# MAPPING THE MOON: A DEEP LEARNING STRATEGY FOR LUNAR CRATER DETECTION FROM CHANDRAYAAN-2 SATELLITE CAPTURES

**MSc Astronomy Thesis**


**By:**

**Kunal Thapar**

**DEPARTMENT OF ASTRONOMY , ASTROPHYSICS AND SPACE ENGINEERING**

**INDIAN INSTITUTE OF TECHNOLOGY INDORE**


**May, 2024**

# MAPPING THE MOON: A DEEP LEARNING STRATEGY FOR LUNAR CRATER DETECTION FROM CHANDRAYAAN-2 SATELLITE CAPTURES

**A THESIS**

*Submitted in partial fulfilment of the*

*requirements for the award of the degree*

***of***

**Master of Science**

***by***

**Kunal Thapar**

**DEPARTMENT OF ASTRONOMY , ASTROPHYSICS AND SPACE ENGINEERING**

**INDIAN INSTITUTE OF TECHNOLOGY INDORE**

**May, 2024**

# INDIAN INSTITUTE OF TECHNOLOGY INDORE

## CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled **MAPPING THE MOON: A DEEP LEARNING STRATEGY FOR LUNAR CRATER DETECTION FROM CHANDRAYAAN 2 SATELLITE CAPTURES** in the partial fulfilment of the requirements for the award of the degree of **MASTER OF SCIENCE** and submitted in the **DEPARTMENT OF ASTRONOMY, ASTROPHYSICS AND SPACE ENGINEERNING, Indian Institute of Technology Indore**, is an authentic record of my own work carried out during the time period from **JULY 2022** to **MAY 2024** under the supervision of **DR. UNMESH KHATI, ASSISTANT PROFESSOR**

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute.

22/5/24

**Signature of the student with date**
**KUNAL THAPAR**

---

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Unmesh 22/05/2024

Signature of the Supervisor of
M.Sc. thesis #1 (with date)
**Dr. Unmesh Khati**

Signature of the Supervisor of
M.Sc. thesis #2 (with date)
**(NAME OF SUPERVISOR)**

---

**Kunal Thapar** has successfully given his/her M.Sc. Oral Examination held on **24 April 2024**.

Unmesh 22/05/2024

Signature(s) of Supervisor(s) of MSc thesis
Date:

Programme Coordinator, M.Sc.
Date: 23/05/2024

S. Das

Convenor, DPGC
Date: 24/05/2024

Head, DAASE
Date:

---

# ACKNOWLEDGEMENT

I would like to express my deepest gratitude to my supervisor, Dr. Unmesh Khati, for granting me the opportunity to undertake this project. His invaluable guidance and the freedom he afforded me when necessary were crucial in moving this project forward. I am grateful for his patience and understanding, especially during challenging times.

I extend my thanks to the Department of Astronomy, Astrophysics, and Space Engineering for providing me with the unique opportunity to explore and experience the field of astronomy in a way that no other place could. I will fondly remember those captivating astronomy classes that never made me check the time.

I would like to take a moment to appreciate those who have supported me behind the scenes. Firstly, my heartfelt thanks go to my parents and my sister, whose belief in me and understanding of my passion for astronomy have been unwavering. Their immense love and support, especially from my mother, have been a constant source of motivation throughout these two years.

A big thanks to Mr Hamish D'Souza for his assistance and guidance with my project.

I am also grateful to my dear friends here – G Akash, Sanjay Kumar Yadav, Sparsh Arya, Ashad Ahmad, Hemanth Potluri, Praneeth Avasarala – for making my time in this campus truly amazing. Their guidance and support was invaluable.

I would also like to thank Sayan Doloi, Sai Kumar Reddy, Abhinav Narayan, Naveen Sharma and Vishal for their help. I would also like to take this moment to thank my friends back home - M Pravin, Steve and Shiva for their help and support.

Finally, I express my sincerest gratitude to myself for the relentless determination and hard work invested in completing this thesis. I am proud of the effort and dedication I have demonstrated throughout this academic endeavor.

# ABSTRACT

The moon has its geological information sustained for billions of years now. With the absence of an atmosphere and no tectonic activities, surface features like craters, rilles, pits, etc. are well preserved. Craters are one of the most predominant features of the lunar surface. Crater counting serves as a crucial methodology, providing key insights into lunar surface evolution, impact history, and other geological processes. With the development of computer vision and image processing, crater counting has switched from the traditional manual counting, which is time-consuming and prone to error, to using deep learning methods with the help of object detection models which has become much faster and precise in just a few years. This project harnesses the power of Convolutional Neural Networks (CNNs) to detect lunar surface craters using satellite images from the Orbiter High-Resolution Camera (OHRC) aboard Chandrayaan 2. The methodology involves image labelling, model training,obtaining weight function during validation, and subsequent testing of the model on unlabelled images. The project leverages YOLOv8's adept architecture, which is designed to be fast and accurate, making it an excellent choice for a wide range of object detection, instance segmentation, image classification and pose estimation tasks. For the project, more than 750 images taken from the OHRC were manually labelled as crater. Precision and recall were the two metrics used to measure the model's efficiency. It was observed that as the number of epochs, there were specific epoch numbers where the model's training performance show higher results as compared to higher or lower epochs.

# Contents

# Chapter 1

# Introduction

The moon's surface has a lot to say about the evolution and alteration of our solar system. From earth, the moon does look like a peaceful and quite place, seems to be having earth-like features but the closer we go to the surface, the more we realize that the surface has had a long and violent history of heavy bombardments with asteroid and comet impacts in the past and in great numbers, introducing imbalance in the core and leading to violent volcanic eruptions, leading to the formation of craters of different sizes and shapes that we know today.[Bottke and Norman [2017]]

The lunar surface remains relatively unaltered by recent geological activity, attributed to the enduring cold and rigidity of the lunar crust and mantle, absence of an atmosphere, surface water and alteration due to species evolution, practically untouched for much of its lifetime, assumed to be close to 4.5 billion years [Heiken et al. [1991]]. Unlike Earth, the Moon lacks convective internal mass transport, leading to the absence of geological effects such as volcanism, uplift, and faulting that shape and reshape Earth's surfaces[Wilhelms et al. [1987]]. This has led to the surface filled with surface features, some dating as old as the bombardment era. Craters can be said to be the most abundant feature of the surface, size scaling from millimeters to thousands of kilometers.

Planetary geologists have identified the significance of crater studies in unveiling planetary evolution mysteries. The Moon's stable yet heavily cratered surface serves as evidence of continuous bombardment by external objects. Correlations between crater density and surface age, along with crater counts, enable estimates of lunar surface relative age [Arya et al. [2012]]. Craters offer insights into impact direction, impactor population, and surface properties, undergoing degradation due to gravity and surface processes[Basilevskii [1976]]. The study of lunar craters contributes to understanding lunar regolith formation and evolution, providing essential information for determining landing sites and guiding rovers independently of Earth communication or Earth's global positioning system (GPS).

Further as humanity is reaching the age of technological and engineering advancements, plans for permanent settlements on the surface of the moon has been a serious topic of discussion [Rugani et al. [2021]]. Scientists, geologists and engineers are taking this as a challenge to push their limits and understand how humans can live outside of earth. While this is undeniably ambitious, it surely raises questions about its practicality and potential challenges that may affect the quality of life for future lunar residents. Nevertheless, In lights of the increasing gravity and consideration accorded to the prospect of establishing permanent lunar settlements by the scientific and engineering communities, there arises an imperative for the comprehensive study and meticulous mapping of lunar craters. Given that lunar craters stand as the predominant surface feature, the precision and thoroughness of crater detection assume paramount importance.

Determining crucial surface information of planetary surface is difficult as it requires physically collecting the samples and working on it, leaving us to find alternative methods. Remote sensing is one of the efficient examples. Crater counting also is considered to be a method which can determine the age of the surface just by counting the number of craters in the vicinity [Neukum et al. [1975]].Manual crater counting was the very first introduced method. This, however, has proven to not only be laborious and time consuming task, but is also prone to human error. As computer vision has emerged, along with deep learning, algorithms have been made for various tasks like object detection, instance segmentation, pose estimation, etc. This technology has been utilized for detection of craters on the surface of planetary surfaces.

Over time, researchers have devised diverse automated crater detection algorithms (CDAs) designed to accelerate this process, as higher-resolution data are at hand. These automated methods have largely evolved in line with the prevailing computer science techniques of their respective time periods. They encompass a wide array of computer vision methods, and more recently, machine learning approaches have gained increasing popularity in this field.

The aim of this thesis project is to utilize these advancements in computer vision and object detection method, along with high quality satellite images going hand in hand to detect craters of the lunar surface in a more efficient, easier and faster way as compared to manual counting of craters, which is, in many ways, prone to error and is a rather frustrating and an impractical process.

# Chapter 2

# Crater - Types and formation

We will first try to understand the very object that we are trying to detect - Craters. We will look through its formation and structure and the various types of them based on their nature of formation. We will also familiarize ourselves with the factors that affect the formation of these craters because these factors does make a great impact on the formation of these craters. Finally, since the lunar surface is predominantly filled with impact craters, we will look through the possible theory behind their formation.

## Crater

Simply put, A crater is a concave feature or a depression on the surface of a terrestrial planet due to the act of the release of immense energy on the planetary surface "Release of immense energy" must be emphasized here as it serves as a gateway to understand the various ways in which a crater can be formed. Craters can be divided into various types based on their nature of formation, or mode of this release of immense energy:

- Endogenic craters, also called volcanic craters, they are formed when volcanos erupt on a planetary surface. The tremendous pressure generated by a volcanic eruption propels magma, gases, and volcanic materials through a vent, resulting in the formation of a central conduit. The amassed materials give rise to a volcanic cone, complete with a bowl-shaped hollow at its apex, known as the volcanic crater.
- Glacier craters, also known as ice cauldrons, are geological features that form in glacial environments. They are typically caused by the drainage of water from beneath a glacier, leading to the melting of ice and the collapse of the overlying glacier. As the glacier retreats, these depressions are left behind, often filled with water, sediment, or ice, and they can have a variety of shapes and sizes.

- Explosion craters are formed by powerful explosions, often from the detonation of explosive devices or meteorite impacts. The explosion creates a sudden release of energy that excavates the surrounding terrain, producing a bowl-shaped depression.
- Impact craters, which are depressions that take shape when an asteroid or comet, referred to as a meteorite if it reaches the planetary surface, collides with a planet's surface. This collision not only leads to the creation of a crater but also transmits its kinetic energy through shock waves that penetrate the solid structure of the planet, causing deformations in both the surface and the planet's interior

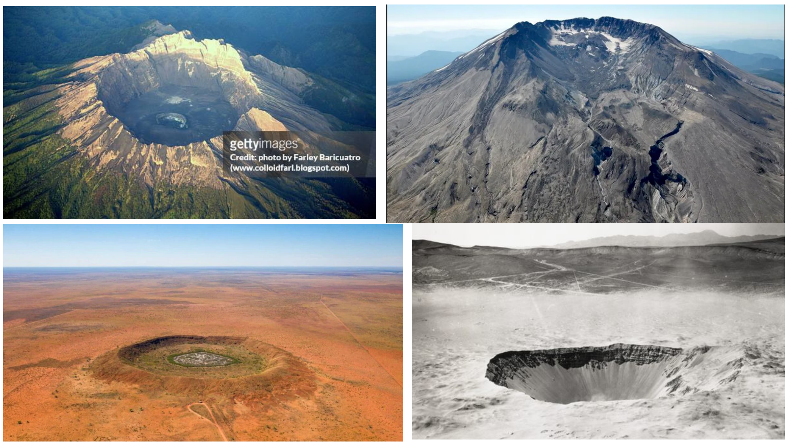



**Figure 2.1:** Clockwise: Endogenic,Glacier,Explosion and Impact crater

# Impact craters

Impact craters, as discussed, is the most predominantly found crater on the surface of the moon, and the most abundant surface feature as well, which is why understanding of the creation is important.
The ideal structure of an impact characteristic takes the form of a concave depression, known as a crater, encircled by an elevated rim, but we will come to why that's not always the case. The shapes of impact craters alter in relation to their diameter, measured from one rim to the other. As the diameter increases, they become comparatively shallower and exhibit more intricate rim and floor features, including the emergence of central peaks and rings. These morphological variations are observed in craters across all terrestrial planets and moons.

