Applying Artificial Intelligence to Support the Detection and Treatment of Melasma

Van Lam Ho[1*], Vu Tuan Anh[2], Tran Xuan Viet[2].

[1]*Faculty of Information Technology, Quy Nhon University, Binh Dinh, Vietnam.*

[2]*Quyhoa National Leprosy Dermatology Hospital, Binh Dinh, Vietnam*

E-mail: hovanlam@qnu.edu.vn, drvtavn@gmail.com, thstranxuanviet@gmail.com.

Abstracts: This study, we propose a solution to apply artificial intelligence to assist in detecting whether a person may have melasma or not through data sets related to information about a person's daily activities, then If we detect a person with a high likelihood of having melasma, we will apply machine learning to diagnose the type of melasma through a photo taken of that person. Through a machine learning model of predicting and diagnosing a person's melasma, we also suggest relevant prevention and treatment options based on the disease's prevention and treatment regimen. Our method build predict Melasma model based on Catboost machine learning algorithm on users' data combined with medical practice data commu-nity by dermatologists to predict the disease and make some necessary recommendations in the patient screening. Based on our dataset, we have statistically described the data characteristics as well as the correlated data parameters that may cause Melasma. The method using for diagnosing melasma disease based on machine learning algorithms with input data being facial images. we built a machine learning model for diagnosing melasma to detect melasma objects to support dermatologists in predicting the risk of melasma in a person after entering his/her facial image. Our dataset of facial images combined with the expertise of melasma experts to classify different types of melasma. We used YOLO V8 with machine learning algorithms to detect melasma objects to build a diagnostic model for whether a patient has melasma and with which type of melasma such as central melasma, butterfly-shaped melasma, or mandibular melasma.

Keywords: Cat Boost algorithm, Melasma disease, Machine learning algorithm, Prediction Melasma model, Objects detection

1. INTRODUCTION

A Melasma is an acquired hypermelanosis with complex etiology and pathogenesis. The primary lesion of the disease is macules and/or dark brown, symmetrical patches in sunexposed areas. Common sites of infection are the cheeks, upper lip, chin, and forehead. Though the disease is benign, it greatly affects the psychology and aesthetics of the patients [1]. In women, the disease can be idiopathic or related to pregnancy [2].

About melasma are based on clinical indications with the following characteristics [24]:

- Melasma patches on both cheeks, no itch, no burning, no scab.

- Hyperpigmented macules in other areas such as eyebrows, private areas, chin, and nose.

Melasma is classified based on clinical [27], [28]:

- Central melasma: including cheeks, eyebrows, upper lip, nose, and chin.

- Butterfly-shaped melasma: localized hyperpigmentation in cheeks, and nose.

- Mandibular melasma: related to the lower jaw area.

To effectively treat melasma, a combination of three factors is required: broad-spectrum sun protection, topical opacifying agents, and elimination of calculated risks [29], [30]. In any case, melasma treatments are still limited and a challenge in dermatology [31]. Therefore, it is extremely important to predict and process random variables [32], [33], [35], [37].

Artificial intelligence (AI) plays an increasing role in medicine and healthcare by leveraging computer control, machine learning computing, and accessibility of massive data from treatment, examination, and health records of wearable devices.

In other words, Machine Learning is a field of Artificial Intelligence, machine learning means that computers can think on their own like humans. Another approach argues that machine learning is a method of drawing lines that represents the relationship of a data set [4], [5]. Combining the expertise of dermatologists with people's public information on Melasma [3], we used data analysis techniques to show correlations features of the data: descriptive analysis, visualisation data may help experts and people easily monitor the possibility of having Melasma through input data of a person's daily information. From the results of the analysis, we built a Catboost machine learning model with the input data of expertise of dermatologists specializing in Melasma combined with patients' community information, so that computer can support to predict whether a person is infected to Melasma or not. The prediction model was built based on CatBoost algorithm, a machine learning algorithm that is evaluated to have numerous advantages [6], [7]. We built a machine learning model using the CatBoost algorithm to predict whether a person is at risk of getting Melasma and how much the probability is. The adjustment of parameters to optimize the model is also done in this paper through analyzing some properties of the model such as confusion matrix, ROC-AUC curve, Precision-Recall curve and data variables that affect the predictive model.

We also used the facial data of patients who go for melasma examination at Quyhoa National Dermatology Hospital in combination with machine learning algo-rithms to build a diagnostic model of melasma to support the diagnosis of doctors in the hospital and can be transferred to other hospitals in need.

Combining the expertise of dermatologists with a dataset of images of melasma patients, we built the model to analyze melasma types with the input of the patient's facial images. The result is the conclusion that the individual has melasma or not, and what type of melasma?

The machine learning model is built based on YOLO V8 and is considered to have many advantages [26]. This study also focuses on adjusting the parameters to optimize the model by analyzing some properties of the model such as the confusion matrix, Precision-Recall curve, and data variables affecting the detection of melasma objects. That is to approach several Model Evaluation methods to evaluate the results obtained from the model, evaluate whether the model has achieved the set goals or not, analyze the achieved criteria of the model, and make decisions to use the results of the analysis in practice.

2. MATERIEL AND METHODS

2.1. Literature Review

Today, artificial intelligence is applied to many fields of our lives and healthcare is one of them. In this field, the application of artificial intelligence has brought many good results.

Healthcare Expo using AI is growing at a rate of 40% and is predicted to reach $6.6 billion by 2021 [1]. The abundance of healthcare information is making a difference to drive the development of AI applications that ensure enhanced knowledge for healthcare. Massive healthcare-related data is accessible from sources such as Electronic Recovery Records (EMRs) and health screening devices, and images from patients are analyzed and processed to use in the diagnosis and treatment of the patient.

The rise of AI in the age of massive information could help doctors, especially radiologists, with the precision tools to move forward. AI is well-suited to handling monotonous workflows, monitoring massive amounts of information, and providing lessons in rechoosing ways to eliminate mistakes.

AI will become a regular part of radiologists' lives and make their work more efficient and accurate [21], [35]. It is likely that in the next ten years, long-term therapeutic images will be preanalyzed by an AI device that has been tested by radiologists recently. This tool will perform common reading tasks such as measuring, dividing, and

recording.

Machine Learning is a field of Artificial Intelligence, which is a technique that helps computers learn on their own without setting up decision rules. Normally, a computer program needs rules to be able to execute a certain task, but with machine learning, computers can automatically execute the task upon receiving input data.

AI applied in medicine to support diagnosis and treatment will bring many benefits to patients and doctors [20], [22], [23], [36]. In this study, we used data on people's daily activities and image data available at the hospital to support doctors and people in predicting a person's likelihood of developing melasma. If we predict that this person is likely to be sick, we use artificial intelligence to support the diagnosis and subsequent treatment of the person through the image data provided by the person. We use web-based applications to perform prediction and diagnosis models combined with recommendations from specialists to help people prevent and treat diseases.

