CSCI 1430 Final Project Report:
Shine Bright like a Diamond!

Shine Bright like a Diamond: Nicholas Fah-Sang, Bryce Jones, Suyash Kothari, Zhirui Li
Brown University
Dates of Mentor Meetings: Checkin 1: 11/21/2022; Checkin 2: 12/06/2022

Abstract

Grading diamonds by clarity–the quality of a diamond, which is determined by its lack of inclusions or imperfections–is typically an intensive process of the diamond supply chain. This paper attempts to automate the clarity grading process of diamonds using extracted interest points from a top-view image of a diamond and a deep neural network to then assort the input diamonds into one of the six clarity grades that range from FL (Flawless) to I3 (Included). We use an open-source Kaggle dataset [3] of diamonds for a neural network with architecture based on VGG-16 with additional dropout functions to prevent overfitting. This model consistently achieves an accuracy of X after extracting interest points on each diamond where light intensities drastically change, and an accuracy of Y without such feature extractions.

1. Introduction

The Gemological Institute of America popularized the categories by which diamonds are assorted and certified: color, clarity, cut and carat (weight), each category determining the price of a diamond. While it is a trivial task to classify diamonds by carat, color, and cut, classifying by clarity remains a difficult part of the diamond supply chain that requires human graders.

2. Related Work

For background information on diamonds and diamond clarity, we used information from the Gemological Institute of America. [1]

Previous research on the evaluation of networks of increasing depth using an architecture with very small (3x3) convolution filters conducted at Cornell University was used when creating our Neural Net approach. [2]

Information from previous computer vision analysis on diamonds performed by the Optica Publishing Group was used for insight on mapping inclusions with information on light refraction & intensity to support it. [4]

3. Method

Approach 1: NN without extracted features
Approach 2: NN with extracted features

3.1. Pre-Processing

The dataset contains eight CSV files–one per diamond cut (cushion, emerald, heart, marquise, oval, pear, princess and round). Each of these CSV's contains data including measurements, colors, polishes, clarity and a unique identifier. Corresponding to every diamond is an image, contained in a separate directory. For the purposes of this project, we extracted the diamond images (JPEGs) and their corresponding clarities (Strings).

A minority of diamonds had faulty images or had irregularly-sized images. We removed these from the dataset.

3.2. Neural Network Design

The Neural Network used in this project is based off of the VGG-16 architecture. It consists of five blocks of convolutional filter layers, each with ReLU activation. Each block is followed by a dropout and max pooling layer to prevent overfitting. We chose to use a categorical cross-entropy loss function and the standard tensorflow ADAM optimizer.

3.3. Feature Extraction

Reflections and facets are structurally different from inclusions. Specifically, regions of reflection contain a small number of variations in intensity in all directions, whereas regions of inclusions contain a large number of variations in intensity. [4]

We based our feature extraction methods on these principals, which means that to find the inclusions in our diamonds we would need to find the areas where there was the largest variation in intensity in all directions. This can be expressed by (Eq. 1):

$$E(u,v) = \sum_{x,y} w(x,y)[I(x+u,y+v) - I(x,y)]^2 \quad (1)$$

where $w(x,y)$ is the window function to give weight to the pixels. This can either be a rectangular window or a Gaussian window, in our testing we used a Gaussian window. $I(x+u,y+v)$ is the shifted intensity, and $I(x,y)$ is the intensity of a pixel. [4]

Applying the Taylor Expansion to this equation yields (Eq. 2) and (Eq. 3)

$$E(u,v) \approx [uv]M[\frac{u}{v}] \quad (2)$$

$$M = \sum_{x,y} w(x,y)[\frac{I_x I_x}{I_x I_y} \frac{I_x I_y}{I_y I_y}] \quad (3)$$

where I_x and I_y are the image derivatives in the x and y directions.

Finally a score is given to determine if the window contains an intersection or not (Eq. 4):

$$R = det(M) - k(trace(M))^2 \quad (4)$$

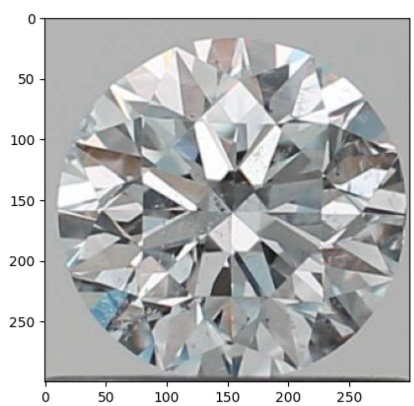

Figure 1. Round shaped diamond with "I_1" clarity rating

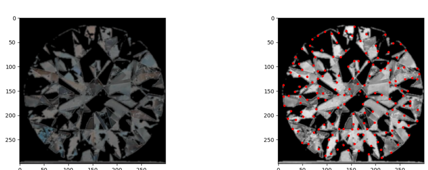

Figure 2. *Left:* Round shaped "I_1" diamond pre-processed using conventional CV methods. *Right:* Same diamond with features plotted

4. Results

The VGG model had underwhelming success in classifying the diamonds, both with and without feature extraction. While it had higher accuracy than random selection of categories, it still averaged at around 40%.