The formation of large impact craters and the release of immense energy involved in the process cannot be replicated in any laboratory setting. Fortunately, throughout recorded human history, no such structure has been created. Our understanding is solely derived from indirect sources, encompassing theoretical and experimental studies focused on the analysis of shock

**Figure 2.2:** Meteor crater in Arizona,USA(Courtesy: USGS.gov)

waves.
Before we move further, A basic understanding of shockwaves is necessary.
When a meteorite hurtles towards a planet, it carries with it immense kinetic energy. Upon impact, this energy is transferred to the planetary surface, generating formidable pressure waves known as **shock waves** . These shock waves, akin to the blast waves from an explosion, propagate through the solid material, causing a cascade of effects. The dynamics of the impact, including the meteorite's speed and mass, play a significant role in determining the intensity and reach of the shock waves. The larger and faster the meteorite, the more powerful the shock waves will be. As the shock waves traverse the planetary surface, they cause the solid material to deform, fracture, and undergo structural alterations. This process is distinct from ordinary geological processes, such as erosion or plate tectonics, as it involves the sudden application of extreme pressure.

The impact of shock waves extends beyond their immediate effects. They play a pivotal role in shaping the morphology of impact craters, those immense scars that dot the surfaces of many planets and moons. In addition, shock waves can modify geological structures, influencing the distribution of minerals and the formation of rock formations. The influence of shock waves extends beyond the physical structure of a planet. They can also affect the planet's atmosphere, potentially altering climate patterns and contributing to the development of life. For instance, some theories suggest that the Chicxulub impact, which is thought to have led to the extinction of the dinosaurs, may have triggered global climate changes.

Studying shock waves resulting from meteorite impacts provides valuable insights into the geological history and processes of celestial bodies. By analyzing the effects of shock waves, scientists can gain a better understanding of the impact cratering process, the evolution of planetary surfaces, and the potential for life in the cosmos. Ongoing research and exploration

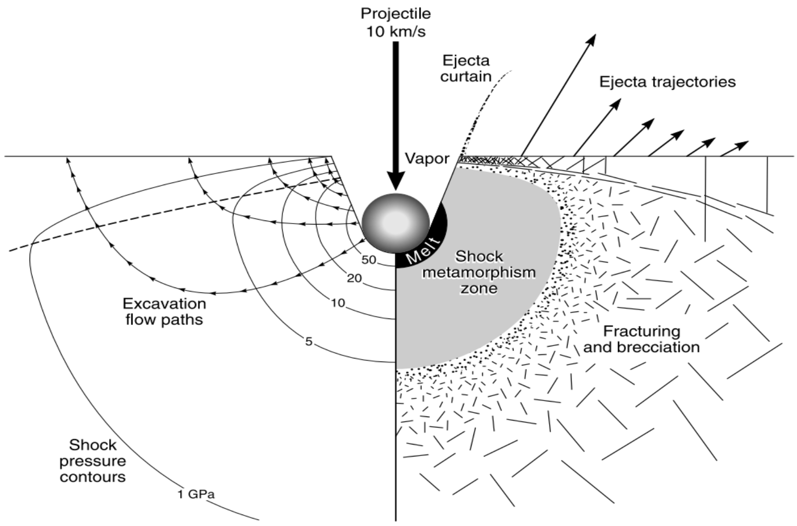



**Figure 2.3:** Contact/compression stage where the projectile touches the planetary surface and severe shock waves (theoretical values of pressures calculated in in GPa as shown) are emanating from the interface and extending outward into the target (Image courtesy: French [1998])

efforts at impact sites worldwide continue to enrich our knowledge of these extraterrestrial events and their broader implications for planetary dynamics.

## Crater formation stages

Theoretically, there are three stages of crater forming process(As explained by French [1998] and Wilhelms et al. [1987]):

- **Contact andcompression stage**: This stage starts when the projectile ( formal word for a meteorite striking on the planetary surface) makes contact with the surface of the planet. The fundamental characteristics of how kinetic energy is transformed into shock waves have been established through both experimental investigations and theoretical research (French [1998].

  From these examinations, it is possible to envision the impact area as surrounded by a sequence of concentric, approximately hemispherical regions of shockwaves. Each of these zones is defined by a specific range of maximum shock pressures, as depicted in Figure 2.3, and is characterized by a unique set of modifications induced by the shock-waves in the rocks. At the immediate impact location, maximum shock wave pressures may surpass 100 gigapascals (GPa), equating to 1000 kilobars (kbar), under typical cosmic encounter velocities. This exceptional pressure magnitude could result in the total liquefaction or vaporization of the impacting object and a substantial portion of the adjacent target rockFrench [1998]]. Upon penetrating approximately two diameters into the target, the initially spherical projectile undergoes substantial transformation, nearly disintegrating into a state of molten and vaporized matter. Shock waves originating from the boundary between the projectile and target rapidly decrease in peak pressure (depicted in GPa on the left side of the cavity), creating concentric, nearly hemispherical regions that exhibit distinct shock-related effects (visible on the right side of the cavity).

From the initial interface extending outward, these regions encompass the following phenomena:

- intense melting (exceeding 50 GPa) leading to the creation of a substantial molten mass
- Deformation introduced by shocks (ranging from 5 to 50 GPa)
- fracturing and the formation of brecciated material (with pressures varying from 1 to 5 GPa).

The subsequent stage involves two crucial processes: (1) the expulsion of significant near-surface fragments and smaller ejecta, indicated by upward-pointing arrows above the ground surface; (2) subsurface movement of target material, leading to the formation of a transient crater, as denoted by arrow paths intersecting isobars on the left side. As the shock waves emanate into the target rocks, there is a swift reduction in energy as they move away from the impact site. This dissipation is governed by two primary factors: (1) the shock front's expansion covering an increasingly larger hemispherical region as the radial distance increases, resulting in an overall decrease in energy density; (2) additional energy is dissipated within the target rocks through processes such as heating, deformation, and acceleration.[French [1998]].
The contact and compression stage is a brief event, lasting only a few seconds, even in the case of substantial impactors. For example, it takes just 2 seconds for a 50-kilometer-diameter projectile traveling at 25 kilometers per second (km/s), and less than 0.01 seconds for a 100-meter-diameter object traveling at the same speed.

- **Excavation stage**:In this stage the formation of the actual impact crater is orchestrated by intricate interactions involving the expansion of shock waves and the pre-existing ground surface. The shock waves permeate through the rocks, causing enduring deformation and brecciation (the process where rocks undergo exposure to intense heat, melting, and subsequent reformation). These shock waves subsequently return to their point of origin.
  The shock waves, upon traveling upward and reaching the initial ground surface, undergo a transformation, with some of them being reflected downward as rarefaction waves (release waves). Think of it as if the surface is exerting a reaction force as a result of the collision of the projectile which can be thought as the action force. In the near-surface area where the stresses within the tensional release wave surpass the mechanical strength of the target rocks, the release wave results in the fracturing and fragmentation

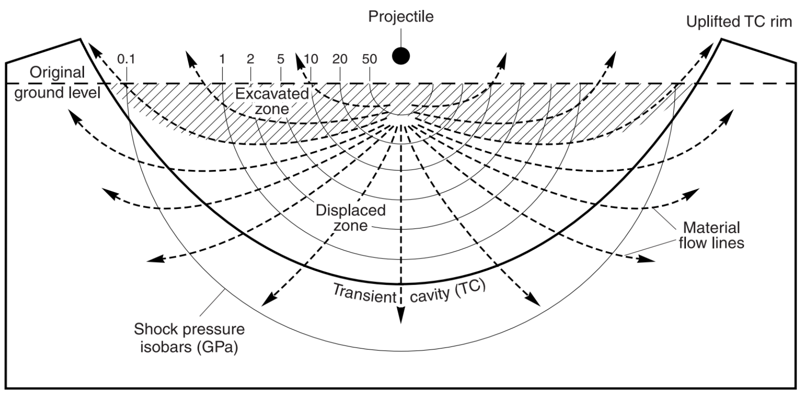



**Figure 2.4:** In simpler terms, we represent the initial pressure peak generated by shocks (measured in GigaPascals, abbreviated as GPa) near the impact site as hemispherical isobars. The complex interaction among the shock wave, the ground surface, and the ensuing rarefaction wave results in an outward excavation process (indicated by dashed arrows), culminating in the creation of the temporary crater. (Image courtesy: French [1998])

of the target rock, as illustrated in Figure 2.3. This reflective process also converts a portion of the initial shock-wave energy into kinetic energy, causing the involved rock to accelerate outward. A significant portion of this material is propelled as individual fragments moving at high speeds, leaving an empty void on the target point.
The process of excavation, although longer compared to the contact/compression stage, remains relatively brief in geological terms. For instance, if the near-surface excavation flow maintains an average velocity of at least 1 kilometer per second (km/s), a transient crater with a diameter of 200 kilometers can be excavated in less than 2 minutes. More detailed calculations, as described in Melosh (1989), page 123, indicate that the excavation of a 1-kilometer-diameter crater (e.g., Barringer Meteor Crater in Arizona) takes approximately 6 seconds, while a crater with a diameter of 200 kilometers requires only about 90 seconds.

- **Modification stage:** Shock waves at this point have lost so much energy that it no longer has any effect on the crater. This is the stage where other factors such as gravity and rock mechanics takes over and 'modifies' the crater, as the name suggests. One simple way to put this is that the modification stage ends “when things stop falling”. That being said, 'things' on craters never seize to stop 'falling' as factors such as an impact in the near vicinity also goes through the same process explained above, shock waves from the impact will surely have an effect on the crater and will lead to falling of material. Ejecta from the newly formed impact will surely fall on other craters leading to addition of material inside a crater, modifying it. Other factors such as seismology(if any) can have an effect on the crater. The modification stage, in my case, never ends.

### Factors that affect crater formation:

Planetary surfaces host a multitude of crater types, each displaying distinct shapes and dimensions. The presence of various crater types can be attributed to a range of factors:

- Gravitational pull of the planet: The more the pull of the planet, the more kinetic energy a projectile attains and the more intense will be the impact. An asteroid coming at 11km/s will have a much greater effect on the surface as compared to an asteroid hitting the lunar surface at 2km/s.

- Impact site geology: The strength and composition of the target material play a crucial role. Hard and solid materials may resist excavation, leading to shallower craters, while softer, more easily fragmented materials can result in deeper craters.

- Angle of impact: Projectiles which fall perpendicular to the surface of the planet make a very hemispherical and perfect crater, while projectiles falling on the surface with an angle generate a much shallower and disfigured

- Atmospheric interactions: The engagement of projectile with the atmosphere significantly influences the ultimate result of the projectile impact and the subsequent appearance of the crater. This interaction results in the combustion of a section of the projectile, diminishing its size. Additionally, it induces a slight deceleration of the projectile, leading to an impact with slightly reduced kinetic energy, although the observable effect of atmospheric drag on the projectile might not be particularly noticeable.