2.2. Solutions

The data set used in training and testing the machine learning model was collected from the community through a survey from many diffirent areas in Vietnam for prediction Melasma. The data set includes 17 data fields containing information of persons to be checked and medical practice information with a total of 1503 recorded samples organized as .csv files. [3]

Descriptive analysis from the data set helps us to get some more information such as variables of clinical characteristics data and daily habit factors for Melasma. Through this data descriptive analysis, we also obtained clinical characteristics-independent variables such as age, occupation, ethnicity, comorbidities, family history as well as habit-independent variables such as sun exposure, cosmetic use, pregnancy, and oral contraceptive use. This makes the training, evaluation and model correction process more effective.

In terms of family economic status, the analysis results also indicate that the proportion of poor and near-poor patients with Melasma is higher than that of the non-poor group. From the age information in the data set, we have the age distribution of people who are likely to have Melasma in Figure 1. It is shown in the Figure 1 that 37-47-year-old subjects have a high probability of Melasma.

The analysis also shows that the percentage of Melasma patients who are pregnant is higher than that of Melasma patients who are not. A multivariate logistic regression analysis indicates that those with a history of Melasma during pregnancy have a 3,3 times higher risk of Melasma compared with those without a history of Melasma during pregnancy as presented in Figure 2.

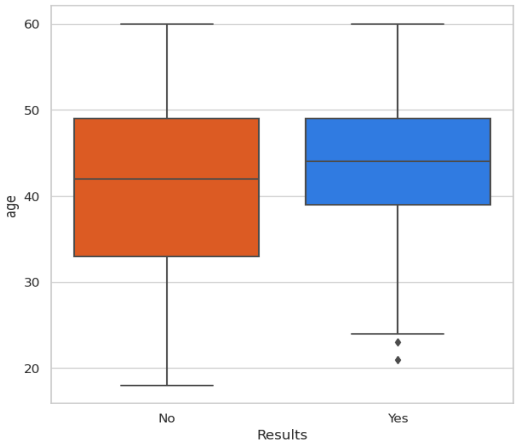


Figure 1. Distribution of Melasma by age

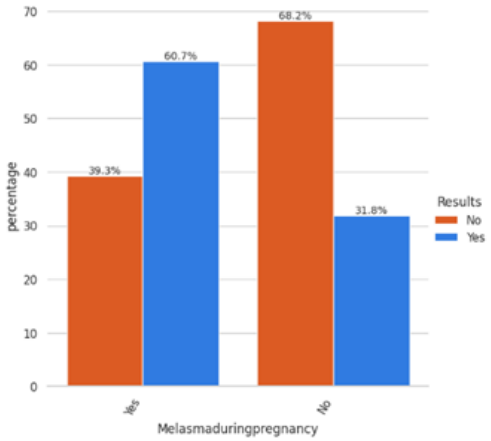


Figure 2. Distribution of Melasma by history of Melasma during pregnancy

Occupation is also indicated to have an influence on possibility of Melasma as in Figure 3. Cosmetics use is another influential factor of Melasma.

The dataset include 17 features: age, Location, Occupation, Ethnicity, Religion, Education, FamilyEconomy, Marriage, Numberofpregnancies, Numberofbirths, Birthcontrolpills, Familyhistory, Melasmaduringpregnancy, Monthofpregnancy, Numberofhoursofsunlightexposure, Usingcosmetics, Results and analysed their correlation so that we have more insignt about the dataset in Figure 4. Correlation analysis used to study the strength of a relationship and possible connections between features in dataset.

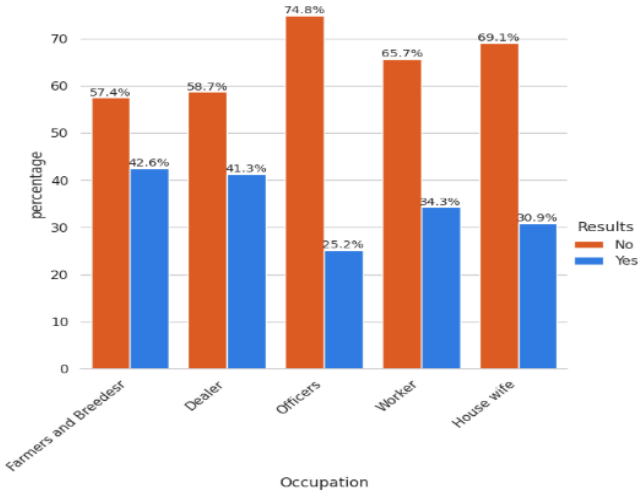


Figure 3. Distribution of Melasma by occupation

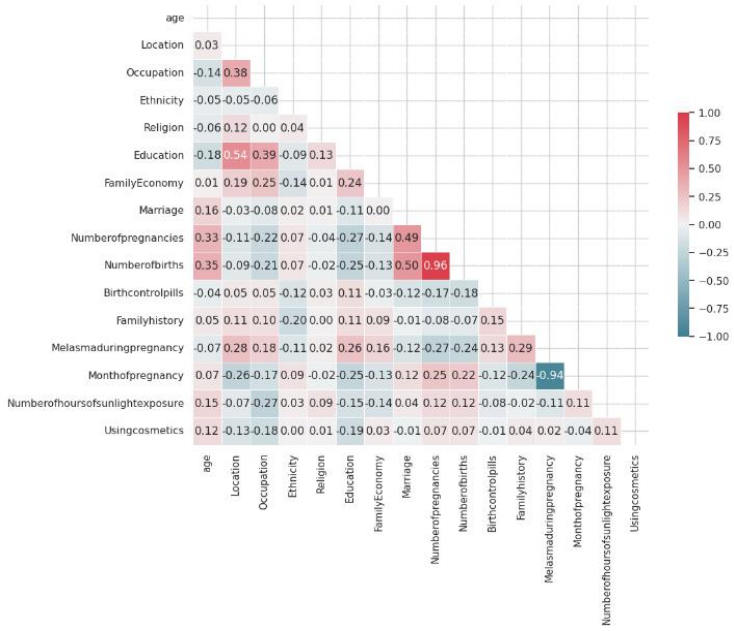


Figure 4. Correlation analysis in the data set

Through correlation analysis we get ranking of correlation coefficients; Example: Pair (Numberofpregnancies, Numberofbirths) is 96% or pair (Melasmaduringpregnancy, Monthofpregnancy) is 94% and so on, from there it helps us to make accurate assessments and provide solutions to upgrade our Catboost machine learning model.

From the dataset and the results of the dataset analysis above, to predict whether a person is infected to Melasma or not, we chose Catboost algorithm for machine learning model with the input data of expertise of dermatologists specializing in Melasma combined with patients' community information. Catboost is short for eXtreme Gradient Boosting. It is an efficient and scalable implementation of gradient boosting framework. It has several features:

1. Speed: Catboost can automatically do parallel computation.

2. Input Type: Catboost takes several types of input data: Dense Matrix, Sparse Matrix, Data File.

3. Sparsity: Catboost accepts sparse input for both tree booster and linear booster, and is optimized for sparse input.

4. Customization: Catboost supports customized objective function and evaluation function.

5. Performance: Catboost has better performance on several different datasets.

In next section we will present how Catboost actions and apply to our dataset to solve prediction problem whether a person is infected to Melasma or not.

