Ophthalmopathy Disease Detection

Siraphat Boonchan
Asia Institute Technology
Bangkok, Thailand
st121415@ait.asia

Phanchayapit Chanyapisit
Asia Institute Technology
Bangkok, Thailand
st121531@ait.asia

Arkawat Thammanitayakul
Asia Institute Technology
Bangkok, Thailand
st121709@ait.asia

Natnaree Sornkongdang
Asia Institute Technology
Bangkok, Thailand
st121730@ait.asia

ABSTRACT

Globally technology access to medicine extensively, 1 billion people have a vision impairment that could have been prevented or has yet to be addressed. These 1 billion people include those with moderate or severe distance vision impairment or blindness due to unaddressed refractive error (123.7 million), cataract (65.2 million), glaucoma (6.9 million), respectively. Furthermore, those people have a high chance of invisible vision that is caused by unaddressed presbyopia (826 million). As we can see that the most popular of the anterior eye diseases are cataract and glaucoma, which can be detected by using the images. However, we defined that Pterygium is a disease on the anterior eye. Therefore, we will include the Pterygium in this paper and detect it by using a convolutional neural network (CNN) to classify these eye diseases. The current method of ophthalmopathy classification must be diagnosed by doctors, which is time-consuming in queues and dispose of human error. Therefore, creating the Mathematics model by using deep learning techniques can take the role of ophthalmopathy classification and provide useful information to remind early severity of the stage.

1 MOTIVATION

Base on the information from the Bureau of Information Office of the Permanent Secretary of Ministry of Public Health of Thailand, more than 200,000 eye disorder patients have been surgeries in 2018 [1]. Eye disorder includes Pinguecula, Cataract, and Pterygium which are the reason for the loss of vision.

Pinguecula is a yellowish patch or bumps on the conjunctiva, near the cornea. It is a change in the normal tissue that results in a deposit of protein, fat, and/or calcium.

Cataract is a clouding of the lens in the eye that affects vision. Most cataracts are related to aging. Cataracts are very common in older people.

Pterygium is a triangular-shaped growth of fleshy tissue on the white of the eye that eventually extends over the cornea. This growth may remain small or grow large enough to interfere with vision.

For example, people may not notice that they have a cataract. But over time, these symptoms can make the vision blurry, hazy, or less colorful which may cause trouble in reading or doing other everyday activities. Surgery is recommended when symptoms prevent the doing of daily activities. However, surgery has a high cost and risk. If patients did not get treatment in time, it could be leading to legal blindness or even total blindness. Thus, there will be more beneficial for patients to restore vision easier and cost cheaper to get back to view clearly.

To classify these eye diseases which are Pinguecula, Cataract, and Pterygium, patients have to meet doctors or experts to classify the problem. The purpose of this study is to use the knowledge of programming techniques and machine learning to help in a real-world problem. The research was conducted with samples of Pinguecula, Cataract, and Pterygium. The idea of this project is to help eye disorder patients to primary classify if they truly have eye disorder, to save time for medical personnel, and to save medical cost.

2 LITERATURE REVIEWS

2.1 Tournament Based Ranking CNN for the Cataract grading

There are many types of research aim at Cataract. This research aims to rank a convolution neural network (CNN) for the state of Cataract. To aid Cataract grading, there are several approaches to the automation of cataract grading. However, this research did not consider the characteristics of Cataract and medical dataset. Based on this research, The dataset is images of each state of cataract which had a specified beginning from Slit-Lamp (Medical equipment).

They have six states: start from state one to state six which the opacity gets worse:They had six strategies to measure the performance of their model. Firstly, the tournament structure is based on AUC. Secondly, the tournament structure is based on balancing the number of images. Thirdly, tournament structure based on Balancing the number of classes. Fourthly, tournament structure based on Balancing the number of classes. Fifthly, training, and Prediction based on Tournament structure. Sixthly, binary deep convolutional neural network. For each strategy, the main point is that they divided class sets into two subsets and classify binary classification. Then, the model classification through into subsets [2].

The result is that an imbalanced number of images in each class can lead to making mistakes to predict the outcome. This is because the model did not have enough images in some classes which have small images and get higher accuracy in large classes that have large images.

2.2 Semantic segmentation of colour eye images for improving iris segmentation

In the last years, there have many types of research which in Iris segmentation under visible spectrum (VIS) topic has been gaining attention, due to the interest in iris recognition in a-distance and non-cooperative environments such as blur, off-axis, occlusions, specular re- flections among the others.

This paper will separate the iris region from the non-iris regions on images affected by the above factors. The main method is boundary-based which generally uses techniques for detecting derangement of the eye by specular reflections. After that this research will use algorithms with the detection of non-iris regions. The general iris-segmentation algorithm is a low-level feature for locating the iris region. However, they choose semantic segmentation on color eye images method to improve the iris segmentation.

The first step, using Felzenszwalb's segmentation algorithm. This algorithm presents Region Adjacency Graph (RAG), where each vertex represents a region and each edge represents the adjacency between regions. This research using varying algorithm parameters to avoid under and over-fitting problems. Then trained all-region in the datasets.

Next, create a hierarchy on top of each initial segmentation to improve the classification method. Then apply a criterion by contracting the edges of the graph, joining the regions connected to the edges for semantic information. For small information, it will use Canny edge detection to form an important edge of the images.

After the new segmentation level is created, they operate the classification process on the level and create a new level on top of this one. This process can be repeated until no further contractions can be performed. Then this research have the segmentation hierarchy and the first classification of all levels, a refinement is operated through a hierarchical Markov Random Field (MRF)

Within the segmentation hierarchy, an MRF is built per level. Each MRF will receive information from the MRFs computed in adjacent levels in the hierarchy. In this way, it is possible to link spatial information within each level of the segmentation and the hierarchical information between levels. Furthermore, one MRF per segmentation ensures a smoothing in terms of region labeling, by taking into account the spatial relations among the regions. However, some result of classification is wrong so, they will put annotated with more relative clauses at the end of the optimization process [3].

