Index N0: CT-01-29

Convolutional Neural Network for Classification and Value Estimation of Selected Gemstones in Sri Lanka

Samali Amarasekara (1st Author)
Apple Research and Development Centre, Department of Computer Science, Faculty of Applied Sciences, University of Sri Jayewardenepura,
Nugegoda, Sri Lanka.
samali.r.amarasekara@gmail.com

Ravinda Meegama (2nd Author)
Apple Research and Development Centre, Department of Computer Science, Faculty of Applied Sciences, University of Sri Jayewardenepura,
Nugegoda, Sri Lanka.
rgn@sjp.ac.lk

***Abstract*— Gemstone classification and value estimation are considered to be tedious tasks encountered in the gem industry all over the world. This happens due to colour variations found in the same gem type which is often difficult to detect with the naked eye. This paper presents a machine learning approach to automatically classify the gem type by using an image and also to estimate the value of the stone using a few measurements. The proposed technique uses a microscopic image of a gemstone which is taken using a gemological microscope. A Convolutional Neural Network (CNN) is trained to classify gem type while features such as type, colour palette, shape and weight are used to estimate the value of the. This work creates a system that is capable of classifying and estimating the value of four types of gemstones, namely, Blue Sapphire, Yellow Sapphire, Amethyst and Cat's eye. The results indicate that the proposed technique managed to classify the gemstones with the highest accuracy of 87% for yellow sapphires and 77% for blue sapphires. The yellow sapphires produced the highest accuracy in colour categorization which can be attributed to the high contrast of the images available. As such, it can be concluded that the quality of the original image is important in correctly identifying the exact colour of a gemstone.**

Keywords - Gem classification, gem value estimation, artificial intelligence, machine learning

I. INTRODUCTION

The gem industry in Sri Lanka is reported to be existent for over 2500 years consisting of about thirty popular and rare gem stones [1]. Over the past decades, only people who have gathered sufficient experience from their generations in this industry can categorize and make a valuation by examining the gem using the naked eye and a few tools and therefore,the value of a gem is highly dependent on the person who examines the stone and often varies by a substantial margin when valuations are made by different persons. The National Gem and Jewelry Authority [2] is a state institute in Sri Lanka that provides a credible valuation for a gem and clients from other countries depend on this entity to obtain the correct value of a gem.

While some gemstones such as blue sapphires, pink sapphires, yellow sapphires, ruby, spinel and chrysoberyl show extreme colour variations from each other, gemstones such as blue sapphires, blue spinels, tanzanites and aquamarines show minor variations of the same colour. As shown in Fig. 1, the blue sapphires have about seven colour categories [3]. This problem of varying valuations can be eliminated by modern technology and automated tools that utilizes machine learning techniques to classify and make a valuation for a particular gemstone.

II. RECENT WORK

A labeling technique for amber gemstones have been suggested where image processing and machine learning are used for identification [4]. Amber pieces are identified and labeled to one of 30 colour classes or one of 20 geometric shape classes. Experimental results have shown that the technique is effective for inorganic shape classification within a selected geometric shape seven if there is high ambiguity between organic shapes.

A method is presented for detecting synthetic gemstones using image processing where the gemstone is illuminated using a laser beam to produce a reflection/refraction pattern on a screen [5]. Then, an image of the reflection pattern is captured to determine the plurality of white spots. Further, the coordinates of these white spots are used to determine whether the gemstone is synthetic or not.

A study is carried out to identify three types of gemstone, namely, ruby, sapphire and emerald using the HSV colour space, image processing techniques and an Artificial Neural

Network (ANN) based backpropagation algorithm that learns by examples [6].

A machine vision system is developed to automate grading of opals, referred to as gemological digital analyzer (GDA) which analyzes the opal images to extract a summary of the flash, body tone, and other characteristics to automatically classify the opal into one of several opal classes [7]. The experimental results have shown that the opals in the training set could be classified with over a 90% correct classification rate.