Gradient boosting is a powerful machine learning technique that achieves state-of-the-art results in a variety of practical tasks. It has remained the primary method for learning problems with heterogeneous features, noisy data, and complex dependencies: weather forecasting, recommendation systems, and many others [6], [7], [8], [18]. Gradient boosting is essentially a process of constructing an ensemble predictor by performing gradient descent in a functional space. It is backed by solid theoretical results that explain how strong predictors can be built by iteratively combining weaker models in a greedy manner [6].

Assume the method observe a dataset of examples D = {(x_k, y_k)}, k=1..n, where x_k = (x_k^1,..., x_k^m) is a random vector of m features and y_k ∈ R is a target, which can be either binary or a numerical response. Examples (x_k, y_k) are independent and identically distributed according to some unknown distribution P(•, •). The goal of a learning task is to train a function F: $R^m \rightarrow R$ which minimizes the expected loss L(F) := EL(y, F(x)). Here L(:, •) is a smooth loss function and (x, y) is a test example sampled from P independently of the training set D.

Catboost is an implementation of gradient boosting, which uses binary decision trees as base predictors. A decision tree [4], [6], [10] is a model built by a recursive partition of the feature space R^m into several disjoint regions according to the values of some splitting attributes a. Attributes are usually binary variables that identify that some feature x_k exceeds some threshold t, that is, a = $1_{\{xk>t\}}$, where x_k is either numerical or binary feature, in the latter case t = 0.5. Each final region is assigned to a value, which is an estimate of the response y in the region for the regression task or the predicted class label in the case of classification problem. Analysis of prediction shift the method formally analyze the problem of prediction shift in a simple case of a regression task with the quadratic loss function $L(y, \hat{y}) = (y - \hat{y})^2$.

Algorithm: Building a tree in Catboost
input: $M, \{(x_i, y_i)\}_{i=1}^{n}, \alpha, L, \{\sigma_i\}_{i=1}^{s}, Mode$
$grad \leftarrow CalcGradient\ (L, M, y);$
$r \leftarrow random\ (1, s);$
if *Mode = Plain* **then**
  $G \leftarrow (grad_r(i) for\ i = 1..n);$
if *Mode = Ordered* **then**
  $G \leftarrow (grad_{r,\sigma_r(i)-1}(i) for\ i = 1..n);$
$T \leftarrow empty\ tree;$
foreach *step of top-down procedure* **do**
  foreach *candidate split c* **do**
    $T_c \leftarrow add\ split\ c\ to\ T;$
    $\boldsymbol{if}\ Mode = Plain\ \boldsymbol{then}$
      $\Delta(i) \leftarrow avg(grad_r(p)\ for$
      $p: leaf_r(p) = leaf_r(i))\ for\ i = 1..n;$
    $\boldsymbol{if}\ Mode = Ordered\ \boldsymbol{then}$
      $\Delta(i) \leftarrow avg(grad_{r,\sigma_r(i)-1}(p)\ \boldsymbol{for}$
      $p: leaf_r(p) = leaf_r(i), \sigma_r(p) < \sigma_r(i)$
      $\boldsymbol{for}\ i = 1..n;$
    $loss(T_c) \leftarrow \cos(\Delta, G)$
  $T \leftarrow argmin_{T_c}(loss(T_c))$
$\boldsymbol{if}\ Mode = Plain\ \boldsymbol{then}$
  $M_{r'}(i) \leftarrow M_{r'}(i) - \alpha avg(grad_{r'}(p)\ for$
  $p: leaf_{r'}(p) = leaf_{r'}(i))\ for\ r' = 1..s, i = 1..n;$

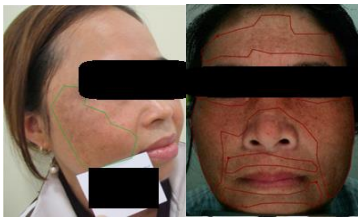
$\boldsymbol{if}\ Mode = Ordered\ \boldsymbol{then}$
  $M_{r',j}(i) \leftarrow M_{r',j}(i) - \alpha\ avg(grad_{r',j}(p)\ for$
  $p: leaf_{r'}(p) = leaf_{r'}(i), \sigma_{r'}(p) \leq j\ for\ r'$
  $= 1..s, i = 1..n, j \geq \sigma_{r'}(i) - 1$
return *T,M*

At the start, Catboost generates s + 1 independent random permutations of the training dataset. The permutations $\sigma_1,...,\sigma_s$ are used for evaluation of splits that define tree structures (i.e., the internal nodes), while σ_0 serves for choosing the leaf values b_j of the obtained trees. Using only one permutation may increase the variance of the final model predictions, while several permutations allow us to reduce this effect in a way we further describe. Building a tree in Catboost, base predictors are oblivious decision trees [9], [16] also called decision tables [6]. Term oblivious means that the same splitting criterion is used across an entire level of the tree. Such trees are balanced, less prone to overfitting, and allow speeding up execution at testing time significantly. The procedure of building a tree in Catboost is described in Algorithm.

In the Ordered boosting mode, during the learning process, the method maintain the supporting models $M_{r,j}$, where $M_{r,j(i)}$ is the current prediction for the i-th example based on the first j examples in the permutation σ_r. At each iteration t of the algorithm, we sample a random permutation σr from $\{\sigma_1,..., \sigma_s\}$ and construct a tree Tt on the basis of it.

Addional, about dataset for diagnosing melasma, we use a set of information on melasma patients including 1,624 photographs of melasma patients, taken from the front. The dataset is collected from many sources, and the brightness of the images is different to meet the needs of different qualities of the information set.

Images selected from the image dataset will help classify patients with central, butterfly-shaped, mandibular melasma and no melasma as in Figure 5



The dataset will be labeled:
0: central body,
1: butterfly body,
2: mandibular body,
3: no melasma

Figure 5. Dataset of images will be labeled

To label the images, we use the AnyLabeling application to manually label them, moreover, the experts who are dermatologists of Quyhoa National Leprosy - Dermatology Hospital commented on the images of melasma and no melasma, and what type of melasma (central, butterfly-shaped, mandibular melasma). Each photo can have one or more names in the title depending on the type of melasma and the area. There is also a melasma cream photo in the set, which is marked with the corresponding melasma area depending on the patient. In the problem using the YOLO model, we save the annotation file as .txt.

Example on label: 0 0.7142857142857143 0.46256531158404485 0.6296992481203008 0.5246909646998853 0.524436090225564 0.5453995157384988 0.47180451127819556 0.5549573085255511 0.5206766917293234 0.6218618580349179 0.6146616541353384 0.633012616286479 0.637218045112782 0.6935453039378108 0.6484962406015037 0.7700076462342297 0.7180451127819549 0.8018669555244042 0.7913533834586466 0.7986810245953868 0.8571428571428571 0.785937300879317 0.8984962406015037 0.6632789601121449 0.8609022556390977 0.578851790493182 0.7838345864661654 0.4753090353001147

This is a object label which is set of <x1> <y1> <x2> <y2> … <xn> <yn>. In which: <xi> <yi> are the coordinates of the vertices in the polygon delineating the object respectively. These values have been renormalized; therefore, the value is always in the range [0,1]. Object class is an index to mark classes.