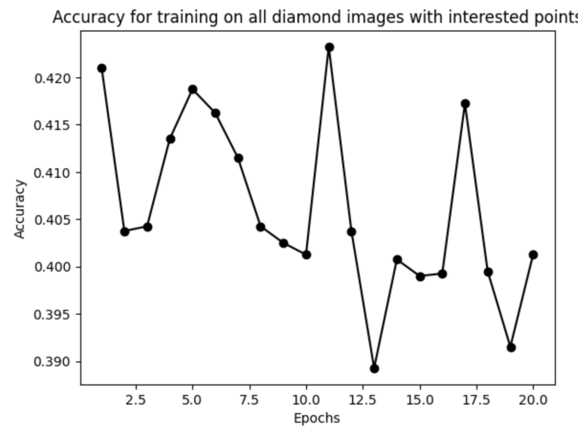

Figure 3. Graph of Accuracy over 20 epochs

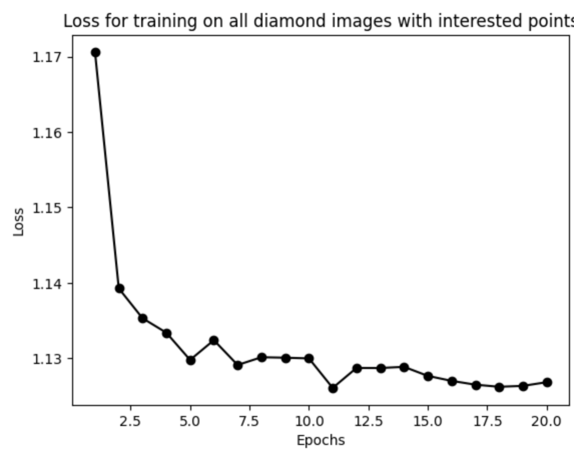

Figure 4. Graph of Loss over 20 epochs

4.1. Technical Discussion

We decided that including dropout functions wasn't necessary: the model consistently had low accuracy while training, meaning that it wasn't over-fitting to the training data. To the contrary, we believe that the task is probably too complicated due to the variety in shapes, color, and cut present in the training data. We attempted to decrease the complexity of the task by only including images of round-shaped diamonds. This resulted in a higher accuracy during the training process (around 60%), but the resulting model was less accurate when tested.

Initially, we used dropout functions at the end of each block to reduce over-fitting. However, the model's low accuracy with the testing data demonstrated that it was not over-fitting, likely due to the difficulty of the task at hand. As such, we removed these.

4.2. Societal Discussion

The broader social impact of automated diamond grading by clarity is vast. For jewelers, procuring diamonds that have skipped the human part of the grading supply chain would ultimately reduce the cost of diamonds, increasing potential profits. For end-consumers, automating diamond assortment could have two consequences. The first is that consumers can be more confident of their diamonds' value as using a model can reduce the chance of human error. The second is that it could backfire–part of the luxury of having a diamond is that it is rare, difficult to procure and sort. If it is perceived to be easy to buy diamonds, it could degrade consumers' perception of how luxurious or exclusive a diamond is. Most significantly, the impact of automated diamond grading by clarity is very consequential. Human gemologists have been a crucial component of the grading process; many would be made redundant with the success of models that can assort diamonds by clarity automatically. Many jobs may be lost. Given that profits–for better or for worse–will ultimately be the prevailing motivation to develop automatic grading systems, it is essential that the industry protects its gemologists either by repurposing them to manage and maintain the computer vision models for classification, or by transitioning their roles to be jewelers, enabling them to gain capital rather than to simply be the means of production.

5. Conclusion

We think this exploration was successful in developing our understanding of how to combine classical computer vision methods with deep learning using CNNs. Getting results that indicated an accuracy more than twice as good as chance was good, but not sufficient to be used in industry and to ultimately replace human gemologists. That being said, it is definitely technically feasible to automate grading diamonds by clarity because the grading process involves inspecting the diamond using a jewlers' loop along the top surface of the diamond; in theory, top-view images of diamonds should therefore be sufficient.

5.1. Possible Improvements

If we were to continue attempting to improve the accuracy of the model, the most impactful change we could make would be using a better dataset. The set we used from Kaggle had many issues that likely reduced the accuracy of the model. There were very few higher-quality diamonds, meaning that the model was skewed towards predicting many inclusions. Some diamond cuts (EG cushion shaped diamonds) were underrepresented in the dataset as well. What's more, many of the diamonds were positioned differently in the images or had different lighting conditions. Consistent lighting conditions in particular would have made it much easier to recognize inclusions. Additionally, it was difficult to verify if the extracted interest points were indeed inclusions. Datasets with diamonds with their inclusions marked on them would be helpful for verification. We would also recommend generating more data if there is insufficient data (shears, rotations etc.).

References

[1] GIA. What is diamond clarity: The 4cs of diamond quality, Dec 2019. 1

[2] Simonyan K. and Zisserman A. Very deep convolutional networks for large-scale image recognition, 2014. 1

[3] H. Lakhani. Natural diamonds (prices + images), 2022. 1

[4] Wenjin Wang and Lilong Cai. Inclusion extraction from diamond clarity images based on the analysis of diamond optical properties. 1, 2

Appendix

Team contributions

Please describe in one paragraph per team member what each of you contributed to the project.

Nicholas Fah-Sang Worked in tandem with Suyash on preprocessing, background removal, and feature plotting for diamond images. Implementation of the classical CV Methods side of the project, and creation of feature plotted image sets for DL Model training.

Bryce Jones Worked on the convolutional neural network with Zhirui. Designed the initial model and helped to tweak to achieve better accuracy. Wrote part of the report and assisted with the poster as well.

Suyash Kothari Worked with Nicholas on classical computer vision methods to extract feature points on diamonds. This involved working through methodologies described in papers to identify intersections that indicated diamond inclusions. I also found the original dataset and papers used to conduct the project.

Zhirui Li Preprocessed the diamond dataset by extracting corresponding images and labels. Constructed a CNN-based neural network to classify raw diamond images and diamond images with interest points into six clarity levels.