MDPI

Article

Automatic Gemstone Classification Using Computer Vision

Bona Hiu Yan Chow *[ID] **and Constantino Carlos Reyes-Aldasoro** [ID]

Department of Computer Science, School of Mathematics, Computer Sciences and Engineering,
City, University of London, London EC1V 0HB, UK; reyes@city.ac.uk
* Correspondence: bona.chow@city.ac.uk

Abstract: This paper presents a computer-vision-based methodology for automatic image-based classification of 2042 training images and 284 unseen (test) images divided into 68 categories of gemstones. A series of feature extraction techniques (33 including colour histograms in the RGB, HSV and CIELAB space, local binary pattern, Haralick texture and grey-level co-occurrence matrix properties) were used in combination with different machine-learning algorithms (Logistic Regression, Linear Discriminant Analysis, K-Nearest Neighbour, Decision Tree, Random Forest, Naive Bayes and Support Vector Machine). Deep-learning classification with ResNet-18 and ResNet-50 was also investigated. The optimal combination was provided by a Random Forest algorithm with the RGB eight-bin colour histogram and local binary pattern features, with an accuracy of 69.4% on unseen images; the algorithms required 0.0165 s to process the 284 test images. These results were compared against three expert gemmologists with at least 5 years of experience in gemstone identification, who obtained accuracies between 42.6% and 66.9% and took 42–175 min to classify the test images. As expected, the human experts took much longer than the computer vision algorithms, which in addition provided, albeit marginal, higher accuracy. Although these experiments included a relatively low number of images, the superiority of computer vision over humans is in line with what has been reported in other areas of study, and it is encouraging to further explore the application in gemmology and related areas.

Keywords: gemstone; segmentation; machine learning; computer vision; human vision

1. Introduction

Accurate gemstone classification is crucial to the gem and jewellery trade as the identification is an important first step in the evaluation and appraisal of any gem [1]. Currently, the identity of a gemstone is determined using a combination of visual observation and spectrochemical analysis [2]. Through careful observation of gemstones with the unaided eye and under magnification, gemmologists detect visual characteristics, such as colour, transparency, lustre, fractures, cleavages, inclusions, pleochroism, phenomenon and birefringence, to facilitate the separation of gemstones [3]. This is a difficult process, as many gems share colour and characteristics, as illustrated in Figure 1, which displays a sample of 500 gems distributed among 87 different categories. Identification and classification are often accompanied by the use of gemmological tools, which include refractometers [4], polariscopes and conoscopes [5], handheld spectroscopes [6], dichroscopes [7] and ultraviolet light [8] to probe the optical properties of gemstones. Measuring physical properties such as specific gravity [9] (also known as relative density) provides additional information related to the identity of a gemstone. With the emergence of new synthetic gemstones and treatment techniques, increasingly complex instruments with powerful spectroscopic, fluorescent or chemical analysing abilities have been introduced into gemmological laboratories [2]. Such instruments include infrared spectrometers [10], Raman and luminescence spectrometers [11–13], ultraviolet–visible spectrometers [14], cathodoluminescence [15], energy-dispersive X-ray fluorescence spectrometers [4,16], laser ablation-inductively coupled plasma-mass spectrometers [17] and fluorescence spectrometers [18].
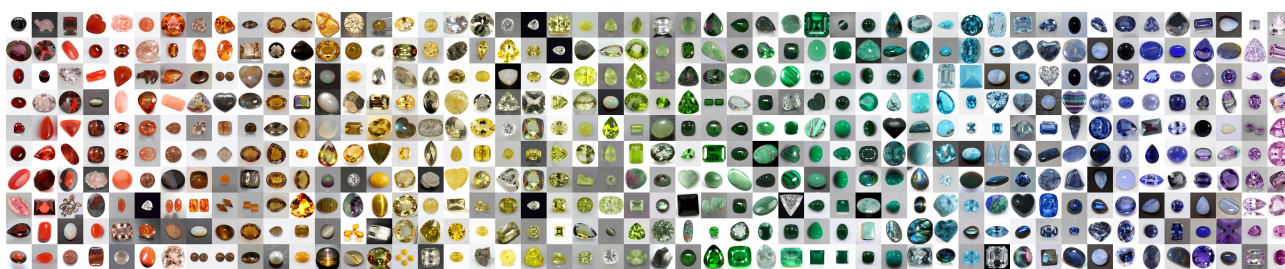
Figure 1. Five hundred images of gemstones selected from eighty-seven different categories, from Almandine to Zoisite. The images are arranged by hue to illustrate the difficulty of identifying the gems by visual inspection.

Yet, the identification is still difficult and time consuming, and not all laboratories have access to these sophisticated instruments; thus, the identification through automatic techniques based solely on images is attractive. In recent years, computers and algorithms have evolved significantly, and image processing and computer vision tasks are commonplace in many areas such as medical imaging, manufacturing and security. In geological sciences, computer vision algorithms have been developed for classifying mineral grains [19–24] and rocks [25–29]. Thompson et al. [19] segmented microscopic thin sections using edge detection and achieved a test accuracy up to 93.53% when classifying 10 different minerals with an Artificial Neural Network trained on extracted colour and texture features. The reported accuracy was likely inflated, as the same examples of Biotite were used for training and testing. Baykan and Yilmaz [20] developed an Artificial Neural Network to separate five minerals using the Red–Green–Blue (RGB) values of pixels in manually segmented thin sections as the input and attained an accuracy of 89.53%. Izadi et al. [21] segmented thin sections using incremental clustering and performed mineral classification using a cascade approach. An Artificial Neural Network was first used to differentiate 23 types of minerals and glass based on pixel colours, and only those minerals exhibiting similar colours under both plane- and cross-polarised light were passed into a second Artificial Neural Network for concurrent colour and texture analysis. This resulted in an overall accuracy of 93.81%. Borges and de Aguiar [22] demonstrated that simple machine-learning algorithms—K-Nearest Neighbour and Decision Tree—were capable of classifying minerals in microscopic thin sections based on colour and texture with a high average accuracy of 94.11–97.71% using two datasets with four and seventeen mineral types. Maitre et al. [23] segmented microscopic images containing eight mineral types and background by simple linear iterative clustering and classified them using three machine-learning algorithms, namely K-Nearest Neighbour, Random Forest and Decision Tree, based on colour features in the RGB, Hue–Saturation–Value (HSV) and CIELAB space. The Random Forest algorithm produced the highest accuracy of 82%. Zhang et al. [24] investigated the classification of four minerals with six different algorithms, namely Logistic Regression, Support Vector Machine, Random Forest, K-Nearest Neighbour, Multilayer Perceptron and Naive Bayes, using features extracted from microscopic images with Inception-v3. Support Vector Machine was identified as the single algorithm yielding the highest accuracy (90.6%). Stacking Support Vector Machine, Logistic Regression and Multilayer Perceptron models further improved the accuracy by 0.3%. Despite the robustness of computer vision systems in mineral recognition, only one study on automatic identification of gemstone images [30] has been reported to date, to the best of the authors' knowledge. A per-class accuracy of 75–100% was attained for classifying unseen Ruby, Blue Sapphire and Emerald images based on the Hue channel of the HSV colour space using an Artificial Neural Network. It should be noted that Rubies, Blue Sapphires and Emeralds have very distinctive colours and as such are relatively easy to distinguish from each other, much easier than gems of similar colours such as Topaz and Aquamarine. Other computer vision research in the context of gemmology focused mostly on gemstone evaluation [31–33] and recognition [34,35]. Robust computer vision systems for grading the colour of Amber [36], Jadeite Jade [37],

Opal [38] and Pearls [39] were developed. Zhang and Guo [37] proposed a system that can be developed into a tool for measuring gemstone colour.

In this paper, we present a computer vision approach for automatic image-based classification of 68 categories of gemstones. We first introduce the image dataset and describe the feature extraction techniques and the classification algorithms in Section 2. Section 3 presents an evaluation and comparison of the results with reference to an expert group. In Section 4, a discussion of the experimental findings is provided. Finally, conclusions and ideas for further work are presented in Section 5.

2. Materials and Methods

2.1. Materials

A total of 2326 images of gemstones were obtained from Kaggle [40] (accessed on 27 April 2021) and analysed in this work. A sample of these is illustrated in Figure 2. The images were grouped into categories, and for this work, the following 68 classes were selected for analysis: *Alexandrite, Almandine, Amazonite, Amber, Amethyst, Ametrine, Andradite, Aquamarine, Aventurine Green, Aventurine Yellow, Benitoite, Beryl Golden, Bixbite, Bloodstone, Blue Lace Agate, Carnelian, Chalcedony, Chalcedony Blue, Chrome Diopside, Chrysoberyl, Chrysocolla, Chrysoprase, Citrine, Coral, Diamond, Diaspore, Dumortierite, Emerald, Fluorite, Hessonite, Iolite, Jasper, Kunzite, Kyanite, Lapis Lazuli, Malachite, Onyx Black, Onyx Green, Onyx Red, Peridot, Prehnite, Pyrite, Pyrope, Quartz Beer, Quartz Lemon, Quartz Rutilated, Quartz Smoky, Rhodochrosite, Rhodolite, Rhodonite, Ruby, Sapphire Blue, Sapphire Pink, Sapphire Purple, Sapphire Yellow, Serpentine, Sodalite, Spessartite, Sphene, Sunstone, Tanzanite, Tigers Eye, Topaz, Tourmaline, Tsavorite, Turquoise, Zircon* and *Zoisite*. The images were acquired under very different conditions of illumination and background colours, as can be appreciated in Figure 1. The dimensions spread across a wide range: heights between 93 px and 3055 px and widths between 89 px and 3947 px. Post-processing may have been applied to some of the images, i.e., cropped or processed in Photoshop, but these details were not available. A total of 2042 images were used for training and 284 images were reserved for testing. For each class, 24–44 training images and 4–5 test images were available. The original Kaggle dataset consists of 3219 images distributed into 87 classes, but some of these were discarded according to the criteria described in Appendix A.