An invention to automate gem evaluation and reporting is presented that uses predefined criteria for the estimation [8]. The relationship between the selling price of diamonds and their weight in carat is investigated by collecting more than 100,000 certified data about diamonds from the site: www.infodiamond.com [9].

To estimate the correct value of a gem, we mostly consider the category, colour, clarity, cut, carat, rarity and weight [10]. However, as of today, only a few experiments have been carried out in the gem industry in Sri Lanka for gemstone identification and value estimation. The research presented in this paper proposes a method to identify, categorize the colour and subsequently, compute a value for a given gemstone.

III. METHODOLOGY

Although Sri Lanka has approximately twenty gem types, it is difficult to obtain high-resolution images of good quality stones.The scope of this research is limited to identifying only four types of gem types, which are, Amethyst, Blue Sapphire, Cat's eye and Yellow Sapphire as depicted in Table I. The most tedious task of this research is to collect a data set that can effectively be used to train the neural network models. The microscopic images, often given with a verification issued to clients who wish to purchase the stones, were collected from the National Gem and Jewelry Authority of Sri Lanka. This image data is used to obtain accurate information about the colour, shape, weight and price of each stone.

Fig. 2 shows different colour palettes attributed to each gem type. In this study, the images are separated into three colour palettes as dark, middle and light.

The flow chart in Fig. 3 provides the stages involved in training CNN and the regression model using Python programming and Keras libraries. Some researchers have performed color categorization of the amber gemstone using pruned decision tree classifier in which the mean, standard deviation, kurtosis, and skewness are calculated on amber pixels from grayscale and HSV color spaces were selected for the classification [20]. This technique requires a pipeline of routines starting from pre-processing and subsequent morphological operations to extract features of interest and as such, the final output is heavily biased towards image processing algorithms used in the said sub-tasks resulting in an overall accuracy of about 72%.

Fig. 1: Color variations of blue colour gems (top) and blue sapphire (bottom)

The following steps are used to train the CNN model11 given in Fig. 4 to obtain four classes of gemstones.

i. Images having a 2D resolution of 64x64 are used for training.
ii. Applied MaxPoolingwith a pool size of (2,2).
iii. Applied fully connected layer of 128 neurons and activation Relu (Rectified Linear Unit)
iv. Applied another fully connected layer of 4 neurons and activation Softmax.
v. Applied image augmentation methods with (shear, zoom and scale).
vi. Trained the model for 250 epochs.

Table 1. Four types of gemstones and the collected sample size of each.

Class	Image	Gem name	# of images
1		Amethyst	82
2		Blue Sapphire	85
3		Cat's eye	65
4		Yellow Sapphire	80

Fig. 2: Classified colour palette of gem types

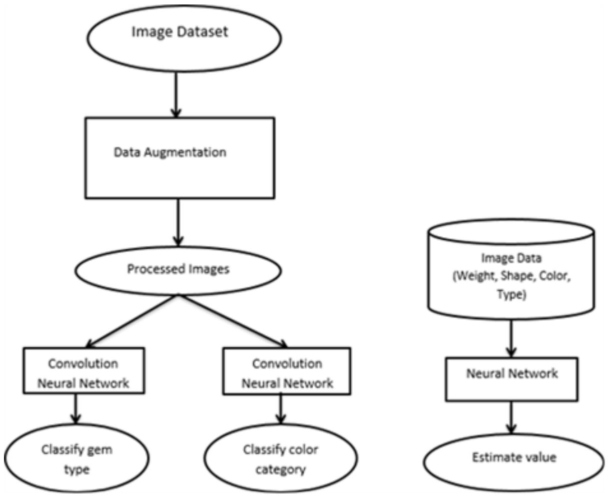

Fig. 3. Flow chart of training the neural network used to train the model.

After the images are classified into 4 classes, it is required to perform colour categorization for each class to determine whether each image belongs to dark, mid or light colour. The following steps are used to train the CNN model [12] as in Fig. 5 for colour categorization:

i. Images having a 2D resolution of 64x64 are used for training.
ii. Applied MaxPooling with a pool size of (2,2).
iii. Applied fully connected layer of 128 neurons and activation Relu.
iv. Applied another fully connected layer of 3 neurons and activation Softmax.
v. Applied image augmentation method with (shear, zoom and scale).
vi. Trained the model for 250 epochs.

