Diamond Grading

Shamit Savant
Mumbai, India
Email: shamitsavant1919@gmail.com

***Abstract*—Since the microchip, machine learning and Artificial Learning (AI) is the next biggest innovation.Technologies such as neural networks are paving the way towards artificial intelligence breakthroughs. People are now more productive, healthier and happier than before with machine learning technology tools. Industry leaders claim that machine learning is ushering in another phase of the industrial revolution.**

Unlike the previous revolution that harnessed mechanical and physical strength, the AI-driven revolution will harness cognitive and mental ability.

I. Introduction

Diamond clarity refers to the absence of tiny, natural inclusions (imperfections) inside a diamond or on its surface. Almost all diamonds contain their own unique inclusions due to their natural formation process.

The 4Cs have long been the accepted standard for grading the rarity and value of diamonds. Yet diamond grading is an imprecise art, subject to the natural inconsistencies of the human eye when observing extremely fine differences in diamond color, and the location and appearance of inclusions, which form the basis for the clarity grade.

Diamond clarity is one of the most important criteria for assessing the quality of diamonds. It describes the absence of tiny, natural inclusions inside a diamond or on its surface. Almost all diamonds contain unique inclusions because they are natural products of carbon formed deep within the earth and exposed to tremendous heat and pressure. Although inclusions tend to be microscopic, they interfere with the path of light traveling through a diamond, resulting in less brilliance inside the diamond. Diamond clarity is the most complex criterion to assess and classify automatically. This criterion is normally graded by competent and experienced appraisers with the help of a 10X magnification placed under adequate lighting. This manual process is quite time-consuming and takes an enormous amount of concentration. Also, the diamonds are likely to be misgraded because it is difficult for all appraisers to form a consistent evaluation and to have the same level of visual acuity. Their assessments and grading results depend on their experience and a general consensus from the laboratory rather than on physical measurements.

Many kinds of impurities have already been identified by diamond laboratories such as Gemology institute of America (GIA), listed as follows:

Crystal:

Nowadays, most people learn about the quality of a diamond by referring to the GIA grading report, which represents the highest standard of reliability and consistency in the world. The report gives an assessment of diamond clarity with a plotted clarity diagram drawn by the appraiser. Significant clarity characteristics present in the diamond that influence the quality are shown in the plotted diagram, such as feather, needle and pinpoint inclusion. However, inclusions of the same type can come in thousands of shapes and sizes and have wide-ranging transparency levels. Their appearance under the same light source can also be different. Therefore, it would be more accurate to show the real appearance of inclusions rather than using symbols to substitute for the inclusions. There is a high demand in the gemological industry to replace manual diamond clarity evaluation methods with an automated system.

In this paper I have tried to make a Deep learning model which will find these inclusions and comment on the overall purity of the diamond.

II. Project Requirements

So for the above project we need prior knowledge of machine learning and Deep learning. For this our mentor suggested 2 courses on Coursera namely:

1) Neural Network and Deep Learning
2) Convolutional Neural Networks

A. *Neural Network and Deep Learning*

In this course, I learned the foundations of deep learning. I understood the major technology trends driving Deep Learning, I was able to build, train and apply fully connected deep neural networks and also know how to implement efficient (vectorized) neural networks. This course helped me understand the key parameters in a neural network's architecture and also taught how Deep Learning actually works, rather than presenting only a cursory or surface-level description. After completing it, I was able to apply deep learning to a your own applications.

B. *Convolutional Neural Networks*

In this course, I learned how to build convolutional neural networks and apply it to image data. Thanks to deep learning, computer vision is working far better, and this is enabling numerous exciting applications ranging from safe autonomous driving, to accurate face recognition, to automatic reading of radiology images. I understood how to build a convolutional neural network, including recent variations such as residual networks. I also learned how to apply convolutional networks to visual detection and recognition tasks. This course taught me how to use neural style transfer to generate art and be able to apply these algorithms to a variety of image, video, and other 2D or 3D data.

III. In-depth details of Diamonds

A. *Optical properties of the pure diamond*

Diamond contains the lowest mass element capable of forming a highly symmetrical, tightly coupled crystal lattice. This structure causes the diamond itself to have a unique set of optical properties. An in-depth understanding of how electromagnetic radiation interacts with the diamond itself and internal inclusions is critical to analyzing the physical properties of the diamond and the effect of the inclusions on light propagation. When a light source shines on a diamond, three important optical phenomena-reflection, absorption and transmission must be considered. The diamond clarity grading studied in this paper is performed to extract the inclusions seen by the naked eye. Therefore, this section focuses on the diamond optical properties in the visible light range with frequencies from 390 nm to 700 nm.

