ARTIFICIAL INTELLIGENCE POWERED DYNAMIC SKIN ANALYSIS APPLICATION

Yang Yu Han[1]*
Mohammad Kamrul Hasan[2]

[1,2] *Faculty of Information Technology & Science, National University of Malaysia, 43600 UKM Bangi,, Selangor Darul Ehsan, Malaysia*

ABSTRACT

The skin, body's outermost covering, shields the body from heat, light and injury. However, due to the atmospheric effect and sleeping disorders, it causes skin problems like acne. The original skin analysis and diagnosis method is manually operated and dependent upon the findings and experiences of expert physicians. However, due to the incubation period of skin problems, some of them cannot be found directly with the naked eye. Therefore, the early stage of skin condition diagnosis can significantly identify the issues for the treatment. AI-Powered dynamic skin analysis application is proposed that analyses skin problems by using real-time object detection algorithms. The proposed application considered the YOLOv5 model using the acne dataset. The proposed application can help pinpoint a person’s skin condition by using a large dataset and detecting skin problems has the ability to suggest the actions that need to be taken to preserve or enhance a person's skin condition. The main objective of the proposed application is to design skin analysis system to help users detect skin problems and give treatment advice. The study was performed using 786 images of the acne problem user dataset. The YOLOv5 model's best performance was evaluated in terms of accuracy, precision, recall and F1 score. The result suggests that the proposed application achieved 99.7% detection accuracy. Consequently, the model described is reliable and can be used to do acne detection to help people know the skin conditions. The impact of this work enhances the awareness of skin analysis, and prevention of skin problems with a minimal cost of skin analysis facilities.

Keywords: Skin analysis, YOLOv5, Artificial intelligence, Deep learning algorithm, Real-time object detection

Introduction

Acne, a prevalent dermatological condition affecting millions worldwide, poses significant challenges in diagnosis and treatment. Timely and accurate detection of acne lesions is crucial for effective management and prevention of potential complications. With the rapid advancement of technology and the growing interest in computer vision and image analysis, the development of automated acne detection systems has gained substantial attention in recent years. These systems aim to provide dermatologists, skincare professionals, and individuals with an efficient and objective tool for assessing acne severity, tracking progress, and guiding treatment decisions.

Nowadays, teenage girls and young women in particular frequently experience skin problems. Acne is the most frequent skin problem people experienced. Having skin problems not only harms one's physical health but also leads to psychological issues. Because acne can significantly impact how individuals perceive themselves. The presence of visible blemishes on the face or other parts of the body may lead to feelings of embarrassment. These negative self-perceptions can erode self-esteem and confidence, affecting one's overall psychological well-being. The skin also serves as the body's biggest defence against illness, thus proper skincare is essential. So, maintaining healthy skin helps to keep this barrier robust. Another factor that leads to skin issues frequently appearing is the absence of a platform to provide people with skin analysis and recommend skincare products with appropriate components. Because skin problems may vary over time and are influenced by the environment, it can be difficult to determine. So, to assist patients in progressively minimising their skin issues, the development of this skin analysis application is crucial.