In conclusion, this paper illustrates the segmentation algorithm for eye images which in under VIS. Moreover, the paper shows more on different classes of eye images and using annotation of eye images of 9 semantics classes. However, We see that this project can be expanded semantic on the diseases of the eye to prevent disease occurring in the eyes which is the most common in Thailand.

2.3 Automated Pterygium Detection in Anterior Segment Photographed Images using Deep Convolutional Neural Network

This paper demonstrates the automated detection of pterygium disease in anterior segmented photographed images. In the Asian continent, the pterygium popularity rate is high among the population indicates that 1 out of 10 adults in the age range of 21 and above potentially suffer from this disease. Normally, physical initial considerations detect the presence of pterygium by using a slit lamp camera these techniques have a positive impact on medical by using anterior segment photographed image (ASPI).

In the first method, They preparation ASPI data by converting to grayscale and rescaled the image to 50x50 pixels, by using the bilinear interpolation method. However, a rescale image needs training a smaller number of parameters only needs a small amount of training data and several iterations.

Next, set hyperparameter values which consisted of some epochs varied with the minimum of 5 until 25, with increment of 5, the learning rate of 0.0001, 0.001 and 0.1 with a fixed convolutional layer of 3 and batch size of 100.

Then, the proposed DCNN architecture was used before the pterygium and normal ASPI were classified on the data output. A sequential type of model was generated, and this API model enabled layer-by-layer modeling of problems that have occurred. The loss function is binary-cross-entropy because the classification in the research involved only two classes. The type of optimization used was 'Adam (Adaptive Moment Optimisation) optimizer' because of its efficient computing, minimal memory usage, and suitability for static data [4].

In conclusion, the processes will meet the needs of this analysis, the value of sensitivity, specificity, accuracy, F1 scores, and the Matthew correlation coefficient (MCC) were calculated according to the classification performed in the previous step. the research of pterygium detection method using the DCNN architecture with the proposed hyperparameter combination successfully. Our team sees that we can perform more than one of the diseases occurring in the eyes with can be seen or use the ordinary digital cameras image which also indicates in Asian people.

2.4 Automatic Cataract Grading Methods based on Deep Learning

Another interesting research of Hongyan Zhang et al. developed an algorithm and platform to diagnose and grade cataracts automatically based on fundus images.

The first step is preprocessing that they performed SGRIF to improve the contrast of retina images by simply taking specification to get unified rectangular fundus images.

The second step is extracting high-level features and texture features by using ResNet18, which uses 17 residual convolution layers to obtain semantic information to getting 512-dimensional high-level feature extraction. The deep neural network ResNet18 composes of 3 modules including shallow, remaining, and pool modules. The shallow module consists of many convolutional layers that are used to extract features of fundus images. The residual module used ReLU as an activation function by adding the element by element. The shortcut connection and the residual unit are combined as the output layer of the residual block. The pooling module is used for data dimension reduction and over-fitting prevention. Besides, they manually perform extraction of the texture feature of cataract fundus images by using Gray level co-occurrence matrix (GLCM) to compose a 72-dimensional feature dataset.

The final step is using a support vector machine as a base-learner to classify texture features and high-level features and is procuring 12 probability values for each fundus image [5].

We have tried to use the Gray level co-occurrence matrix (GLCM) technique to extract features and trained features extracted by GLCM through a support vector machine. We achieve not good accuracy; our assumption is this technique may not be suitable for our dataset.

2.5 Artificial intelligence manages congenital cataract with individualized prediction and telehealth computing

Artificial intelligence manages congenital cataract with individualized prediction and telehealth computing

The researcher explores the feasibility of applying AI to improve the quality of follow-up care by applying Bayesian and deep-learning algorithms to create CC-Guardian with functions of prediction module that dispatching module that schedules individual follow-up based on the prediction results, and telehealth module with the using of slit-lamp images that makes intervention decisions in each follow-up examination.

Researcher performed a comparison among naive Bayes with random forest and neural network. Naive Bayes was constructed to score function to perform binary classifications. Random forest was developed by constructing a multitude of decision trees at training time and outputting the mode of the classes for classification.

The prediction module is trained to identify high-risk patients likely to suffer complications. The dispatching module is responsible for scheduling individual follow-up based on the prediction module results. In the telehealth module, a clinical decision regarding further treatment is made after each telehealth examination based on follow-up images and result values.

Researcher investigated the efficiency and cost-efficient benefits which include time, travel distance and expenditure before and after the use of system. For complication detection, prediction with an advanced schedule and accurate detection was considered a timelier intervention, and the time difference of intervention before and after the use of the system was calculated. If system can predict and detect successfully, this individual will have an additional visit and be detected at the visit.

The cost-efficient benefits were calculated based on the assumption that all follow-up visits using our agent were conducted via telehealth computing [6]. Travel distance savings was defined as the round-trip distance savings between the distance traveled from the patient's home address to the telehealth site and the distance the patient would have traveled. Time savings was defined as the round-trip time savings for travel to the telehealth sites compared to travel to an in-person consultation.

This research show that machine learning and AI can help reach high benefit on medical sector. It also shows the benefit of saving time and travel distance savings. However, this research using the medical camera called slit-lamp camera to capture the image and feed it to the model. As the slit-lamp is available in medical sector and highly cost and number of available sample is limited, we decide not to use this sample as the input.

3 SYSTEM ARCHITECTURE OVERVIEW

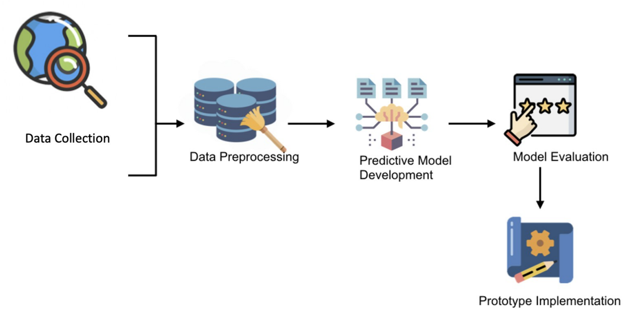

Figure 1: Performing process

Deep learning is another way to perform recognition of Ophthalmopathy. As shown in Figure 1. First of all, to collect the data from an internet data collection.