When light radiation passes from the air into the diamond, the light is reflected, transmitted and scattered at the interface between the two media because of their different index of refraction. The refractive index of diamond is very high at 2.417, and is also dispersive with a coefficient of 0.044. Fig. 2 shows an example of how a light ray passes from air into a diamond and is finally refracted out into the air again. When the light ray is incident at point 1, it is reflected and simultaneously refracts inside and travels through the diamond lattice. Then the light ray is reflected from different surfaces inside the diamond, such as points 2 and 3, because their incident angles are higher than the critical angle c. Finally, the light is refracted out when the incident angle at the diamond surface is lower than the critical angle, such as incident angle 4. Ei is the electric field amplitude of the incident light wave, which is partially reflected and transmitted at the air-diamond interface. Er and Et are the amplitudes of reflected and transmitted waves. The critical angle of the diamond (c) can be calculated according to Snell's law:

$$\theta_c = sin^{-1}\frac{nt}{ni} \tag{1}$$

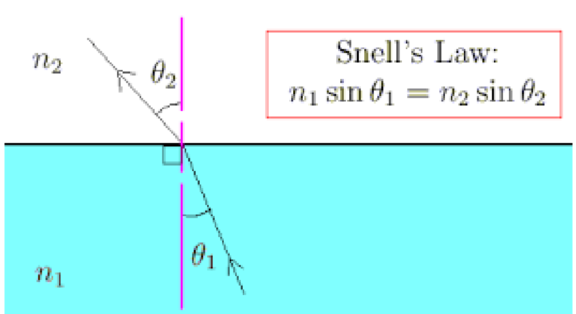

Fig. 1: Snell's law

where ni and nt are respectively the indices of refraction for diamond and air. The reflection and transmission coefficients can be derived as follows:

B. *Optical properties of diamond inclusions*

However, the pure diamond does not exist because natural diamonds are crystals formed deep within the earth under high pressure and extreme heat. During the formation process, slight irregularities and another crystal are developed within the diamond crystalline structure, which represent their natural birthmarks and are collectively known as diamond inclusions. The inclusions affect not only the appearance of the diamond, but also the light propagation inside. Inclusions can take many forms depending on their formation process. They have a large diameter range and they may be present individually or in tight clusters. In order to clearly understand the effect of inclusions on the light transmission within diamonds, diamond inclusions are divided into various types. The first type is the liquid or amorphous inclusion comprising atoms that are closely aggregated but not crystallized. Such impurities would cause a diffuse amorphous ring or irregular shape to be superimposed on the diamond pattern. Therefore, when the diamond is illuminated under dark background field, diffuse reflection occurs on the surface of diamond impurities. Therefore, the reflection of diamond inclusions is usually faceted, but shapeless. The second type of diamond inclusion is the gas-like distribution of single atoms, because air may be captured within a diamond crystal during its formation. The optical effect of such an inclusion is strong small-angle scattering. The light intensity drops in a way that depends on the atomic scattering factor of the inclusions. This optical effect can be attributed to the random distribution of Fe, Si, Ca or other elements, possibly replacing one or more carbon atoms. The atoms forming the inclusions produce an intermediate energy level, which gives rise to more energetic electrons that could jump across the gap to the conduction band and absorb the light rays due to the opaque minerals. By analyzing the diamond formation and the physical properties of the pure diamond and diamond inclusions, it can be concluded that two features can be applied to distinguish the inclusions in a diamond according to their optical properties. One is a wide discrepancy between the absorption rate of the diamond itself and that of the inclusions. All the electrons in the diamond valence band and a photon with at least 5.45 eV of energy are needed to push an electron to the conduction band, so that no photon would be absorbed by the visible light. On the other hand, the random distribution of the opaque minerals such as Fe, Si, Ca or other elements enables absorption by visible light due to their narrow band gaps. The other feature is that the texture and irregular shape of the diamond inclusions result in a difference between light transmission through the inclusions and that through the pure diamond according to Fresnel's equations.

IV. THE 4 C'S OF DIAMONDS: EVALUATING DIAMOND QUALITY

Each of the 4 C's (Cut, Color, Clarity and Carat) play a role in a diamond's beauty, though it is difficult to decipher one component by itself. As a comprehensive whole, the 4 C's interact with one another within each diamond. As a general rule, we suggest a high quality Cut above all else—as this greatly impacts a diamond's beauty and brilliance—while balancing a fine line on Color and Clarity to get the best bang for your buck. With the right Cut, both a 0.9 Carat Round Brilliant (K Color, VS2 Clarity) and a 1.5 Carat Round Brilliant (I Color, SI1 Clarity) can be stunning.