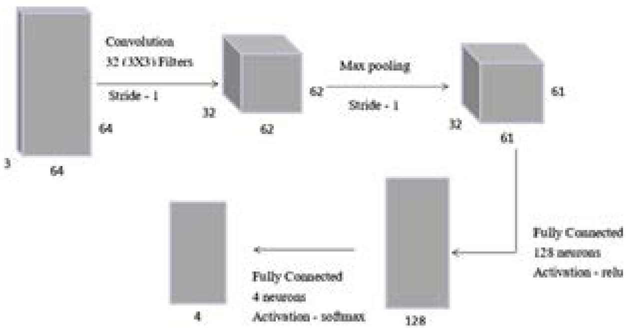

Fig. 4. CNN model structure for gem classification into four classes

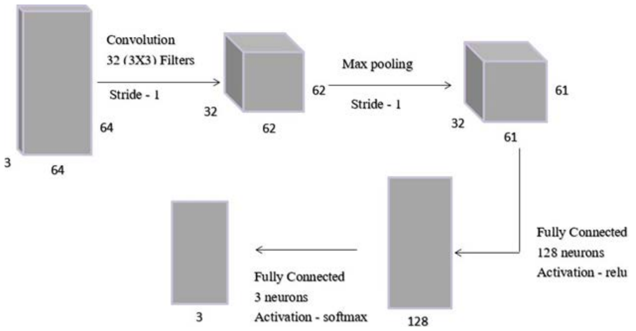

Fig. 5. CNN model structure for gem categorization into three colours.

To estimate the value, the type, shape, colour category and the weight of the gem are considered as inputs to the data pre-processor that uses Min-MaxScaler and Label encoder. Using a quarter of the data randomly for testing, the regression model with one hidden layer of 128 neurons and ReLU activation function is used for the inputs while the output layer is made up of 1 neuron and ReLU activation function [13]. As our task is a multiclass classification problem, we used a confusion matrix to determine incorrect classifications [14].

IV. RESULTS AND DISCUSSION

A confusion matrix, as shown in Table II, is used to measure the performance of the classification model where the actual and predicated colour variation is indicated. In this study, a quarter of the dataset is randomly used for testing the model for 993 epochs. A confusion matrix is specifically used as the output may fall into several classes. As seen in Table II, gemstones with dark colour texture managed to come up somewhat closer to the predicted values.

The 30% testing dataset was selected randomly before starting the training. The accuracy of colour categorization in each of the four gemstone types is given in Table III. It is observed variations in accuracy over different gem types. The highest accuracy is exhibited for yellow sapphires while the blue sapphires exhibit the lowest. This can be attributed to the high contrast of the images available for yellow sapphires than others. The cat’s eye gemstone too is uniquely identifiable from its sparkling spot and thus, produced a similar accuracy in colour categorization to the yellow sapphire.

The main objective of this study is to identify gem type and estimate the value. Related works on this study were not found for gemstones, especially for colour classification in yellow and blue sapphires. Our machine learning approach yielded good results in colour categorization of amethyst, blue and yellow sapphires while an accuracy of 77% was obtained for the cat’s eye.

Table 2. Confusion matrix of gem categorization based on color (n indicates sample size).

Gem type	Colour variation (actual)	Colour variation (Predicated)		
		Dark	Middle	Light
Blue sapphire (n = 30)	Dark	6	2	0
	Middle	0	10	1
	Light	0	2	9
Cat’s eye (n = 22)	Dark	3	1	1
	Middle	2	3	1
	Light	0	0	11
Yellow sapphire (n = 44)	Dark	13	0	0
	Middle	0	24	0
	Light	0	0	17
Amethyst (n = 30)	Dark	9	0	0
	Middle	0	6	4
	Light	0	0	11

Table 3. Accuracy of color categorization in four gem types.