Most images contained a single gemstone, but a small number of images portrayed multiple gemstones (Figure 2). The gemstones exhibited various colours, shapes and cutting styles. The top of the gemstones were featured in most images. It should be noted that whilst some gemstones such as Malachite and Zoisite were readily recognised by their colours and patterns, it would be challenging to separate some gemstones such as Emerald and Tsavorite based solely on colour.

(**a**) Amber	(**b**) Benitoite	(**c**) Coral	(**d**) Emerald	(**e**) Kyanite	(**f**) Malachite
(**g**) Onyx Red	(**h**) Pyrope	(**i**) Rhodolite	(**j**) Topaz	(**k**) Tsavorite	(**l**) Zoisite

Figure 2. Twelve representative images of the gemstones in this work. It can be noticed that some gems, such as Malachite and Onyx Red, can be readily recognised by the unique colours, whereas it would be challenging to distinguish between Emerald and Tsavorite.

2.2. Methods

The framework consisted of data acquisition, background segmentation, feature extraction, construction of the machine-learning classifiers and evaluation (Figure 3).

Data acquisition	→	Background segmentation	→	Feature extraction	→	Classifier construction	→	Evaluation

Figure 3. The computer vision framework followed in this work: data acquisition, background segmentation, feature extraction, construction of machine-learning classifiers and evaluation.

All the algorithms used in this work were coded in Python 3.7.9. and are freely available via GitHub (https://github.com/hybchow/gems accessed on 31 December 2021). With the exception of transfer learning, the scripts were executed on a MacBook Pro equipped with a 2.3 GHz Intel Core i5 processor. Transfer learning was implemented on a virtual NVIDIA Tesla K80 Graphics Processing Unit (GPU) with two workers in Google Colaboratory. Experimental findings were visualised using Python and Tableau.

2.2.1. Background Segmentation

We used Otsu thresholding [41] to automatically extract gemstones from the background. Otsu thresholding is a non-parametric, unsupervised segmentation method that operates by maximising the variance between background and foreground intensities. The binary masks created by application of Otsu thresholding to either grey-level intensity or the Saturation channel of the HSV space were compared visually (Figure 4). An image was regarded as well-segmented upon satisfying these criteria: (1) the background was completely removed, and (2) the majority of the gemstone was extracted. The approach yielding more well-segmented training images for a gemstone class was adopted for segmenting the test images in the same class. All test images were retained regardless of segmentation quality.

The pipeline to obtain the background through intensity-based Otsu thresholding was as follows: 1. Converted images to grayscale; 2. Applied Gaussian smoothing with sigma of 2 to the grayscale image; 3. Applied Otsu thresholding to grey-level intensity to create a binary mask; 4. Flipped the mask if the average intensity of 20×20 px regions from each corner of image is higher than average of the entire image, i.e., the background had a higher intensity than the gemstone, so that the gemstone instead of the background was extracted; 5. Filled holes in the mask; 6. Applied binary closing to the mask using a disc-shaped structuring element with a radius of 8 px; 7. Removed objects smaller than 301 px in size; 8. Filled holes in the mask; 9. Applied binary erosion to the mask using a square-shaped structuring element of 2×2 px; 10. Filled holes in the mask; 11. Applied the mask to the original image by setting the pixels identified as background to a value of zero.

The pipeline to obtain the background through Otsu thresholding to the Saturation channel was very similar: 1. Converted image from the RGB to HSV space and extracted the Saturation channel of HSV; 2. Applied Gaussian smoothing with a sigma of 5 to the Saturation channel of HSV; 3. Applied Otsu thresholding to Saturation channel of HSV to create a binary mask; 4. Flipped the mask if the average Saturation of 20×20 px regions from each corner of image is higher than average of the entire image, i.e., the background had a higher Saturation than the gemstone, so that the gemstone instead of the background was extracted; 5. Filled holes in the mask; 6. Applied binary closing to the mask using a disc-shaped structuring element with a radius of 9 px; 7. Removed objects smaller than 301 px in size; 8. Filled holes in the mask; 9. Applied binary erosion to the mask using a square-shaped structuring element of 2×2 px; 10. Filled holes in the mask; 11. Applied the mask to the original image by setting the pixels identified as background to a value of zero.
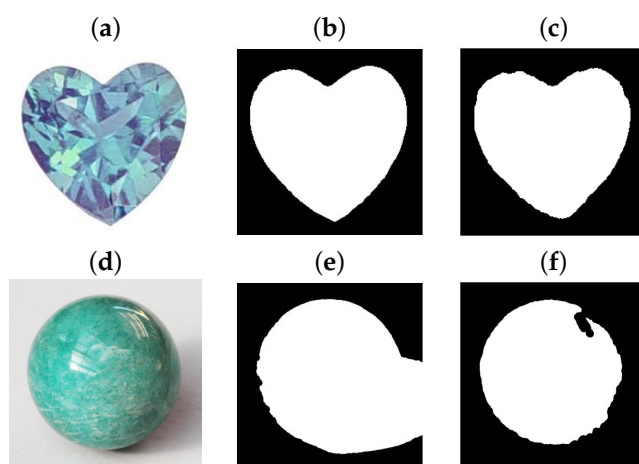
Figure 4. Illustration of the background segmentation for two representative images of Alexandrite (**top row**) and Amazonite (**bottom row**). (**a**,**d**) Original images. (**b**,**e**) Mask through intensity-based Otsu thresholding. (**c**,**f**) Mask through Otsu thresholding based on the Saturation channel of the HSV space. It should be noted that the purely intensity-based mask is less accurate than the Saturation-based mask.

2.2.2. Feature Extraction

Feature extraction is one of the essential processes of computer vision and image-processing tasks [42]. Feature extraction can be understood in many ways: low-level extraction, which focuses on edges, colours, textures, shapes, regions and other characteristics of an image, sometimes extracted through transforms such as Fourier or Discrete Cosine Transform [43–45] ; high-level extraction that jumps to understanding or behaviours [46] and also to the reduction of dimensionality, which sometimes is accomplished by selecting a reduced set of features or measurements from the data [47].

A total of 33 features based on colour and texture were extracted from masked images for posterior classification. These features were: *colour of the non-background K-means cluster in the RGB, HSV or CIELAB space, 4- and/or 8-bin histogram(s) in the RGB, HSV or CIELAB space, combination of the RGB or HSV 4- or 8-bin histogram and Haralick texture, combination of the RGB 8-bin histogram and one grey-level co-occurrence matrix (GLCM) property from correlation, dissimilarity, energy, angular second moment (ASM), contrast or homogeneity, combination of the RGB or HSV 4- or 8-bin histogram and the local binary pattern (LBP) with 8 points at radius 1, combination of the RGB 8-bin histogram and the LBP with 8 points at radius 3, combination of the RGB 8-bin histogram and the LBP with 16 or 24 points at radius 1 or 3, combination of the RGB 4- and 8-bin histograms and the LBP with 8 points at radius 1* and combination of the RGB 4- and 8-bin histograms and the LBPs with 8 points at radius 1 and 24 points at radius 3.

Colour feature extraction was performed in three colour spaces: RGB, HSV and CIELAB colour space. The RGB space is conceptualised by human trichromatic colour vision and describes colours by the additive combination of orthogonal red, green and blue components [48]. HSV is based on human intuition [49], closely related to the artistic ideas of hue tint and shade [50] and provides excellent discrimination for highly saturated areas [51,52].

HSV is represented by a hexacone, where Saturation is the horizontal axis, Hue can be either a circular angle or a value with the horizontal axis, and Value is the vertical axis [53]. CIELAB is designed to represent perceptual uniformity (the colour difference matches that perceived by humans) [49]. Unlike the RGB and HSV space, CIELAB is device-independent [49]. Figure 5 illustrates the three colour spaces for all the images of the training data and Figure 6 illustrates the median Hue and Saturation per gemstone.
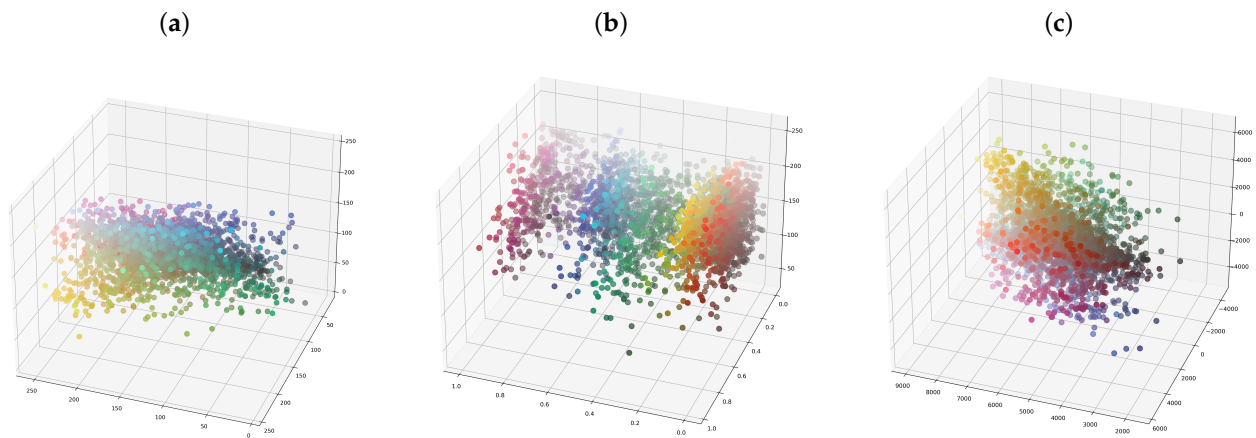
(**a**)   (**b**)   (**c**)