Next is the data preprocessing step, in this step, there are 4 main types of methodology including data augmentation and eye detection. For data augmentation, it will help to increase the number of datasets to be a larger one by blur, horizontal flip, random noise, and rotate the image. For eye detection, in a real situation, when people try to cap a picture of an eye, it may have other objects in that frame. Hence, object detection will be going to detect which part of the image is an eye and to crop only that detect part. Moreover, eye detection will help to increase the accuracy of the result as well. After the image was prepared, it will go to the prediction step. This step will use deep learning network architecture to predict and recognize the orchid species.

Then, it will need to evaluate the model in order to test that model whether it ready for use or not. This is an important step because it can tell the accuracy of each model. Moreover, it can use this to compare the accuracy to another model whether it better than the current model or not. Then, it will find the reason and solve the cons of model to be the best one.

Next is the prototype implementation step. This step will create a design pattern for the system in order to show both logically and physically with the interface presented. Finally, the users will test the entire system to improve the usability of the system interface.

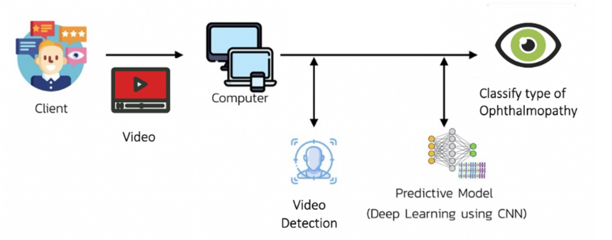

Figure 2: The Architecture Interaction between User and System

The architecture of the interaction between a user and the systems shown in Figure 2. The user will use the system via our platform. The users use the system by opening a camera on their computer. The system will use the video detection to detect only eye area as shown in Figure 3 and predictive model, which was taught to learn the set of training image, to recognize type of Ophthalmopathy from the input image that is captured as a frame and to show the type of Ophthalmopathy as shown in Figure 4.

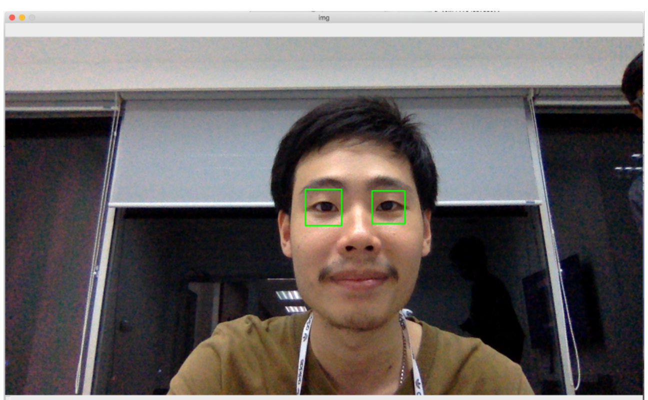

Figure 3: Video detection

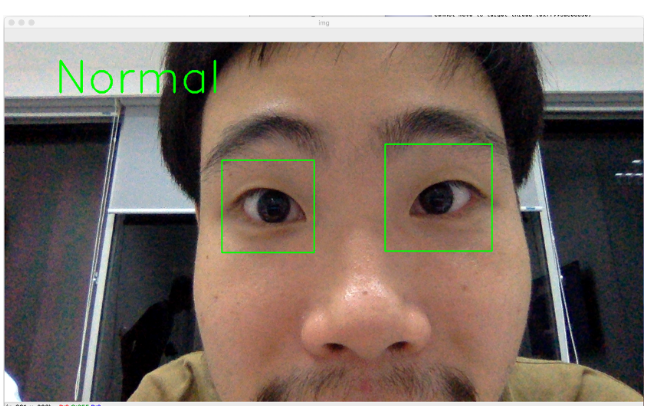

Figure 4: The Output of Predictive Model

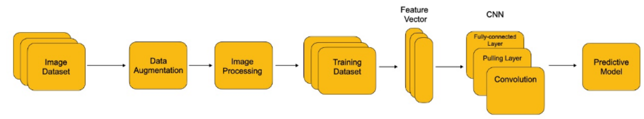

Figure 5: The Architecture of the Training process

The architecture is shown in the Figure 5 shows method to train the predictive model for classifying type of Ophthalmopathy. At the beginning, the system receives the set of the original image and use data augmentation to increase the number of images in training dataset because the number of images the set can increase the accuracy of result in deep learning. After that, we do the image processing with the augmented dataset and merge both data to be image dataset for training model. Feature Vector used to contain the multiple feature elements, which show a pixel or overall of the object image, as a vector. It is used in convolution neural network to train the data image to find the predictive model for classifying type of Ophthalmopathy as shown in Figure 3 and 4.

4 METHODOLOGY

4.1 Data Augmentation

Data augmentation [7][8][9] is a technique to increase the number of different images from the less image dataset because increasing image numbers can increase the percentage of precision and reduce the chance of mistakes in recognition by using deep learning algorithms. The quality of the result of deep learning algorithms depends on the quantity and diversity training dataset in some cases. The popular techniques are Flip, Rotation, Scale, Crop, Translation, and Gaussian Noise [10].

4.1.1 Flip.

There are 2 methods, which are the horizontal flip and a vertical flip. The example of a flip image in horizontal flip and vertical flip is shown in Figure 6.