For the problem with many labels, we assign the same labels with the label order agreed upon in advance. The reason is that the file annotations only save the label's index (0, 1, 3, 4...), not the label name. After labeling is complete, we put the annotation file and the corresponding image in the same folder.

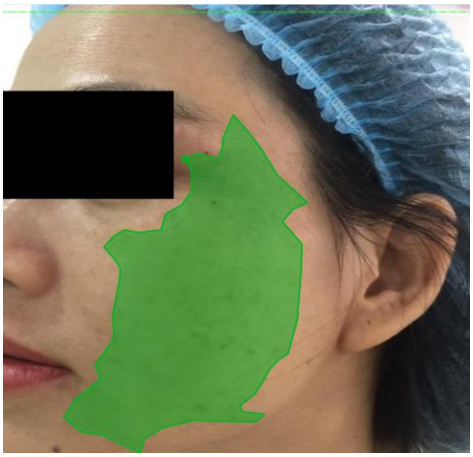


Figure 6. Label for images

The machine learning algorithm we use for the problem of diagnosing melasma is YOLOv8 [26]. YOLOv8 is a computer vision algorithm used for object detection. This model is developed based on previous versions of YOLO (You Only Look Once - Identify objects with one look). Additionally, YOLOv8 is designed for real-time object detection and can process images in a few milliseconds. YOLOv8 is an object recognition model based on a convolutional neural network (CNN) developed by Joseph Redmon and his research team at the University of Washington. YOLOv8 is an upgraded version of YOLOv7, with faster and more accurate object recognition. This is achieved through a number of innovations, including feature pyramid networks, spatial attention modules, and advanced data augmentation techniques.

The YOLOv8 model uses a darknet-53 architecture neural network to extract image features and apply the YOLOv8 object recognition algorithm on those features with main features of Yolov8 such as:

- Enhance the model by adding distributed channels to speed up calculations.

- Use the Attention technique to improve the model's object recognition ability.

- Apply new training methods to increase convergence speed.

- Use new neural network architecture: Use YOLOv4 architecture as a basis to increase model performance and accuracy.

- Integrated mechanism to automatically adjust the size increase ratio of the input image (AutoScale).

- Video Supervision Support: The model is capable of detecting and monitoring objects in videos and making continuous predictions across the entire video.

- Integrating Ensemble technology.

- AutoAnchor: Improves object detection at various scales.

3. RESULTS AND DISCUSSIONS

The A benefit of using gradient boosting is that after the boosted trees are constructed, it is relatively straightforward to retrieve importance scores for each feature. Generally, importance provides a score that indicates how useful or valuable each feature was in the construction of the boosted decision trees within the model. The more an attribute is used to make key decisions with decision trees, the higher its relative importance. This

importance is calculated explicitly for each attribute in the dataset, allowing attributes to be ranked and compared to each other. Importance is calculated for a single decision tree by the amount that each attribute split point improves the performance measure, weighted by the number of observations the node is responsible for. [6], [7], [8], [18]

Input data to built and train the model of prediction melasma is the dataset of a study on clinical characteristics and some factors related to Melasma in women in 2023 provided by Quyhoa National Leprosy Dermatology Hospital [3] with of 1503 data samples. The goal is to find the outcome variable (y = Results; non-infected = 0, infected = 1). The findings show that there are 569 Melasma infected cases and 934 non-infected cases. Gradient boosting is one of the most powerful techniques for building predictive models and we have used Catboost to build our predictive model with 80% dataset used for training set and 20% for testing set.

Based on Catboost feature, it is Catboost that gives good results without parameter adjustment, which can reduce the minimum time spent on parameter tuning because Catboost provides excellent results with default parameters. We build a Catboost classifier (Catboost Classifier) machine learning model with 10000 iterations and use logloss and rest parameters using Catboost default parameters. Running on our dataset trains the model 80% of the data and evaluates the model every training epoch 20% test. The feature importances are then averaged across all of the the decision trees within our model and showed in the Figure 7.

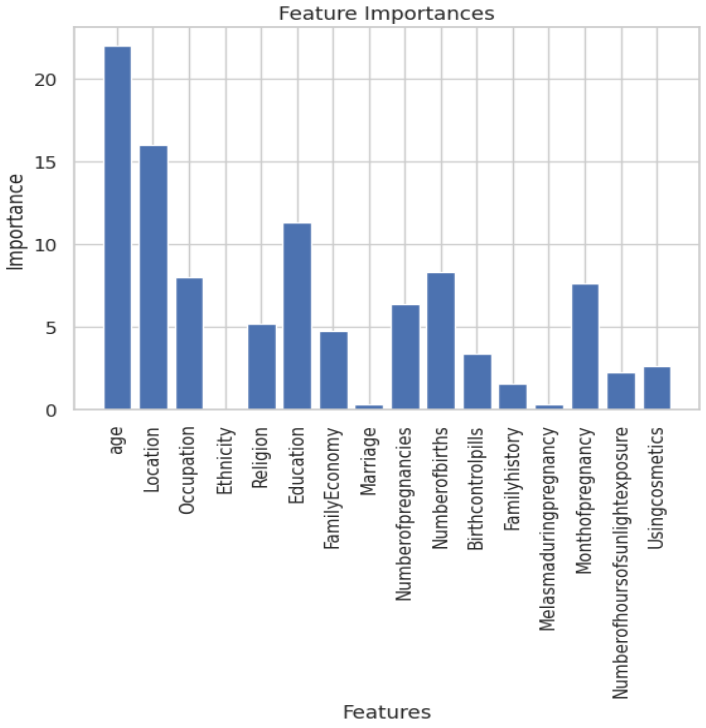


Figure 7. Importance level of features affecting the outcome

Important (influential) variables on the outcome (infected or non-infected) are: Age (25,9%), Location (14.7%), Education (11%), Occupation (8,4%), Numberofbirths (7,7%), Monthofpregnancy (7%), Numberofpregnancies (5,9%), Re-ligion (5,08%),…which have little effect on the outcome.

To evaluate our predictive model using Catboost for this melasma dataset, we approach some ways to evaluate the machine learning model's performance as below:

Accuracy: 83,6%

Learning Curves: We can retrieve the performance of the model on the evaluation dataset and plot it to get insight into how learning unfolded while training. We can then use these collected performance measures to create a line plot and gain further insight into how the model behaved on train and test datasets over training epochs.

Figure 8 Shows the logarithmic loss of the Catboost model for each epoch on the our training and test datasets.

Confusion Matrix: A confusion matrix is a correlation between the predictions of a model and the actual class labels of the data points. Our predictive model using Catboost algorithm for 1503 records of melasma dataset has Confusion Matrix in Figure 9.