In the literature review, a comprehensive analysis of the existing body of research on acne detection systems is presented. The objective is to evaluate the advancements made in this field, identify the methodologies employed, and highlight the challenges faced by researchers in achieving accurate and reliable results. By synthesizing the findings from a wide range of studies, aim to provide a valuable resource for both researchers and practitioners interested in developing or implementing acne detection application. AI-Powered skin analysis is mostly used to analyse facial skin by using image processing, and object detection method. The AI will look for signs of problems by analysing skin problems. The method of analysis and diagnosis are manually operated, dependent upon the findings and experiences of expert physicians. However, a wrong diagnosis, late diagnosis, or ignorance of diagnosis may result in adverse or life-threatening effects. The main target of this study is to design a computer application using artificial intelligence to enhance the diagnostic process of skin disease and bridge the gap between diagnosis and cure. The process of selecting the corresponding human skin pixel area in the image is known as the dynamic skin analysis system. The comprehensive assessment of acne is made possible using a camera to take pictures of the skin on the face and image processing technology. The meaning of developing the skin analysis system is to help users understand their skin conditions because many people blindly use skin care products in their daily life, and they do not have a comprehensive understanding of their skin conditions. If it happens to be neutral skin with good skin texture, then users will pay attention to the proper use of skin care products it will not have skin problems. The major goal of this study is to provide a flexible skin image analysis system with a variety of practical skin image analysis features. Additionally, it is made to be adaptable so that additional new features may be introduced in the future. The functions that are now in use include skin live image analysis, skin classifications, skin condition analysis, and skin image retrieval utilizing the Gabor wavelet transform, PCA (principal components analysis), and GLCM (greylevel co-occurrence matrix). The skin classification tool is the most important classification of skin and skin images. The image can distinguish between normal skin and problematic skin. It can also be used for deep learning neural networks (such as AlexNet (Jerry Wei, 2019), GoogLeNet (Richmond Alake, 2020) VGG19 (Melisa Bardhi ,2021) and ResNet101 (QingSheng Jiang et.el, 2019) for migration learning to achieve realistically. Convolutional neural networks are a sort of combination of biology, math and computer science, but these networks have been the most influential innovations in the field of computer vision and artificial intelligence (Laurence Van Elegem, 2022). The architecture of AlexNet consists of eight layers: five convolutional layers and three fully connected layers. But this isn't what makes AlexNet special. The three different AlexNet-based CNN systems considered are model 1, transfer-learned AlexNet; model 2, AlexNet-SVM; and model 3, AlexNet-KNN. Both SVM and KNN approaches used the error-correcting output codes (ECOC) method, whose main function is to decompose a multiclass classification problem into several binary ones, in classifying deep features according to their class label. GoogLeNet is a convolutional neural network that is 22 layers deep (Richmond Alake, 2020). GoogLeNet achieved this success with a 6.67% top-5 error rate in this contest. Instead of applying arbitrary convolutions like AlexNet, GoogLeNet has inception modules involving 1×11×1, 3×33×3, and 5×55×5 convolution sublayers with a 3×33×3 max pooling operation block acting as parallel arbitrary operations. For reducing calculation loss, the 1×11×1 convolution operation applies before the parallel convolution operations. However, the 1×11×1 convolution sublayer is placed after the max pooling layer inside the inception module. VGG19 was first developed, which is an enhanced version of VGG16(Beverly Hills MD, 2018). VGG takes in a 224x224 pixel RGB image. For the ImageNet competition, the authors cropped out the centre 224x224 patch in each image to keep the input image size consistent (Jerry Wei ,219). Since the VGG-19 model takes a colour image as input, a 3-channel image is created by assigning (red, green, and blue) channels, respectively. All regional images (from the refined region proposal) are cropped from the 3-channel image and scaled to 224×224×3 for VGG-19 training and testing. In this way, subtle changes over time are reflected in this 3-channel image and featured in the adapted VGG-19 model. The majority of academics nationally and internationally have recently conducted a great deal of research and inspection on skin detection technology and skin detection system, and have also made notable advancements, however, there are still a lot of unresolved issues. At the moment, skin problem detection accuracy has been challenged. Under uncooperative circumstances, occlusion is a very serious issue for skin state collection. When the test subject is wearing glasses, a mask, or other accessories, the collected face is insufficient, making it impossible for the skin detection system to conduct a global analysis. This impacts the next steps of feature extraction and recognition and can even result in the skin detection algorithm failing (Nojun Kwak, 2008). Therefore, it is necessary to mark the facial features before skin analysis, but the changes around the extracted feature points present a challenge. Even a small amount of occlusion won't have a significant impact. The image quality is the second issue. The image quality varies because of the various portrait acquisition tools. How to accurately detect skin for skin images with low resolution and poor quality is a problem worth thinking about. Since images with the same size and high definition are commonly used in skin detection, the problem of image quality can be solved. However, if go back to the real problem, the situation will become more complicated due to the different camera equipment used in skin detection, so need to optimize it. The software is for skin sensors that analyse things like moisture, wrinkles, and acnes. It may come in a variety of shapes and sizes, including cameras, sensors, smart mirrors, movies, facial imaging booths, smartphone attachments, and more. In tandem with digitization comes the demand for personalization. Consumers are increasingly dissatisfied with existing skin analysis solutions and look to electronics, gadgets, and even genetic tests to provide quantifiable results using scientific measurements. Furthermore, an emerging trend is hyper-personalized skin care, which can range from bespoke bottles formulated for a specific person, daily customized skin care based on current skin condition and local weather, to 3D face masks that address specific problem areas on the face(Nadia Tsao, 2018).Such products invariably rely on skin sensors to measure parameters such as skin moisture, wrinkles, sebum, elasticity, redness, dark spots, and more to provide the necessary data for personalization. Image recognition comes under the banner of computer vision which involves visual search, semantic segmentation, and identification of objects from images. The bottom line of image recognition is to come up with an algorithm that takes an image as an input and interprets it while designating labels and classes to that image. Most of the image classification algorithms such as bag-of-words, support vector machines (SVM), face landmark estimation, K-nearest neighbours (KNN), and logistic regression are used for image recognition also. Another algorithm Recurrent Neural Network (RNN) performs complicated image recognition tasks, for instance, writing descriptions of the image.

The main objective of this project is to design and develop an intelligent application for identifying skin problems and also analyse skin images using a designed intelligent application to help users know their skin condition better in the easiest way. Through this project, develop an application which can detect acne problems, and show the accuracy of the detection. And after detection can display the severity of the acne problem and give treatment advice are hoped.

The scope of this project is there's no fixed group of people who use this application, regardless of whether men or women, adults or teenagers can use this software for skin analysis. However, minors are recommended to be accompanied by their parents testing to prevent blind treatment after testing, blindly following the suggestions given by the system. The scope of the application also needs to make sure users are under sufficient light and the images uploaded to the system are clear and contain full face, or else may affect detection accuracy. And also, the application can only detect acne problems, if there's a more complicated skin problem, this application cannot do the detection.

This skin analysis application by creating a user-friendly application, increasing the accessibility of acne detection to a broader audience. People can conveniently use it from the comfort of their homes without the need to visit a dermatologist physically. And also the application provides a cost-effective alternative to in-person consultations with dermatologists. This can be particularly beneficial for those who may not have easy access to healthcare facilities or are looking for a preliminary assessment. This project achieved use YOLOv5 model to do real-time object detection. Users can not only upload images to do detection but also can use the front camera to take pictures to detect any time. after detection, the user can see the severity of the acne problem, which helps the user to know the skin conditions better. The detection accuracy will also be shown, making users can see the accuracy helps make this application more reliable.

To develop this application, a YOLOv5 model was developed. The first step is data preprocessing. Data preprocessing is a critical step in building any machine learning model, including YOLOv5-based skin analysis systems. The first step is data collection. The dataset it gets from online open resources are all authorised by the uploader. Need to ensure that the dataset has a diverse representation of skin problems to improve the model's generalization. Then next it's data cleaning, checking for any corrupted or invalid images in the dataset and removing them. Verify that the bounding box annotations are accurate and aligned properly with the corresponding skin regions. Next is to split the dataset into training, validation, and testing sets. A common split could be 70% for training, 15% for validation, and 15% for testing. Need to ensure that images from the same patient or subject do not appear in different sets to avoid data leakage. Then resize and normalize the image. Resize all images to a consistent resolution suitable for YOLOv5. The size is 416x416 pixels for the datasets. Normalize the pixel values of the images to a range of 0 to 1 or to have zero mean and unit variance. The last is to create a text file that contains the names of the classes in the dataset, one class name per line. This file will be used to map class indices to class names during training and inference. Next step is data augmentation. Data augmentation is a crucial technique to increase the diversity of training data, improve model generalization, and reduce overfitting. In this project, the data augmentation method "flip" data augmentation method was used. The "Flip" method for data augmentation involves flipping images horizontally and/or vertically. In the context of skin analysis using YOLOv5, horizontal and vertical flipping can be particularly useful as skin problems can appear in various orientations. Lastly it’s transfer learning.