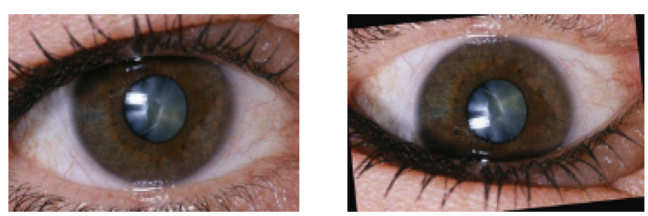

Figure 6: Vertical Flip Images

4.1.2 Noise. Noise is the random changing of color or shading of the pixel in the image. There are 4 main types of noise following:

4.1.1 Gaussian Noise Gaussian Noise is statistical noise known as a Gaussian Distribution, which has a probability density function equal to the normal distribution

4.1.2 Impulse Noise has 3 sub-types

4.1.2.1 Salt Noise is adding the random pixel with white or 255-pixel value.

4.1.2.2 Pepper Noise is adding the random pixel with black or 0-pixel value.

4.1.2.3 Salt and Pepper Noise is adding the random pixel with both white or 255-pixel value and black or 0-pixel value as shown in Figure 7.

4.1.3 Poisson Noise The appearance of this noise is seen because of the applied mathematics nature of electromagnetic waves like x-rays, visible light, and gamma rays and is also called quantum (photon) noise or short noise.

4.1.4 Speckle Noise is a granular noise that inherently exists in an image and degrades its quality. Speckle noise is often generated by multiplying random component values with totally different pixels of an image.

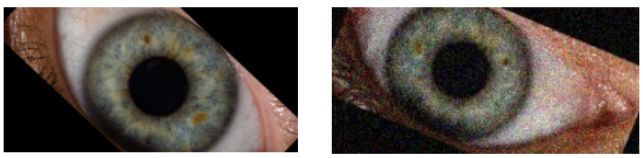

Figure 7: Example of Noise Which are Gaussian Noise and Salt and Pepper Noise

4.1.3 Blur.

The technique takes neighboring pixels and averages them in terms of reducing details and implementing what it recognizes as a blur. The different amounts of blur are the numbers of pixels to be included. It will evaluate a spread from a single pixel in both horizontal X and vertical Y as the standard deviation, which the size of standard deviation means hoe images are blurred as shown in Figure 8.

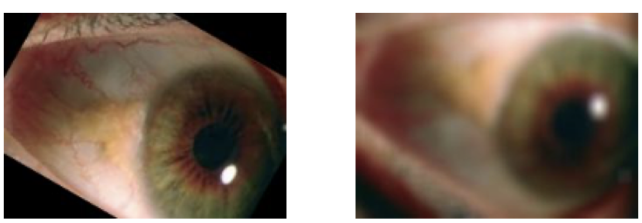

Figure 8: Blur Image

4.1.4 Rotation.

The technique rotates the image from several angles, which may not be preserved if the image does not rotate at the right angle. For instance, the rectangle which should rotate it by 180 degrees can preserve the image size because the size of the final image will change as shown in Figure 9.

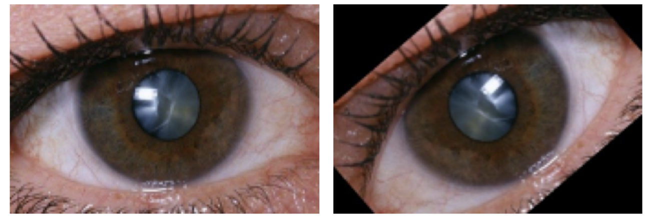

Figure 9: Rotate Image

4.2 Face Detection

In term of object detection, we use tools that called Haar feature-based cascade. This tool is an effective object detection method proposed by Paul Viola and Michael Jones in their paper, It is a machine learning-based approach where a cascade function is trained from a lot of positive and negative images. It is used to detect objects in other images. Figure 10 is an image to show the outcome of this method [11].

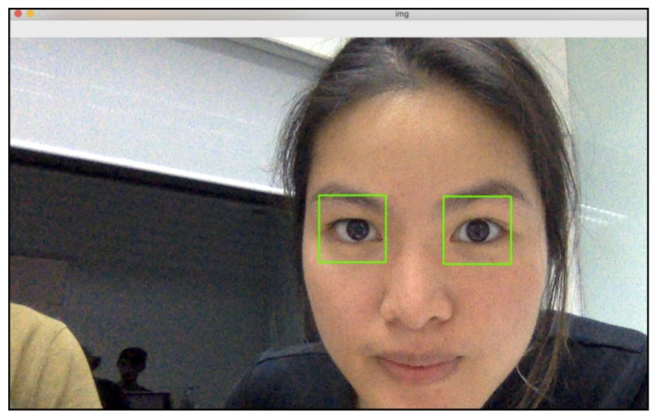

Figure 10: The Output of Face Detection

4.3 Convolutional Neural Network (CNN)

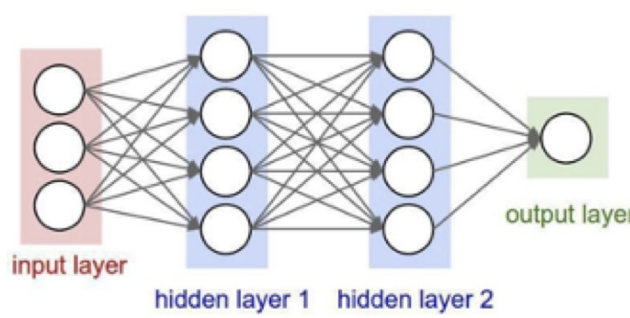

Figure 11: Convolutional Neural Networks [12]

Convolutional Neural Networks [12] [13] [14] as shown in Figure 11, which is a deep learning technique, is a network to simulate the human vision. Since human is normally visible things as several small areas, Convolutional Neural Networks will act as a compiler, who assemble those small areas to be represented in the understandable perspective. By the responsibility of the Convolutional Layer is calculating resultant layers (Feature Map), which usually consist of more than 1 layer.