The 4C's of diamonds are as follows:

1) Cut
2) Colour
3) Clarity
4) Carat

Fig. 2: 4 C of a diamond

The quality of a diamond is determined by the 4C's: Cut, Color, Clarity and Carat. These four parameters of Diamond's Quality are the main components, but the the only, of its beauty and structure.

The GIA and the AGS are the prominent institutions with a sophisticated grading system for evaluating diamond quality. They are the most consistent entities and the ones we recommend gaining a diamond certificate from. Each of the C's are graded on a scale, and can be evaluated for quality. Though some universal terminology and standard grading exists, it does vary by lab entity.

Gradings of the 4 C's help determine the value of a diamond and indicate its quality. Diamond sellers often set their prices based on grading reports. Knowing the basics of these gradings is helpful when comparing two similar diamonds, but what remains most important is how the diamond appears to the naked eye—and how attractive the diamond is overall. In this sense, having a foundational understanding of the 4 C's is imperative as a buyer, so that you can avoid spending your budget on a component that will go unnoticed.

A. *Cut*

The 'Cut' is perhaps the most important aspect of a diamond quality that impacts a diamond's beauty. Diamond Cut specifically refers to the quality of a diamond's angles, proportions, symmetrical facets, brilliance, fire, scintillation and finishing details. These factors directly impact a diamond's ability to sparkle, along with its overall aesthetic appeal. For example take a look at this beautiful diamond and this horrific choice. The only issue differentiating the two is the cut.

The GIA grades Diamond Cut on the scale of Ideal, Excellent, Very Good, Good, Fair and Poor. The Ideal and Excellent grades, depending on Diamond Shape, signify proportions and angles cut for maximum brilliance and fire.

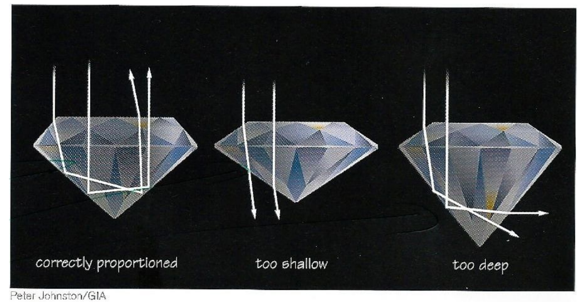

Fig. 3: Diamond Reflection

Cuts vary significantly among diamonds and diamond cutters. At times, a cutter may aim for maximum Carat weight, leaving the diamond too deep or too shallow for optimal light reflection. Other times a diamond may be cut to minimize the number of inclusions, improving its Clarity, but forgoing maximum sparkle. Even an Ideal cut diamond may have a yellow tint that is too noticeable and detracts from the gem's beauty.

More importantly, though, is ensuring Cut is a focal point of your diamond selection. Even a pristine 2 Carat Diamond with no blemishes or color tinting can be dull if it's not cut exceptionally well. Cut is the biggest indicator of beauty, and should be made priority over the other C's. As an example, this 1.50 Carat Round Brilliant is graded well for each "C" but lacks a vivacious sparkle.

It's essential to note that a top grade designation, like Excellent, doesn't necessarily indicate an outstanding diamond cut. Almost 55% of all diamonds sold online are Excellent cuts. Some are stunning, while others are mediocre. An example of an exquisite Excellent cut diamond is this 1.5 Carat Round Brilliant from James Allen.

Also, you might have heard about triple ex diamonds (excellent cut, polish and symmetry). A lot of people think that these diamonds are worth the premium some jewelers are charging but the reality is a bit different. You can read more in our Triple Excellent article that explains it in more depth.

Because Cut is so important to a diamond's fundamental beauty, it's crucial to review a diamond's Cut carefully and ask for the eye of an expert.

B. *Colour*

Diamond Color is graded in terms of how white or colorless a diamond is. The GIA grades diamonds from D to Z, with D being the most colorless, and Z containing noticeable brown or yellow tint. The pricing of diamonds usually reflects these grades—sometimes significantly. In most cases, the naked eye cannot tell the difference between two adjacent color graded diamonds, though the price difference may be significant. The most critical aspect with Color is to determine if it appears colorless in relation to its setting. You also want to be certain that a diamond is clear of any tinting that takes away or interferes with white and colored light reflections. The K Color in this Cushion Cut Diamond, for example, distracts from the sparkle of the diamond while this I Color Diamond is radiant. Brilliance, or sparkle, is created from the way the diamond is cut. It is not advantageous to purchase a diamond that distracts from this important principal characteristic. As a general recommendation, review each diamond closely and ask for the assistance of an expert. This is the best way to ensure you're not paying for a feature (i.e. too high of Color grade) that will go unnoticed, or purchasing a diamond that distracts or interferes with light reflection. Note: Certain colored diamonds are valued stones, like a fancy pink or green diamond. Color grades for these diamonds are distinctly different than traditional "white" diamonds.