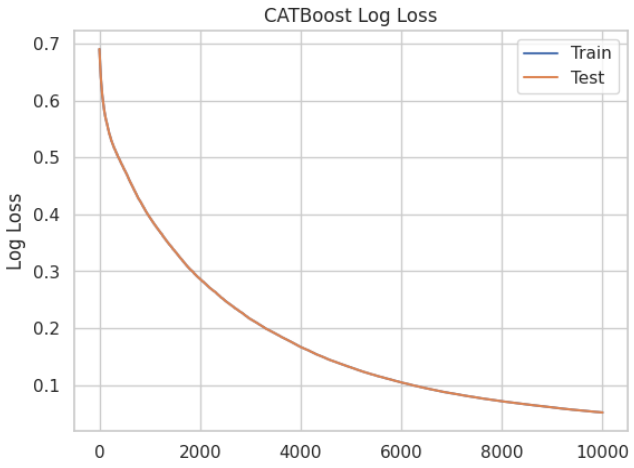


Figure 8. Catboost Learning Curve Log Loss

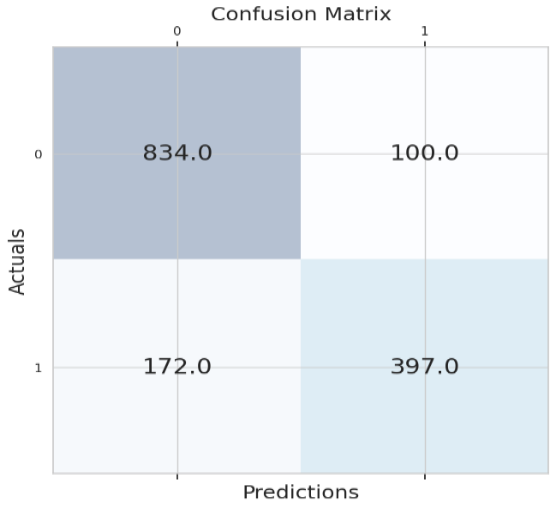


Figure 9. Confusion matrix

In this: Positive (P): Observation is positive (eg. is infected). Negative (N): Observation is not positive (eg. is not infected). True Positive (TP): Outcome where the model correctly predicts the positive class (834). True Negative (TN): Outcome where the model correctly predicts the negative class (397). False Positive (FP): Also called a type 1 error, an outcome where the model incorrectly predicts the positive class when it is actually negative (100). False Negative (FN): Also called a type 2 error, an outcome where the model incorrectly predicts the negative class when it is actually positive (172).

Accuracy is what its literal meaning says, a measure of how accurate your model is: Accuracy = Correct Predictions / Total Predictions.

By using confusion matrix, Accuracy = (TP + TN)/(TP+TN+FP+FN). In our predictive model using Catboost algorithm Accuracy = (834+397) / (834+397+100+172) = 83,6%.

Precision-Recall Curves

Precision is a ratio of the number of true positives divided by the sum of the true positives and false positives. It describes how good a model is at predicting the positive class. Precision is referred to as the positive predictive value.

Precision (non-infected) = 91%

Precision (infected) = 79%

Recall is calculated as the ratio of the number of true positives divided by the sum of the true positives and the false negatives

Recall (non-infected) = 93%

Recall (infected) = 75%

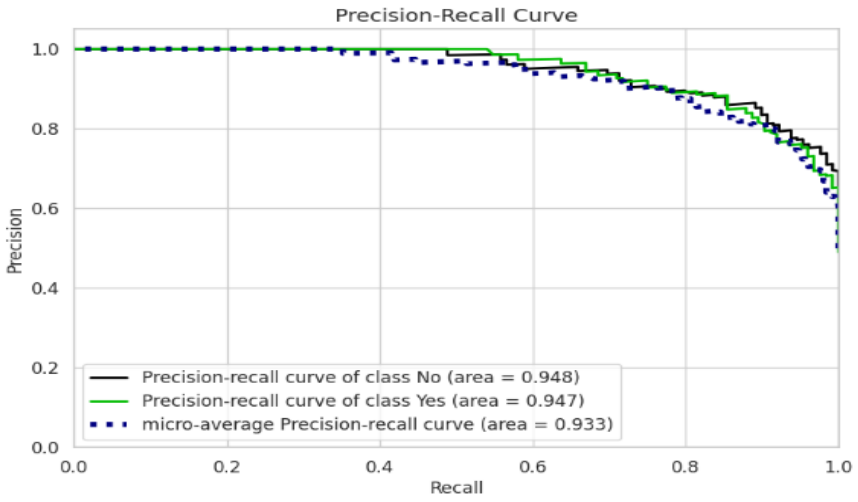


Figure 10. Precision-Recall curve

F-Measure

F1-score (non-infected) = 86%

F1-score (infected) = 82%

ROC-AUC Curves

A useful tool when predicting the probability of a binary outcome is the Receiver Operating Characteristic curve, or ROC curve. It is a plot of the false positive rate (x-axis) versus the true positive rate (y-axis) for a number of different candidate threshold values between 0.0 and 1.0. Put another way, it plots the false alarm rate versus the hit rate. The true positive rate is calculated as the number of true positives divided by the sum of the number of true positives and the number of false negatives. It describes how good the model is at predicting the positive class when the actual outcome is positive.

The ROC-AUC curve below presents the accuracy of the model. ROC for "non-infected" is 0.94 and for "Infected" is 0.94 in Figure 11.

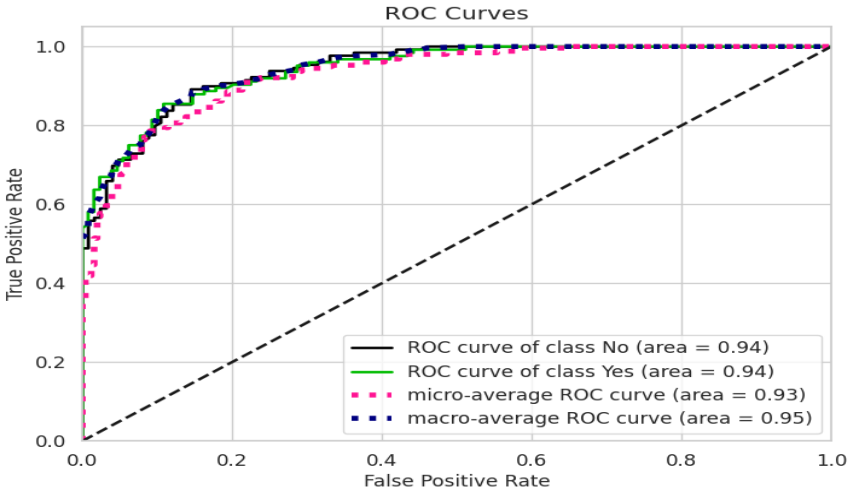


Figure 11. ROC-AUC of model

To detect and classify melasma on face images taken by the camera, we use the machine learning program YOLO V8. This program is prepared to identify and classify the components associated with melasma on the face.

The labels used in the program are: 0: central body; 1: butterfly body; 2: mandibular body. In addition, the program will recognize the label '3' in cases of no facial melasma. When applying this program to the melasma image dataset, it will identify and name the melasma locations on the face. The names given will help determine the appropriate type of melasma for each defined site of the face.