## Summary

Craters are bowl shaped planetary surface feature, achieved when large amount of energy is exerted. Impact craters are formed when an asteroid hit the planetary surface. Upon impact the surface goes through many stages to become the crater we see in the present day. Apart from kinetic energy, there are many other factors that affect the overall dimension of the crater

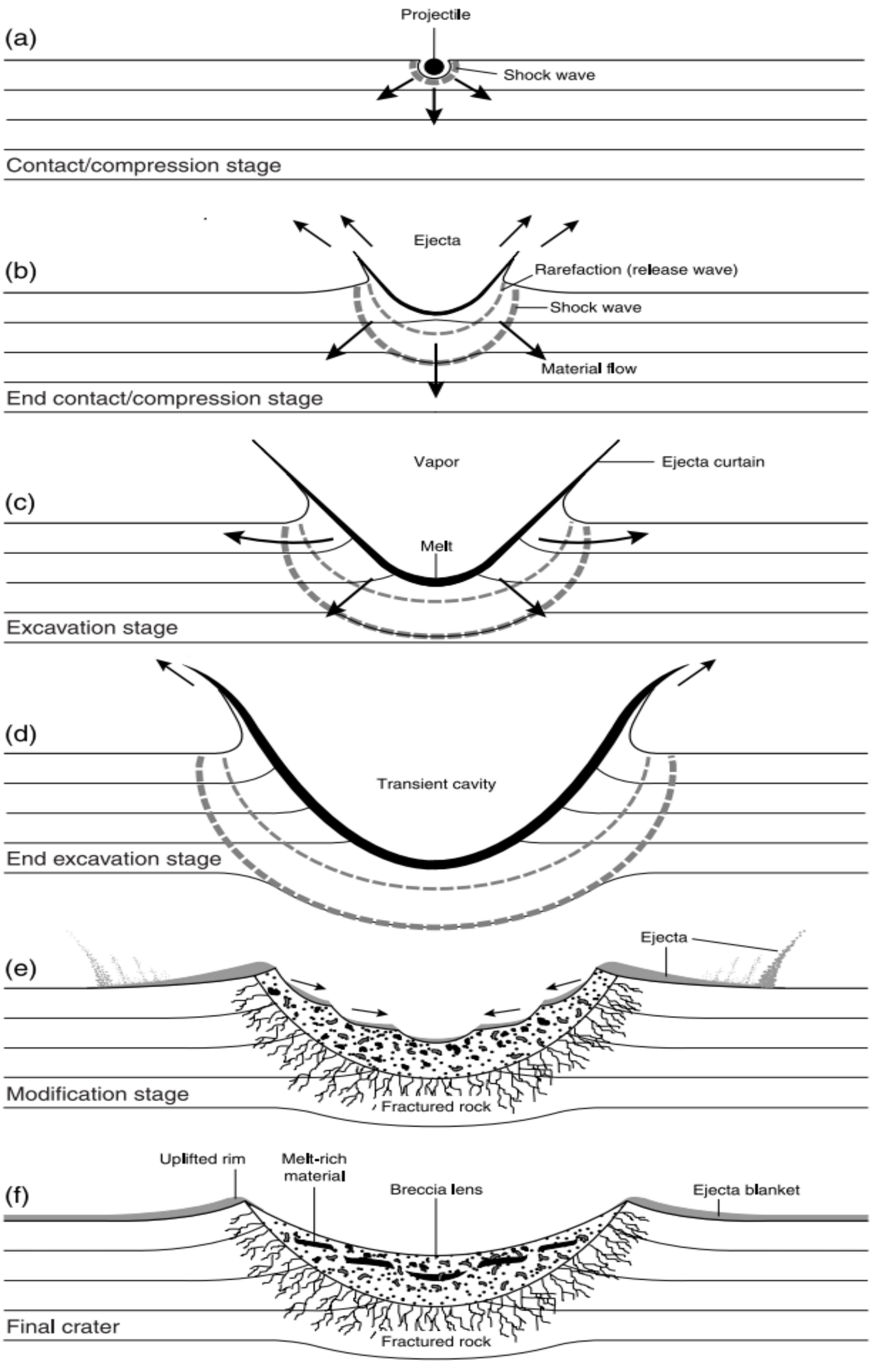



**Figure 2.5:** Crater formation stages(Image courtesy: French [1998])

# Chapter 3

# Neural networks and object detection

Undoubtedly, the influence of artificial intelligence has become pervasive, intricately threading its impact through nearly every sector and discipline, including the research domain. This widespread integration is rightfully attributed to the remarkable potency and adaptability of AI, capable of applications that stretch beyond conventional imagination. In the expansive field of artificial intelligence, neural networks assume a pivotal role, mirroring the nuanced operations of the human brain to re-define the landscape of machine learning. Inspired by the neural connections within our biological nervous system, these computational frameworks have emerged as powerful tools across diverse tasks, ranging from image recognition to natural language processing.
This chapter will delve into the fundamentals of neural networks and their origins. It will explore Convolutional Neural Networks (CNNs) and their role in computer vision applications. An overview of how a CNN operates, utilizing kernels—a mathematical function to extract image information—will be provided. Furthermore, the architecture of CNNs will be examined. Lastly, a comparative analysis will be conducted between the approach to object detection using a conventional CNN architecture and the YOLOv8 model.

Let's understand it this way - A computer, the hardware of the system, acts as a stage set for the computer model, the performer, to perform. Here, the model performs mathematical and statistical methods to simulate the real world to a virtual world.

**A computer model leverages mathematics to simulate real-world scenarios, constructing a virtual environment within the computer. This enables the computer to apply the rules and principles of the real world in a virtual setting, allowing it to effectively tackle and solve problems on our behalf.**

# Introduction

## The neurons and the networks

Curiosity arises when considering the origin of the term 'Neural Network' in the realm of computer science. Let's delve into the details.

**Neurons, the essential units of the nervous system of the human body, are specialized cells responsible for processing and transmitting information throughout the body. These cells communicate via electrical impulses, establishing complex neural networks that form the basis of our thoughts, emotions, and actions.** This complex interaction between neurons is crucial for the various functions of the nervous system, affecting things like bodily processes, thinking abilities, and emotional experiences.

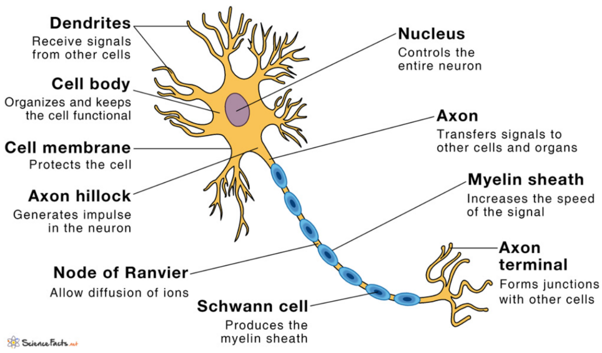



**Figure 3.1:** A Neuron

## Neural network

In the realm of computing, neural networks are the backbone of modern artificial intelligence, drawing inspiration from the intricate workings of the human brain. These networks comprise interconnected artificial neurons organized in layers, each layer contributing to the processing and transmission of information. The neurons within these networks perform complex mathematical computations, transforming input data into meaningful output predictions. Central to the functioning of neural networks is the concept of weighted connections between neurons, which modulate the flow of information and enable the network to learn from experience.

During the training phase, neural networks undergo a process of iterative optimization, adjusting the weights of these connections to minimize prediction errors. This learning process, often facilitated by algorithms like backpropagation, allows the network to adapt and improve its performance over time. Through exposure to vast amounts of labeled data, neural networks can learn to recognize patterns, classify objects, and make predictions with remarkable accuracy.

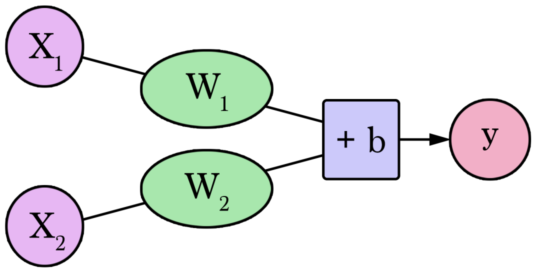



**Figure 3.2:** A simple neural network.Here, X1 and X2 are inputs (can be considered an image for example), W1 and W2 are the weight functions, b is the bias and y is the output

One of the key innovations in neural network architecture is the introduction of deep learning, which involves stacking multiple layers of neurons to create deep neural networks.

Deep learning has revolutionized the field of artificial intelligence, enabling unprecedented

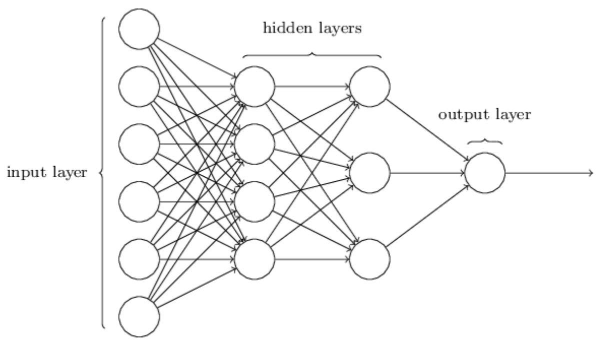



**Figure 3.3:** A

much complicated neural network having multiple inputs and getting only one output. [Image credit ]

breakthroughs in image recognition, natural language processing, and other complex tasks. Convolutional Neural Networks (CNNs), for instance, excel in image analysis by leveraging hierarchical feature extraction, while Recurrent Neural Networks (RNNs) are well-suited for sequential data processing tasks like speech recognition and language modeling.

Moreover, neural networks have found applications across diverse domains, including healthcare, finance, and autonomous systems. In healthcare, neural networks aid in medical image analysis, disease diagnosis, and drug discovery. In finance, they facilitate fraud detection, risk assessment, and algorithmic trading. Meanwhile, in autonomous systems, neural networks power self-driving cars, drones, and robotics.

As neural networks continue to evolve and become more sophisticated, they hold the promise of unlocking new frontiers in artificial intelligence, transforming industries, and reshaping the way we interact with technology.

The concept of transforming the idea of neurons into neural networks likely stemmed from the aspiration to emulate the computational capabilities of the human brain within a computer system. While a neuron consists of a cell body with branching extensions, neural networks introduce the entirety of real-world phenomena to a computer through mathematical formulations. These mathematical concepts facilitate the replication of real-world scenarios within the computer, creating a virtual counterpart of the world we inhabit.

## Image representation

## CNNs and a typical CNN architecture

### Theory

The convolution theorem stands as a cornerstone in image processing and computer vision, profoundly influencing the way signals are manipulated and processed. Its significance lies in providing a theoretical foundation for convolving two functions, whether in the spatial or frequency domain, enabling a wide array of operations crucial for tasks like image filtering, feature extraction, and object recognition.

At its essence, the convolution theorem posits that convolving two functions in the spatial domain equates to multiplying their Fourier transforms in the frequency domain. This concept underpins various computational techniques, facilitating efficient implementations of convolution operations. In image processing, convolutions are extensively used for tasks such as edge detection, noise reduction, and image enhancement, while in computer vision, they serve as the backbone for extracting meaningful features from visual data.

Within the realm of Convolutional Neural Networks (CNNs), the convolution operation emerges as a key mechanism for feature extraction from input images. Here, the convolution process involves sliding a learnable filter (often referred to as a kernel) over the input image, computing the dot product of its weights with corresponding pixel values at each position. Mathematically, the convolution operation is represented as:

$$\mathscr{F}\{f * g\} = \{\{t\} \cdot \}\{t\} \tag{3.1}$$

In this equation, f(t) represents the input image, and g(t) denotes the filter. The filter contains learnable parameters that are optimized during the training process to extract relevant features from the input data. By convolving the input image with the learned filter weights, CNNs can capture hierarchical representations of visual information, enabling them to perform tasks like object detection with remarkable accuracy and efficiency. The convolution theorem thus serves as the bedrock upon which modern image processing and computer vision algorithms are built, facilitating advancements in artificial intelligence and enabling machines to interpret and understand visual data more effectively. As previously explored, the convolution

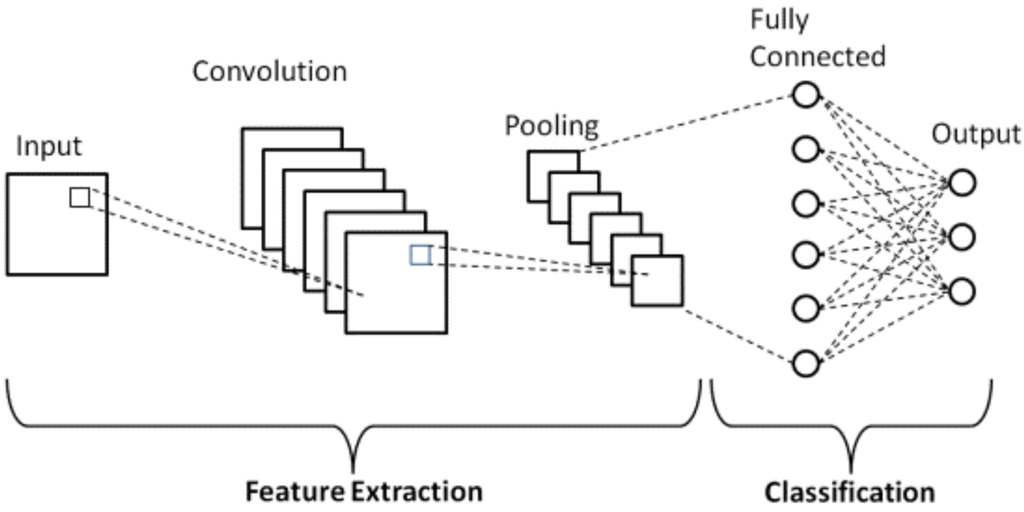



**Figure 3.4:** CNN framework

theorem addresses the process of merging two functions through a ’convolution’ operation, resulting in a single output function. The whole idea boils down to this - There are two functions that we are convolving – the image and the sliding kernel, both having a numerical value. The two functions are ‘convolved’ to get one output, which is also a numerical value. Let’s delve into the mechanics to comprehend how this concept is integrated.
CNNs consist of three distinct types of layers: convolutional layers pooling layers fully-connected layers, which together form the architecture of a CNN. The fundamental functionality of a typical CNN can be delineated into four main stages.