Fig. 4: Diamond Colour

C. *Clarity*

A Diamond's Clarity grade evaluates how clean a diamond is from both inclusions and blemishes. Clarity is graded by the GIA on the following scale:

- FL (Flawless)
- IF (Internally Flawless)
- VVS1 (Very, Very Slightly Included 1)
- VVS2 (Very, Very Slightly Included 2)
- VS1 (Very Slightly Included 1)
- VS2 (Very Slightly Included 2)
- SI1 (Slightly Included 1)
- SI2 (Slightly Included 2)
- I1 (Inclusions 1)
- I2 (Inclusions 2)

To understand diamond clarity, we must first understand how diamonds are created. Natural diamonds are the result of carbon exposed to tremendous heat and pressure deep in the earth. This process can result in a variety of internal characteristics called 'inclusions' and external characteristics called 'blemishes.'

Evaluating diamond clarity involves determining the number, size, relief, nature, and position of these characteristics, as well as how these affect the overall appearance of the stone. If you are trying to determine what is the best clarity for a diamond, remember that no diamond is perfectly pure. But the closer it comes to purity, the better its clarity.

The GIA Diamond Clarity Scale has 6 categories, some of which are divided, for a total of 11 specific grades.

- Flawless (FL) No inclusions and no blemishes visible under 10x magnification
- Internally Flawless (IF) No inclusions visible under 10x magnification
- Very, Very Slightly Included (VVS1 and VVS2) Inclusions so slight they are difficult for a
- skilled grader to see under 10x magnification
- Very Slightly Included (VS1 and VS2) Inclusions are observed with effort under 10x magnification, but can be characterized as minor
- Slightly Included (SI1 and SI2) Inclusions are noticeable under 10x magnification
- Included (I1, I2, and I3) Inclusions are obvious under 10x magnification which may affect transparency and brilliance

Many inclusions and blemishes are too tiny to be seen by anyone other than a trained diamond grader. To the naked eye, a VS1 and an SI2 diamond may look exactly the same, but these diamonds are quite different in terms of overall quality. This is why expert and accurate assessment of diamond clarity is extremely important. Knowing what diamond clarity truly means helps you understand the factors that contribute to diamond quality and price.

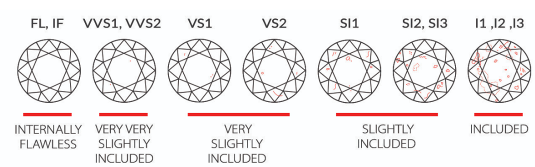

Fig. 5: Diamond Clarity

D. *Clarity*

Often when people hear the term "Carat Weight," they think it refers to the size of the diamond. In actuality, Carat refers to the weight of the diamond, not how large the stone is. A 1 Carat Diamond equals 200 milligrams, or 0.2 grams—and weighs about the same as a quarter of a raisin. Depending on the Diamond's Shape and how it is cut, two 1 Carat Diamonds might be quite different in size. While Carat weight is an element to consider when buying a diamond, the overall appearance and brilliance should carry more importance. For example, a mediocre 1.5 Carat diamond will not shine as brightly—or draw as much attention—as a stunning 1.0 Carat diamond, no matter how much more it weighs. By definition:

Carat is the unit of measurement for the physical weight of diamonds. One carat equals 0.200 grams or 1/5 gram and is subdivided into 100 points. For comparison, in units more familiar in the United States, one carat equals 0.007 ounce avoirdupois.

Carat weight is the most objective of the diamond's 4Cs. All that is required is a precisely balanced scale capable of measuring extremely small weights. In the AGS Laboratories, carat weight is measured using a highly accurate, and calibrated digital scale. Here are some facts about a diamond's weight and price that are important to understand before purchasing. Comparing the value of diamonds by carat weight is like comparing the value of paintings by size. A wall-sized canvas by an unskilled artist may be bigger than a miniature by Rembrandt, but it will not be worth more. Large diamonds are rarer than smaller ones, and as the carat weight increases, the value of the diamond increases as well. However, the increase in value is not proportionate to the size increase.

For example, a 1-carat diamond will cost more than twice that of a ½-carat diamond (assuming Color, Clarity and Cut grade are the same). Weight does not always enhance the value of a diamond, either. Two diamonds of equal weight may be unequal in value, depending upon other determining factors such as Cut, Color and Clarity. In fact, if a diamond is improperly cut, the added weight may serve only to reduce its brilliance.