Transfer learning in YOLOv5 refers to the process of taking a pre-trained YOLOv5 model, which has been trained on a large dataset for the acne detection task, and fine-tuning it on a new, smaller dataset for a different task. By leveraging the knowledge learned from the pre-training on the larger dataset, transfer learning allows the model to achieve better performance and faster convergence on the new task with limited data. In YOLOv5, transfer learning can be easily achieved using the --weights command-line argument or by modifying the model definition in the code. In this project, transfer learning reduced the overall training time. Since the YOLOv5 model is pre-trained on a large dataset, fine-tuning on a smaller acne problem dataset requires fewer training epochs. This can significantly reduce the overall training time compared to training the model from scratch. The most important is, by using transfer learning with YOLOv5, can quickly prototype and deploy the skin analysis application. The pre-trained model provides a solid foundation, allowing us to focus on optimizing hyperparameters and fine-tuning the model for acne detection tasks.

This report is organized in three parts, research methodology, results and discussion, and conclusions. The research methodology will explain the methods and approaches used in conducting the study. It also describes the specific development process model used and explains why the process model is selected. While the results and discussion will display the results of the study and the information obtained and give meaning and conclusion to the study which was conducted. The conclusion section will give an overview of the results and information obtained from the study. It also explains the strength and weaknesses of the project, and how to improve in future work.

Research Methodology

The development process is crucial because it provides the researcher with direction so that the data may be conveniently retrieved, gathered, recognized, and evaluated. Prototype development is the main component of the prototyping methodology, which is used in this study. This approach was selected because it lowers the possibility of having inaccurate user requirements. Additionally, prototyping is a technique for minimizing design flaws and removing failure causes at the early design stage. Additionally, the issue may be found much sooner, which will save future maintenance costs. The proposed solution integrates the artificial intelligence algorithm with the designed system.

For creating the skin analysis application, the data such as raw images without labelling will proceed using standard data pre-processing techniques, and then the data will be trained by using the YOLOv5 model. An interactive application will be approached for interpreting the skin data of users and analysis that simplify the usage and show the skin analysis results including skin problem, detection accuracy and treatment advice. The YOLOv5 model of this project is developed in the Jupyter Notebook environment, i.e. Google Colab platform and anaconda Jupyter Notebook. The first step is to do the data collection. Gather a dataset of skin images which got acne problems to do the skin analysis. Then train a deep-learning object detection model on a publicly available image set to predict acne in photos. The model which trained to achieve precise accuracy with an existing deep neural network architecture and no special preprocessing or hyperparameter evolution. The model was trained on both single-class (acne) detection and multi-class (severity levels 1-3, mild, moderate, severe). The dataset includes images of various acne problems, lesions, or different severity of skin problems. And it is important to ensure that the dataset covers a diverse range of objects and variations in object appearance, pose, and lighting conditions to enable the model to generalize well. Before training the model, need to split the dataset into separate sets for training, validation, and testing. The training set is used to train the model, the validation set helps in monitoring the model's performance during training and tuning hyperparameters, and the testing set is used to evaluate the final model's performance on unseen data. It's important to have a diverse, large, and representative dataset for training a robust model. The next step it’s to do data annotation and annotate the collected images by labelling the regions of interests (ROIs) relevant to the skin analysis task. Try labelling different types of acne like with different colours or shapes. The annotations include bounding box coordinates and corresponding class labels for each ROI. The dataset consists of 783 images that have been labelled using the "labelImg.py" tool, which is commonly used for annotating object detection datasets. The labelling process likely involved drawing bounding boxes around objects of interest in each image and assigning corresponding class labels. The dataset contains images of people from different backgrounds regarding age, region, and ethnicity. Labelling of the dataset involves one class which is acne. After the detection, it was divided into three levels, which are mild, moderate, and severe. The dataset was customized and named a real image-based labelled acne problem dataset by taking pictures of people with and without skin problems and labelling them manually. After annotating the image, the next step is image pre-processing. YOLOv5 requires fixed-size input images. Resize the images while maintaining the aspect ratio. Then resize the images to a specific resolution, which is 608x608 pixels. This step ensures consistent input dimensions for the model. Next is to normalize the pixel values. YOLOv5 commonly need pixel values in the range of 0 to 1 it divides the pixel values by the maximum pixel value to achieve normalization. Next is converting the images to RGB format, which YOLOv5 can recognize. To make sure the diversity and robustness of the training data, the data augmentation technique was used. Because applying flips, and rotations can make sense for the acne detection task. Data augmentation helps the model generalize better and handle variations in real-world data. YOLOv5 uses a specific label format, which includes the class index and bounding box coordinates relative to the image size. So, need to convert the annotation data (bounding box coordinates and class labels) into the YOLOv5-compatible format. For this process, the software was used to label its labelImg.py, which will be explained in detail later. The last step is to organize the pre-processed images and corresponding labels into the dataset structure which is suitable for training YOLOv5. Last it creates text files where each line contains the image file path followed by the labelled data. For model training, train the YOLOv5 model using the pre-processed and annotated dataset. This involves feeding the images and their corresponding annotations into the model and optimizing its parameters to learn to detect and classify the skin features of interest. The training process needs to define the network architecture, set hyperparameters, and perform forward and backwards passes through the network using techniques. Then split the dataset into separate sets for training, validation, and testing to do the model evaluation. The training set is used to train the model, the validation set helps in monitoring the model's performance during training and tuning hyperparameters, and the testing set is used to evaluate the final model's performance on unseen data. Additionally, it's also essential to ensure the accuracy and quality of the annotations in the dataset. So, that's why needed carefully reviewing the labelled images and manually verifying the annotations can help minimize errors and improve the overall performance of the model training. Lastly, when model training and evaluation are done, can deploy the model into the application. To deploy the model, need to integrate the model into a software application with coding by Python. The system can accept images as input and run the YOLOv5 model for object detection and classification. Then it can provide the detection information based on the acne which is detected.