- **Input Image Reception**: In the initial stage, the CNN accepts the input image, interpreting it as a matrix of pixel values representing the raw visual data. This image serves as the foundation upon which subsequent layers will operate, providing the network with the essential information needed for feature extraction and pattern recognition.

- **Convolutional Layers**: As the image progresses through the convolutional layers, a set of learnable filters convolve across the input data, extracting distinctive features such as edges, textures, and shapes. Each filter functions as a feature detector, activating when it encounters specific patterns within the image. Through this process, the convolutional layers capture hierarchical representations of visual information, gradually transforming the raw pixel data into meaningful features.

- **Activation Function**: Following convolution, the feature maps undergo activation functions like the Rectified Linear Unit (ReLU). These functions introduce non-linearity into the network, allowing it to model complex relationships between features. By applying ReLU, for instance, negative pixel values are replaced with zeros, effectively enhancing the network’s ability to detect and represent intricate patterns within the data.

- **Pooling Layers**: Subsequent to activation, the feature maps are subjected to pooling layers, which serve to downsample the spatial dimensions of the data while preserving essential information. Operations like max pooling or average pooling aggregate neighboring pixel values, effectively reducing the size of the feature maps. This downsampling process enhances computational efficiency, prevents overfitting, and ensures that the network focuses on the most salient features.

- **Fully Connected Layers**: The output from the pooling layers is flattened into a one-dimensional vector and passed through fully connected layers. These layers connect every neuron in one layer to every neuron in the subsequent layer, enabling the network to learn high-level representations of the input data. By analyzing the learned features in a holistic manner, fully connected layers facilitate the extraction of complex patterns and relationships essential for the task at hand.

This sequential transformation process enables CNNs to iteratively process the original input using convolutional and downsampling operations, ultimately yielding class scores for classification and regression tasks [O'shea and Nash [2015]].

## Applications in computer vision

The applications of neural networks in computers are diverse, covering tasks such as image and speech recognition, natural language processing, and decision-making. Leveraging their parallel processing capabilities, similar to the parallel information processing in the human brain, neural networks excel at handling complex tasks concurrently.

Although computer-based neural networks emulate certain aspects of human neural function, it is crucial to acknowledge significant differences. The human brain, significantly more intricate and adaptable, possesses capacities like emotional understanding, creativity, and abstract thinking that contemporary artificial neural networks find challenging to replicate. Additionally, the biological brain operates with remarkable energy efficiency, a feat that computer systems have yet to achieve. Despite these disparities, the concept of artificial neural networks, inspired by the efficiency and adaptability of the human brain, continues to propel advancements in machine learning and artificial intelligence.

Focusing on the computer vision aspect of neural networks, object detection is one of the most important applications. Object detection aims to identify and pinpoint the location of objects within images or videos. It's like giving a computer the ability to see and understand the world around it. This complex task benefits from the power of neural networks. Neural networks, inspired by the human brain, can learn from massive datasets of labeled images. By processing these images, they progressively improve their ability to recognize objects and distinguish them from the background. This allows them to not only say ”there’s a cat in the image” but also pinpoint exactly where the cat is located. Neural networks are crucial for object detection because they can handle the immense amount of data and complex variations that exist in the real world, making object detection more accurate and robust.

## Seeing Through Silicon Eyes: Object Detection

The understanding of the concept of objects is different for humans and computers. Humans, when defining an object, can go to the mass, volume, density and the tendency of this ’object’ to occupy space aspects based on their understanding of the real, 3 dimensional world that we live in.
Computers, on the other hand, has a very different observation. With the aid of mathematics, computers construct a virtual world within their systems. The concept of using mathematics to create virtual worlds or simulations is a fundamental idea in computer science and is often discussed in various contexts. The ideology of an object for a computer bears a very different definition and not the same as humans have. One can say that, at least in image processing, an object for a computer appears when there is a difference in the pixel value in an image.

Object detection is essentially teaching computers to see and understand the objects within images and videos. Imagine giving a computer the ability to point out and label the objects it sees in a picture, just like a human can. This feat requires sifting through a massive amount of data and recognizing complex patterns. Here’s where neural networks come in. Neural networks are powerful algorithms that can learn from vast datasets of labeled images. By analyzing these images, they progressively improve their ability to identify objects, differentiate them from the background, and even pinpoint their location. Neural networks are the workhorse behind object detection because they can handle the immense data and variations that exist in the real world. This allows object detection to become more accurate and robust, with applications ranging from self-driving cars to medical image analysis. Before we understand the various ways computer can extraact information from images, we shall first dive into understanding an image and how is it represented, virtually.

## The Language of Pixels: Decoding Image Representation

Digital images serve as a cornerstone in modern communication, entertainment, and technology, encapsulating vast amounts of visual information in a compact, portable format. At the heart of every digital image lies a matrix of pixels, where each pixel represents a single point in the image. These pixels are the fundamental building blocks of digital imagery, with their collective arrangement determining the overall appearance of the image. Each pixel is assigned

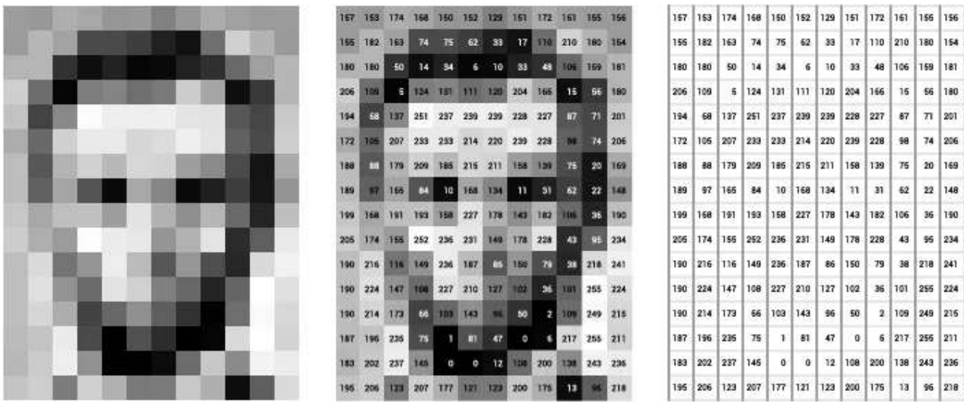


**Figure 3.5:** Pixel representation in an image. Image credits

a specific color value, typically represented in binary format, which denotes its hue, saturation, and brightness. The concept of representing images with pixels stems from the discrete nature of digital data storage and processing, where visual information is quantized into discrete units for efficient handling by computers. Pixels enable the faithful representation of images by discretizing the continuous variation of light intensity across a scene into individual points, allowing for precise encoding and manipulation of visual data.
Kernels utilize this pixel representation in image processing to leverage the localized information encoded in individual pixels. This approach allows kernels to efficiently analyze and manipulate images at the pixel level, enabling precise control over image features and characteristics. By operating directly on pixel values, kernels offer a versatile and adaptable method for tasks such as feature extraction, noise reduction, and image enhancement.

# Image processing techniques and methodology

## Sliding matrix method - concept of a kernel

**CNNs uses the concept of kernels. A kernel can be thought as a square matrix in the form of a square box having small little boxes, each having a numerical value, the values can be same or varied. This box is slided on the picture, or convolved, pixel by pixel to get an output.**

With reference to 3.6, let's understand how image processing is proceeded step by step.

- The method is that the kernel slides on top of the image, the input from the figure above, from left to right.
- As it sits on one position, overlaying on pixel values of the image, the kernel value is multiplied with the red, green and blue values of the particular pixel (The above case

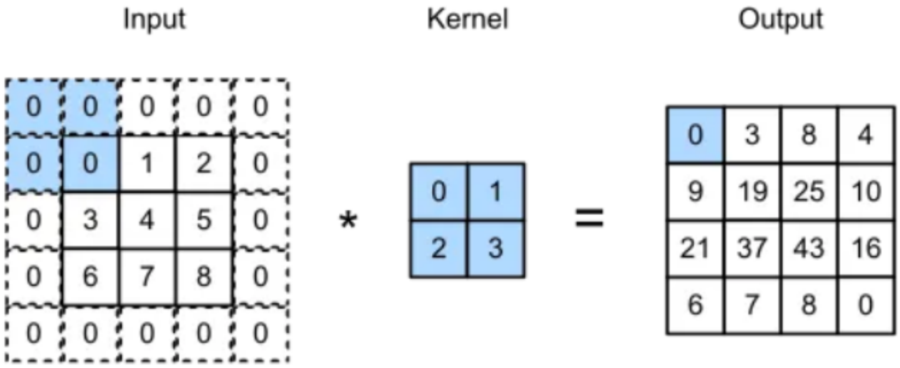



**Figure 3.6:** Visual representation of a kernel. Here, the input is an image and is convolved with a kernel. The kernel slides on top of the image. The output we get gives us edge information of the object in the image. Image credits

does not assume to take the RGB values of the image).

- When the kernel again slides to the next position, it again is multiplied with the respected RGB values and added up, the resultant values representing a pixel. This process goes on until the kernel has covered every single pixel of the image. The resultant image is a blurred image of the original one, and we will come back to why blurred images are more informative to a computer than a clear one to humans.
- Once that is done, all the 9 values of the kernal (multiplied with the RGB values of the respected pixels), they are all added up and resultant value represents one pixel in the output image.

This operation occurs 16 times.

Kernel can take any value based on the user's preference and the picture they are dealing with. In the above case, the kernel is a 2x2 matrix with 4 different boxes, each bearing different values.

However, there can also be cases where kernel values can be the same in every box, giving . This has its own drawbacks. This method collapses at the edges, when more than half of the kernel is outside the image, that is where the pixel values of those outside the image will be zero during convolution, and since all the boxes bear the same value, those boxes going to zero will surely have a dramatic effect on the overall output pixel value.

Which is why, one can consider not giving equal values to all the boxes. Instead, one can consider giving a much larger value in the central box of the kernel than the sides so when the kernel slides to the edges of the image, the effect of the corner values of the edges is not very significant. Graphically, one can imagine this kernel to be like a gaussian sliding on the image left to right.

In any case, the output is a blurred image - How does blurring help the computer understand what and where the object that we want to detect is? There are many reasons for this. The main reasons include:

- Noise reduction: Image can hold a lot of unnecessary information in the form of noise which evolve from too much lighting, instrument noise or imperfections, etc. This can be reduced and the actual object information is only retained during blurring of the image.
- Feature and edge information: The process of blurring exposes the contours and edge features of the image, making the information more straight forward and easily accessible for the machine, thereby reducing the overall complexity of data and making the computer more efficient.

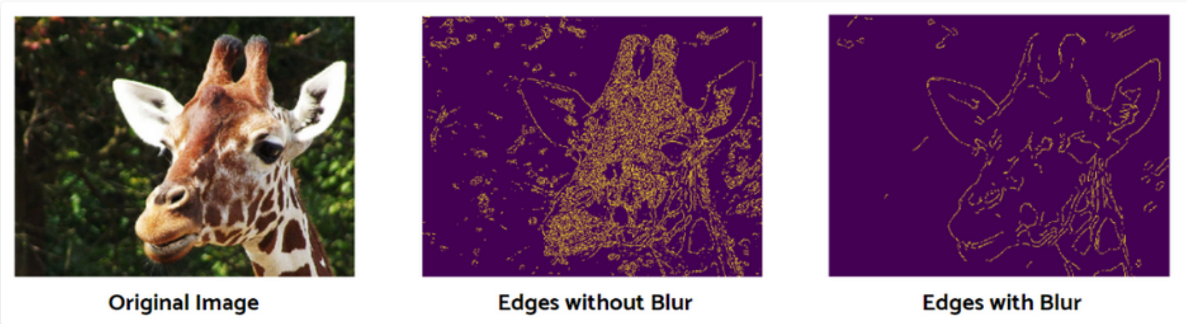



**Figure 3.7:** Edges with and without blurring

## Grid-based feature extraction

Grid-based feature extraction presents several advantages over the traditional kernel sliding technique for object detection, particularly in the context of YOLOv8. Here's a breakdown of the key benefits:

Advantages of Grid-Based Feature Extraction:

- Efficiency: Unlike kernel sliding, which entails applying the kernel filter across the entire image, grid-based extraction divides the image into a fixed grid, significantly reducing computational overhead.
- Localization: Each grid cell is responsible for a specific image area, simplifying object detection by directly associating predictions with particular regions.
- Parallelization: Grid-based techniques are highly parallelizable due to the independence of predictions within each cell, enhancing the efficiency of object detection computations.