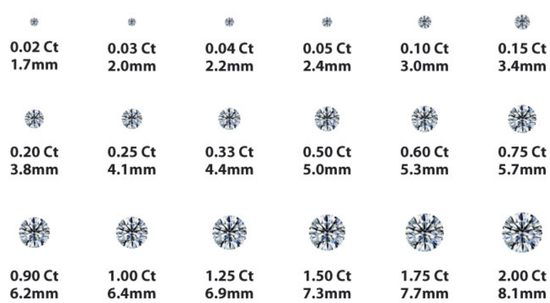

Fig. 6: Diamond Carat weight

V. Project Explaination

So the project can be divided into three major parts:

1) Annotating the data: Both the images and the videos
2) Preparing the model
3) Training the model to finally test it on validation set

A. *Data Annotation*

Data annotation is the process of adding metadata to a dataset. This metadata usually takes the form of tags, which can be added to any type of data, including text, images, and video. Adding comprehensive and consistent tags is a key part of developing a training dataset for machine learning.

In simple words, Annotation in machine learning is the process of labeling data, which could be in the form of text, images, audio, etc. Typically, annotation is done by a human.

Hand Image annotation is marking of images buy humans readable for machines. It is the task that extracts information as classes from the image.

There are plenty of ways to annotate images for computer-vision projects. At a high level, you can simply bucket images into classes, draw tight bounding boxes around objects in images, dot the corners of important entities, or label every individual pixel in a given image.

Millions of annotated photos help an AI identify patterns that help it understand, say, what a human looks like. Eventually the AI grows smart enough to draw boxes around pedestrians on its own.

Data annotation is an indispensable stage of data preprocessing because supervised machine learning models learn to recognize recurring patterns in annotated data. After an algorithm has processed enough annotated data, it can start to recognize the same patterns when presented with new, unannotated data. As a result, data scientists need to use clean, annotated data to train machine learning models.

There are various Data Annotation Types namely:

- **Semantic Annotation** is the task of annotating various concepts within text, such as people, objects, or company names. Machine learning models use semantically annotated data to learn how to categorize new concepts in new texts. This can help to improve search relevance and train chatbots.
- **Image and Video Annotation** comes in a variety of forms, from bounding boxes, which are imaginary boxes drawn on images, to semantic segmentation, where every pixel in an image is assigned a meaning. This label usually helps a machine learning model to recognize the annotated area as a distinct type of object. This type of data often serves as a ground truth for image recognition models that can recognize and block sensitive content, guide autonomous vehicles, or perform facial recognition tasks.

  Similar to image annotation, video annotation often involves adding bounding boxes, polygons, or keypoints to content. This can be done on a frame-by-frame basis, with these frames then stitched together to help track the movement of the annotated object, or in the video itself using a video annotation tool. This type of data also plays an important role in the development of computer vision models for tasks like object tracking and localization.
- **Text Categorization** and content categorization refer to the task of assigning predefined categories to documents. For example, you can tag sentences or paragraphs within a document by topic, or organizing news articles by subject such as domestic, international, sports, or entertainment.
- **Entity Annotation** is the process of labeling unstructured sentences with information so that a machine can read them.

  Within entity annotation, there are a multitude of processes that can be layered to create an understanding of language. An exhaustive list would be too long to reproduce here, but these examples should give a sense of the broad possibilities on offer:

Named entity recognition: Named entity recognition (NER) refers to the classification of named entities present in a body of text. These entities are labeled based on predefined categories such as person, organization, and place. Named entity recognition models add semantic knowledge to your content, making it easy for individuals and systems to quickly identify and understand the subject of any given text.

Entity linking: This is the process of annotating the relationship between two parts of a text. For example, you can tag the company and employee, or person and their hometown as related concepts.

- **Intent Extraction** For chatbots, it's important for the algorithm to accurately determine the user's intent when they type in a query. For example, consider the following queries for a chatbot on a restaurant website:

  I agree to pay the cancellation fee and cancel the reservation.

  How much is the cancellation fee?

  Do you charge a cancellation fee for no-shows?

  All three examples contain the phrase 'cancellation fee', but all have different intents. In the first sentence, the intent is for the chatbot to take an action: cancel the reservation. The second and third sentences share a different intent: to receive more information about the restaurant's cancellation fee policy. If the chatbot can't recognize this, it might cancel the user's restaurant reservation by mistake.

  Intent extraction is the technical solution to the above problem. For intent extraction, we explicitly label user intents in the data on a phrase or sentence level. This way, the algorithm has a library of ways that people phrase certain requests, and the algorithm can begin to extrapolate about new sentences based on that ground truth.

- **Phrase Chunking** consists of tagging parts of speech with their linguistic or grammatical meaning. For example, some machine learning training datasets would require every word to be annotated with its part of speech, such as 'noun' or 'verb'.

  The above example of phrase chunking was created in Brat, the popular annotation tool for natural language processing.