The reason why choose YOLOv5 model it's because YOLOv5 is known for its speed and efficiency in object detection tasks. It can achieve real-time performance on various hardware platforms, including CPUs, GPUs. YOLOv5 has demonstrated competitive accuracy in object detection tasks, it strikes a good balance between speed and precision, which is crucial for user satisfaction in a real-time application. And also, YOLOv5's architecture is relatively simple and easy to understand, making it accessible for developers to implement and customize according to the specific needs. This simplicity also facilitates deployment on various platforms and devices, ensuring wider accessibility for users.

The dataset was collected manually because there's no real acne dataset resource online nowadays, so need to collect the dataset manually. Part of the dataset was collected from Kaggle, but even Kaggle got the dataset, still without labelling. The dataset contains images of people from different backgrounds regarding age, region, and ethnicity. Labelling of the dataset involves one class which is acne. The dataset was customized, so the dataset was named a real image-based labelled acne problem dataset by taking pictures of people with and without skin problems and labelling them manually. First, need to do is clearly define the specific skin analysis tasks want the YOLOv5 model to detect or classify. So, for this project, the main function is to detect acne, so to find a dataset which contains all acne it's important. Then try to look for relevant data sources to collect skin images. This dataset part from the online publicity dataset, and some from clinic real patients. Need to make sure to collect diverse and balanced data. It's better to collect a diverse range of skin images that cover various skin types, ages, genders, and ethnicities. Because it's important to have a balanced dataset that includes examples of different skin conditions or features in sufficient quantities. The selected dataset contains 783 acne images, some of which only have part of their face, so why choose this kind of dataset it's because it can highly improve the model training accuracy with different kinds of source images.

Data annotation is the process of labelling or annotating the regions of interest (ROIs) in an image with corresponding class labels and bounding box coordinates. Data annotation is crucial for training a YOLOv5 model because it provides the ground truth information needed for the model to learn to detect and classify objects accurately. For labelling the image, the software used is labelImg.py, because it supports downloading YOLO format text files for model training. In datasets lacking labels, labels are used to assist identify data components that were wished to train my model to recognize. To create a highly effective model for computer vision, good-quality datasets are necessary. According to the trash in, and garbage out principle, it is crucial to identify pictures appropriately while developing computer vision models.

For data augmentation method selection, since have already got the available dataset and done the image preprocessing. So, the next step is to process data augmentation. Because the object is small, the data augmentation strategy is important. In this project, the purpose is to help the model learn and extract more information about the details of acne through some data augmentation techniques. The mosaic data enhancement is another way of loading data. This method uses four images to train the network and stitch the four images together. Each image has its corresponding target's bounding box, and after stitching the four images together, a new image is obtained, which is randomly cropped out the same size as the input image, but contains the information of the four images, and obtains the corresponding target's bounding box of this image.

Generally, training a YOLOv5 model contains a few steps like, gather a dataset of images and labelling the collected dataset, then exporting the dataset to YOLOv5, training YOLOv5 to recognize the objects in the dataset and evaluating the YOLOv5 model's performance and last is run test inference to view the model at work.

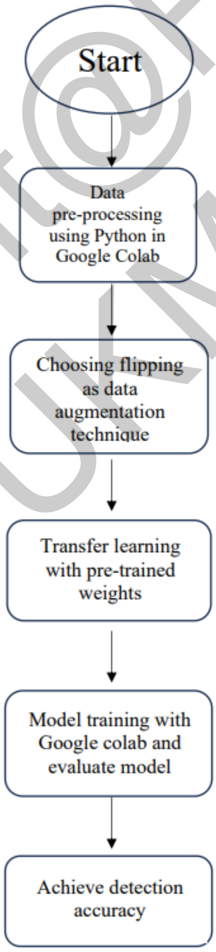

Figure 1 Proposed skin analysis application methodology

Results and discussions

Evaluating the YOLOv5 model is an essential step to assess its performance and determine how well it generalizes to new, unseen data. Evaluation helps to measure the model's accuracy, precision, recall, and other metrics, which are crucial for understanding its strengths and weaknesses.