How YOLOv8 Utilizes Grid-Based Feature Extraction:

- Image Division: YOLOv8 partitions the input image into a fixed-size grid, with each cell serving as a focal point for object detection.
- Feature Extraction Backbone: YOLOv8 employs a convolutional neural network (CNN) to extract informative features from different regions of the image.
- Anchor-Free Prediction: YOLOv8 adopts an anchor-free approach, predicting the object's center point within each grid cell instead of predefined anchor boxes.
- Bounding Box Offsets: YOLOv8 predicts offsets relative to the cell's center, reducing the number of predictions per cell compared to full bounding box prediction.
- Confidence Scores: YOLOv8 utilizes confidence scores for each predicted bounding box to assess the model's confidence in object existence and accuracy.
- Class Probabilities: YOLOv8 predicts the probability of different object classes within each grid cell, facilitating object classification.

Benefits of Grid-Based Approach in YOLOv8 includes:

- Improved Generalizability: By eliminating predefined anchor boxes, YOLOv8 adapts better to objects of various sizes and shapes.
- Faster Speed: Predicting offsets reduces the number of predictions per cell, potentially leading to faster processing during object detection.
- Efficient Information Utilization: The grid structure allows YOLOv8 to efficiently associate extracted features with specific image regions, enhancing overall performance.

In conclusion, grid-based feature extraction offers a computationally efficient and well-localized approach for object detection, with YOLOv8 effectively leveraging this technique to achieve a balance between speed, accuracy, and generalizability for various object detection tasks.

# Examples of grid-based detection technique

## You Only Look Once (YOLO)

YOLO revolutionized object detection by offering a real-time end-to-end solution. Its acronym, "You Only Look Once," encapsulates its groundbreaking approach of achieving detection in a single network pass [Terven and Cordova-Esparza [2023]]. This differed significantly from previous methods, which relied on either sliding windows followed by extensive classifier iterations per image or more complex two-step processes. Unlike Fast R-CNN [Girshick [2015]],

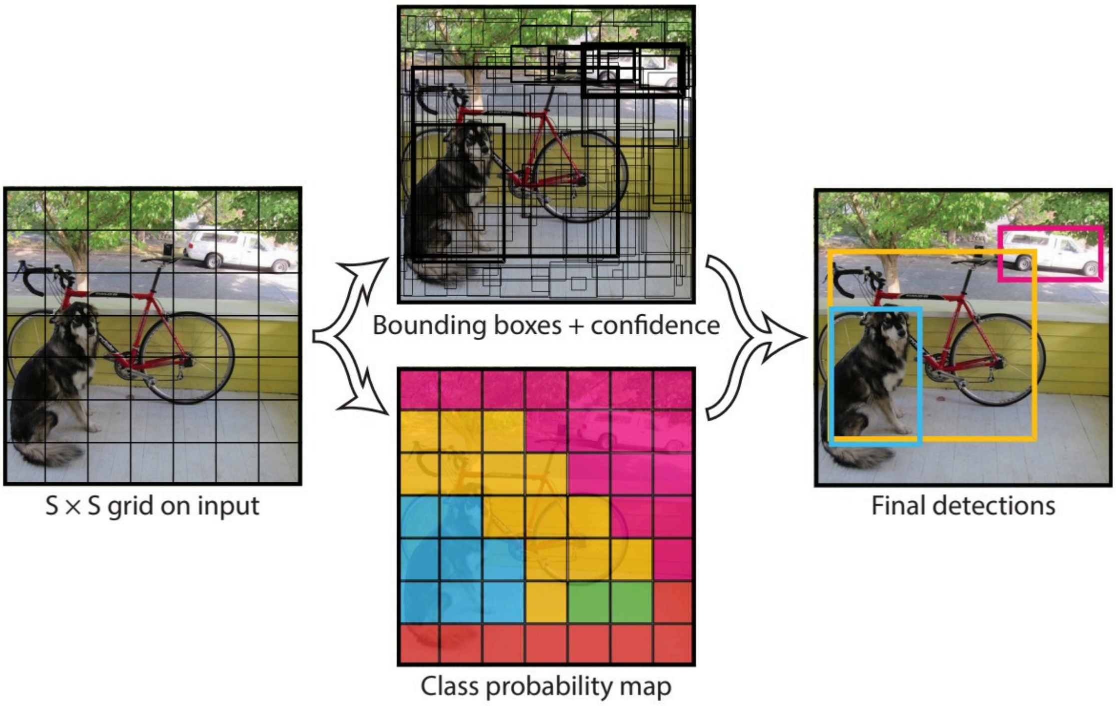



**Figure 3.8:** Grid-based detection method. An image is divided to small cubes. Every box detects a portion of the object. This is how YOLO algorithm works.

which employed separate outputs for classification probabilities and bounding box regression, YOLO utilized a simpler output structure based solely on regression for predicting detection outputs.

YOLOv8 emerges as a cutting-edge object detection model to pinpoint and localize objects within images and videos. YOLOv8's unique architecture positions it as an ideal choice for real-time applications, including self-driving cars and video surveillance systems, where the simultaneous pursuit of speed and accuracy is essential.

### Architecture

**Backbone**: Serves as the feature extractor, the backbone is tasked with extracting significant features from the input data. The operations are as follows

- Initially captures basic patterns like edges and textures in the early layers.
- Exhibits multiple scales of representation, capturing features at various levels of abstraction.
- Ultimately furnishes a comprehensive, hierarchical representation of the input.

**Neck**: Functioning as an intermediary between the backbone and the head, the neck conducts feature fusion and contextual integration operations. Essentially, the neck constructs fea-

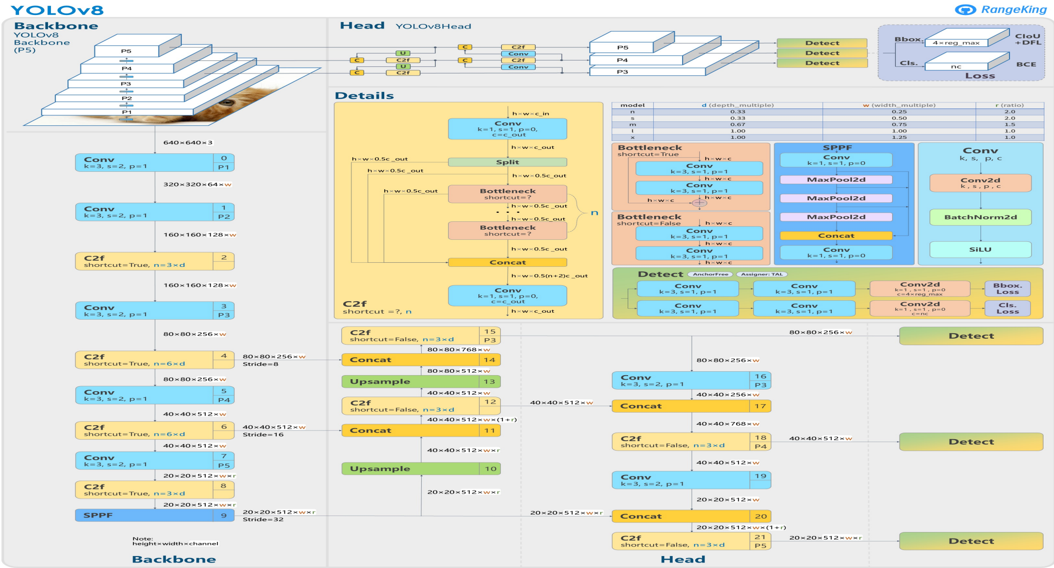



**Figure 3.9:** YOLOv8 internal architecture Image credit

ture pyramids by consolidating feature maps acquired from the backbone, thereby collecting feature maps from different stages of the backbone. The operations are as follows

- Executes concatenation or fusion of features from diverse scales, ensuring the network's capability to detect objects of varying sizes.
- Integrates contextual cues to enhance detection precision by considering the broader scene context.
- Decreases spatial resolution and dimensionality to streamline computation, thereby enhancing speed, albeit potentially compromising model quality.

**Head**: Serving as the final component of the network, the head is responsible for generating the network's outputs, including bounding boxes and confidence scores for object detection. The operations are as follows:

- Generates bounding boxes corresponding to potential objects in the image.
- Assigns confidence scores to each bounding box to signify the likelihood of object presence.
- Categorizes objects within the bounding boxes according to their respective classes.

### Why YOLOv8

- YOLOv8 exhibits robust accuracy metrics on the COCO dataset. For instance, the YOLOv8m variant, categorized as the medium model, achieves a commendable 50.2 % mean Average Precision (mAP) when evaluated on COCO. When scrutinized on the

Roboflow 100 dataset, specifically designed to assess model performance across diverse task-specific domains, YOLOv8 surpasses YOLOv5 by a significant margin. Detailed insights into this performance contrast are elucidated in our subsequent performance analysis within this article.

- Moreover, the inclusion of developer-centric features within YOLOv8 is noteworthy. Unlike other models where tasks are distributed across multiple Python files for execution, YOLOv8 streamlines the training process with a cohesive Command-Line Interface (CLI). Additionally, it offers a Python package that facilitates smoother coding experiences compared to its predecessors.

- The YOLO community's influence is substantial when contemplating model selection. Many adept practitioners in the realm of computer vision are familiar with YOLO and its operational mechanics, with ample online resources available for guidance. Despite the relative novelty of YOLOv8 at the time of this writing, numerous online tutorials and guides are already accessible to aid users in navigating its implementation effectively.

## Metrics of measuring model effeciency

The efficacy of these models is often assessed through key metrics such as **precision** and **recall**.

- **Precision**: Precision measures how accurate the model is when it predicts something as a positive, positive meaning it has detected something which is actually there. It answers the question: Of all the things the model said were positive, how many were actually positive? High precision means the model is careful and doesn't often make mistakes in labeling positives.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives + False Positives}} \tag{3.2}$$

- **Recall**: Recall gauges the model's ability to capture all the actual positives in the dataset. It answers the question: Of all the actual positive things, how many did the model find? High recall indicates the model is good at finding all the relevant instances.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives + False Negatives}} \tag{3.3}$$

Where:

- **True Positives**: True positives are instances where the model correctly identified something as positive, and it was indeed present.

- **False Positives**: False positives are instances where the model mistakenly identified something as positive, but it wasn't there.

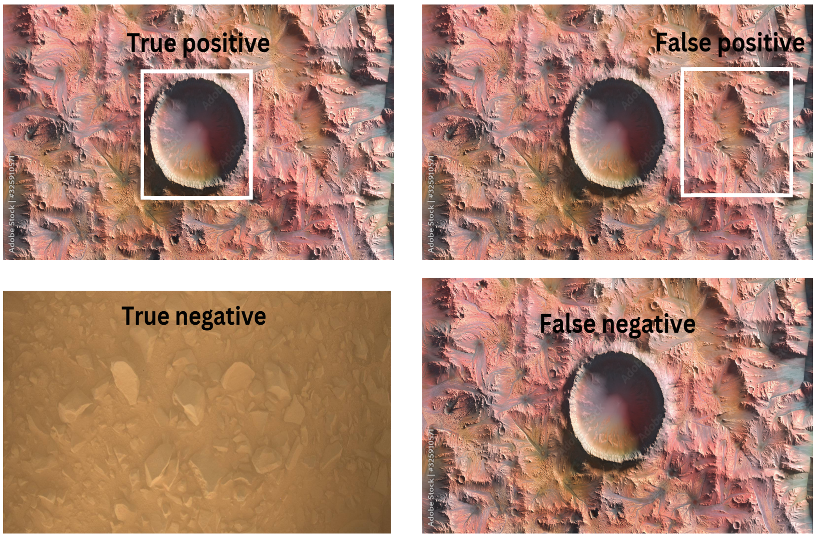



**Figure 3.10:** Visual representation of what true and false positives and negatives will look like. Crater image courtesy: Mars crater

- **True Negatives**: True negatives are instances where the model correctly identified something as negative, and it was indeed negative.
- **False Negatives**: False negatives are instances where the model mistakenly identified something as negative, but it was actually positive.