In this project I used the Image and Video Annotation. I started with diamond images of all sorts of inclusions in them and started to annotate them one by one.

This was one of the most time consuming part of the project and tool almost half of the entire project time. The reason for this much time was the manual working on the diamonds.Some of the diamonds were easy to annotate (like the one below):

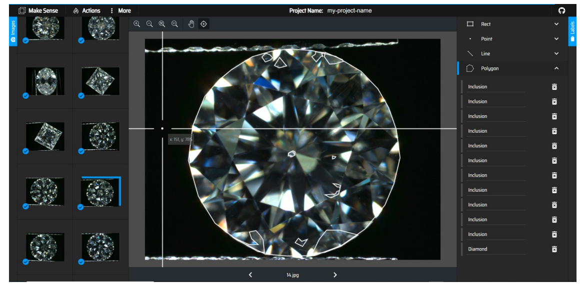

Fig. 7: Diamond Annotation

And some of the diamonds were very easy to annotate (like the one shown below):

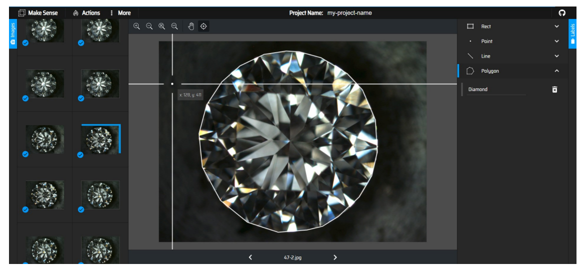

Fig. 8: Diamond Annotation

But most of the diamonds were difficult to annotate (like the one shown below):

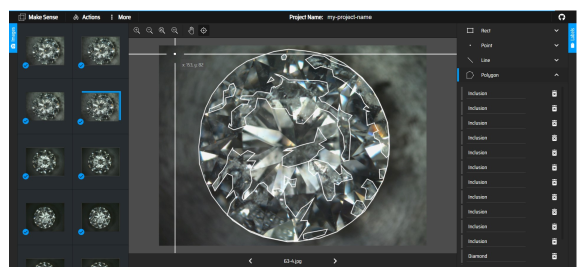

Fig. 9: Diamond Annotation

After manually doing the annotations of the images, I had to now switch it to the annotations of the videos. This needed a bit of coding, I used a github repository and made a few modifications so that the entire process of splitting the video into individual frames and then stacking them up and converting them into a single clear diamond image would be done with a single running of code.

The code was made in three files wherein two files were used as libraries in the final output python file.The three files are namely:

- split.py
- FocusStack.py
- main.py

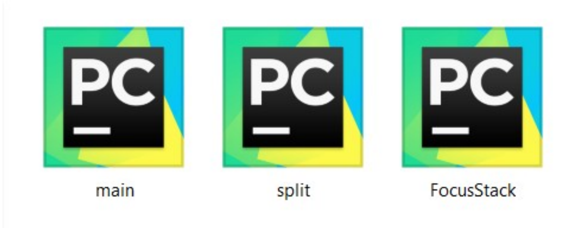

Fig. 10: Files

As the name suggests the code split.py splits the video into individual frames using opencv library. The code in split.py is shown below:

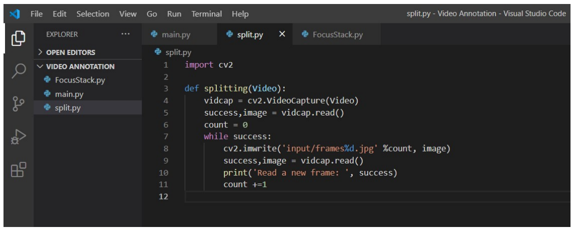

Fig. 11: split.py

Again as the name suggests the code FocusStack.py stacks the individual frames generated by split,py and stacks them together to form a clear diamond image, which can later be used in the dataset for either training or testing. The code of FocusStack is a bit lengthy thus I am providing the Github repository link in the references

Finally we have the main.py which combines both the split.py and FocusStack.py. It takes in the name of the video as input and outputs the clear diamond image after splitting and stacking the video. The code of main.py is shown below:

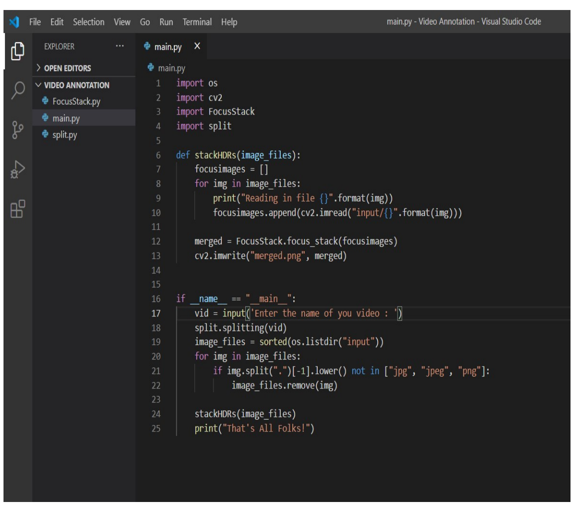

Fig. 12: main.py

B. *Preparing of model*

For this project we could have used various models like Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), YOLO etc. But we decided to go with Masked-RCNN.