Based on the results of the model, we can distinguish the facial sites affected by melasma and relate them by comparing label names. Specifically, "central body" (name 0) is compared to the phrase 'central site', "butterfly body" (name 1) is compared to 'butterfly-shaped site', “mandibular body” (name 2) is compared to 'chin site'. In the absence of melasma recognized, the term 'no melasma' may be used. This approach which improves understanding of the specific melasma condition makes the difference in classification and better telediagnosis.

The model was built on Google Collab with hardware-accelerated Python3 and hardware-accelerated GPU, Type NVIDIA A100-SXM4-40GB, 40514MiB.

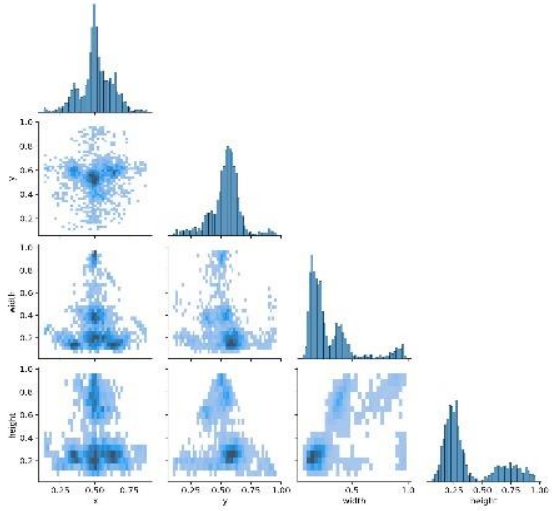


Figure 12. Description of melasma object detection data

The number of images in the training dataset is 1,624. The number of images in the testing dataset is 176. The image size is 640 pixels and epochs are 50 times. The datasets include 4 types of labels mentioned above.

```
      Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
      49/50      2.31G      1.224      1.016      1.351         18        640: 100%|██████████| 102/102 [00:57<00:00,  1.78it/s]
                 Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 6/6 [00:01<00:00,  4.13it/s]
                   all        176        215      0.76      0.76      0.822      0.516

      Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
      50/50      2.21G      1.243      1.024      1.364         21        640: 100%|██████████| 102/102 [00:55<00:00,  1.84it/s]
                 Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 6/6 [00:04<00:00,  1.32it/s]
                   all        176        215      0.784      0.779      0.841      0.523

50 epochs completed in 0.861 hours.
Optimizer stripped from runs/detect/train4/weights/last.pt, 6.2MB
Optimizer stripped from runs/detect/train4/weights/best.pt, 6.2MB

Validating runs/detect/train4/weights/best.pt...
Ultralytics YOLOv8.0.109 🚀 Python-3.10.11 torch-2.0.1+cu118 CUDA:0 (NVIDIA A100-SXM4-40GB, 40514MiB)
Model summary (fused): 168 layers, 3006428 parameters, 0 gradients
                 Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 6/6 [00:03<00:00,  1.66it/s]
                   all        176        215      0.776      0.779      0.839       0.52
           CentralBody        176         36      0.882      0.639      0.758      0.332
         ButterflyBody        176         56      0.774      0.786      0.847       0.39
         MandibularBody       176         10      0.598        0.7      0.776       0.55
              NoMelasma       176        113      0.849      0.992      0.976       0.81
Speed: 1.0ms preprocess, 0.4ms inference, 0.0ms loss, 1.3ms postprocess per image
Results saved to runs/detect/train4
```

Figure 13. Results of training model

An object detection model is based on Precision, Recall, AP, and mAP50-95 parameters. mAP will be the measurement criterion of the object detection model through IoU, Precision, Recall, Precision-Recall Curve, AUC, AP, and mAP metrics.

Precision: Evaluate the reliability of the given conclusions (how many % of the model's conclusions are correct)

Recall: Evaluate the model's ability to find ground truths (how many % positive samples the model recognizes)

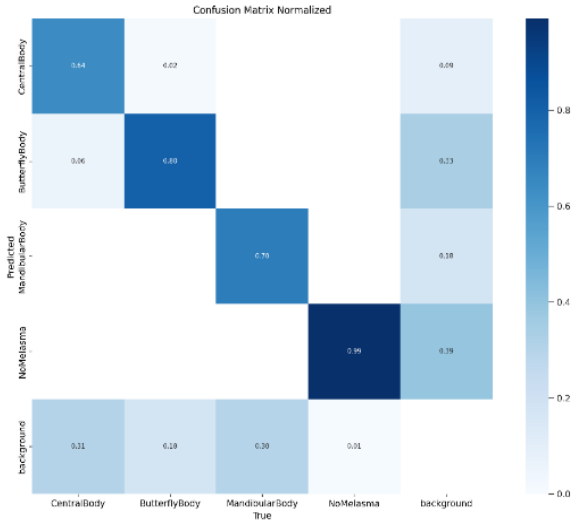


Figure 14. Confusion Matrix

IoU (Intersection over Unit): Measure the overlap between the ground truth bounding box and the bounding box that the model predicts. The Figure 15 shows that the reliability rating of the given conclusion has an accuracy of 1 to 0.86.

mAP=mean Average Precision is the average of the AP values of different classes. The larger the mAP, the better the model. The relationship between precision-recall helps mAP evaluate the accuracy of the classification task.

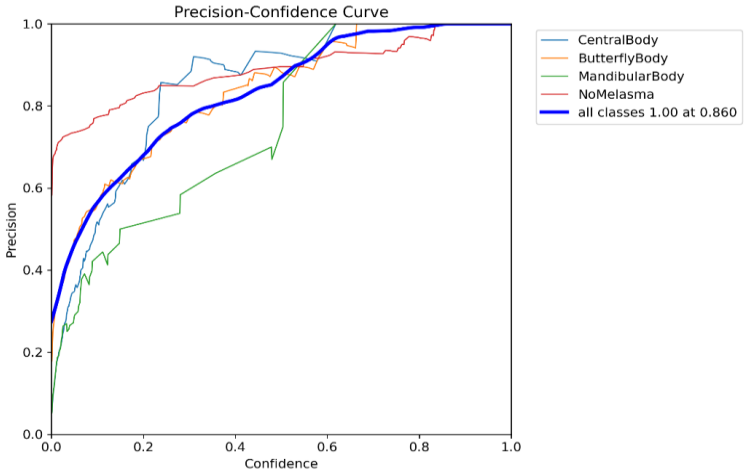


Figure 15. Evaluation of Precision

Doing precision-recall changes when the IoU threshold changes (what class is the threshold to predict a box). Therefore, at a given IoU value, it is possible to measure/compare the goodness of the models.

According to Figure 9, the mAP50 score at epochs=50 is 0.839 and the mAP50-95 score is 0.52, and the precision score is 0.776, indicating that the model has detected melasma objects quite well.