Convolutional Neural Networks consists of 3 layers as the followings:

1. Convolutional Layer: This is the most important layer, which composes of filters and feature maps, will act as a feature-extractor, which each feature has a different weight based on the importance of input. This convolutional layer consists of the small size of filters, which were convoluted with the input volume where a given spatial place to learn the features. For every learnable filter, there had a formation through the width and height of the network of 2 dimensions feature maps as the filter slides and had the dot product computation of its entries and the input. After that, there had a combination of the reciprocal 2-dimension feature maps to build up the final output volume, which had the entries of learnable filters that considered only a definite spatial neighborhood in the input image.

2. Pooling Layer: This is the next layer of the Convolutional Layer, which is used as an intermediate layer in the network. It aims to perform sub-sampling or compress the incoming volume following the spatial dimension to reduce the size of the Feature Map. For instance, it downs the sample input volume size of 64x64x12 to 32x32x12. Therefore, it reduces the former layer of feature maps procured from the diverse filters to condense the over-fitting and computations of the network.

3. Fully Connected Layer: This is the Hidden Layer and Output Layer of Convolutional Neural Networks. It officiates for training and classifying objects. The Fully-Connected layer in Convolutional Neural Network contains neurons, which are fully connected to the neurons in the former layer as in Artificial Neural Network. This Fully-Connected layer is frequently collected as the final layer with SOFTMAX in the Convolutional Neural Network as its activation function for multiple class classification problems [15]; therefore, it is obligatory to forecast the final class label of the input image.

4.3.1 *ResNet18.*

The ResNet18 architecture includes 5 convolutional layers as shown in Figure 11. ResNet18 deep neural network is pre-trained on the collection of scene and object images imposed by the ILSVRC 2015 challenge. The output of a given layer is linearity that is used as a feature vector. The size of feature vector results in the computational cost and size; therefore, the dimensionality can be reduced by a reduction technique such as Principal Component Analysis [16].

Layer Name	Output Size	ResNet-18
conv1	$112 \times 112 \times 64$	7×7, 64, stride 2
conv2_x	$56 \times 56 \times 64$	3×3 max pool, stride 2 $\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$
conv3_x	$28 \times 28 \times 128$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$
conv4_x	$14 \times 14 \times 256$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$
conv5_x	$7 \times 7 \times 512$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$
average pool	$1 \times 1 \times 512$	7×7 average pool
fully connected	1000	512×1000 fully connections
softmax	1000	

Figure 12: ResNet18 Architecture

4.3.2 *ResNet101.*

The ResNet101 as shown in Figure 13 consists of 10 convolutional layers that we can load a pre-trained version if pre-trained network that is pre-trained on a collection of million images that can classify 1000 object categories from the ImageNet database. According to rich feature representations for a wide range of images, the input size of an image is 224x224 [17].

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$
conv3_x	28×28	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times8$
conv4_x	14×14	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times23$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times36$
conv5_x	7×7	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

Figure 13: ResNet101 Architecture

4.3.3 *Our Network.*

In neural network, Convolution Neural Network (CNN) is an important model which has been recognized to do images recognition, classifiable images. The model is shown in Figure 14, based on our project, the model has 3 convolutional layer for feature learning and 4 dense layers for classification. The output layers have 4 unit for each class (normal, cataract, pterygium, pinguecula).

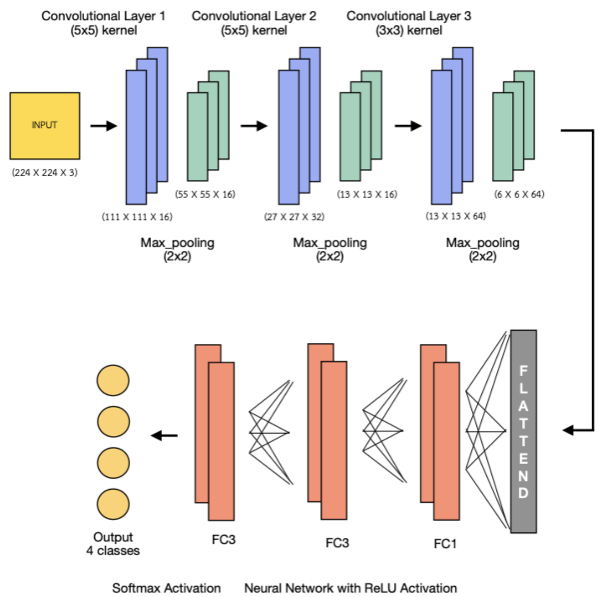

Figure 14: Our own network Architecture

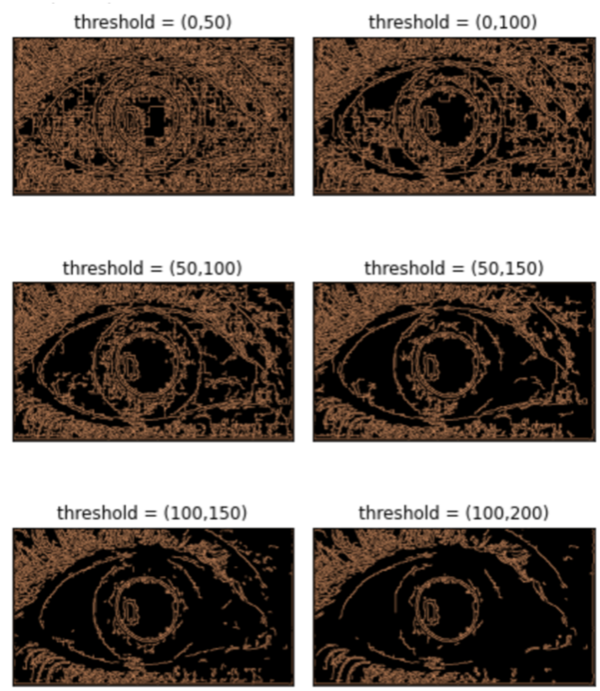

Figure 15: Images of Canny edge detection of normal eye

5 EXPERIMENTS

5.1 Experiment for preprocessing

5.1.1 Canny edge detector.

The Canny edge detector is an edge detection operator that uses a multi-stage algorithm to detect a wide range of edges in images.We used Canny algorithm for detect images which composed of noise reduction.We applying Gaussian blur to smooth it.The threshold is a change point value for screening where the edges should be.The process of this method, compute the boundary probability of each point is calculated [18]. Then the value is taken into account by looking at the minimum and maximum filter values: threshold is (0,50),(0,100),(50,100),(50,150),(100,150), and (100,200).The result from this filter will be the only two values: the edge and the non-edge only. The margins are 255 and non-margins are 0, the resulting data type is always unit8 as shown in Figure 15.

