Vol.2 No.1 February 2025: 15-21
International Journal of Computer Science and Humanitarian AI (IJCSHAI)
e-ISSN: 3064-4372
DOI: 10.21512/ijcshaijournal.v2i1.12418

Smoker Melanosis Classification Using Oral Photographic Feature Extraction Based On K-Nearest Neighbor

I Gede Maha Prastya Budhi[1*], Nada Fitrieyatul Hikmah[2], Tri Arief Sardjono[3]

[1-3]Department of Biomedical Engineering,
Institut Teknologi Sepuluh Nopember,
Surabaya, Indonesia 60111
mahaprastya41@gmail.com; nadafh@its.ac.id;
sardjono@bme.its.ac.id

Abstract— **Smoking is one of the causes of various diseases in the body. Smoking can also cause abnormal conditions that are pathological and physiological in the oral cavity, one of which is smoker melanosis. The clinical picture of pigmentation smoker melanosis is the presence of scattered brown spots with a diameter of less than 1 cm and is most often located on the gingiva. The data was taken using the oral photograph image capture method using a 12MP resolution camera, provided that the object distance from the camera was 6 cm and the flash was on. This analysis utilized the Gingiva Pigmentation Index (GPI) classification system proposed by Hedin, which compares the pigmented area, and Dummett's Oral Colour Index (DOPI), which assesses the density of pigmentation. In this study, the classification process was carried out with the KNN algorithm using features from digital image processing in the segmentation area, the average value of the red, green, and blue colour levels. The classification process uses the nearest neighbour value of 3 and a p-value of 2 to measure the distance to the nearest neighbor using the Minkowski distance formula. The results of the test data accuracy (1.0) with F1 scores for each class for test data DOPI 0 = 1.0, DOPI 1 = 1.0, DOPI 2 = 1.0, DOPI 3 = 1.0. Meanwhile, the classification process can use more up-to-date methods, such as CNN to improve classification accuracy.**

Keywords— KNN, Gingiva, Pigmentation, Image Processing, Smoker Melanosis

I. INTRODUCTION

The World Health Organization's (WHO) Framework Convention on Tobacco Control (FCTC) guidelines include public education on the dangers of tobacco use, with over 100 countries mandating the inclusion of written health warnings Smokers' perceptions of the dangers of tobacco use have changed as a result of these policies, and quitting attempts have increased and smokers' and their communities' health outcomes have improved [1][2][3].

Tobacco use is linked to systemic diseases such as cardiovascular disease, lung disorders, and various types of cancer. It has been established that smoking is harmful, particularly to women and children. Cigarette smoking hurts the oral cavity. It has also been proven that it causes diseases such as oral cancer, periodontitis, leukoplakia, and other oral lesions [4].

Smoking is one of the causes of various diseases in the body. In addition to causing systematic effects, smoking can also cause the appearance of various abnormal conditions that are pathological and physiological in the oral cavity, one of which is smoker melanosis. The clinical picture of smoker melanosis pigmentation is the presence of brown patches that spread out with a diameter of less than 1 cm and are located most often in the gingiva [4][5][6][7].

The gingiva is the part of the oral mucosa that surrounds and attaches to the teeth and alveolar bone. On the surface of the oral cavity, the gingiva extends from the marginal peak of the gingiva to the linkage of the gingiva mucosa. The gingiva consists of three parts, namely, the gingiva free, fixed and interdental. The free gingiva is characterized by the presence of a gingiva sulcus which is an area that does not glue between the free gingiva and the tooth. The fixed gingiva has a junctional epithelium where the apical region of the gingiva is not bound to the base of the tooth and also binds the neck of the tooth. Finally, the interdental area is between two adjacent teeth below their point of contact [8][9].

Smoker melanosis is a characteristic change in the colour of the oral mucosa exposed to cigarette smoke and is the main result of melanin deposition in the basal cell layer of the mucosa. Smoker melanosis is a disorder in the oral cavity that is not dangerous, but if left unchecked, it will interfere with aesthetics. 25–31% of smokers get smoker melanosis, which is much more common during the first year of smoking. The longer a person smokes, the more widely distributed the coloration gets. This implies that a smoker's risk increases with the length of time they smoke [4][7][10].