# Chapter 4

# Data

Object detection models rely heavily on data availability, with the model's effectiveness closely tied to the quality and quantity of the data provided. While the adage "more data leads to better performance" holds true in many cases, the importance of data quality cannot be overstated. In this project, both the quantity and quality of the data were meticulously considered. Establishing the concept of crater formation, along with its associated uncertainties, size variations, and shapes, was paramount. Craters can range in size from minute to immense, presenting a significant challenge for detection algorithms when dealing with such diverse dimensions.

This era is often referred to as a "data-rich" generation, thanks to advancements in technology leading to a surge in data generation and cataloging. Notably, satellite imagery has undergone substantial enhancements, with improvements in cameras across various wavelengths. Optical cameras onboard satellites now provide clearer and more detailed images, offering a distinct advantage in generating comprehensive datasets for object detection models. This chapter will delve into the significance of satellite imagery, elucidating how these advancements contribute to the enhancement of our object detection model.

## Orbiter High Resolution Camera

The Orbiter High Resolution Camera (OHRC) on Chandrayaan-2 boasts impressive specifications for detailed lunar imaging. It achieves a remarkable ground sampling distance (GSD) of 0.25 meters, meaning each pixel in the captured image corresponds to a mere quarter of a meter on the Moon's surface. While offering exceptional close-ups, it can also capture wider swaths of 3 kilometers at an orbital altitude of 100 kilometers. This makes OHRC a valuable tool for studying both fine-scale features and broader lunar terrain.

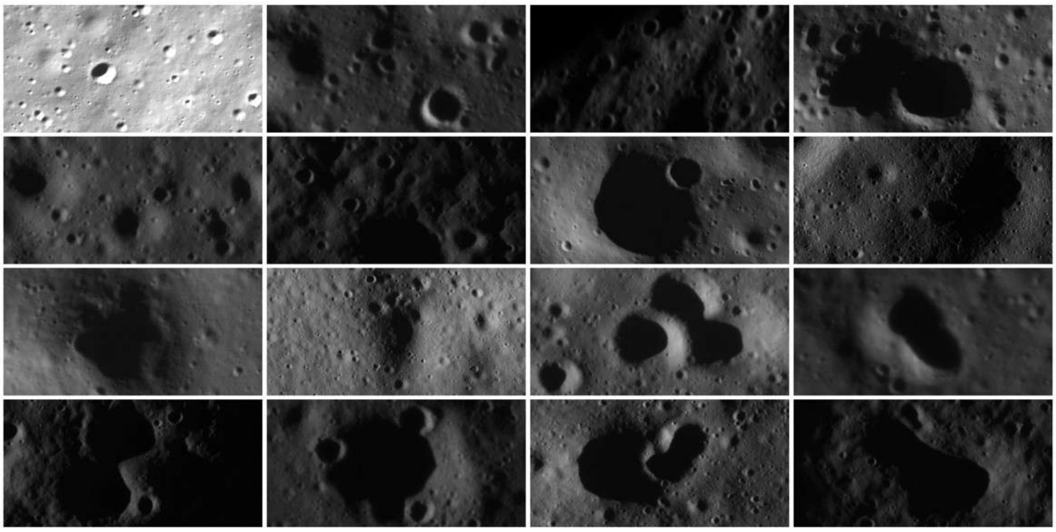


**Figure 4.1:** Some images used to generate the dataset. The data was taken from here

### Comparison between different optical cameras

| Cameras | Resolution (m/pixel) |
|---|---|
| OHRC (Chandrayaan-2) | 0.25 |
| LRO WAC (Wide Angle Camera) | 100 |
| LRO NAC (Narrow Angle Camera) | 0.5 |

- OHRC - Orbiter High Resolution Camera
- LRO - Lunar Reconnaissance Orbiter Camera (By NASA)
- WAC - Wide Angle Camera
- NAC - Narrow Angle Camera

# Process for data generation

The satellite data was downloaded from the ISDA website which holds all the data acquired from all ISRO missions. The data, when downloaded, comes in a zip files.
The data, once downloaded, comes in a zip format. Once unzipped, there are four folders - browse, data, geometry and miscellaneous. In the data folder, one will come across folders calibrated and another folder bearing the date at which the data was recorded in yyyymmdd format. Once entered, there will be two files - one a .img file and an xml file, which contains the metadata of the satellite image taken.
The .img file cannot be opened as there are not many applications available to open the file. Converting it to TIF format will help us view the image as applications like QGIS can open tif files. Pyhton is used to convert .img files to tif file.

Once that is done, the next thing is to convert the image to png or jpg, a much lighter version of image formats. If we convert the entire tif file to png, it will reduce the image clarity so much that it will be useless. Hence, the plan was to zoom in at various places in the tif file in QGIS, go to project, import/export, export map to image and save only the region being viewed in QGIS as image formats much lower than tif.

**For the project, more than 750 images were manually saved in this fashion and labelled using labelling software.**

# Data labelling

Once data was collected, manual labelling of it started. Labelimg was used. LabelImg is a user-friendly open-source tool designed for annotating images with bounding boxes, making it indispensable for tasks like object detection in machine learning projects. It is opened using python.

Green rectangular boxes were used to cover the object of interest in the image, i.e., craters. All

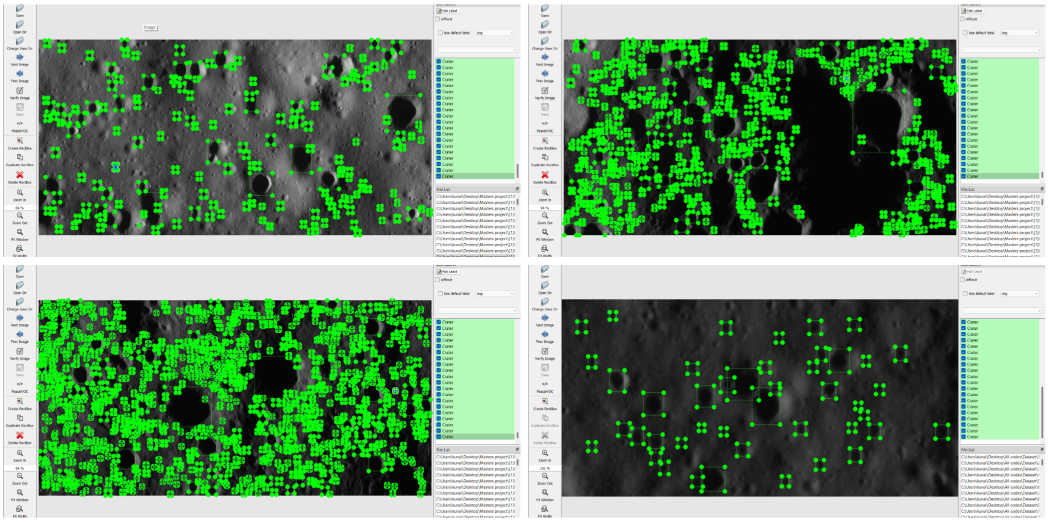


**Figure 4.2:** The proecess of labelling the images using Labelimg

the labels were carefully labelled ”Crater”. More than 750 images were labelled in the same fashion.

# Splitting of data for the model

Now that we have briefly discussed about the various models available online and the type of data that is going to be used, The data will be stratigically split into three important subsets:

- **Training**: This category of data forms the foundation of a machine learning algorithm. The data scientist supplies input data to the algorithm, each paired with an anticipated output. The model iteratively examines the data, gaining insights into its patterns and dynamics, and subsequently refines itself to fulfill its designated objectives.

- **Validation**: Throughout the training process, validation data introduces novel information to the model that it hasn't encountered previously. This data serves as an initial examination of the model's predictive abilities on unseen instances, enabling data scientists to assess its performance with fresh data. While not universally employed, validation data can offer valuable insights for optimizing hyperparameters, influencing the model's data evaluation mechanisms.

- **Testing**: Upon constructing the model, the testing data reaffirms its ability to generate accurate predictions. If training and validation data incorporate labels for monitoring model performance metrics, it is advisable for testing data to be unlabeled. The testing dataset offers a conclusive, real-world assessment of an unseen data subset, serving to validate the efficacy of the machine learning algorithm's training process.

It is always advisable to split the data in such a way that more number of images are in the training dataset. More amount of data going in training the model will familiarize the model with the targeted object much better. Especially in the case of crater detection, where complexity of crater formation is already introduced, more images will introduce the model to more and more complexity in terms of visuals.

# Chapter 5

# Results & Discussion

The chapter will showcase the results obtained so far.

## Stage 1 results

### Results from YOLOv8

The initial stage was just literature reading and familiarizing with the topic. With limited knowledge on object detection and remote sensing, a small trial detection project was done to gain some practical insights. Craters are divided into three major types - simple, complex and ring basin craters (French [1998]). I decided to make a model project which could not only detect craters, but also labels these craters based on their type. To expedite the implementation process, an existing Python codebase for YOLOv8, openly accessible online, was leveraged. To label the data, I used a website called Roboflow. Roboflow is a comprehensive platform for computer vision development that provides tools for training, deploying, and managing computer vision models. It offers a user-friendly interface, a variety of pre-trained models, and a robust set of tools for annotating and managing datasets. The input data for training, validation and testing was images taken from multiple missions, including LROC and a few images I had randomly found online. 150 images were taken and manually labeled with the help of box features in-built in the roboflow website. The labels were then saved in a project folder and the data is split. In this case, data was almost equally split while it could detect many craters, It could not detect large craters, the reason being that there weren't many images with large craters on it. Moreover, with very limited labelled and equal distribution of data, the model, although very powerful and efficient, could not detect a few small craters.

**The process took over a month. Through this experience, Few minor insights were observed:**

- Paramount importance of data was known—more data equated to greater model effi-

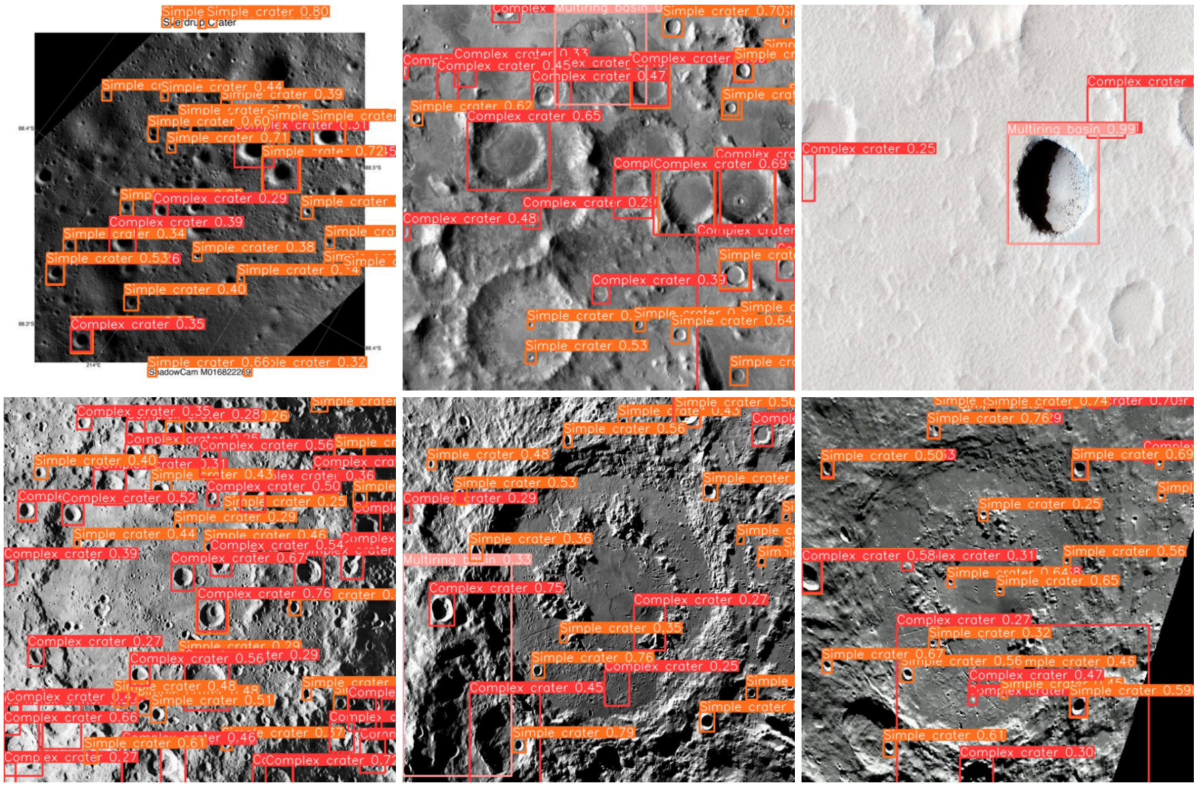


**Figure 5.1:** Results obtained by YOLOv8 with 'Simple crater', 'Complex crater' and 'Multi-ring basin crater' as labels

ciency. The true prowess of a model is intricately tied to the volume of available data, and machines exhibit an insatiable appetite for it.