From the result, Canny edge detection of the images of eye is not work for preparation data processes because in our group project collect the diseases of eye in difference four type with tissue growth on the cornea of the eye.However we need to focus on detecting tissue but Canny edge detection strong edge with also concluding disturb environment.

5.1.2 Image detection using Histogram.

Histogram is considered as a graph or plot which is related to frequency of pixels in Gray Scale Image with pixel values (ranging from 0 to 255). Using threshold values to divide or filter portions of an image. Considering the brightness in the picture,It is generally used to determine the brightness of a single colour. If it's a colour image, convert it to black and white first. Therefore, in this chapter, black and white images are used or open the prepared image mainly in black and white mode [19]. We use function cv2.threshold() for filter image by considered the brightness as shown in the Figure 16
.

Faction	Value	Meaning
cv2.threshold_BINARY	0	If above threshold will be 255 where the rest are 0
cv2.threshold_BINARY_INV	1	If above threshold will be 0. where the rest are 255.
cv2.threshold_TRUNC	2	If it is higher than the threshold, it becomes the turning point. The rest remain the same.
cv2.threshold_TOZERO	3	If it is lower than the threshold, it becomes 0. The rest remain the same.
cv2.threshold_TOZERO_INV	4	If it is higher than the threshold, it becomes 0. The rest remain the same.
cv2.threshold_OTSU	8	Calculate automatic threshold without different from the input threshold.

Figure 16: Function of the threshold

The result of this check the generated by the first value of the modifier, which is usually required with the threshold for which the seed is loaded cv2.THRESH OTSU Which will automatically correct the turning point value.Create an array with values from 0-255 and filter at the value [10,60,110,160,210] and see the results of the methods as shown in the Figure 17 and 18.

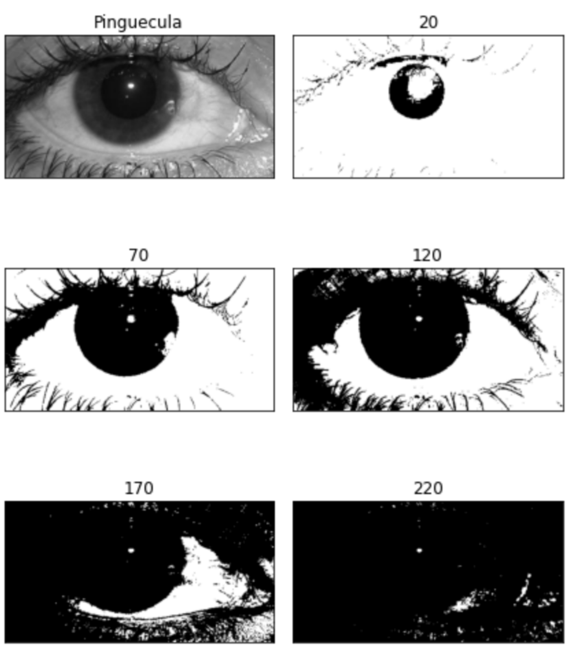

Figure 17: Image detection using Histogram of normal eye

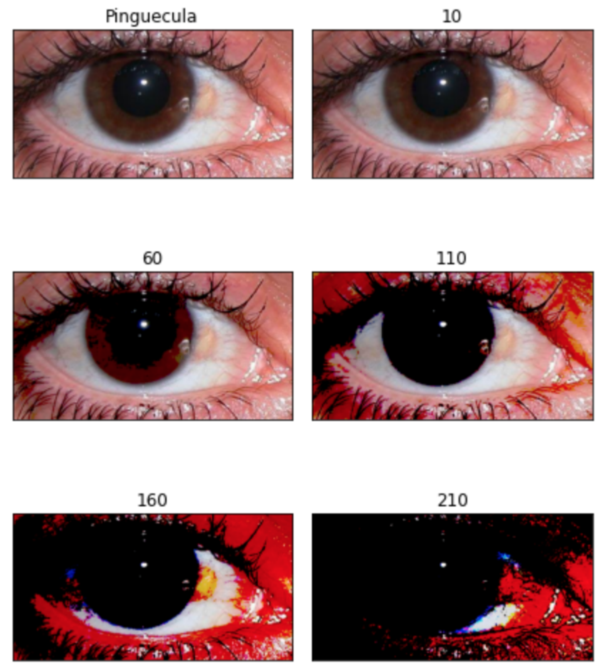

Figure 18: Image detection using Histogram of Pinguecula eye

From the result , we can compare that using Histogram detection on the Pinguecula for two difference way figure11 convert to gray scale and another one try to not convert to gray scale which focusing on the yellowish patch or bump on the conjunctiva, near the cornea but the colour is not too difference so the result is not good enough for detecting.

5.2 Experiment for deep neural model

5.2.1 Train a model by using Gray level co-occurrence matrix (GLCM) features through models of sklearn.

In this experiment, we have tried to train a model on the training dataset that consists of 4 categories of eyes including normal, Pinguecula, Pterygium, and Cataract eye. There are roughly 100 to 150 images per category, then Data Augmentation is applied to increase the number of images per category in original training dataset by flip, noise addition, blur, and rotation techniques; therefore, there are totally 4000 images. Then, we have tried to train a model on the training dataset that consists of 6 texture features including contrast, dissimilarity, homogeneity, energy, correlation, and angular second moment, which are six gray level co-occurrence matrices for images with 0, 30, 60, 90, 120, and 150 degree.