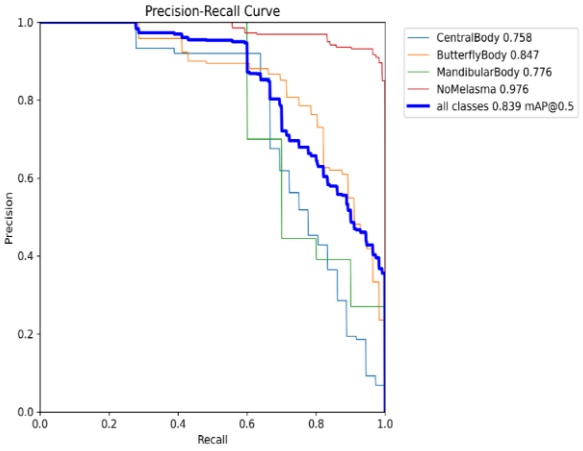


Figure 16. Evaluation of the model's ability to search all labels

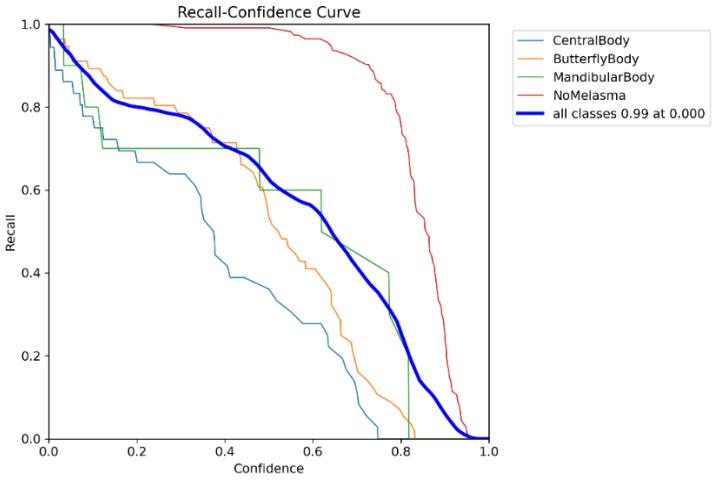


Figure 17. Evaluation of the model's reliability

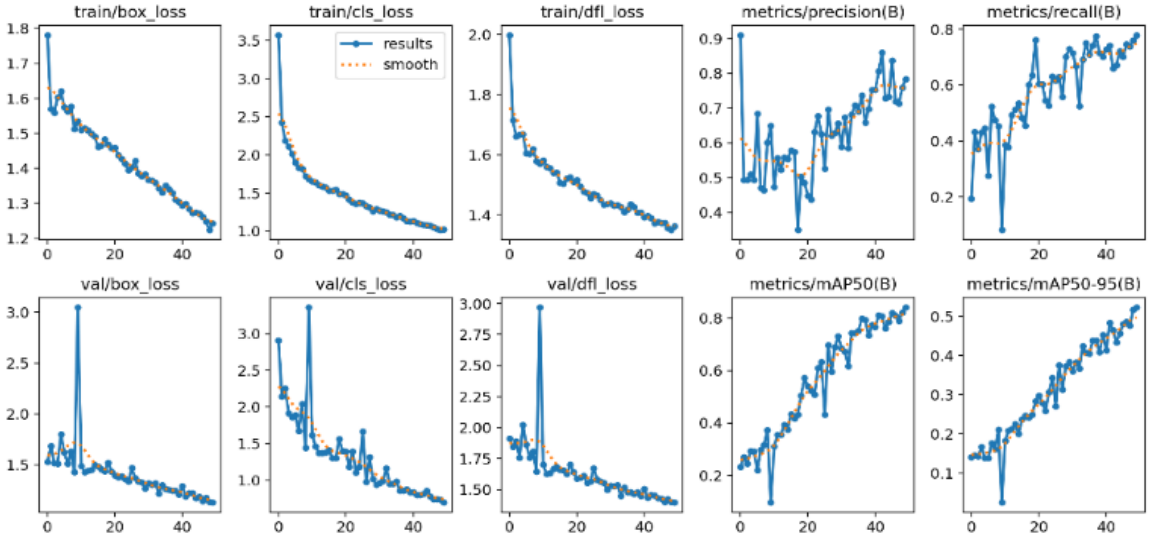


Figure 18. Results of model evaluation

The illustrations in Figure 19 when comparing the detection results between the training dataset and the testing dataset show that the YOLO V8 model gives quite accurate results.

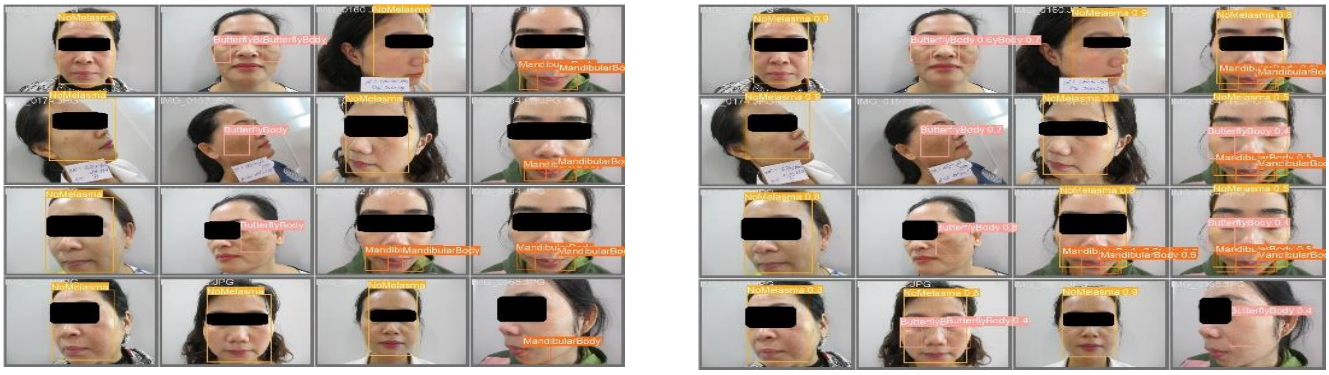


Figure 19. Labeled training dataset and the testing dataset

CONCLUSIONS

The study, we present steps of a data analysis process in practice and build a machine learning model using the Catboost algorithm to predict the possibility of a user getting infected to Melasma. With this approach, the proposed method has exploited existing community data together with data collected through surveys to help the machine learning model have hight accurate prediction results, which assists in the prevention, diagnosis and treatment of the disease, thereby helping to reduce the cost of treatment.

After that, if we detect that people are likely to have Melasma, we will use the machine learning model built from YOLO V8 with inputting people's photos. The model will support doctors and people to diagnose type of Melasma. The machine learning model has a prediction accuracy higher than 83.9 %, therefore, it helps prevent, diagnose, and treat the disease as well as reduce treatment costs.

The machine learning model that predicts the risk and diagnose of melasma are packaged and embedded into the web application to help users know the possibility of being infected to Melasma, and provide users with habits that may cause Melasma, so that the users can prevent it. [39], [40]. Dermatologists use the application to contact patients and evaluate support to upgrade the model through expertise and practical results. The application updates the data once collecting enough new data (the model is set to be retrained each time 100 new data are inputted) to enhance the accuracy of the model.