- The absence of a reliable and readable dataset led to manual labeling of images, which turns out to be a very tedious task, demanding substantial patience and time. The laborious process of labeling 150 images consumed almost two weeks

Geological knowledge of a crater is very crucial. An experienced geologist can only differentiate between a simple, complex and multiring basin crater. With very limited knowledge and with no experience, the results here will not be scientifically right, and so the accuracy of the model calling it a simple, complex or multi-ring basin crater will most certainly have unnecessary uncertainties that cannot be solved. That being said, the project then focused only on the detection of craters.

A brand new dataset was then produced where there was only one label, crater. The dataset was prepared only from satellite images taken from Orbiter High Resolution Camera (OHRC) on-board Chandrayaan 2. 200 instances were taken from the data and labelled using labelimg, a labelling software [more information here]. The dataset was then loaded in roboflow and data was split in 60 % training, 30 % validation and the rest 10 % was used in testing.

The model was run for 10, 20, 50 and 100 epochs. The precision and recall of the epochs were

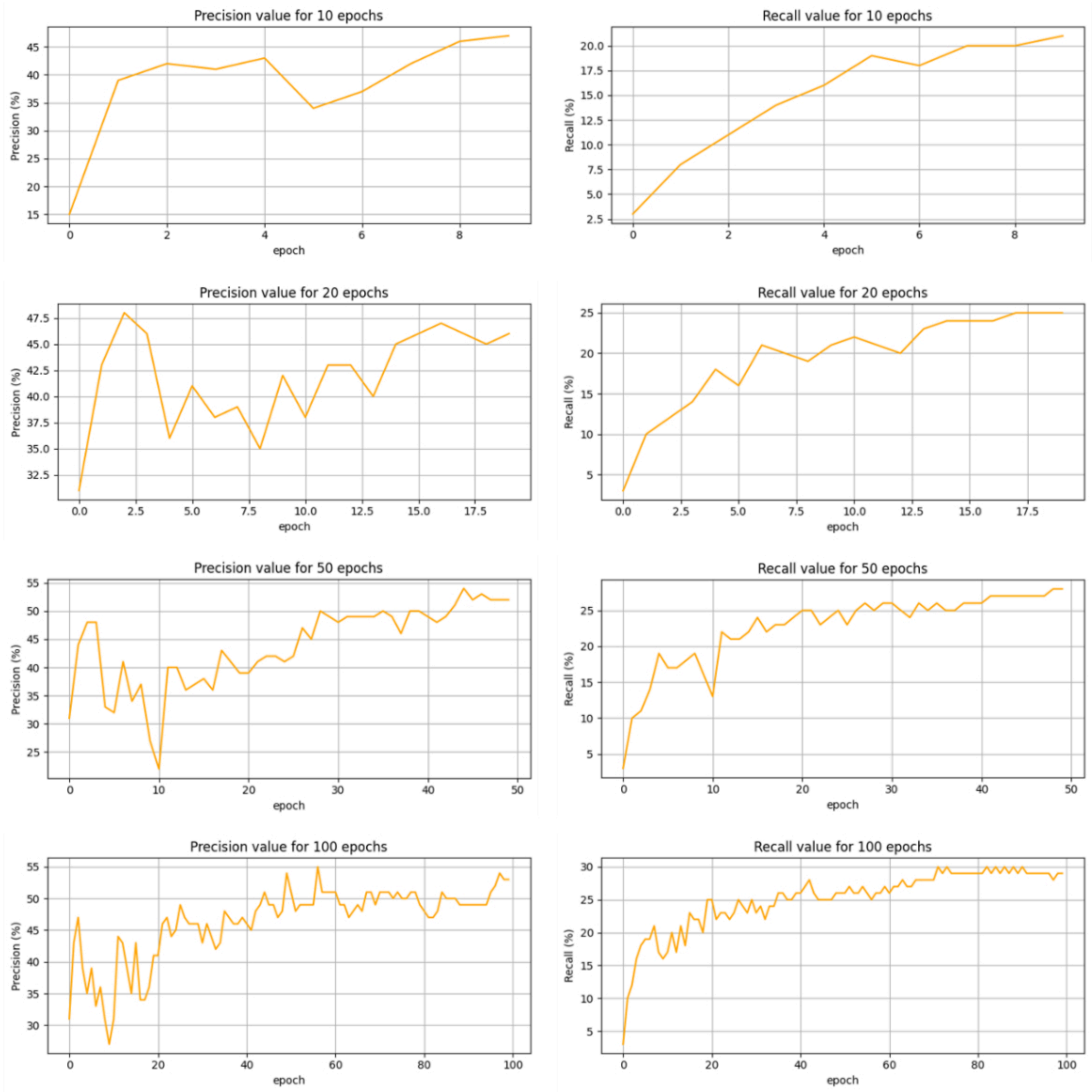



**Figure 5.2:** Results obtained when 200 images, labelled 'Crater' were used

obtained. Subsequently, visual results were also obtained.

# Stage 2 results

## YOLOv8

In stage one of the project, 200 labelled images were split to train, validate and test the images. In stage two, over 750 labelled images were used to for the same. We will see how splitting of images makes an impact on the results. Over time, as more number of images were included in the model,

The YOLOv8 model resized the images to 640x640 for training in almost all the results ob-

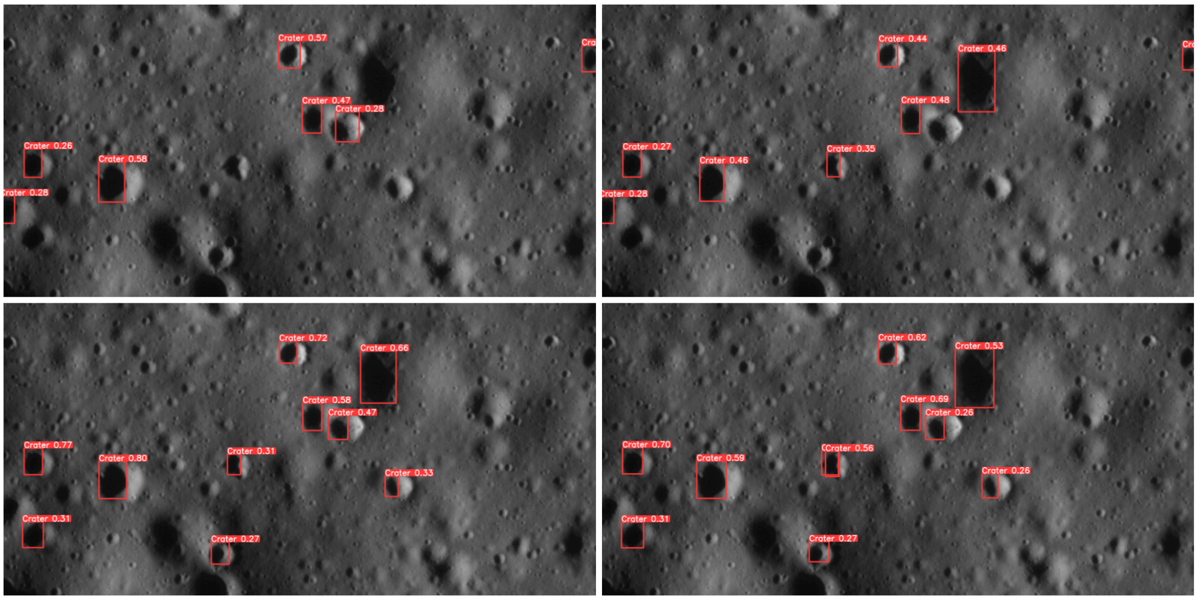



**Figure 5.3:** Results obtained from YOLOv8 after training with 200 images. Clockwise: 10, 20, 50 and 100 epochs. It can be observed that the detection of the model increases as the number of epochs increases

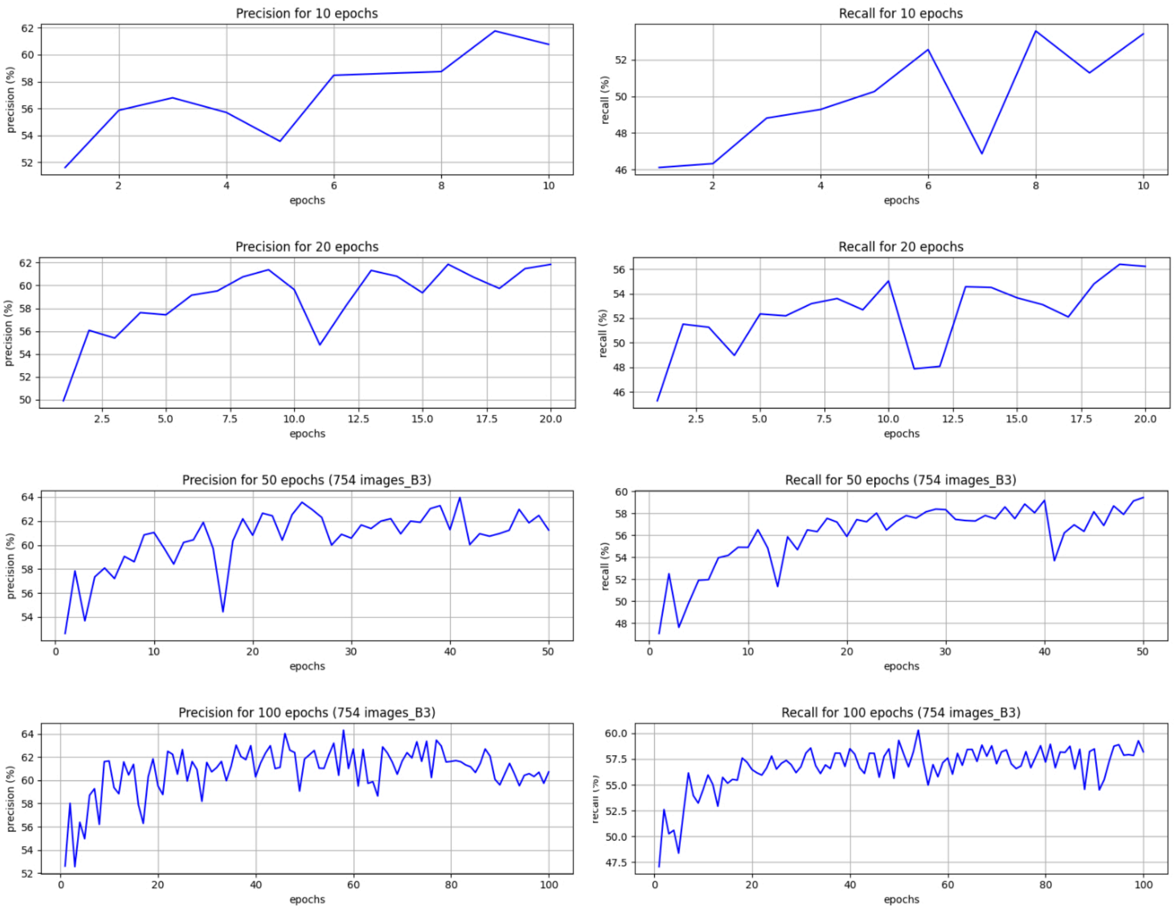



**Figure 5.4:** Top to bottom: The precision and recall observed for 10, 20, 50 and 100 epochs. The precision and recall value attains maximum value between 40-50 epochs, but seem to slightly dip when higher epochs are run.