After that, we train this are six gray level co-occurrence matrices as shown in Table 1 through 10 models from sklearn as shown in Table 2 including K-Nearest Neighbors, linear Support Vector Machine, Radial Basis Function Support Vector Machine, Gaussian Process Classifier, Decision Tree, Random Forest Classifier, Neural

Network, AdaBoost, Gaussian Naïve Bayes, and Quadratic Discriminant Analysis, which we obtained accuracies of 36.2, 21.3, 40.4, 36.2, 42.6, 44.7, 34.0, 29.8, 31.9, and 29.8 respectively.

5.2.2 Train model by using ResNet18 deep neural network.
In this experiment, we have tried to train a model on the training dataset that consists of 4 categories of eyes including normal, Pinguecula, Pterygium, and Cataract eye. There are roughly 100 to 150 images per category, then Data Augmentation is applied to increase the number of images per category in original training dataset by flip, noise addition, blur, and rotation techniques; therefore, there are totally 4000 images is learned by the ResNet18 deep neural network. The accuracy from training the ResNet18 on a training dataset is 84.78 as shown in Figure 19.

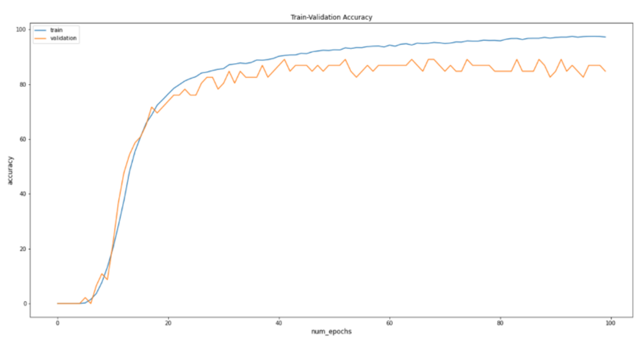

Figure 19: Accuracy of training and validation by ResNet18

5.2.3 Train model by using ResNet101 deep neural network.
According to the experiment in section 1, the accuracy needs to be improved. Hence, we decided to train the training dataset on the larger deep neural network, ResNet101, to improve the classification performance of a model. The accuracy from training the ResNet101 on a training dataset is improved by 2 percent, which is 86.96 as shown in Figure 20.

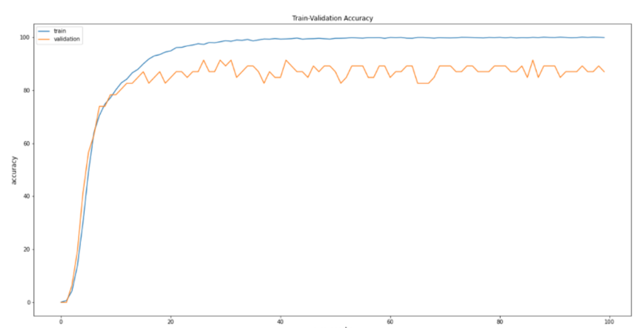

Figure 20: Accuracy of training and validation by ResNet101

5.2.4 Train model by using ResNet18 deep neural network.
We trained a model on the dataset as mentioned above. Then, after images through the model. The accuracy on train and validation datasets has significantly improved as shown in Figure 21, which are 98 percent on training set, 93 percent on validation set, and 80 percent on test set.

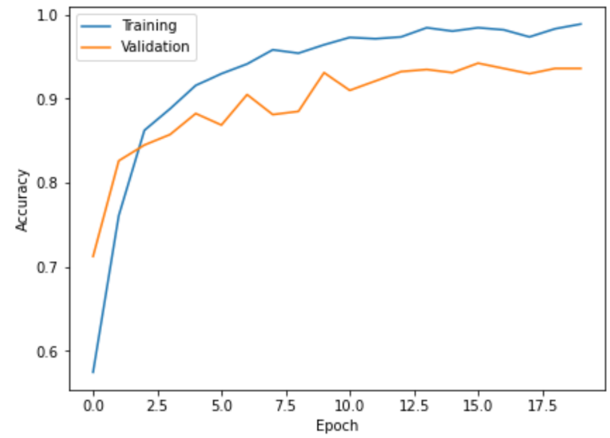

Figure 21: Accuracy of training and validation by Our Network

6 DISCUSSION

In the experiment, we used Canny algorithm for detect images and applied Gaussian blur to make it smooth. It can calculate the boundary probability of each point. Our project focus on detecting tissue. On the other hand, Canny algorithm detects strong edge with disturb environment.

We also tried image detection using Histogram which is normally use to determine the brightness of a single colour. However, the result from Histogram show very few difference of the yellowish patch or bump on the conjunctiva on the difference of diseases of eye.

Originally we had 100 to 150 images for each catagory so we did the augmentation to increase sample to totally 4000 images. Then We used Convolution Neural Network (CNN) as a model because CNN has good results on doing images recognition with the accuracy 98 percent of the training set and 93 percent of the validation set. We also did classifiable images with using of ResNet18 and ResNet101. The ResNet101 returns the accuracy of 99.8 percent of training set and 97 percent of validation set which is more than the the accuracy of Resnet18 which is 78 percent for both training and validation set.

The results of the research indicate that this application can detect the eye health problem which can be classify into 4 classed including normal, Pinguecula, Pterygium and Cataract eye with the accuracy of 80 percent of the unknown test set.

The generalizability of the results is limited by the quality of the taken eye images. The model need to get the clearly vision of both white and black eye. If the test image does not contain wrong condition and lots of noise, it can reach to wrong classification.

7 FUTURE SCOPE

As this project has been developed base on detecting and classification the normal, Pinguecula, Pterygium and Cataract eye on the test samples, the working future can be considered in many ways.