However, in order to get higher accuracy for the models, it is necessary to collect community data of many individuals and from many different regions, though it would require a lot of effort, time and expense.

REFERENCES

[1] Nguyen Van Thuong, “Melasma disease”, Dermatology Pathology, Volume 1, Medical Publishing, 143-148, 2017.

[2] Ratna Rajaratnam Asad Salim, Eva Soos Domanne (2014), "Melasma", Evidence-based Dermatology. Third Edition.

[3] Quyhoa National Leprosy Dermatology Hospital, Binh Dinh, Vietnam, “Dataset of Studying on clinical characteristics and some factors related to melasma in women in 2023”, 2023.

[4] Panesar, “Machine Learning and AI for Healthcare”, Arjun Panesar, 2019.

[5] Dhar, V., “Data science and prediction”. Communications of the ACM. 56 (12): 64. 2013. doi:10.1145/2500499.

[6] Liudmila Prokhorenkova .et. al, “CatBoost: unbiased boosting with categorical features”, arXiv:1706.09516v5 [cs.LG] 20 Jan 2019.

[7] Abdullahi A. Ibrahim .et .al, “Comparison of the CatBoost Classifier with other Machine Learning Methods”, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 11, No. 11, 2020.

[8] John T. Hancock and Taghi M. Khoshgoftaar “CatBoost for big data: an interdis-ciplinar review”, Hancock and Khoshgoftaar J Big Data (2020), 7:94 https://doi.org/10.1186/s40537-020-00369-8.

[9] Ramraj S, Nishant Uzir, Sunil R and Shatadeep Banerjee, “Experimenting XGBoost Algorithm for Prediction and Classifi cation of Diff erent Datasets”, Interna-tional Journal of Control Theory and Applications, Volume 9, Number 40, 2016.

[10] Tianqi Chen, Carlos Guestrin, “XGBoost : A scalable tree boosting system”, March 9, 2016, arXiv:1603.02754 [cs.LG].

[11] Zhiyuan He, Danchen Lin1, Thomas Lau1, and Mike Wu1, “Gradient Boosting Machine: A Survey”, arXiv:1908.06951v1 [stat.ML] 19 Aug 2019.

[12] R. Caruana and A. Niculescu-Mizil. An empirical comparison of supervised learn-ing algorithms. In Proceedings of the 23rd international conference on Machine learning, pages 161-168. ACM, 2006.

[13] J. Friedman, T. Hastie, and R. Tibshirani. The elements of statistical learning, volume 1. Springer series in statistics New York, 2001.

[14] J. H. Friedman. Greedy function approximation: a gradient boosting machine. An-nals of statistics, pages 1189-1232, 2001.

[15] Gulin, I. Kuralenok, and D. Pavlov. Winning the transfer learning track of ya-hoo!'s learning to rank challenge with yetirank. In Yahoo! Learning to Rank Chal-lenge, pages 63-76, 2011.

[16] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu. Lightgbm: A highly efficient gradient boosting decision tree. In Advances in Neural Information Processing Systems, pages 3149-3157, 2017.

[17] J. Langford, L. Li, and T. Zhang. Sparse online learning via truncated gradient. Journal of Machine Learning Research, 10(Mar):777-801, 2009.

[18] https://github.com/catboost/catboost, January 6, 2024.

[19] https://catboost.ai/en/concepts/python-reference_catboost, January 6, 2024.Jordan MI, Mitchell TM, (2015). Machine learning: Trends, perspectives, and prospects. Sci (NY), 249: 255–60.

[20] Meskó B: Artificial Intelligence is the Stethoscope of the 21st Century. (2017).

[21] Recht M, Bryan RN. Artificial intelligence: threat or boon to radiologists? Journal of the American College of Radiology. 2017;14(11):1476-80.

[22] Houssami N, Lee CI, Buist DS, Tao D. Artificial intelligence for breast cancer screening: opportunity or hype? The Breast. 2017;36:31-3.

[23] Magazine E, Roach L. Starting With Retina. Artificial Intelligence. 2017.

[24] Salim A, Rajaratnam R, Domanne ES. Melasma. Evidence-Based Dermatology. 2014:470-85.

[25] Dhar V. Data science and prediction. Communications of the ACM. 2013;56(12):64-73.

[26] Solawetz J, Francesco: What is YOLOv8? The Ultimate Guide. (2023). Accessed 27/05/2023.

[27] 9. Balkrishnan R, Mcmichael A.J, F.T.Camacho, F.Saltzberg, T.S.Housman, S.Grummer, et al. Development and validation of health-related quality of life instrument for women with melasma. British Journal of Dermatology. 2003;149:572-7.

[28] Katsambas, al e. Melasma. Classification and treatment. Journal of the European Academy of Dermatology and Venereology. 1995;4:217-33.

[29] Lynde C.B., J.N. Kraft M, C.W. Lynde M, FRCPC,. Topical Treatments for Melasma and Postinflammatory Hyperpigmentation. Skin Therapy Letter. 2006;11(9):1-12.

[30] Yuri T. Jadotte, Schwartz RA. Treatment of Melasma. Evidence-based dermatology. 2011.

[31] Noh TK, Choi SJ, Chung BY, Kang JS, Lee JH, Lee MW, et al. Inflammatory features of melasma lesions in Asian skin. The Journal of Dermatology. 2014;41(9):788-94. doi 10.1111/1346-8138.12573.

[32] Ortonne JP, I Arellano, M Berneburg, T Cestari, H Chan, P Grimes, et al. A global survey of the role of ultraviolet radiation and hormonal influences in the development of melasma. Journal of the European Academy of Dermatology and Venereology: JEADV. 2009;23(11):1254-62. doi: 10.1111/j.1468-3083.2009.03295.x.

[33] Saumya Panda. Agenda for Future Research in Melasma: QUO VADIS? Journal of Pigmentary Disorders. 2014;1(5):1:5. doi 10.4172/JPD.1000e103.

[34] Rath S: YOLOv8 Ultralytics: State-of-the-Art YOLO Models.

[35] Ho Van Lam, Vu Tuan Anh, Pham Thi Hoang Bich Diu, Tran Xuan Viet. Appling machine learning to predict Melasma. International Journal of Computer Science and Information Security (IJCSIS), Vol. 19, No. 11, November 2021.

[36] Mohamed A. Kassem et al. Machine Learning and Deep Learning Methods for Skin Lesion Classification and Diagnosis: A Systematic Review. Diagnostics 2021, 11, 1390.

[37] Lin Liu et al. An Intelligent Diagnostic Model for Melasma Based on Deep Learning and Multimode Image Input. Dermatol Ther (Heidelb) (2023) 13:569–579.

[38] Amirreza Mahbod, Isabella Ellinger. Special Issue on "Advances in Skin Lesion Image Analysis Using Machine Learning Approaches. Diagnostics 2022, 12, 1928.

[39] http://112.78.3.35:5000/home, Retrieved January 6, 2024.

[40] http://112.78.3.35:5000/uploads, Retrieved January 6, 2024.

DOI: https://doi.org/10.15379/ijmst.v11i1.3568