Figure 5. Illustration of the central colour characteristics for all the images of the training data represented in the three colour spaces. (**a**) Red–Green–Blue (RGB). (**b**) Hue–Saturation–Value (HSV). (**c**) CIELAB. It should be noted that, visually, HSV provides better discrimination as the colours are ranked, in this case along the horizontal axis. HSV can also be displayed in polar plots as Hue is a circular property.
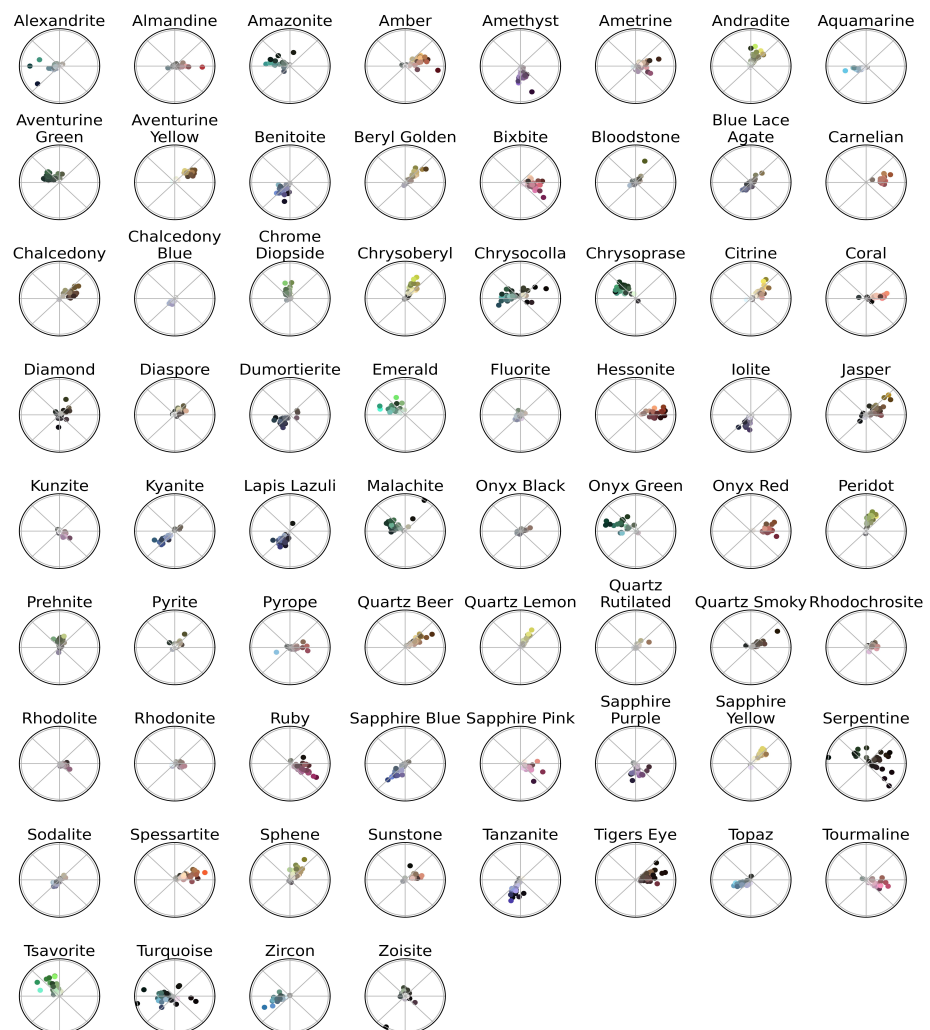

Figure 6. Distribution of median Hue and Saturation of gemstone images aggregated by class and displayed as polar scatter plots. It should be noted that the distribution of several gemstones is very similar, e.g., Emerald and Tsavorite.

　　　K-means clustering and colour histograms were used to extract colour from the gemstones (Figure 7). K-means clustering was performed to divide the pixels in each image into groups based on colour such that the difference between the groups was maximised and the variation within each group was minimised [54]. For simplicity, we assumed that each image consisted only of two colours representing the gemstone and the background and that the centre of the background colour cluster had a smaller sum of Red, Green and Blue values in RGB, a smaller sum of Hue and Saturation components in HSV or a lower Luminosity in CIELAB than the equivalent of the gemstone colour cluster. Colour histograms are three-dimensional arrays representing the counts of pixels in each colour space component in individual images [55,56]. We investigated the use of either 4-bin, 8-bin or a combination of both 4- and 8-bin colour histograms. The background pixels having a value of zero in the masked images were neglected when constructing the colour histograms.
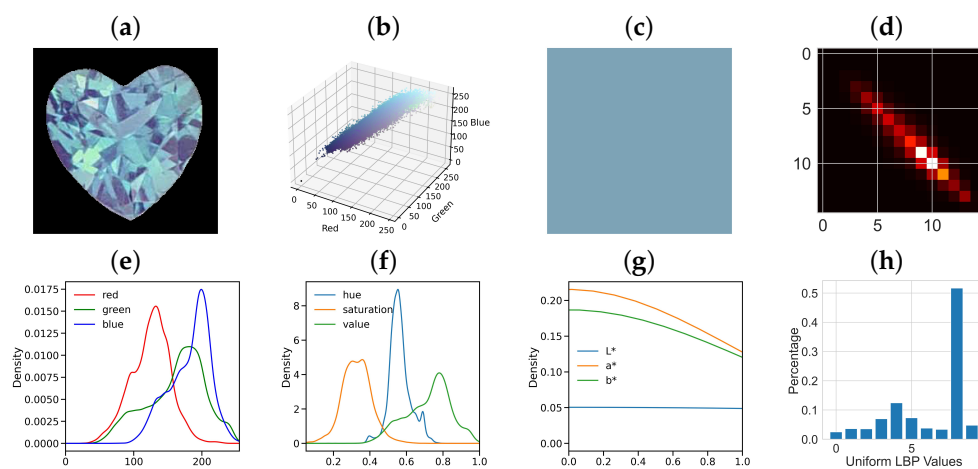

Figure 7. Illustration of colour and texture features extracted from a representative image. (**a**) Masked image of Alexandrite. (**b**) Scatter plot of the RGB values of all pixels. (**c**) RGB colour of the centre of the non-background K-means cluster. (**d**) GLCM matrix with an offset of 1 and angle of 0 (**e**) RGB histogram. (**f**) HSV histogram. (**g**) CIELAB histogram. (**h**) LBP histogram.

　　　Textural features extracted from images have been widely employed in the past [57–61] in areas such as crystallography [62], stratigraphy [63], the natural stone industry [64] and medical imaging [65–67]. Besides chromatic characteristics, texture seems to be one of the most attractive features to use to discriminate between gemstones, e.g., the gems of Figure 2d,f,k,l have similar hues, but different textures. Texture features were extracted using either the LBP, the Haralick texture or the selected GLCM property. The LBP, proposed by Ojala et al. [68], is a grey-scale-invariant texture-extraction method that binarises the grey levels of neighbouring pixels relative to a central reference pixel. Only LBPs with either 8 points, 16 points or 24 points at a radius of 1 or 3 were explored in this work. Haralick [57] developed the GLCM to describe the spatial dependencies of grey levels of neighbouring pixels and derived 14 properties from the GLCM, which later became known as the Haralick texture. Only six individual properties, namely correlation, energy, dissimilarity, homogeneity, contrast and ASM, for the GLCM with an offset of 1 and angle of 0, were investigated. When computing the Haralick texture and GLCM, the background pixels were ignored. However, the background pixels were not explicitly disregarded when extracting the LBP. With little variation in intensity, the background would contribute only to the lowest LBP values, which was expected to have little impact on the image classification.

　　　Following feature extraction, the Synthetic Minority Oversampling Technique (SMOTE) with the four nearest neighbours (k = 4) was applied to equalise the class proportions. SMOTE is a technique that generates "synthetic" training data for non-majority classes by interpolation [69].

2.2.3. Machine-Learning Algorithms

Seven supervised machine-learning algorithms, namely Logistic Regression, Linear Discriminant Analysis, K-Nearest Neighbour, Decision Tree, Random Forest, Naive Bayes and Support Vector Machine, were investigated.

Logistic Regression

Logistic Regression [70] yields probabilistic predictions for each class based on a non-linear transformation of the input features. In multi-class classification, the softmax function is applied to the input features. The parameter C is the inverse of the regularisation strength.

Linear Discriminant Analysis

Linear Discriminant Analysis [71] works by reducing the dimensionality of the input features such that the difference between classes is maximised, whilst minimising intra-class variance. A linear decision boundary can thus be drawn to separate any two classes. In this study, Linear Discriminant Analysis with the least squares formulation was applied. Shrinkage is a regularisation parameter that takes values between 0 and 1 and may improve accuracy in cases where the number of training samples is small compared to the number of features.

K-Nearest Neighbour

K-Nearest Neighbour [72] is a non-parametric classification algorithm based on simple majority voting by a predefined number of nearest (most similar) neighbours.

Decision Tree

Decision Tree [73] operates by sequentially dividing data into smaller subsets based on one of the criteria (features) until each subset contains the most homogeneous (lowest Gini impurity) collection of data possible. The three parameters optimised in this work were the maximum depth of the tree, the maximum number of features to consider at each split and the minimum number of samples required in a leaf node.

Random Forest

Random Forests, first proposed by Breiman [74], are based on voting by an ensemble of independent, dissimilar decision trees. Each individual tree is constructed using a random selection of training data with replacement (bootstrap) such that the correlation between trees is reduced. The data are recursively partitioned at each node using the feature from a randomly selected subset, which results in the most homogeneous collection of samples. Three parameters, namely the number of estimators (trees), the maximum depth of the tree and the minimum number of samples required in each leaf node, were optimised. For a comprehensive description of Random Forests, the reader is directed to the book by Criminisi and Shoton [75].

Naive Bayes

Naive Bayes [76] is based on Bayes Theorem and makes probabilistic predictions using the probabilities for each class and the likelihood probabilities of the features given the class. The smoothing factor was assigned the default value of 1×10^{-9}.

Support Vector Machine

Support Vector Machine [77] is a binary classifier that separates two classes by maximising the margin between them. To solve a multi-class problem, a series of binary Support Vector Machine classifiers is constructed, each separating a single class from the remaining classes ("One-versus-Rest"). Three parameters, namely kernel type, regularisation parameter C and the kernel coefficient *gamma*, were optimised.

Parameter Optimisation

A 5-fold cross-validated grid search [78] was performed to find the optimal parameters using negative cross-entropy loss as the scoring metric. The ranges of the parameters specified for the grid search are listed in Table 1.