Base on the machine learning method, the scope of work can increase the classification to detect the state of each type of abnormal eye problem, Pinguecula, Pterygium and Cataract by teaching with more classification. In addition, every model want to have as much sample as posible. If we can get more sample with correctly classification data, the model will probably get more accuracy on both training and testing data. Moreover, Deep learning technique is an approach to improve the training accuracy. The streamlined training process is required. It has been observed that there has been no order technique that works best on any given issue. Thus we have to keep learning the new techniques and applying these techniques to model to see the progress of increasing accuracy.

For now, the program run base on python code using a build-in camera on Laptop, the interested developer can use the model of this paper to develop an application such as a mobile application.

On any ways, there will be lots of benefits for the healthcare section and patients and also for machine learning section.

ACKNOWLEDGMENTS

This work was supported by the [...] Research Fund of [...] (Number [...]). Additional funding was provided by [...] and [...]. We also thank [...] for contributing [...].

REFERENCES

[1] Who.int. 2020. Vision Impairment And Blindness. [online] Available at: <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment> [Accessed 1 December 2020].

[2] kim, D., Jun, T., Eom, Y., Kim, C. and Kim, D., 2019. Tournament Based Ranking CNN for the Cataract grading. 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),. Copy bibliography citation Copy in-text citation Check for grammar

[3] Roig, D., Morales-González, A. and Garea-Llano, E., 2018. Semantic Segmentation of Color Eye Images for Improving Iris Segmentation. Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, pp.466-474.

[4] Saad, A., 2019. Automated Pterygium Detection in Anterior Segment Photographed Images using Deep Convolutional Neural Network. International Journal of Advanced Trends in Computer Science and Engineering, 8(1.6), pp.225-232.

[5] Zhang, H., Niu, K., Xiong, Y., Yang, W., He, Z. and Song, H., 2019. Automatic cataract grading methods based on deep learning. Computer Methods and Programs in Biomedicine, 182, p.104978.

[6] Long, E., Chen, J., Wu, X., Liu, Z., Wang, L., Jiang, J., Li, W., Zhu, Y., Chen, C., Lin, Z., Li, J., Li, X., Chen, H., Guo, C., Zhao, L., Nie, D., Liu, X., Liu, X., Dong, Z., Yun, B., Wei, W., Xu, F., Lv, J., Li, M., Ling, S., Zhong, L., Chen, J., Zheng, Q., Zhang, L., Xiang, Y., Tan, G., Huang, K., Xiang, Y., Lin, D., Zhang, X., Dongye, M., Wang, D., Chen, W., Liu, X., Lin, H. and Liu, Y., 2020. Artificial intelligence manages congenital cataract with individualized prediction and telehealth computing. npj Digital Medicine, 3(1).

[6] "Data augmentation: boost your image dataset with few lines of Python", Medium, 2019. [Online]. Available:"https://medium.com/@thimblot/data-augmentation-boost-your-image-dataset-with-few-lines-of-python-155c2dc1baec.com"[Accessed: 1 Dec 2020].

[7] Mikołajczyk, A., Grochowski, M. (2018). Data augmentation for improving deep learning in image classification problem. 2018 International Interdisciplinary PhD Workshop (IIPhDW), 117–122. IEEE.

[8] Mikołajczyk, A., Grochowski, M. (2018). Data augmentation for improving deep learning in image classification problem. 2018 International Interdisciplinary PhD Workshop (IIPhDW), 117–122. IEEE.

[9] "Data Augmentation | How to use Deep Learning when you have Limited Data", AI Machine Learning Blog, 2019. [Online]. Available:"https://nanonets.com/blog/data-augmentation-how-to-use-deep-learning-when-you-have-limited-data-part-2/"[Accessed: 30 November 2020].

[10] Opencv-python-tutroals.read thedocs.io. 2020. FaceDetection Using Haar Cascades — Opencv-Python Tutorials 1 Documentation. [online] Available at: <https://opencv-python-tutroals.readthedocs.io/en/latest/pytutorials/pyobjdetect/pyfacedetection/ pyfacedetection.html> [Accessed 30 November 2020].

[11] Nattawat Raksaard and Olarik Surinta. (2018). Comparative Study Between Local Descriptors and Deep Learning for Silk Pattern Image Retrieval. [online] Journal.msu.ac.th. Available at: http://journal.msu.ac.th/upload/articles/article239299979.pdf [Accessed 28 November 2020].

[12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Communications of the ACM, vol. 60, no. 6, pp. 84–90, 2017.

[13] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural networks, vol. 61, pp. 85-117, 2015.

[14] I. Gogul and V. S. Kumar, "Flower species recognition system using convolution neural networks and transfer learning," In 2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN), 2017, pp. 1-6: IEEE.

[15] I. Gogul and V. S. Kumar, "Flower species recognition system using convolution neural networks and transfer learning," In 2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN), 2017, pp. 1-6: IEEE.

[16] 2020. [online] Available at: "https://www.researchgate.net/figure/ResNet-18-Architecturetbl1322476121" [Accessed 5 December 2020].

[17] Medium. 2020. Understanding And Implementing Architectures Of Resnet And Resnext For State-Of-The-Art Image.... [online] Available at: "https://medium.com/@14prakash/understanding-and-implementing-architectures-of-resnet-and-resnext-for-state-of-the-art-image-cf51669e1624"[Accessed 5 December 2020].

[18] S. Sahir, “Canny Edge Detection Step by Step in Python-Computer Vision,” Medium, 27-Jan-2019. [Online]. Available at: "https://towardsdatascience.com/canny-edge-detection-step-by-step-in-python-computer-vision-b49c3a2d8123" [Accessed: 5 December 2020].

[19] GeeksforGeeks. 2020. Opencv Python Program To Analyze An Image Using Histogram - Geeksforgeeks. [online] Available at: <https://www.geeksforgeeks.org/opencv-python-program-analyze-image-using-histogram/> [Accessed 5 December 2020].