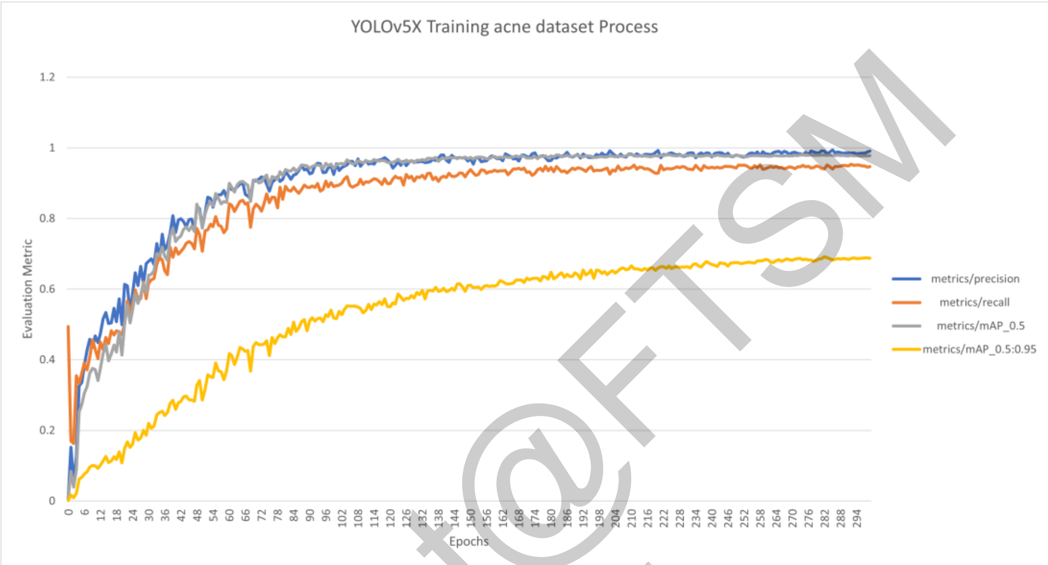

Figure 2 Yolov5X model training process line chart

Figure 2 shows the Yolov5 model training process. The blue line stands for precision, the orange line stands for recall, grey and yellow stand for mAP_0.5 and mAP_0.5:0.95 separately. the model was trained 300 epochs with times of 0.089 hours,and the accuracy of mAP@.5 is 0.995 and mAP@.5: .95 is 0.813. In this dataset, there’re 786 images in the evaluation set contain instances of any class. There are a total of 3144 ground truth labels (objects) across all images. Precision (P) is 0.995, indicating that 99.5% of the predicted instances for all classes are true positives. Recall (R) is 1, indicating that all ground truth instances for all classes are correctly detected. The mean average precision at the IoU threshold 0.5 (mAP@.5) is 0.995, suggesting high precision and recall at this threshold for all classes. The mean average precision across IoU thresholds 0.5 to 0.95 (mAP@.5:.95) is 0.813. Overall, the results indicate that the model has achieved high precision and recall for both the "all" class and the “acne" class. The mean average precision scores indicate the model's overall detection performance at different IoU thresholds, with mAP@.5 being particularly high. Figure 3 will show model annotation training results. The top left pictures show the classes, which means how many pictures got for one class. Since there's only one class 'acne' in my dataset, so there's only one red class, and the number of istances is 4000. Because to do data augmentation as stated above. The top right picture shows the size and number of labels boxes. it shows that there are around 2000 to 3000 bounding boxes in my dataset, and the size of them it's almost the same, it's all centred on the centre. Because the size of the images in the dataset has been trimmed into the same size when doing image preprocessing, the location of the annotation it's almost going to be the same. The bottom left image it's the coordinates of the centre point of the bounding boxes. (x, y) represents the coordinates of the top-left corner of the bounding box. The bottom right image shows the labels' width and height, and (w, h) represents the width and height of the bounding box.

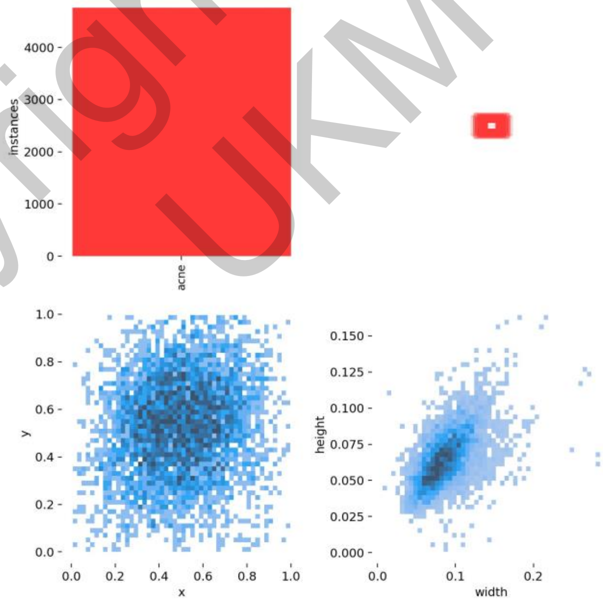

Figure 3 Model annotation training results

Next will show the precision curve. The precision rate indicates the proportion of examples that are classified as positive examples that are positive examples. The reason why needs the precision of the model it's it helps developers measure the model's ability to classify positive samples. When calculating the accuracy of the model, should consider both positive and negative samples of the classification. When a model correctly classifies most of the positive samples and has many false positive samples, the model is called a high-recall low-precision model. The accuracy of a machine learning model depends on both negative and positive samples. In accuracy, should consider all positive samples that are classified as correctly or falsely positive. Figure 4 below its precision curve which is the relationship between precision and confidence.

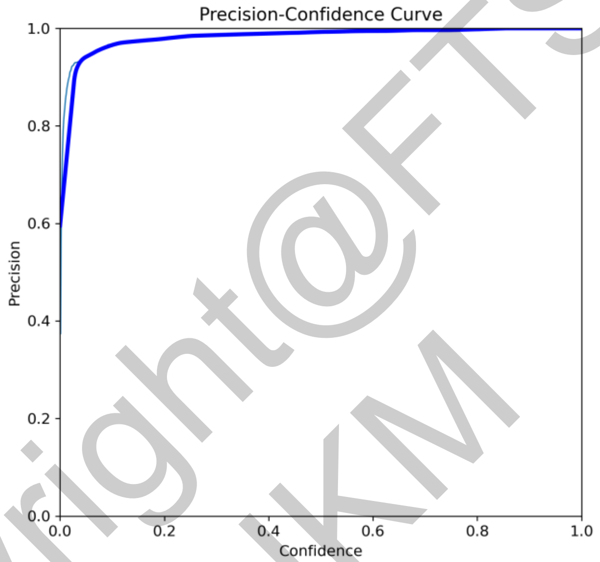

Figure 4 Relationship between precision and Confidence

Figure 4 means when the confidence level is set to a certain value, the accuracy rate of each category recognition. When the confidence is greater, the category detection is more accurate. This is also easy to understand. Only when the confidence level is high can it be judged to be a certain category. But in this case, some categories with low confidence will be missed. For example, when running the program, even if the target is acne, the model predicts that it is also acne, but the confidence level given to it is only 70%. When the confidence level is set to 80%, it is considered acne, and the target will be ignored. The system with high recall but low precision returns many results, but most of its predicted labels are incorrect when compared to the training labels. A system with high precision but low recall is just the opposite, returning very few results, but most of its predicted labels are correct when compared to the training labels. An ideal system with high precision and high recall will return many results, with all results labelled correctly. When the classes are severely unbalanced, precision recall is a helpful indicator of prediction success. In information retrieval, recall measures the quantity of relevant results returned, whereas precision measures the relevancy of the findings. The precision-recall curve shows the trade-off between precision and recall at different thresholds. A low false positive rate is associated with good accuracy, whereas a low false negative rate is associated with strong recall. great recall and great accuracy are both indicated by a large area under the curve. High results produced by the classifier with high accuracy and primarily high recall are indicated by high scores for both.