Table 1. Parameters specified for the 5-fold cross-validated grid search for the seven machine-learning algorithms and for the ResNets are listed. Note that a single parameter value was assigned to Naive Bayes, whereas three hyperparameters of the Decision Tree, Random Forest and Support Vector Machine algorithms were optimised.

Algorithm	Range of Parameters
Logistic Regression	"C": [0.001,0.01,0.1,1,10]
Linear Discriminant Analysis	"solver": "lsqr"; "shrinkage": [0,0.5,1]
K-Nearest Neighbour	"n_neighbors": [3,5,7,9]
Decision Tree	"max_depth":[10,None]; "max_features": [3,5,7,9]; "min_samples_leaf": [3,5,7,9]
Random Forest	"n_estimators": [50,100]; "max_depth": [3,5,7,9]; "min_samples_leaf": [3,5,7,9]
Naive Bayes	"var_smoothing": 1×10^{-9}
Support Vector Machine	"estimator__kernel": ["linear", "poly", "rbf", "sigmoid"]; "estimator__C": [1,10,100]; "estimator__gamma": [0.1,0.01]
ResNet	Number of layers: 18 or 50; Training images:RandomResizedCrop(224), RandomHorizontalFlip, RandomVerticalFlip and Normalize; Test images: Resize(256), CenterCrop(224) and Normalize; "batch_size": 16; "max_epochs": 25; "criterion": torch.nn.CrossEntropyLoss; "lr": 0.001; "optimizer": torch.optim.SGD; "optimizer__momentum": 0.9; "iterator_train__num_workers": 2; "iterator_valid__num_workers": 2; "iterator_train__shuffle": True; "callbacks": LRScheduler(policy = "StepLR", step_size = 7,gamma = 0.1), Checkpoint(monitor = "valid_acc_best"), Freezer(lambda x: not x.startswith("model.fc"))

2.2.4. Convolutional Neural Networks and Transfer Learning

Recent years have been dominated by the advances in the areas of *deep learning* [79]. Deep learning can be considered as a branch of machine learning where large amounts of input data and their corresponding labels are provided to a model, also known as the architecture or network, which will then learn the characteristics or representations intrinsic to the data in order to classify or regress the data [80]. These architectures have a large number of layers, thus considered *deep*, and a very large number of parameters between these layers. Deep learning has provided incredible results, perhaps the most significant being that related to the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [81].

One of the limitations of deep learning is the need for a large amount of input data, which can be minimised through training of the architecture with data from a different context and then fine-tuning later on. Transfer learning [82] is a popular and efficient approach for image classification [83–85], in which a model applies the knowledge acquired from one task to another.

In this work, Microsoft's Residual Network (ResNet) [86], the winner of the 2015 ILSVRC, was selected for analysis. ResNet is characterised by a deep architecture with shortcut connections between non-adjacent convolutional layers. In this work, 18- and 50-layer ResNets pre-trained on the ImageNet dataset were applied to the gemstone classifi-

cation. For compatibility with the ResNets, the gemstone images were cropped and resized to 224 × 224 px and each batch of 16 images was normalised using the mean and standard deviation of the ImageNet dataset. Slightly differing approaches were used to process the training and test images. For the training images, a random portion was cropped and resized to 224 × 224 px and data augmentation in the form of random horizontal or vertical flip was applied. Weighted random sampling was applied when grouping the training images into batches, so as to eliminate the effect of class imbalance. The test images were first resized to 256 × 256 px, and the centre portion was cropped to produce images of 224 × 224 px. When training the ResNet models, the weights of the neurons in the convolutional layers were frozen and only those in the final fully-connected layer were adapted. The learning rate was initially set at 0.001 and was scheduled to decay by a factor of 0.1 every seven epochs. The optimiser Stochastic Gradient Descent (SGD) with a momentum of 0.9 and the cross-entropy loss function were selected. Shuffling was applied when presenting batches of training images to the models. Within a maximum epoch of 25 epochs, the model with the highest 5-fold stratified cross-validation accuracy was regarded as the final model.

2.2.5. Evaluation

The algorithms were evaluated and compared in terms of: accuracy, top-5 accuracy, training time and test computation time. Accuracy [87] is the proportion of correct predictions (true negatives and true positives) out of all predictions (false positives, false negatives, true positives and true negatives). Top-k accuracy [88] is similar to accuracy, the difference being each algorithm is allowed a number of guesses instead of a single guess for each prediction and the prediction is regarded as correct when one of the guesses matches the true label. The time required by a machine-learning classifier to perform a grid search on the most important parameters to optimise the algorithm [89] or that required by ResNet to complete 25 training epochs is regarded as the training computation time. The total time a classifier required to make predictions on all 284 test images was recorded as the test computation time. Both the training and test computation time of the machine learning classifiers was measured using a MacBook Pro equipped with a 2.3 GHz Intel Core i5 processor, whereas for ResNets, a virtual NVIDIA Tesla K80 GPU provided by Google Colaboratory was used.

2.2.6. Expert Group

The expert group consisted of three gemmologists with both Graduate Gemologist of Gemological Institute of America (GIA) and Fellowship of the Gemmological Association of Great Britain qualifications and 5–8 years of experience in gemstone identification. The performances of the algorithms and the expert group were evaluated and compared in terms of accuracy and time requirement for the classification of the 284 unseen images. Confusion matrices [90] were used to visualise the counts of predictions for all combinations of true and predicted labels.

3. Results

3.1. Background Segmentation

The segmentation of the backgrounds was assessed visually. Saturation-based Otsu thresholding provided better segmentation results for the training images of 52 classes, whereas grey-level Otsu thresholding produced better segmentation results for the images of 16 classes. The 52 classes segmented using Saturation-based Otsu thresholding were *Amazonite, Amber, Amethyst, Ametrine, Andradite, Aventurine Green, Aventurine Yellow, Benitoite, Beryl Golden, Bixbite, Carnelian, Chalcedony, Chalcedony Blue, Chrysoberyl, Chrysocolla, Chrysoprase, Citrine, Coral, Diaspore, Dumortierite, Emerald, Hessonite, Iolite, Jasper, Kunzite, Kyanite, Lapis Lazuli, Malachite, Onyx Green, Onyx Red, Peridot, Prehnite, Pyrite, Quartz Lemon, Quartz Smoky, Rhodochrosite, Ruby, Sapphire Pink, Sapphire Purple, Sapphire Yellow, Serpentine, Spessartite, Sphene, Sunstone, Tanzanite, Tigers Eye, Topaz, Tourmaline, Tsavorite, Turquoise,*

Zircon and *Zoisite*. The remaining 16 classes segmented with grey-level Otsu thresholding were *Alexandrite, Almandine, Aquamarine, Bloodstone, Blue Lace Agate, Chrome Diopside, Diamond, Fluorite, Onyx Black, Pyrope, Quartz Beer, Quartz Rutilated, Rhodolite, Rhodonite, Sapphire Blue* and *Sodalite*. It was estimated that Saturation-based Otsu thresholding yielded an average of 77% well-segmented test images per class, whereas grey-level Otsu thresholding produced an average of 76% well-segmented test images per class.

3.2. Feature and Algorithm Comparison

Seven different machine-learning algorithms were compared, each with 33 different feature-extraction methodologies, which provided a total of $7 \times 33 = 231$ combinations. Deep-learning classification with ResNet-18 and ResNet-50 was also investigated. The algorithm that provided the highest accuracy on the unseen images was based on Random Forest using the RGB eight-bin colour histogram and local binary pattern with eight points at radius one with an accuracy of 69.4% and required a test time of 0.0165 s. This accuracy was better than the expert group, who achieved an accuracy of 42.6–66.9% in 42–175 min.

Table 2 shows the results for each machine-learning algorithm. The test accuracy and time requirement of each gemmologist in the expert group are listed in Table 3. Results for the most accurate algorithm for each feature extraction method are presented in Table 4. Figure 8 displays the test accuracy of all combinations, grouped by algorithm on the horizontal axis. Each combination is shown as a filled circle, with the colour corresponding to the feature extraction method. In addition, the distribution of the results per algorithm is summarised as a boxplot with black horizontal lines corresponding to maximum/minimum values, a grey box between the 25th and 75th percentiles and a change of grey tone at the median of the distribution. In some cases (e.g., Decision Tree), some values are considered outliers and fall outside the boxplots. A continuous horizontal line at the test accuracy of 66.9% indicates the performance of the most accurate gemmologist.

Table 2. Accuracy, top-5 accuracy, training and test time of the most accurate classifier for each machine-learning algorithm. Random Forest was the most accurate algorithm. Support Vector Machine required a significantly longer training and test time than the other algorithms. It should be noted that unlike the other algorithms, the ResNets were trained and tested on a Graphics Processing Unit (GPU) instead of a Central Processing Unit (CPU).

