagery than merely capturing photos. A numerical representation of the energy reflected across a variety of wavelengths is present in each pixel. Additionally, a variety of measurements and analytical conclusions can be drawn from this data. This paper aims to provide a philosophical and practical knowledge of the theories that underpin the subject by introducing satellite image generation and preprocessing. In order to obtain important insights from pictures and other satellite-taken images of the Earth, it is necessary to properly arrange the conceptual knowledge regarding satellite image production and preprocessing.

REFERENCES

[1] Pandey, Pratibha, Kranti Kumar Dewangan, and Deepak Kumar Dewangan. "Enhancing the quality of satellite images by preprocessing and contrast enhancement." *2017 international conference on communication and signal processing (ICCSP)*. IEEE, 2017.

[2] Ranghetti, Luigi, et al. "'sen2r': An R toolbox for automatically downloading and preprocessing Sentinel-2 satellite data." *Computers & Geosciences* 139 (2020): 104473..

[3] Manohar, N., et al. "Classification of satellite images." *Information and Communication Technology for Intelligent Systems: Proceedings of ICTIS 2020, Volume 1*. Springer Singapore, 2021.

[4] Vivone, Gemine, et al. "A benchmarking protocol for pansharpening: Dataset, preprocessing, and quality assessment." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 14 (2021): 6102-6118.

[5] Guo, Yanan, et al. "Cloud detection for satellite imagery using attention-based U-Net convolutional neural network." *Symmetry* 12.6 (2020): 1056.

[6] Ghorai, Debabrata, and Manik Mahapatra. "Extracting shoreline from satellite imagery for GIS analysis." *Remote Sensing in Earth Systems Sciences* 3.1-2 (2020): 13-22.

[7] Asokan, Anju, et al. "Image processing techniques for analysis of satellite images for historical maps classification—An

overview." *Applied Sciences* 10.12 (2020): 4207.