Gem type	Accuracy of colour categorization
Amethyst	80%
Blue sapphire	77%
Cat’s eye	86%
Yellow sapphire	87%

As gemstones are often transparent, human experts use color coverage, brilliance and dispersion to make an accurate evaluation and color categorization [15]. As there is no clear standard to obtain a valuation, certain gemstones are heat-treated to obtain color variations. Such treatment requires expertise knowledge in advance so that proper parameters can be set before heating because even a small error can make a yellow sapphire turn orange making it a raw stone that does not have any value. Availability of such precise treatment methods can be the reason for obtaining higher accuracy in colour categorization for yellow sapphires [16].

Colors of blue and yellow gemstones can also be evaluated by using fluorescence spectroscopy [17]. Using short and long pass filters, the gemstones can be screened for color categorization in both treated and un-treated stones. However, as spectroscopy results in color-bias, it is difficult to make a general identification rate for each gemstone type other than simply stating identification rates for individual colors for each gemstone. Also, Cat's eye and Amethyst gemstone types do not respond to fluorescence spectroscopy to produce a color separation resulting in less than 30% accuracy for classification.

Reflectance spectroscopy, coupled with an artificial neural network (ANN), can also be used for the identification of certain types of gemstones [18,19]. The reflectance spectral information of a gemstone can be fed into an ANN during the training stage. An inherent limitation of using such spectroscopy-based methods is that identification of gems of the same type, whether natural or synthetic, is not effective thus producing poor results. For instance, the yellow and blue sapphire exhibit several shades of the primary color and in such cases, using the reflectance spectroscopy will not guarantee an accurate result in the classification process.

The blue sapphires are produced across the world in a variety of colour saturations. The actual color of a natural blue sapphire is so complex that even with a human expert with considerable experience, classification tends to be a tedious task [21].

While a majority of blue sapphires produced, show a very dark colour that hides the actual radiant nature of the stone, the Sri Lankan blue sapphires exhibit bright colour due to their low iron content. The high iron content of the blue sapphires can result in low accuracy in colour categorization in the experiments. A previous study had used K-means algorithm for color categorization of blue sapphires producing accuracies of 98.4%, 97.9%, 97.8 %, 98.7 % and 99.0%, respectively of blue, cyan, light blue, dark blue, and gray blue. However, it does not produce an overall color categorization for the blue gemstone [22].

As our machine learning technique utilizes the colour pixel values of the images, images with low contrast produced lower accuracy in colour categorization. As such, the requirement to have high contrast images to obtain high accuracy results is a limitation of the proposed methodology. Further, the biggest obstacle that was encountered during the research is collecting approximately 300 high-resolution images of gems. As images of all the colour pallets of each gem are not available, the gem colours are divided into three categories as dark, medium and light using expert knowledge.

V. CONCLUSION

To arrive at an approximate estimate of a gemstone, an appraiser would manually review and analyze each gemstone separately and generate a report indicating the value. As this process is time-consuming, the proposed technique allows us to automate this task of estimating the value using a machine learning approach. Although there are more than seventy varieties of gems in Sri Lanka, approximately twenty types of stones are found commonly in the industry. For experiments conducted in this paper, only four gem types are considered which are widely available in the local industry. The quality of the original image is important in correctly identifying the exact colour of a gemstone. Extending the present research for the identification and value estimation of other gem types is possible with a large number of data sets having good quality images.