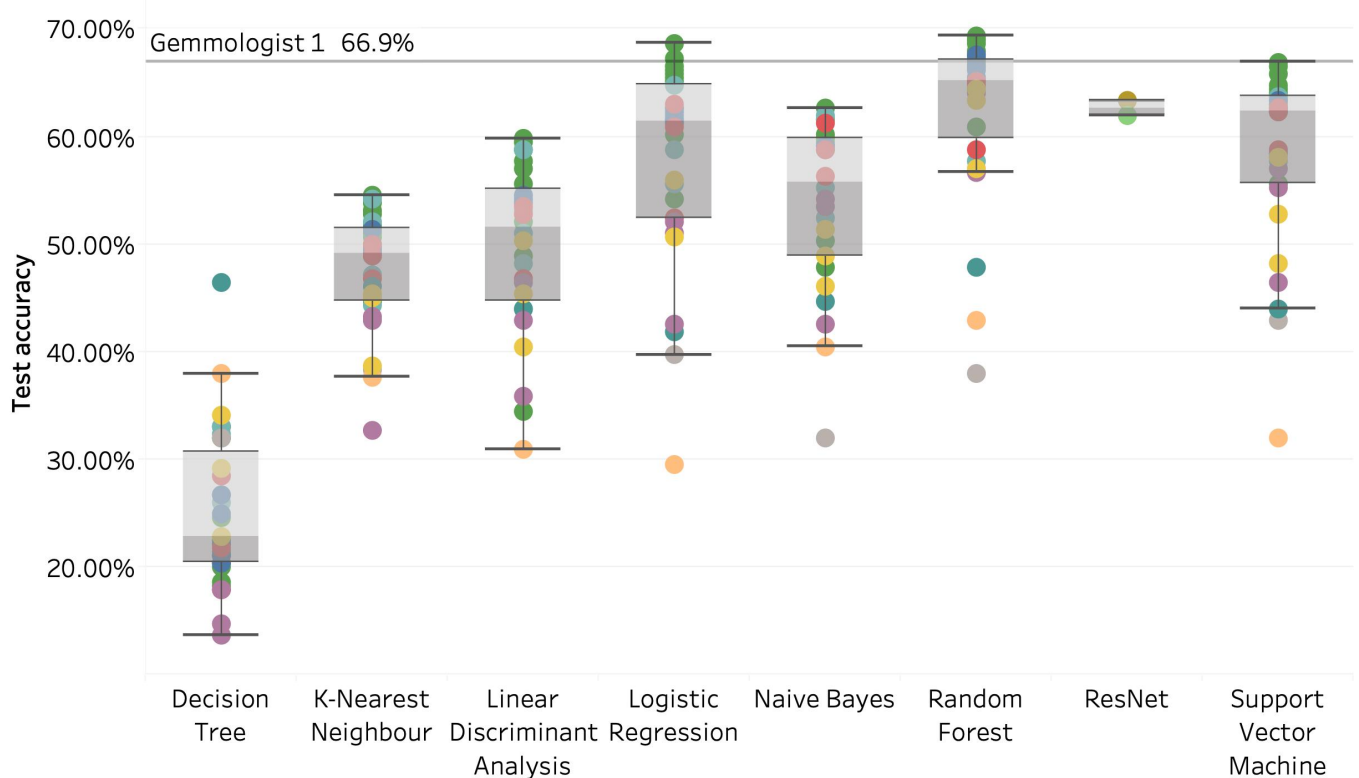| Algorithm	Accuracy	Top-5 Accuracy	Training Time in Seconds	Test Time in Seconds
Random Forest	69.4%	94.4%	39.81	0.0165
Logistic Regression	68.7%	92.6%	17.79	0.0008
Support Vector Machine	66.9%	86.3%	1881.36	0.5459
ResNet50	63.4%	91.5%	449.09	4.5244
Naive Bayes	62.7%	77.8%	0.54	0.0281
ResNet18	62.0%	89.4%	293.05	2.2119
Linear Discriminant Analysis	59.9%	94.0%	3.71	0.0007
K-Nearest Neighbour	54.6%	85.9%	1.09	0.0479
Decision Tree	46.5%	73.9%	0.56	0.0002

Table 3. Accuracy and test time of three expert gemmologists classifying the 284 unseen images varied significantly.

Expert	Accuracy	Test Time
Gemmologist 1	66.9%	175 min or 10,500 s
Gemmologist 2	46.8%	97 min or 5820 s
Gemmologist 3	42.6%	42 min or 2520 s

Table 4. Accuracy, top-5 accuracy, training and test time of the most accurate classifier for each feature extraction method. The system based on the RGB 8-bin histogram and the LBP with 8 points at radius 1 yielded the highest accuracy of 69.4%. The highest top-5 accuracy of 96.5% was attained by the system using the RGB 4 and 8-bin histograms and the LBP with 8 points at radius 1. It should be noted that unlike the other algorithms, the ResNets were trained and tested on a Graphics Processing Unit (GPU) instead of a Central Processing Unit (CPU).

Method	Accuracy	Top-5 Accuracy	Training (s)	Test (s)
RGB 8-bin hist. and LBP, 8 points, radius 1	69.4%	94.4%	39.81	0.0165
RGB 4-bin hist. and LBP, 8 points, radius 1	69.0%	93.7%	33.78	0.0164
RGB 4/8-bin hist. and LBP, 8 points, radius 1	68.7%	96.5%	60.73	0.0181
RGB 8-bin hist. and LBP, 16 points, radius 1	68.7%	95.4%	43.33	0.0168
RGB 4/8-bin hist. and LBP, 8 points, radius 1 & 24 points, radius 3	68.7%	92.6%	17.79	0.0008
RGB 8-bin hist. and GLCM correlation	67.6%	94.4%	38.42	0.0162
RGB 8-bin hist. and LBP, 8 points, radius 3	67.6%	94.4%	78.96	0.0176
RGB 8-bin hist. and GLCM dissimilarity	67.3%	94.0%	43.60	0.0190
RGB 8-bin hist. and LBP, 24 points, radius 3	66.9%	94.7%	66.56	0.0169
RGB 8-bin hist. and LBP, 24 points, radius 1	66.9%	86.3%	1881.36	0.5459
RGB 8-bin hist. and GLCM energy	66.5%	96.1%	37.92	0.0164
HSV 8-bin hist. and LBP, 8 points, radius 1	66.5%	93.3%	82.35	0.0484
RGB 8-bin hist. and GLCM ASM	66.2%	94.7%	39.65	0.0164
RGB 8-bin hist. and LBP, 16 points, radius 3	65.8%	92.6%	14.97	0.0008
RGB 8-bin hist. and GLCM contrast	65.5%	96.1%	48.90	0.0164
RGB 4-bin hist. and Haralick texture	65.5%	95.1%	77.00	0.0262
HSV 8-bin hist. and Haralick texture	65.5%	93.7%	55.76	0.0100
RGB 8-bin hist. and Haralick texture	65.5%	93.7%	69.33	0.0176
RGB 8-bin hist. and GLCM homogen.	65.1%	95.1%	38.59	0.0165
RGB 4 and 8-bin hist.	65.1%	94.0%	45.88	0.0163
RGB 4-bin hist.	64.8%	95.4%	26.42	0.0164
HSV 4 and 8-bin hist.	64.4%	93.3%	57.64	0.0166
CIELAB 4 and 8-bin hist.	64.1%	94.0%	31.49	0.0196
CIELAB 8-bin hist.	64.1%	93.7%	26.91	0.0165
HSV 8-bin hist.	63.4%	95.1%	52.25	0.0165
ResNet50	63.4%	91.5%	449.09	4.5244
RGB 8-bin hist.	62.7%	87.0%	1387.63	0.4420
ResNet18	62.0%	89.4%	293.05	2.2119
HSV 4-bin hist. and LBP, 8 points, radius 1	60.9%	91.2%	46.68	0.0176
HSV 4-bin hist. and Haralick texture	57.7%	87.3%	580.56	0.3129
HSV 4-bin hist.	57.0%	88.7%	32.88	0.0167
CIELAB 4-bin hist.	56.7%	91.5%	21.12	0.0163
CIELAB non-background K-means cluster centre	47.9%	87.7%	20.95	0.0180
RGB non-background K-means cluster centre	44.0%	86.3%	0.17	0.0002
HSV non-background K-means cluster centre	43.0%	81.3%	17.57	0.0165

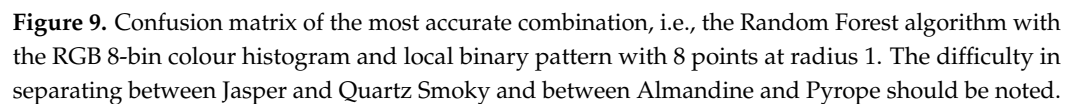## Test accuracy of classifiers aggregated by algorithms

Feature extraction method type

- ◼ HSV colour histogram
- ◼ HSV of the non-background K-means cluster centre
- ◼ LAB colour histogram
- ◼ LAB of the non-background K-means cluster centre
- ◼ ResNet18
- ◼ ResNet50
- ◼ RGB colour histogram
- ◼ RGB colour histogram and GLCM property
- ◼ RGB of the non-background K-means cluster centre
- ◼ RGB/HSV colour histogram and Haralick texture
- ◼ RGB/HSV colour histogram and local binary pattern

Figure 8. Results of all the combinations of algorithms and feature extraction methodologies grouped by algorithm. Each combination is represented by a coloured circle, and the summary of the distribution per algorithm is displayed as a statistical boxplot (for an explanation, see the text). The most accurate results corresponded to Random Forest and Logistic Regression, both of which surpassed the best gemmologist in the expert group.