Next will show the relationship between recall and confidence in Figure 5.

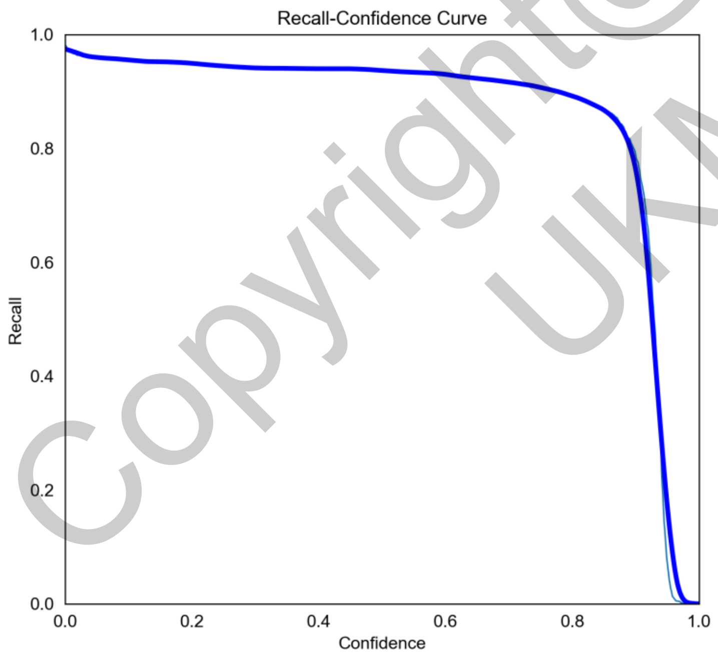

Figure 5 Relationship between recall and confidence

Precision and Recall are usually a pair of contradictory performance metrics. The higher the Precision, the lower the Recall. The reason is, if want to improve the Precision, that is, the positive examples predicted by the two classifiers are as true as possible, then need to increase the threshold for the two classifiers to predict positive examples. For example, before predicting positive examples as long as the samples with a confidence level of 0.5 were marked as positive examples, now need to increase the confidence level to 0.7 before marking them as positive examples, to ensure that the positive examples selected by the binary classifier are more likely to be positive examples. The real positive example: this goal is exactly the opposite of improving Recall. If want to improve Recall, that is, the two classifiers can pick out the real positive examples as much as possible, then it is necessary to lower the threshold for the two classifiers to predict positive examples, such as the previous prediction. For example, as long as the sample with a confidence level of 0.5 is marked as a real positive example, then if it is reduced to 0.3, will mark it as a positive example, to ensure that the binary classifier selects as many as possible true examples. The algorithm assigns a confidence level to each target.

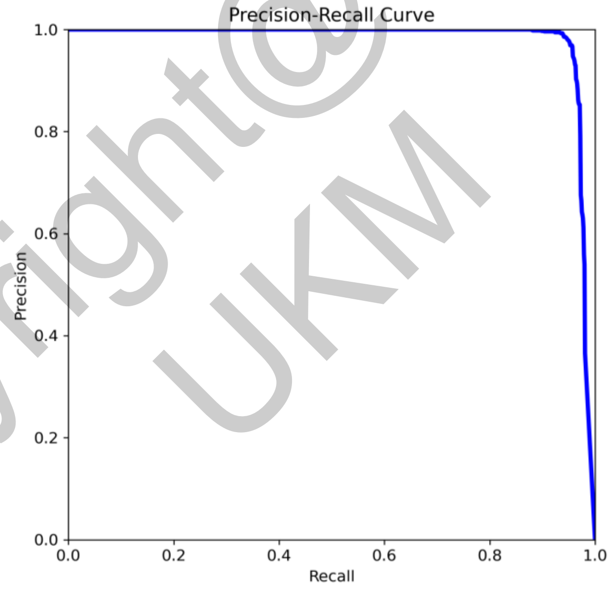

Figure 6 Relationship between recall and precision

The last part of the model training result analysis will be the summary results analysis. The detail will be shown in Figure 7.