REFERENCES

[1] J.W. Herath, “Geology and occurrence of gems in Sri Lanka,” The Journal of the National Science Council of Sri Lanka, vol. 12, no. 2, pp. 257-271, 1984.
[2] National Gem and Jewelry Authority of Sri Lanka, Available: http://www.ngja.gov.lk/en/.
[3] S. De Meo, A. Plutino and A. Rizzi, “Assessing Color of Gemstones,” Color Research & Application, vol. 45, no. 2, pp. 224-234, 2020.
[4] F. Emmanuel and R.M. George, “An update on color in gems. Part 3: Colors caused by band gaps and physical phenomena,” Gems & Gemology, vol. 24, no. 2, pp. 81-102, 1998.
[5] A.T. Collins, “The detection of colour-enhanced and synthetic gem diamonds by optical spectroscopy,” Diamond and related materials, vol. 12, pp. 1976-1983, 2003.
[6] X. Yang, “Application and Development of Digital Image Processing Technology in Jewelry Design,” International Conference on Machine Learning and Big Data Analytics for IoT Security and Privacy. Springer, Cham, 2020.
[7] D. Wang et. al., “Automated Opal Grading by Imaging and Statistical Learning,” IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 46, no. 2, pp. 185–201, 2016.
[8] P.J. Malnekoff, “Automated gemstone evaluation system,” 2011. US6304853B1161 [Online]. Available from: https://patents.google.com/patent/US6304853B1/en [Accessed:11-Feb-2020]
[9] N.G. Vaillant and W. François-Charles, “Understanding diamond pricing using unconditional quantile regressions,” Hal archives, 2013, halshs-00853384f. https://halshs.archives-ouvertes.fr/halshs-00853384/
[10] M. Januz, “Application of Neural Networks in Regression Problems with Qualitative Data on an Example of Gemstone Valuation,” Scientific Conference on Social Sciences and Arts SGEM, 2018. DOI: 10.5593/sgemsocial2018/1.5/S05.010
[11] T. Guo, J. Dong, H. Li and T. Gao, “Simple convolutional neural network on image classification,” In: IEEE 2nd International Conference on Big Data Analysis (ICBDA), pp. 721-724, 2017, DOI: 10.1109/ICBDA.2017.8078730.
[12] X. Lei, H. Pan X. and Huang, “A Dilated CNN Model for Image Classification,” IEEE Access, vol. 7, pp: 124087-124095, 2019.

[13] J. Xu, Z. Yuanjian and M. Duoqian, “Three-way confusion matrix for classification: A measure driven view,” Information sciences. vol. 507, pp. 772-794, 2020.

[14] I. Maula, et al., “Development of a Gemstone Type Identification System Based on HSV Space Colour Using an Artificial Neural Network Back Propagation Algorithm,” In: International Conference on Science and Technology, and Technology for Development, 2017.

[15] M. Silvia, A. Plutino and A. Rizzi, “Assessing Color of Gemstones,” Color Research & Application, vol. 45, pp. 224-234, 2020.

[16] K. Kanyarat, et al, “Heat Treating Blue Sapphire of Pailin, Cambodia,” Chian Mai Journal of Science, vol. 45, no. 1, pp. 588-600, 2018.

[17] T. Tsung-Han and U. D’Haenens-Johansson, “Rapid Gemstone Screening and Identification Using Fluorescence Spectroscopy,” Applied Optics, vol. 60, no. 12, pp. 3412-3421, 2021.

[18] Q. Jun-Ting, L. Qiu, and H. Mu, “Feasibility of Gem Identification Using Reflectance Spectra Coupled with Artificial Intelligence,” Spectroscopy Letters, vol. 52, no. 9, pp. 520-532, 2019.

[19] D. Kun, et al., “Price Estimation Model of Jade Bracelet Based on RBF Neural Network,” IOP Conference Series: Materials Science and Engineering, vol. 81, no. 81(1), 2018.

[20] S. Saulius, A. Lipnickas and K. Rimkus, “Automatic Amber Gemstones Identification by Color and Shape Visual Properties,” Engineering Applications of Artificial Intelligence, vol. 37, pp. 258-267, 2015.

[21] J. Han, Y. Guo and S. Liu, “Environmental Issues on Color Quality Evaluation of Blue Sapphire based on GemdialogueTM Color Comparison Charts,” Ekoloji, vol. 106, pp. 1365-1376, 2018.

[22] P. Aaron, “Coexisting rubies and blue sapphires from major world deposits: A brief review of their mineralogical properties,” Minerals, vol. 10, no. 5, pp. 472, 2020.