Confusion matrices of the most accurate algorithm and the best gemmologist are displayed in Figures 9 and 10.
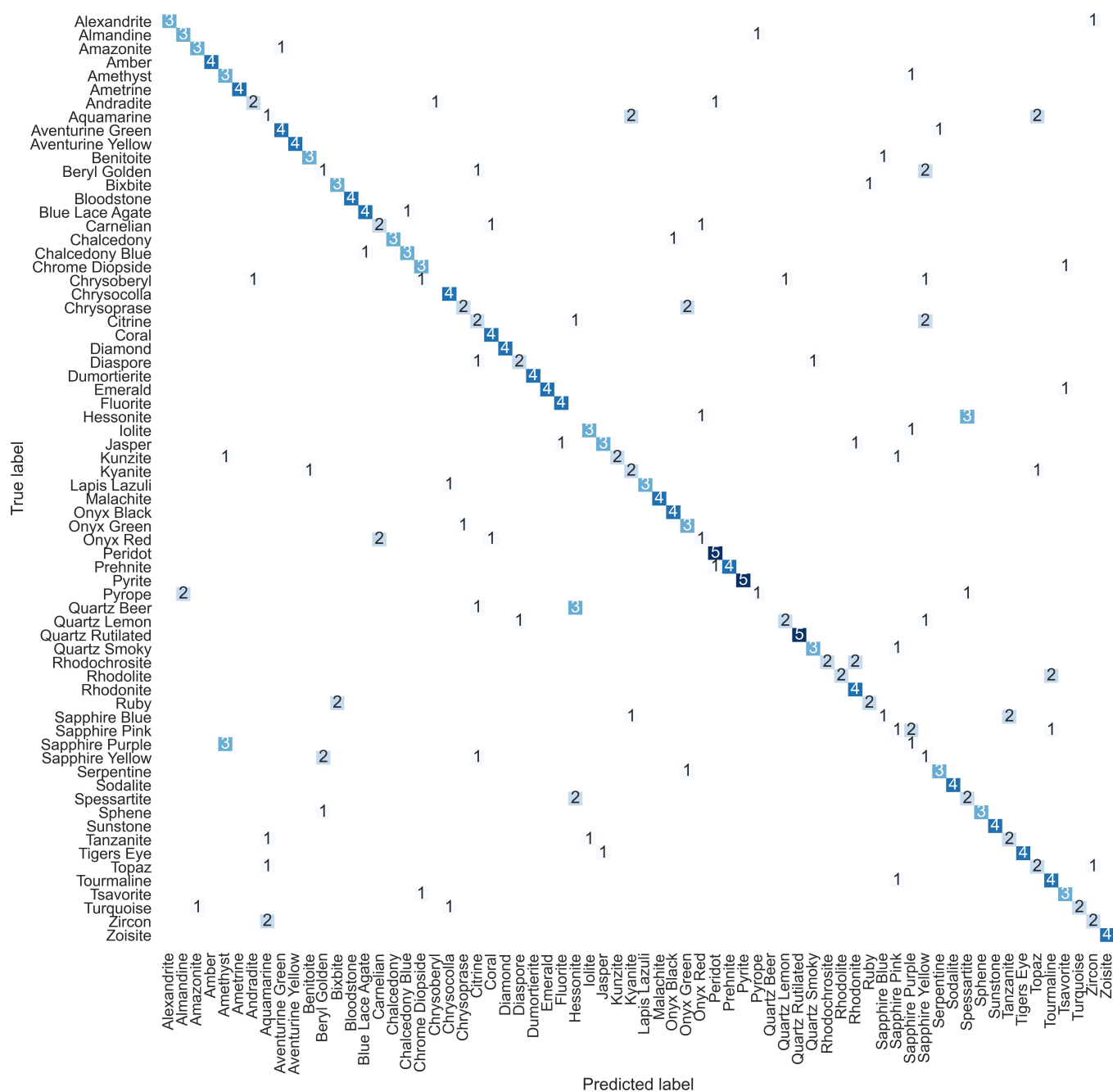
Figure 9. Confusion matrix of the most accurate combination, i.e., the Random Forest algorithm with the RGB 8-bin colour histogram and local binary pattern with 8 points at radius 1. The difficulty in separating between Jasper and Quartz Smoky and between Almandine and Pyrope should be noted.

Figure 10. Confusion matrix of the best gemmologist revealing the poorest ability in distinguishing between similarly coloured gemstones, namely Sapphire Purple and Amethyst; Hessonite and Spessartite; Quartz Beer and Hessonite.

4. Discussion

The background segmentation results revealed that Saturation-based Otsu thresholding was more efficient to separate vividly coloured gemstones from neural background and shadows, whereas grey-level Otsu thresholding yielded better results for multi-colour or low-Saturation gemstones. A challenge encountered was the separation of some transparent, light-coloured gems such as Aquamarine from the background, as the background colour showed through the gemstone. One possible solution, provided infrared images are available, would be to perform segmentation using camera images and infrared images simultaneously, as a gemstone and its background are expected to exhibit distinct transmission characteristics [91].

The choice of the feature extraction technique and machine-learning algorithm both contributed significantly to the robustness of the results obtained, as exemplified in Figure 8 and Table 4. The most accurate combination was based on Random Forest and the RGB eight-bin histogram and the LBP with eight points at radius one, which provided an accuracy of 69.4% within 0.0165 s. Another characteristic of the Random Forest results was that the distribution was more compact than those of Logistic Regression and SVMs, even taking into account the three results considered as outliers. This was more robust than the most accurate gemmologist in the expert group (test accuracy up to 66.9% in 175 min or 10,500 s) for the classification of 68 categories of gemstone images in terms of both accuracy and time.

The superiority in terms of time was expected as the algorithms would be very quick at processing, whilst humans take time to look at the images, observe the characteristics and then decide the category of gemstone to which they belong. The superior performance in terms of classification, albeit by a small margin (69.4% vs. 66.9%), was interesting and can be observed in the confusion matrices shown in Figures 9 and 10. In both matrices, the diagonal, which corresponds to the correct predictions, contains the majority of the predictions, whilst the errors are shown outside the diagonal. Some of the incorrect predictions were common to both the expert and the algorithm (e.g., Number 2 on the bottom left corresponding to Zircon and Aquamarine), but others were not shared and only incorrectly predicted by one or the other (e.g., both cases at the bottom left, Amber/Spessartite or Amazonite/Turquoise). Upon further comparison of the most accurate combination (Figure 9) with the expert (Figure 10), it was discovered that the algorithm had a stronger ability to separate gemstones of similar colours. Nonetheless, gemmologists readily recognised gemstones displaying unique colour patterns, such as the green bandings on Malachite, which was not distinguished by the algorithms. Exploring alternative colour and texture feature extraction techniques, such as colour-scale-invariant feature transform descriptors [92], may improve the accuracy. Furthermore, the widespread variation between the three experts should be noticed, with the lower two closer to the worst results of most algorithms. Yet, the highest gemmologist devoted much more time than the other two. It should also be mentioned that a baseline result, that is a random selection of a class for any given image, would be 1/68 = 1.47%.

The highest top-five accuracy of 96.5% was attained by the Random Forest model using the RGB four- and eight-bin histograms and the LBP with eight points at radius one. Colour histograms were significantly more effective than K-means clustering in extracting colour features. The RGB space in general yielded higher accuracy than the HSV or CIELAB space. The lower accuracy for the HSV space may be attributed to the cyclic nature of the Hue channel. There was inconsistency in the colour features extracted from some red gemstones such asOnyx Red, as the hue of some pixels was close to zero and others were close to one. The incorporation of texture features, in particular local binary pattern, enhanced the capability of the algorithms to differentiate between gemstones of similar colours. The training time was shortest for systems using non-background K-means cluster colours, followed by those using the four-bin colour histograms, whereas those based on the eight-bin colour histograms required the longest training time. The addition of texture features or the simultaneous use of both four- and eight-bin colour histograms did not significantly lengthen the training time. Amongst all machine-learning algorithms, Random Forest yielded the highest accuracy. The training and test time appeared significantly longer for Support Vector Machine (up to 1943.02 s and 0.5918 s, respectively) than for other algorithms (up to 83.26 s and 0.1215 s, respectively). Nonetheless, the test time was remarkably shorter than that required by the gemmologists (a minimum of 42 min or 2520 s).

One surprising result was the lower performance of both ResNet architectures against Random Forest, Logistic Regression and Support Vector Machine. The lower performance may be due to the limited number of training images that were employed. Additionally, it should be noted that the comparison is not exactly like-with-like as the ResNets were trained directly with the images and the other machine-learning algorithms were trained on

the extracted features. This may imply that the images would be harder to discriminate than the extracted features. Furthermore, the ResNets were pre-trained on the ImageNet dataset, and only the final fully-connected layer was adapted for gemstone image classification, whereas the machine-learning classifiers were tailored for this task.

One major limitation to is that it is impossible to identify gemstones that are not in the predefined categories. It is expected that the system would be incapable of separating natural from synthetic gemstones, which share the same optical properties. Additional data, such as the refractive index, specific gravity and spectroscopic, fluorescent or chemical data from laboratory instruments [2], need to be incorporated for the system to perform more complex gemstone analysis. With the constant emergence of novel gemstone colour treatments, it is essential to upgrade the system from time to time for real-world applications.

Although a limited range of gemstones was included, the findings provided a viable proof-of-concept that computer vision can be applied to gemstone classification.

5. Conclusions

To the best of the authors' knowledge, this was the first study that compared the performance of a computer-vision-based methodology against trained gemmologists on image-based classification of as many as 68 classes of gemstones. The number of classes is important, as it should be considered that a random guess would provide a 1/68 or 1.47% accuracy. In turn, a human expert provided an accuracy of 66.9%, which was outperformed by the best computer vision approach with 69.4%. Whilst the difference in accuracy is relatively small, the difference in time was of several orders of magnitude, as could have been expected. Thus, it was demonstrated that computer vision approaches can be successfully implemented for image-based gemstone classification. Whilst one of the experts provided a high accuracy, the other two experts reported much lower levels of accuracy (42.6%, 46.8%), which were well below the median values of all algorithms except Decision Tree.

In addition to the superior results of the computer vision approach in terms of accuracy and time, this approach does not require sample preparation or destruction of the materials, as is sometimes performed by gemmologists when identifying gemstones.

Future work should consider: (a) including a larger range of gemstones, potentially adding those that were discarded in this study and others; this addition could drive the accuracy down as some gems could share features; (b) considering more images for training, validation and testing; this would impact the training of the ResNet architecture and could potentially improve the performance of the architecture; (c) considering other deep-learning architectures; (d) considering, besides expert gemmologists, humans that could be trained to recognise a smaller range of gems, i.e., distinguishing between Emeralds and Tsavorites; (e) A final development could consider to apply computer vision techniques similar to those developed in this work to investigate if it is possible to distinguish between high-quality gemstones, low-quality gemstones and even counterfeits.

Author Contributions: Conceptualisation, B.H.Y.C. and C.C.R.-A.; methodology, B.H.Y.C. and C.C.R.-A.; software, B.H.Y.C.; validation, B.H.Y.C.; formal analysis, B.H.Y.C. and C.C.R.-A.; investigation, B.H.Y.C.; resources, B.H.Y.C.; data curation, B.H.Y.C.; writing—original draft preparation, B.H.Y.C.; writing—review and editing, C.C.R.-A.; visualisation, B.H.Y.C. and C.C.R.-A.; supervision, C.C.R.-A.; project administration, C.C.R.-A. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available from Kaggle (https://www.kaggle.com/lsind18/gemstones-images accessed on 28 December 2021).