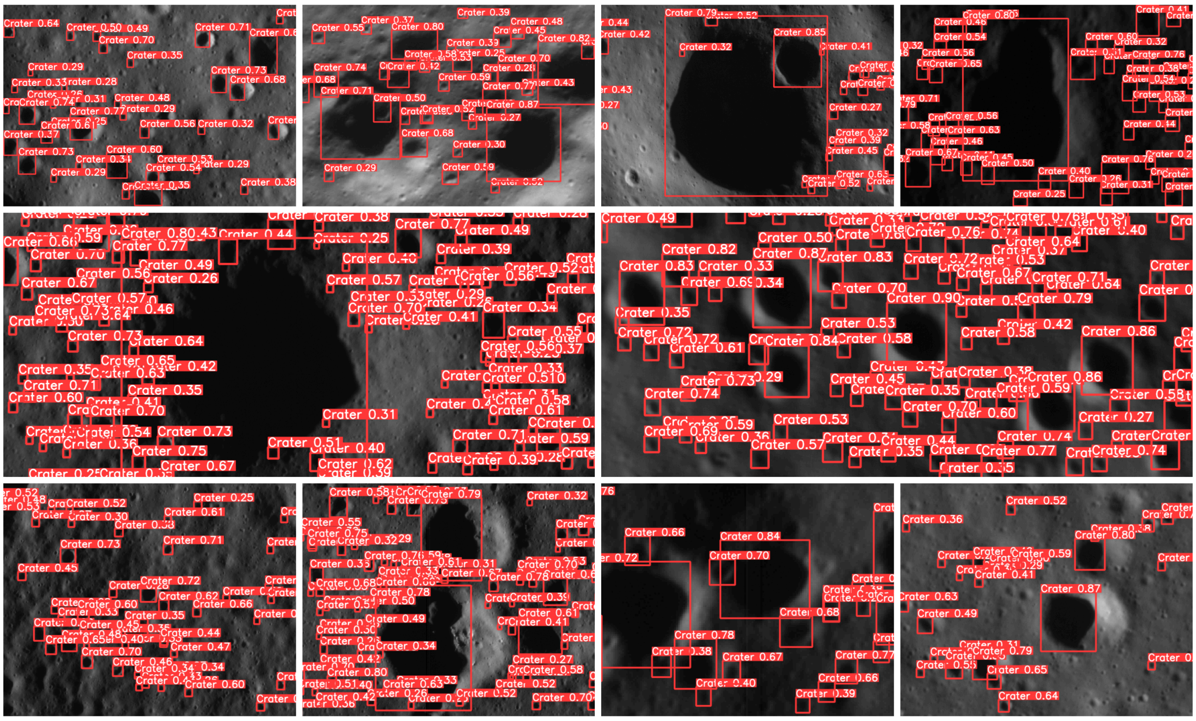


**Figure 5.5:** Results obtained from YOLOv8 model for 40 epochs. The images were resized to 720x720. With 754 images, 70% was used for training, 20% for validation and 10% for testing.

tained in stage 1. In stage 2, the same model was made to resize the image to 720x720. Significant changes and impactful results were obtained in the latter case.
It was observed that

## Discussion

Significant changes in the precision and recall observed primarily happened because of the following reasons:

- Quantity and quality of dataset: Precision and recall predominantly depends on the dataset. As precision is the measure of how precisely the model has detected a crater against the ground truth and recall is a measure of how many craters it could identify against the ground truth, both metrics will see a rise or fall based on how good the dataset used for training is. If the training data is under labelled, i.e., when we have an image and we do not seem to label a few craters but are present in the image, there is a chance that it may confuse the model and that can lead to false positives and a less efficient output, thereby leading to a lower precision and recall value.
  It was observed that, during the first stage, as the images were not labelled well, it led to incomplete and inefficient training of the model. Moreover, if the training dataset is incomplete (missing annotations for some craters), it could potentially lead to a lower recall value. This is because the model has not been exposed to certain variations or instances

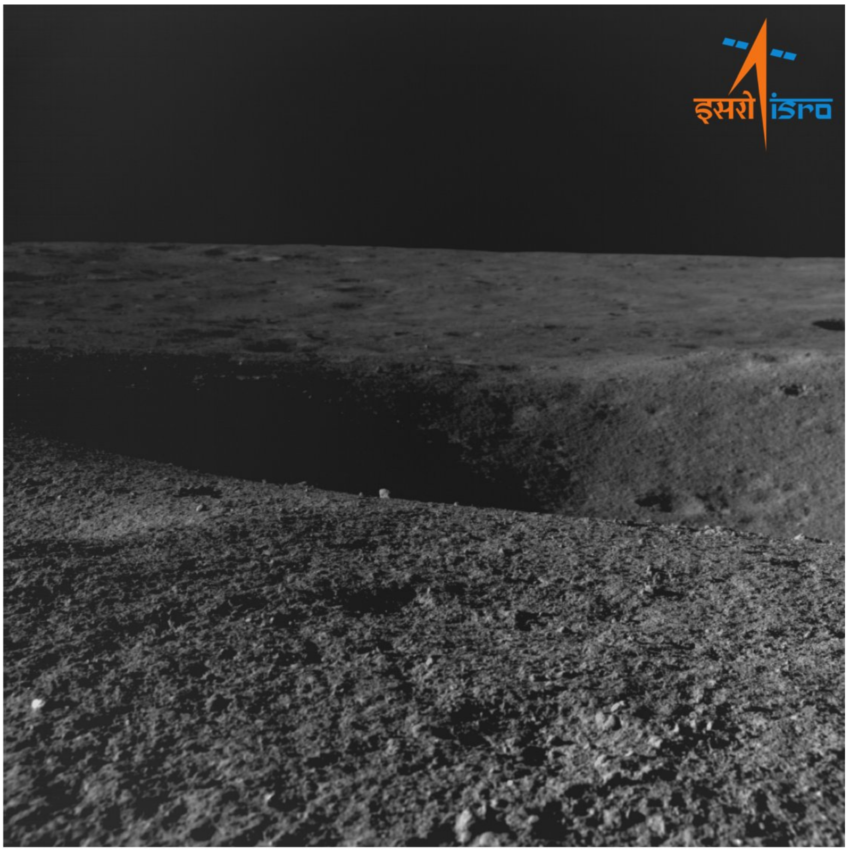
of craters during training. As a result, when the model encounters similar, unseen craters during testing or deployment, it may struggle to detect them, leading to false negatives and reducing the recall, thereby affecting the precision value as well. At the same time, less number of images will make the model less knowledgeable. During the first stage of the project, 200 images were used for training. In the second stage, more than 750 images were used for training. The image selection strategy was such that different types of craters based on its shape and design were chosen. Among the 800 images so chosen, 200 images were non-crater images in order to diversify the dataset.

- Crater complexity: Object detection models heavily depend on the targeted object. The way CNNs, or any given general neural network, works is it constitutes edges and features of the targeted object to detected them. While object detection models can detect almost anything and everything if given enough dataset for training, it can detect with very high precision and recall. Any object has its own common chronological features, regardless of its variation when it comes to design, colour, length and width, etc. For example, when we take a water bottle, the few parameters mentioned, i.e., design, height, colour may be different but the chronological description of the object remains the same – A lid at the top followed by a neck, a cylindrical body and a curve at the bottom. Almost every object which can be detected by an object detection model seem to have a chronological list of features which makes it easier for the model to understand and detect the targeted object. The same thing could not be said for craters, however. Crater can come in different sizes, shapes and even depth. The complexity and factors involved in the generation of a crater – Angle of impact, regional geology of the lunar surface, velocity of the projectile (which results in shockwave propagation inside the surface and heat generation on the surface resulting in melting or permanent deformation of rocks hit by the asteroid), post-impact processes such as quakes, etc. With these many ingredients for crater generation, one can understand the complexity and creativity nature has to offer. The variations involved results in craters of varying structures, which makes it difficult for the algorithm to distinguish between a crater and a non-crater. This problem was met by carefully choosing those images for the dataset which had different sizes, shapes and overall design of a crater to familiarize the model with the degree of complexity involved.

- Affect of illumination changes the whole dimension of the crater. This poses a great challenge when the dataset consists of optical images. Due to a change in the angle of incidence of the sun on the surface, a crater's actual size and shape will get affected. While optical images are abundant, easy to access due to its low usage of story and a good start for beginners, the limitation of data, such as surface topology, incident angle of the sun, leads to many uncertainties and a ineffective result.

# Conclusion

It is well established that determination and detection of surface feature is important for various purposes - to understand the geological evolution of the surface, understand the surrounding effects on the surface, to track a much better and safer place to land landers and rovers and to keep them from getting close to these surface features.
Landers and rovers encounter very harsh environments and are delicate. When we talk about surface features like craters, they can come in very different shapes and sizes, with varying depth. A lander or rover encountering these craters is critical. For example, the Pragyan rover of Chandrayaan 3 encountered a 4m crater.




**The crater that the Chandrayaan-3 Rover encountered on August 27, 2023, as seen by the Navigation Camera.**

**Figure 5.6:** The 4m crater encountered by Pragyaan rover

Detailed surface analysis has a potential to stop these uncertain events to occur. However, manual analysis of the surface from my experience is impractical and very time consuming. Moreover, apart from craters, terrestrial planetary surface is home to other surface features like boulders, rilles, pits, etc. Detection of each surface feature, which bears its own characteristics is not only challenging, but poses a great deal of potential of creating such a precise and en-

riching dataset that it will make mapping of the moon a lot more simpler.
Machine learning methods have proved to be a helping hand in this regard. The scientific community has understood the importance of deep learning and computer vision methods to counter image processing issues like this. Many papers and journals have been published based on this, for example, Mall and Surkov [2023],Sinha et al. [2024], Aussel and Rüsch [2024], Stoken et al. [2023].
In conclusion, Rigorous working on creating a dataset and tapping the capabilities of machine learning methods for surface feature extraction, although poses a great challenge for computer vision engineers, planetary scientists and geologist, in the bigger picture, the results might bring much more information of the surface than before.

# Bibliography


AS Arya, RP Rajasekhar, Guneshwar Thangjam, Ashwin Gujrati, Amitabh, Sunanda Trivedi, B Gopala Krishna, Ajai, and AS Kiran Kumar. Lunar surface age determination using chandrayaan-1 tmc data. *Current Science*, pages 783–788, 2012.

B Aussel and O Rüsch. Automated mapping of boulders on the lunar surface with deep learning to determine crater ages. *LPI Contributions*, 3040:1834, 2024.

AT Basilevskii. On the evolution rate of small lunar craters. In *In: Lunar Science Conference, 7th, Houston, Tex., March 15-19, 1976, Proceedings. Volume 1.(A77-34651 15-91) New York, Pergamon Press, Inc., 1976, p. 1005-1020.*, volume 7, pages 1005–1020, 1976.

William F Bottke and Marc D Norman. The late heavy bombardment. *Annual Review of Earth and Planetary Sciences*, 45:619–647, 2017.

Bevan M French. Traces of catastrophe: A handbook of shock-metamorphic effects in terrestrial meteorite impact structures. Technical report, 1998.

Ross Girshick. Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 1440–1448, 2015.

Grant Heiken, David Vaniman, and Bevan M French. *Lunar sourcebook: A user's guide to the Moon*. Number 1259. Cup Archive, 1991.

Urs Mall and Yehor Surkov. Are day-night heating cycles a trigger for launching the" stones" on tour? In *EGU General Assembly Conference Abstracts*, pages EGU–10887, 2023.

Gerhard Neukum, B König, H Fechtig, and D Storzer. Cratering in the earth-moon system-consequences for age determination by crater counting. In *In: Lunar Science Conference, 6th, Houston, Tex., March 17-21, 1975, Proceedings. Volume 3.(A78-46741 21-91) New York, Pergamon Press, Inc., 1975, p. 2597-2620. Research supported by the Bundesministerium für Forschung und Technologie and Deutsche Forschungsgemeinschaft.*, volume 6, pages 2597–2620, 1975.

Keiron O'shea and Ryan Nash. An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*, 2015.

Roberto Rugani, Francesco Martelli, Massimiliano Martino, and Giacomo Salvadori. Moon village: Main aspects and open issues in lunar habitat thermoenergetics design. a review. In *2021 IEEE International Conference on Environment and Electrical Engineering and 2021 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*, pages 1–6. IEEE, 2021.

Mimansa Sinha, Sanchita Paul, Mili Ghosh, Sachi Nandan Mohanty, and Radha Mohan Pattanayak. Automated lunar crater identification with chandrayaan-2 tmc-2 images using deep convolutional neural networks. *Scientific Reports*, 14(1):8231, 2024.

AH Stoken, AW Britton, MD Lambert, AA Turner, and MA Rubio. Automated boulder counting: Deep learning for boulder detection and height estimation. *LPI Contributions*, 2806: 2663, 2023.

Juan Terven and Diana Cordova-Esparza. A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. *arXiv preprint arXiv:2304.00501*, 2023.

Don E Wilhelms, John F McCauley, and Newell J Trask. The geologic history of the moon. Technical report, 1987.