Mask RCNN is a deep neural network aimed to solve instance segmentation problem in machine learning or computer vision. In other words, it can separate different objects in a image or a video. You give it a image, it gives you the object bounding boxes, classes and masks.

There are two stages of Mask RCNN. First, it generates proposals about the regions where there might be an object based on the input image. Second, it predicts the class of the object, refines the bounding box and generates a mask in pixel level of the object based on the first stage proposal. Both stages are connected to the backbone structure.

Backbone is a FPN style deep neural network. It consists of a bottom-up pathway , a top-bottom pathway and lateral connections. Bottom-up pathway can be any ConvNet, usually ResNet or VGG, which extracts features from raw images. Top-bottom pathway generates feature pyramid map which is similar in size to bottom-up pathway. Lateral connections are convolution and adding operations between two corresponding levels of the two pathways. FPN outperforms other single ConvNets mainly for the reason that it maintains strong semantically features at various resolution scales.

For the training and running of this model I decided to use google colab.

Colaboratory, or “Colab” for short, is a product from Google Research. Colab allows anybody to write and execute arbitrary python code through the browser, and is especially well suited to machine learning, data analysis and education. More technically, Colab is a hosted Jupyter notebook service that requires no setup to use, while providing free access to computing resources including GPUs.

The model used is the official Masked-RCNN model from the github repository of matterport. The link of this will be given in the references.

The repository is an implementation of Mask R-CNN on Python 3, Keras, and TensorFlow. The model generates bounding boxes and segmentation masks for each instance of an object in the image. It's based on Feature Pyramid Network (FPN) and a ResNet101 backbone.

The repository includes:

Source code of Mask R-CNN built on FPN and ResNet101.

- Training code for MS COCO
- Pre-trained weights for MS COCO
- Jupyter notebooks to visualize the detection pipeline at every step
- ParallelModel class for multi-GPU training
- Evaluation on MS COCO metrics (AP)
- Example of training on your own dataset

The code is documented and designed to be easy to extend.

Now for the start of model training I first had to clone the Github repository in my drive. Once the repository was uploaded the next task was to set up the Google Colab.

For setting up of google colab I had to first mount the drive so that I could access the cloned Github repository. We do this using an option provided in google colab or simply putting the below code in one of the google colab cells.

```
from google.colab import drive
drive.mount('/content/gdrive')
```

Once the Drive was mounted we had to run the model in order for it to train on the annotated dataset that we had prepared. Before the running of model there were a lot of modifications in the original code, in order for it to work with our custom dataset of diamonds.

After that I divided the dataset into two parts the first being train and the second being the val (validation). The code for training the dataset of diamond grading is present in the folder named custom and the file name is custom.py. The custom folder also has another file named predict.py which will be used to mark the inclusions and the diamond in an given image. We also need to keep the .json files along with the images to train the dataset.

Once the entire code was done and ready to be executed and environment was also set it was time to run the code and train the dataset. This takes a lot of time about 45min to an hour and also takes about 2.5GB space of your drive. The last thing to do before running the model for training is to import and install a few packages.

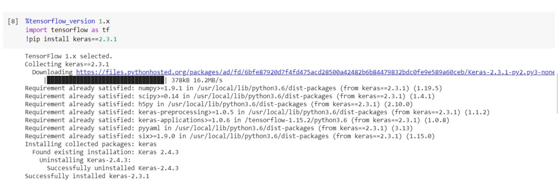

Fig. 13: Tensorflow and Keras installation

The following steps are for the training of the dataset. If you have a GPU machine then it's great otherwise use the Google Colab for training.

Now that everything is ready for the training. Navigate to the Mask RCNN/samples/custom, where we have custom.py.

The default epochs=10, and you can change it in the model.train inside the train function. Run the following command.

python custom.py train –dataset=../../dataset –weights=coco

The above command first downloads the weights of the pre-trained model which was trained on the COCO dataset and then start the training.

You can stop the training at any time and run prediction to see how your model is performing. To resume the training from the last checkpoint we can use the following command.

python custom.py train –dataset=../../dataset –weights=last

This training will keep saving the weights for every checkpoint inside Mask RCNN/logs directory.

The running of the code is shown below. As the epoch is set to 10 it runs the cycle of training 10 times.