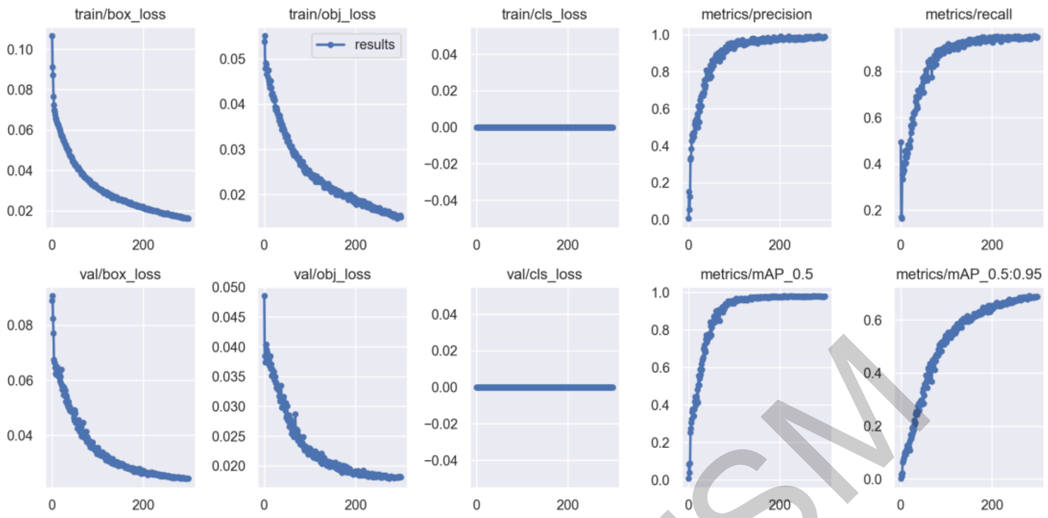

Figure 7 Genaral model training result

Overall, the general training results mainly observe the fluctuation of precision and recall rate. If the fluctuation is not too large, the training effect is better; if the training is better, the figure shows a steady increase. All the above it's the analysis of the model training results, in the next part, will compare the YOLO model which was rained for this project with past work.

This part will compare the YOLOv5 model for acne detection with past studies and other YOLO models. The YOLOv5 model includes 10 individual architectures ,however, usually only consider the first five for research, which are nano, small, medium, large, and X large. The main difference between them is the number of feature extraction modules and convolution kernels, after that, according to the actual situation, it is important to use the YOLO model with its models in the number of neural network parameters. In a previous study, these five YOLOv5 models were compared using the COCO val2017 corpus (Lin et al., 2014; Nelson & Solawetz, 2020).

Table 1. YOLOv5 models comparison on COCOval2017 dataset(Nelson & Solawetz, 2020).

Model	$mAP_{0.5}$	$mAP_{0.5:0.95}$	Training duration[s]
YOLOv5n	46.0	28.4	6.3
YOLOv5s	56.0	37.2	6.4

YOLOv5m	63.9	45.2	8.2
YOLOv5l	67.2	48.8	10.1
YOLOv5x	68.9	50.7	12.1

The main advantages of YOLOv5 over earlier versions are: 1) easy installation that only requires PyTorch and some lightweight Python; and libraries. 2) quick training. 3) functional inference ports on individual photos, batches of images, video feeds, or webcam ports. 4) an intuitive layout that is easy to traverse while working and is based on a typical file folder arrangement. 5) straightforward conversion to mobile platforms like smartphones and tablets with Core ML support. And stand in a developer perspective, YOLOv5 it's easier in training, because it is available in PyTorch using Jupyter notebooks or Google Colab tools, all of these got free GPU.

Better performance and greater object detection accuracy are indicated by a higher mAP. Except for comparing mAP and the need to compare training hours with different YOLOv5 models. The detail will be shown in Table 2 below.

Table 2 Duration of training of YOLOv5 models in hours on the experimental dataset.

Model	Training duration [h]
YOLOv5n	1.17
YOLOv5s	0.83
YOLOv5n	1.83
YOLOv5l	2.83
YOLOv5x	5.34

As seen in Table 2, YOLOv5x training took the longest, lasting almost five hours, and was completed in less than an hour. In general, the length of training grows proportionally to the number of model parameters. However, it is feasible that training of more complex models will take less time than training of simpler models due to the enabled early stopping capability, as has been the case with models. A restriction on free use has been put in place by *Google Colab* (i.e., the halting function). In this sense, all *Google Colab* notebooks have a 90-minute idle timeout and a 12-hour absolute timeout. The instance will shut down automatically if the user doesn't interact with their *Google Colab* notebook for more than 90 minutes.

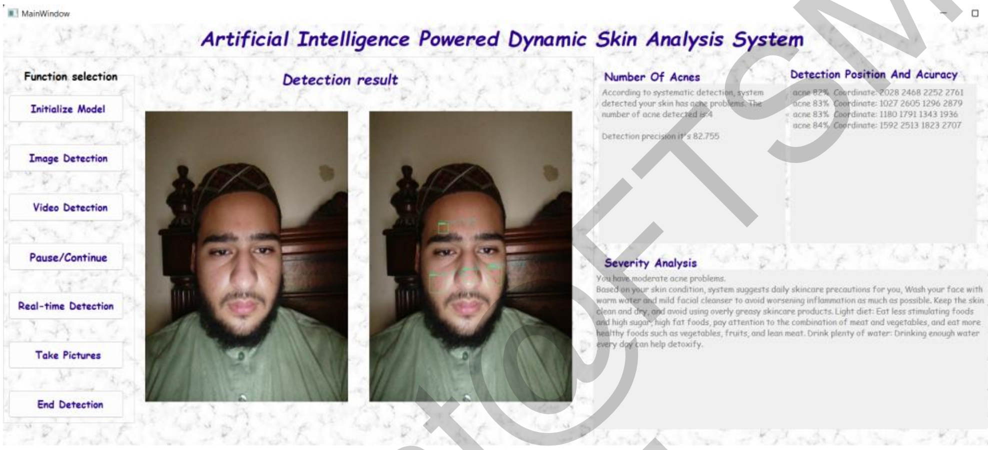

Figure 8 Moderate acne detection

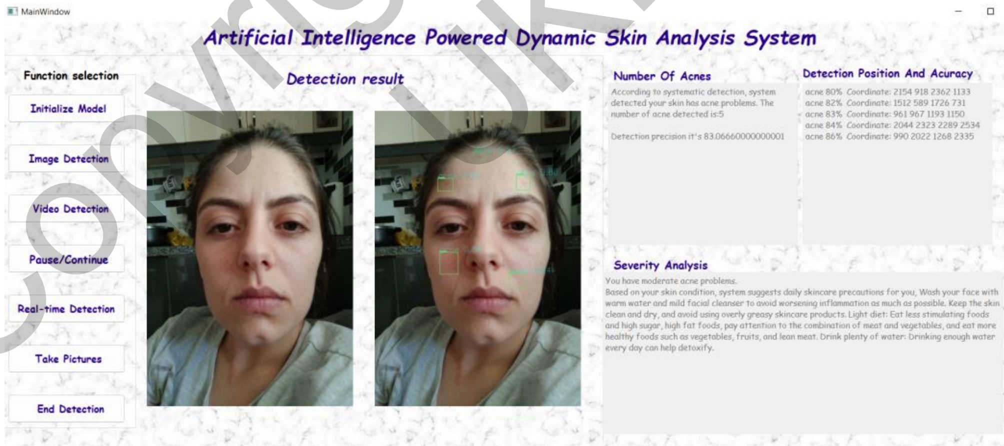

Figure 9 Moderate acne detection (2)