Abbreviations

The following abbreviations are used in this manuscript:

ASM	Angular Second Moment
CPU	Central Processing Unit
GLCM	Grey-Level Co-occurrence Matrix
GPU	Graphics Processing Unit
ILSVRC	ImageNet Large Scale Visual Recognition Challenge
LBP	Local Binary Pattern
ResNet	Microsoft's Residual Network

Appendix A

In the original Kaggle dataset, 87 gemstone categories were present. The category Garnet was removed due to overlapping with Almandine, Pyrope, Rhodolite and Spessartite. The Moonstone images displayed a variety of either orange, white or yellow colour, which was undesirable for the algorithms, and were thus eliminated. Upon background segmentation, poorly segmented training images satisfying either of these conditions were discarded: (1) incomplete removal of background or (2) extraction of only a minor portion of the gemstone. Seventeen classes, namely *Andalusite, Cats Eye, Danburite, Goshenite, Grossular, Hiddenite, Jade, Labradorite, Larimar, Morganite, Opal, Pearl, Quartz Rose, Scapolite, Spinel, Spodumene* and *Variscite*, were removed, as fewer than 24 training images per class were retained.

References

1. Hurrell, K.; Johnson, M.L. *Gemstones: A Complete Color Reference for Precious and Semiprecious Stones of the World*; Chartwell Books: New York, NY, USA, 2016; p. 305.
2. Breeding, C. Developments in Gemstone Analysis Techniques and Instrumentation During the 2000s. *Gems Gemol.* **2010**, *46*, 241–257. [CrossRef]
3. Liddicoat, R.T. Developing the Powers of Observation in Gem Testing. *Gems Gemol.* **1962**, *10*, 291–319.
4. Sturman, D.B. A new approach to the teaching and use of the refractometer. *J. Gemmol.* **2010**, *32*, 74–89. [CrossRef]
5. Devouard, B.; Notari, F. The Identification of Faceted Gemstones: From the Naked Eye to Laboratory Techniques. *Elements* **2009**, *5*, 163–168. [CrossRef]
6. Anderson, B.W.; Payne, J. *The Spectroscope and Gemmology*; Mitchell, R.K., Ed.; GemStone Press: Nashville, TN, USA, 1999.
7. Thibault, N.W. A simple dichroscope. *Am. Mineral.* **1940**, *25*, 88–90.
8. Karampelas, S.; Kiefert, L.; Bersani, D.; Vandenabeele, P. Gem Analysis. In *Gems and Gemmology*; Springer: Cham, Switzerland, 2020; pp. 39–66.
9. Church, A.H. Notes on the Specific Gravity of Precious Stones. *Geol. Mag.* **1875**, *2*, 320–328. [CrossRef]
10. Fritsch, E.; Stockton, C.M. Infrared Spectroscopy in Gem Identification. *Gems Gemol.* **1987**, *23*, 18–26. [CrossRef]
11. Jenkins, A.L.; Larsen, R.A. Gemstone Identification Using Raman Spectroscopy. *Spectroscopy* **2004**, *19*, 20–25.
12. Bersani, D.; Lottici, P.P. Applications of Raman spectroscopy to gemology. *Anal. Bioanal. Chem.* **2010**, *397*, 2631–2646. [CrossRef]
13. Kiefert, L.; Karampelas, S. Use of the Raman spectrometer in gemmological laboratories: Review. *Spectrochim. Acta Part A* **2011**, *80*, 119–124. [CrossRef]
14. He, T. The Applications of Ultraviolet Visible Absorption Spectrum Detection Technology in Gemstone Identification. In Proceedings of the 5th International Conference on Materials Engineering for Advanced Technologies (ICMEAT 2016), Quebec, QC, Canada, 5–6 August 2016; DEStech Publications: Lancaster, PA, USA, 2016; pp. 106–109.
15. Ponahlo, J. Cathodoluminescence as a Tool in Gemstone Identification. In *Cathodoluminescence in Geosciences*; Springer: Berlin, Germany, 2000; pp. 479–500.
16. Hänni, H.A. Advancements in gemmological instrumentation over the last 30 years. *J. Gemmol. Assoc. Hong Kong* **2009**, *30*, 14–16.
17. Abduriyim, A.; Kitawaki, H. Applications of Laser Ablation–Inductively Coupled Plasma–Mass Spectrometry (LA-ICP-MS) To Gemology. *Gems Gemol.* **2006**, *42*, 98–118. [CrossRef]
18. Tsai, T.-H.; D'Haenens-Johansson, U.F.S. Rapid gemstone screening and identification using fluorescence spectroscopy. *Appl. Opt.* **2021**, *60*, 3412–3421. [CrossRef] [PubMed]

19. Thompson, S.; Fueten, F.; Bockus, D. Mineral identification using artificial neural networks and the rotating polarizer stage. *Comput. Geosci.* **2001**, *27*, 1081–1089. [CrossRef]
20. Baykan, N.A.; Yilmaz, N. Mineral identification using color spaces and artificial neural networks. *Comput. Geosci.* **2010**, *36*, 91–97. [CrossRef]
21. Izadi, H.; Sadri, J.; Bayati, M. An intelligent system for mineral identification in thin sections based on a cascade approach. *Comput. Geosci.* **2017**, *99*, 37–49. [CrossRef]
22. Borges, H.P.; de Aguiar, M.S. Mineral Classification Using Machine Learning and Images of Microscopic Rock Thin Section. In Proceedings of the 18th Mexican Conference on Artificial Intelligence, MICAI 2019, Xalapa, Mexico, 28 October–1 November 2019; IEEE: New York, NY, USA, 2019; pp. 63–76.
23. Maitre, J.; Bouchard, K.; Bédard, P. Mineral grains recognition using computer vision and machine learning. *Comput. Geosci.* **2019**, *130*, 84–93. [CrossRef]
24. Zhang, Y.; Li, M.; Han, S.; Ren, Q.; Shi, J. Intelligent Identification for Rock-Mineral Microscopic Images Using Ensemble Machine Learning Algorithms. *Sensors* **2019**, *19*, 3914. [CrossRef]
25. Ślipek, B.; Młynarczuk, M. Application of pattern recognition methods to automatic identification of microscopic images of rocks registered under different polarization and lighting conditions. *Geol. Geophys. Environ.* **2013**, *39*, 373–384. [CrossRef]
26. Chatterjee, S. Vision-based rock-type classification of limestone using multi-class support vector machine. *Appl. Intell.* **2013**, *39*, 14–27. [CrossRef]
27. Młynarczuk, M.; Górszczyk, A.; Ślipek, B. The application of pattern recognition in the automatic classification of microscopic rock images. *Comput. Geosci.* **2013**, *60*, 126–133. [CrossRef]
28. Perez, C.A.; Saravia, J.A.; Navarro, C.F.; Schulz, D.A.; Aravena, C.M.; Galdames, F.J. Rock lithological classification using multi-scale Gabor features from sub-images, and voting with rock contour information. *Int. J. Miner. Process.* **2015**, *144*, 56–64. [CrossRef]
29. Xu, Z.; Ma, W.; Lin, P.; Shi, H.; Pan, D.; Liu, T. Deep learning of rock images for intelligent lithology identification. *Comput. Geosci.* **2021**, *154*, 104799. [CrossRef]
30. Maula, I.; Amrizal, V.; Setianingrum, H.; Hakiem, N. Development of a Gemstone Type Identification System Based on HSV Space Colour Using an Artificial Neural Network Back Propagation Algorithm. In *Advances in Intelligent Systems Research, Proceedings of the International Conference on Science and Technology (ICOSAT 2017), Jakarta, Indonesia, 10 August 2017*; Atlantis Press: Dordrecht, The Netherlands, 2017; pp. 104–109.
31. Ostreika, A.; Pivoras, M.; Misevičius, A.; Skersys, T.; Paulauskas, L. Classification of Objects by Shape Applied to Amber Gemstone Classification. *Appl. Sci.* **2021**, *11*, 1024. [CrossRef]
32. Ostreika, A.; Pivoras, M.; Misevičius, A.; Skersys, T.; Paulauskas, L. Classification of Amber Gemstone Objects by Shape. *Preprints* **2020**, 2020080336. [CrossRef]
33. Rios, C.; Saito, R. Researching of the Deep Neural Network for Amber Gemstone Classification. Master's Thesis, Universitat Politècnica de Catalunya, Barcelona, Spain, 2018.
34. Sinkevičius, S.; Lipnickas, A.; Rimkus, K. Multiclass amber gemstones classification with various segmentation and committee strategies. In Proceedings of the 2013 IEEE 7th International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), Berlin, Germany, 12–14 September 2013; IEEE: New York, NY, USA, 2013; pp. 304–308.
35. Liu, X.; Mao, J. Research on Key Technology of Diamond Particle Detection Based on Machine Vision. In Proceedings of the 2018 2nd International Conference on Electronic Information Technology and Computer Engineering (EITCE 2018), Shanghai, China, 12–14 October 2018; EDP Sciences: Les Ulis, France, 2018; Volume 232, p. 02059.
36. Sinkevičius, S.; Lipnickas, A.; Rimkus, K. Amber Gemstones Sorting By Colour. *Elektron. Ir Elektrotechnika* **2017**, *23*, 10–14. [CrossRef]
37. Zhang, S.; Guo, Y. Measurement of Gem Colour Using a Computer Vision System: A Case Study with Jadeite-Jade. *Minerals* **2021**, *11*, 791. [CrossRef]
38. Wang, D.; Bischof, L.; Lagerstrom, R.; Hilsenstein, V.; Hornabrook, A.; Hornabrook, G. Automated Opal Grading by Imaging and Statistical Learning. *IEEE Trans. Syst. Man Cybern. Syst.* **2016**, *46*, 185–201. [CrossRef]
39. Loesdau, M. Towards a Computer Vision Based Quality Assessment of Tahitian Pearls. Ph.D. Thesis, Université de la Polynésie Française, Puna'auia, French Polynesia, 2017.
40. Gemstones Images. Available online: https://www.kaggle.com/lsind18/gemstones-images (accessed on 27 April 2021).
41. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [CrossRef]
42. Nixon M.S.; Aguado, A.S. (Eds.) *Feature Extraction & Image Processing for Computer Vision*, 3rd ed.; Academic Press: Oxford, UK, 2013.
43. Liu, Y.; Zhou, X.; Ma, W.-Y. Extracting Texture Features from Arbitrary-Shaped Regions for Image Retrieval. In Proceedings of the 2004 IEEE International Conference on Multimedia and Expo (ICME), Taipei, Taiwan, 27–30 June 2004; Volume 3, pp. 1891–1894.
44. Bianconi, F.; Fernández, A.; González, E.; Ribas, F. Texture Classification Through Combination of Sequential Colour Texture Classifiers. In *Progress in Pattern Recognition, Image Analysis and Applications*; Rueda, L., Mery, D., Kittler, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2007; pp. 231–240.
45. Belalia, A.; Belloulata, K.; Kpalma, K. Region-Based Image Retrieval Using Shape-Adaptive DCT. *Int. J. Multimed. Inf. Retr.* **2015**, *4*, 261–276. [CrossRef]
46. Feizi, A. High-Level Feature Extraction for Classification and Person Re-Identification. *IEEE Sens. J.* **2017**, *17*, 7064–7073. [CrossRef]