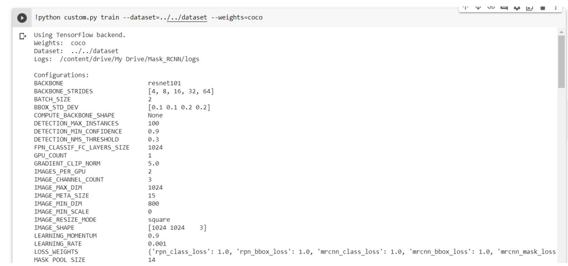

Fig. 14: Training of model

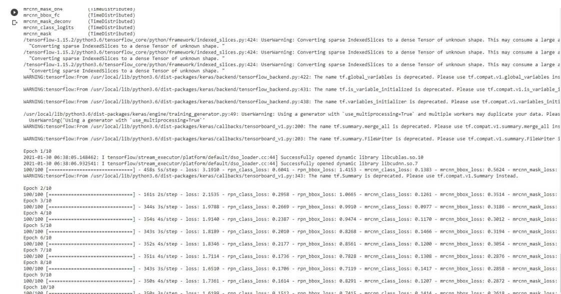

Fig. 15: epoch cycles

Once all the training is done the weights are stored in the logs folder. This helps us as we don't have to run the training model code again and again. The predict.py can be run in two ways.

- The first way is to make inference on the validation set, which picks up an image randomly from validation and run the detection.
- The second way is to run the inference on any arbitrary image.

The predict.py runs as shown below (This is run on an arbitrary image).

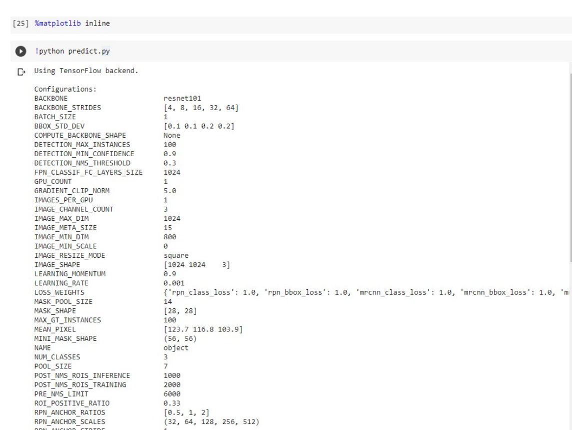

Fig. 16: predict.py

Once we run this code we get the final desired output. that is the image which is annotated by the model itself. Here is the example of how it looks

Fig. 17: Original Image

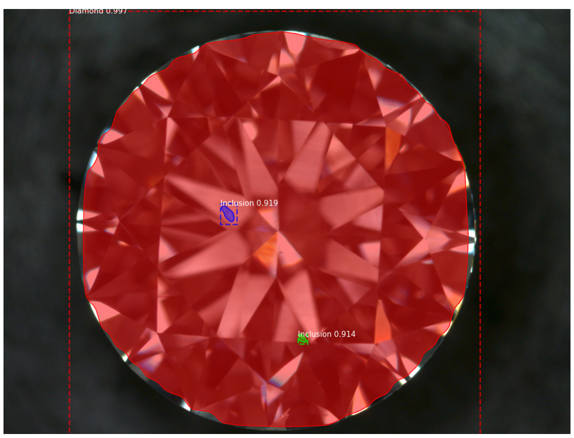

Fig. 18: Model Output

VI. CONCLUSION

This was our main task in this internship program, that is to prepare and train a model which will be able to detect inclusions in a diamond image. The model is trained on a very small dataset (less than 200 images), therefore we cannot expect it to be perfect, but considering the output we can say that the model is doing its job properly.

With increase in dataset and with proper annotated data we can train this model to be feasible for professional use and can actually replace the human intervention in the diamond grading process.

A. *Future Advancements of the project*

In the given time of four weeks we had the main objective of completing the clarity model, but for complete diamond grading we have other factors as well like the cut, colour and carat weight factor. Models for these parts are also possible to make. This will help make the diamond grading process fully independent of human intervention. Diamonds are expensive and need to be graded properly, therefore these models also need to be very accurate in order for them to be deployed in the professional world.

VII. REFERENCES

1) https://www.coursera.org/learn/neural-networks-deep-learning?specialization=deep-learning
2) https://www.coursera.org/learn/convolutional-neural-networks?specialization=deep-learning
3) https://github.com/matterport/Mask RCNN
4) https://www.makesense.ai/
5) https://sarine.com/
6) https://www.gia.edu/diamond-quality-factor
7) https://www.osapublishing.org/oe/fulltext.cfm?uri=oe-27-19-27242i=418763

REFERENCES