As Figures 8 and 9 shown above, it’s the detection result put into the interface, putting a picture into the application, and choosing image detection, can get the precision are 82.755% and 83.066%, which means that all the predicted bounding boxes (object detections) made by the model, approximately 82.755% or 83.066% of them are correctly localized and classified.

Moreover, can see from the interface display that both testers suffer from moderate acne, and the following will show mild acne in Figure 10.

Figure 10 Mild acne detection

As shown in Figure 10, can see that the system detects that the user has mild acne, the number of detections is one, and the accuracy is 70%, which means precision in YOLOv5 is often calculated as the number of true positive detections divided by the total number of detections (true positives + false positives). It represents the proportion of correctly detected objects among all the detections made by the model. so, in this detection, the true positive is 70%.

In future work, this application can detect not only acne problems are hoped, but also other problems, such as wrinkles, freckles, skin moisture, etc. This requires training a more functional model in the next work. Secondly, this system is currently unable to detect overly concentrated acne problems, because during the model training process, the data set is indeed the case, resulting in incomplete model learning. So, the next job is to find more diverse data sets, including acne data sets of different conditions and severities, and to label and train them. Finally, in future work, more than one model of YOLOv5 will be selected for training, and different models will be compared. For example, models such as YOLOv8 and R-CNN may be trained to compare which model is more suitable for acne detection tasks, and eventually, the model with the highest accuracy will be adopted to develop a more accurate and comprehensive skin detection system. In terms of system functions, users can be given more personalized suggestions for different skin problems.

Conclusion

The whole project focuses on training a YOLOv5 model to develop a skin analysis system which can help the user know their skin condition better. With the trained YOLOv5 model, this application achieved high accuracy acne detection task, with the model accuracy of 0.997. This stand for the model training is successful and can achieve a high accuracy acne detection task. The objective of set in the introduction section were fully achieved. The application can accept image as input from user, no matter upload from device or real-time detection, both can achieved a high accuracy detection. The YOLOv5-based acne detection system offers several significant advantages in the field of dermatology. By leveraging state-of-the-art object detection techniques, it provides a accurate and efficient solution for identifying and localizing acne lesions in medical images. The weakness of the application is, it can only detect one skin problem for right now, which is acne problem. And also, the images input into the system must be detected when the light is sufficient, and the acne cannot be very dense. For the model weakness, the YOLOv5 model may have difficulty detecting acne lesions in images with challenging lighting conditions, occlusions, or low image quality. Variations in image resolution, noise, and artefacts can affect the performance of the model, leading to potential false positives or false negatives.

In conclusion, the YOLOv5-based system exhibits a high degree of accuracy in acne detection, thanks to its deep neural network architecture and the extensive training performed on diverse and annotated acne image datasets. This ensures that the system can handle various acne types, including different severities and skin tones, enhancing its applicability across a wide range of patient populations. Moreover, the system's ability to detect acne lesions on both visible and infrared images further expands its potential applications. This versatility enables dermatologists to leverage different imaging modalities, providing a more comprehensive understanding of the underlying acne condition and facilitating more targeted treatment approaches.

Appreciation

First of all, I would like to take this opportunity to say thank you to my supervisor, Dr Mohammad Kamrul Hasan, who has been patient guidance and given instructions and relevant knowledge to me throughout the project. He has contributed various resources to help me to produce this project as well as improve my skill in the Convolutional Neutral Network field. I am also very grateful to the lectures from FTSM who have sprinkled me with knowledge throughout my studies at the National University of Malaysia (UKM). Furthermore, I would like to express my gratitude to my family members who are inspirations and always supports and prays for my success to complete this project and finish my studies. I also want to thank my friends who always give insight and guidance and mutually support each other to complete this project. Finally, I would like to thank the parties who were directly and indirectly involved in the process of producing this project.

Thank you so much.

References

Alshehri, Saleh. 2012. Neural Networks Performance for Skin Detection. *Journal of Emerging Trends in Computing and Information Science. 3. 1582-1585.*

Bochkovskiy, Alexey & Wang, Chien-Yao & Liao, Hong-yuan. 2020. *YOLOv4: Optimal Speed and Accuracy of Object Detection.*

Juwanda, F. S., & Mat Zin, H. 2021. The Development of Skin Analyser for Skin Type and Skin Problem Detection. *Journal of ICT in Education, 8(3), 27–37. https://doi.org/10.37134/jictie.vol8.sp.1.3.2021*.

Kamruzzaman, Sikder & Islam, Md. (2010). *An Algorithm to Extract Rules from Artificial Neural Networks for Medical Diagnosis Problems. Int. J. Inf. Tech.. 12.*

Srishilesh P S. 2022. *Understanding COCO dataset.* https://www.section.io/engineering-education/understanding-coco-dataset/ [18 March 2022].

Sumit Saha. 2018. *A Comprehensive Guide to Convolutional Neural Networks* — the ELI5 way | by Sumit Saha | Towards Data Science. https://towardsdatascience.com/a-comprehensive-guide-to-convolutionalneural-networks-the-eli5-way-3bd2b1164a53 [20 December 2022].

Yang Yu Han (A179276)

Dr. Mohammad Kamrul Hasan

Faculty Information Science and Tchnology,

Universiti Kebangsaan Malaysia