47. Kittler, J. *Feature Selection and Extraction*; Academic Press: New York, NY, USA, 1986; Chapter 3, pp. 59–83.

48. Cheng, H.D.; Jiang, X.H.; Sun, Y.; Wang, J. Color image segmentation: Advances and prospects. *Pattern Recognit.* **2001**, *34*, 2259–2281. [CrossRef]

49. Gevers, T.; Gijsenij, A.; van der Weijer, J.; Geusebroek, J.-M. Color Image Formation. In *Color in Computer Vision*; Kriss, M.A., MacDonald, L.W., Eds.; Wiley: Hoboken, NJ, USA, 2012; pp. 26–45.

50. Gonzalez, R.C.; Woods, R.E. *Digital Image Processing*; Pearson Prentice Hall: Upper Saddle River, NJ, USA, 2008.

51. Angulo, J. Morphological color image simplification by Saturation-controlled regional levelings. *Int. J. Pattern Recognit. Artif. Intell.* **2006**, *20*, 1207–1223. [CrossRef]

52. Reyes-Aldasoro, C.C.; Björndahl, M.A.; Akerman, S.; Ibrahim, J.; Griffiths, M.K.; Tozer, G.M. Online chromatic and scale-space microvessel-tracing analysis for transmitted light optical images. *Microvasc. Res.* **2012**, *84*, 330–339. [CrossRef] [PubMed]

53. Smith, A.R. Color gamut transform pairs. *ACM SIGGRAPH* **1978**, *12*, 12–19. [CrossRef]

54. Hartigan, J.A.; Wong, M.A. A K-Means Clustering Algorithm. *Appl. Stat.* **1979**, *28*, 100–108. [CrossRef]

55. Funt, B.V.; Finlayson, G.D. Color Constant Color Indexing. *IEEE Trans. Pattern Anal. Mach. Intell.* **1995**, *17*, 522–529. [CrossRef]

56. Reyes-Aldasoro, C.C. *Biomedical Image Analysis Recipes in MATLAB: For Life Scientists and Engineers*; Wiley-Blackwell: Chichester, UK, 2015.

57. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural Features for Image Classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *3*, 610–621. [CrossRef]

58. Bigun, J. Multidimensional Orientation Estimation with Applications to Texture Analysis and Optical Flow. *IEEE Trans. Pattern Anal. Mach. Intell.* **1991**, *13*, 775–790. [CrossRef]

59. Bovik, A.C.; Clark, M.; Geisler, W.S. Multichannel Texture Analysis Using Localized Spatial Filters. *IEEE Trans. Pattern Anal. Mach. Intell.* **1990**, *12*, 55–73. [CrossRef]

60. Cross, G.R.; Jain, A.K. Markov Random Field Texture Models. *IEEE Trans. Pattern Anal. Mach. Intell.* **1983**, *5*, 25–39. [CrossRef]

61. Reyes-Aldasoro, C.C.; Bhalerao, A. The Bhattacharyya Space for Feature Selection and Its Application to Texture Segmentation. *Pattern Recogn.* **2006**, *39*, 812–826. [CrossRef]

62. Tai, C.; Baba-Kishi, K. Microtexture Studies of PST and PZT Ceramics and PZT Thin Film by Electron Backscatter Diffraction Patterns. *Textures Microstruct.* **2002**, *35*, 71–86. [CrossRef]

63. Carrillat, A.; Randen, T.; Sonneland, L.; Elvebakk, G. Seismic Stratigraphic Mapping of Carbonate Mounds using 3D Texture Attributes. In Proceedings of the 64th EAGE Conference & Exhibition, Florence, Italy, 27–30 May 2002; European Association of Geoscientists and Engineers: Houten, The Netherlands, 2002.

64. Bianconi, F.; González, E.; Fernández, A.; Saetta, S.A. Automatic Classification of Granite Tiles Through Colour and Texture Features. *Expert Syst. Appl.* **2012**, *39*, 11212–11218. [CrossRef]

65. Reyes Aldasoro, C.C.; Bhalerao, A. Volumetric Texture Segmentation by Discriminant Feature Selection and Multiresolution Classification. *IEEE Trans. Med. Imaging* **2007**, *26*, 1–14. [CrossRef] [PubMed]

66. Kovalev, V.A.; Petrou, M.; Bondar, Y.S. Texture Anisotropy in 3D Images. *IEEE Trans. Image Process.* **1999**, *8*, 346–360. [CrossRef] [PubMed]

67. Kather, J.N.; Weis, C.A.; Bianconi, F.; Melchers, S.M.; Schad, L.R.; Gaiser, T.; Marx, A.; Zollner, F. Multi-class Texture Analysis in Colorectal Cancer Histology. *Sci. Rep.* **2016**, *6*, 27988. [CrossRef] [PubMed]

68. Ojala, T.; Pietikäinen, M.; Harwood, D. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognit.* **1996**, *29*, 51–59. [CrossRef]

69. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic Minority Over-sampling Technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [CrossRef]

70. Bishop, C.M. Linear models for classification. In *Pattern Recognition and Machine Learning*; Jordan, M., Kleinberg, J., Schölkopf, B., Eds.; Springer: New York, NY, USA, 2006; pp. 179–224.

71. Li, T.; Zhu, S.; Ogihara, M. Using discriminant analysis for multi-class classification: An experimental investigation. *Knowl. Inf. Syst.* **2006**, *10*, 453–472. [CrossRef]

72. Cover, T.M.; Hart, P.E. Nearest Neighbor Pattern Classification. *Knowl. Inf. Syst.* **1967**, *13*, 21–27. [CrossRef]

73. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. Review of Classification and Regression Trees. *Biometrics* **1984**, *40*, 874.

74. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

75. Criminisi, A.; Shotton, J. *Decision Forests for Computer Vision and Medical Image Analysis*; Springer: Cham, Switzerland, 2013; 366p.

76. Taheri, S.; Mammadov, M. Learning the Naive Bayes Classifier with Optimization Models. *Rocznik* **2013**, *23*, 787–795. [CrossRef]

77. Crammer, K.; Singer, Y. On the Algorithmic Implementation of Multiclass Kernel-based Vector Machines. *J. Mach. Learn. Res.* **2001**, *2*, 265–292.

78. Feurer, M.; Hutter, F. Hyperparameter Optimization. In *Automated Machine Learning*; Springer: Cham, Switzerland, 2019; pp. 3–33.

79. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016. Available online: http://www.deeplearningbook.org (accessed on 31 December 2021).

80. LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436–444. [CrossRef]

81. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. ImageNet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [CrossRef]

82. Shao, L.; Zhu, F.; Li, X. Transfer Learning for Visual Categorization: A Survey. *IEEE Trans. Neural Netw. Learn. Syst.* **2005**, *26*, 1019–1034. [CrossRef]
83. Rezende, E.; Ruppert, G.; Carvalho, T.; Ramos, F.; de Geus, P. Malicious Software Classification Using Transfer Learning of ResNet-50 Deep Neural Network. In Proceedings of the 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), Cancun, Mexico, 18–21 December 2017; IEEE: New York, NY, USA, 2017; pp. 1011–1014.
84. Reddy, A.S.B.; Juliet, D.S. Transfer Learning with ResNet-50 for Malaria Cell-Image Classification. In Proceedings of the 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 4–6 April 2019; IEEE: New York, NY, USA, 2019; pp. 945–949.
85. Miglani, V.; Bhatia, M. Skin Lesion Classification: A Transfer Learning Approach Using EfficientNets. In Proceedings of the 2020 International Conference on Advanced Machine Learning Technologies and Applications (AMLTA), Jaipur, India, 13–15 February 2019; Springer: Singapore, 2020; pp. 315–324.
86. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 30 June 2016; IEEE: New York, NY, USA, 2016; pp. 770–778.
87. Hossin, M.; Sulaiman, M.N. A Review of Evaluation Metrics For Data Classification Evaluations. *Int. J. Data Min. Knowl. Manag. Process* **2015**, *5*, 1–11.
88. Lapin, M.; Hein, M.; Schiele, B. Top-k Multiclass SVM. *arXiv* **2015**, arXiv:1511.06683.
89. Reif, M.; Shafait, F.; Dengel, A. Prediction of Classifier Training Time Including Parameter Optimization. In *KI 2011: Advances in Artificial Intelligence*; Springer: Berlin, Germany, 2011; pp. 260–271.
90. Tharwat, A. Classification assessment methods. *Appl. Comput. Inform.* **2020**, *17*, 168–192. [CrossRef]
91. Okazawa, A.; Takahada, T.; Harada, T. Simultaneous Transparent and Non-Transparent Object Segmentation With Multispectral Scenes. In Proceedings of the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Macau, China, 3–8 November 2019; IEEE: New York, NY, USA, 2019; pp. 4977–4984.
92. Verma, A.; Banerji, S.; Liu, C. A New Color SIFT Descriptor and Methods for Image Category Classification. In Proceedings of the 2010 IRAST International Congress on Computer Applications and Computational Science (CACS 2010), Singapore, 4–6 December 2010; International Research Alliance for Science and Technology: Singapore, 2010; pp. 819–822.