Clinically the normal gingiva is generally pink. This is due to the presence of blood supply, the thickness and degree of the keratin layer of the epithelium as well as pigment cells. The number of cells determines the gingiva's size, intra-cellular elements, and blood supply. The contours and magnitude of the gingiva vary greatly. This state is influenced by the shape and arrangement of the teeth in their arches,

Cite this article as: *I. G. M. P. Budhi, N. F. Hikmah, and T. A. Sardjono, "Smoker Melanosis Classification Using Oral Photographic Feature Extraction Based On K-Nearest Neighbor", International Journal of Computer Science and Humanitarian Artificial Intelligence (IJCSHAI), vol. 2, no. 1, pp. 15–21, Feb. 2025. DOI: https://doi.org/10.21512/ijcshai.v2i1.12418*

localization and area of proximal contact area and dimensions of both oral and vestibular gingiva embrasure. The gingiva is tightly attached to the lower structure and has no submucosal layer so that the gingiva is immobile and supple. The surface of the gingiva is fixedly spotted like an orange peel. These spots are called stippling [4][11][12].

Pigmentation was analyzed using oral frontal photos taken with non–contact–type dental cameras. All photos were reviewed on a monitor with a liquid crystal display (multisync LCD 2492490WUXi2). Frontal oral photographs were analyzed using GMR for quantitative measurements and Hedin classification for qualitative measurements. GMR determines the presence or absence of gingiva pigmentation at the target site. GMR is used to evaluate the presence or absence of gingiva pigments at the assessment site and is represented as a percentage by dividing the pigmented assessment sites by all assessment locations. Hedin classification uses the classical evaluation method of gingiva pigmentation, and gingiva pigmentation is analyzed qualitatively on a scale of 0 to 4 [13][14][15][16].

This study aims to design a system that can classify the degree of pigmentation of smoker melanosis in the gingiva of smokers. The system uses digital image processing to obtain accurate and fast data. As well as, the classification process combines two modes to produce better accuracy. The expected benefit of this study is the help of dentists in classifying the type of pigmentation of smoker melanosis in the gingiva, and it can also be a learning material regarding smoker melanosis and gingiva image processing.

II. Proposed Method

The research method design is outlined in the block diagrams in Figures 1 and 2. This methodology describes the stages involved in classifying the pigmentation level of smoker's melanosis in the gingiva. Data input is done by capturing images of smoker's melanosis. Data entry was carried out as an image of smoker melanosis. The next stage of the imagery is a cropping process that aims to make the image obtained focused on the gingiva because the pigmentation is in the gingiva. After that, it is continued by carrying out an image processing process that aims to improve image quality to get better data and get a clear and precise melanosis smoker object. Once the image processing is completed, the resulting data will be prepared for the dataset stage. For this dataset, both training and test data are prepared. The features used as classification parameters are derived from extracting pigmentation area and color features from each RGB channel [17][18]. The dataset used in this study comprises general image data. Images of smoker's melanosis on the lips are saved in JPG format, with the dataset divided into different categories, named DOPI 0, DOPI 1, DOPI 2, and DOPI 3 [6][9].

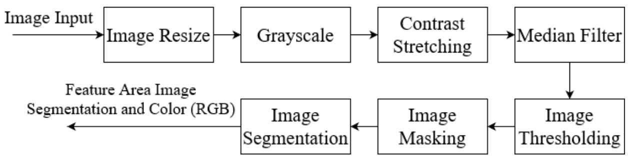

Fig. 1. Image Processing Block Diagram

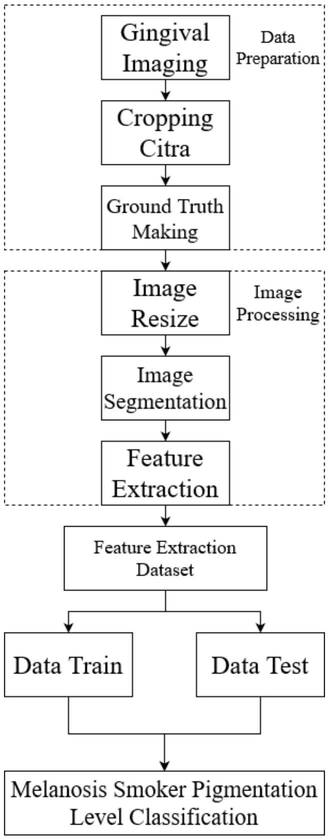

Fig. 2. Functional Diagram of Research Methods

A. Dataset

This process uses an image-taking method using the oral photograph technique[7][15]. In addition, this procedure uses an auxiliary tool in the form of an oral opening tool to help take an image to get a gingiva image of the patient. In the process of taking imagery using a camera with a resolution of 12 Mega Pixels with the condition of taking flash imagery on and the camera distance with objects 6 cm and fixed camera focus as shown in Figures 3 and 4 [19].

Fig. 3. Image Capture Process

Fig. 4. Camera Posis Calibration Process

In the cropping process, the image is manually cropped to focus on the gums (gingiva) and the smoker's melanosis pigmentation [20][14]. In this process, the process of making ground truth is carried out, which aims to become a data train, while the data test is used using data that is not carried out manually by researchers and experts. For the making of ground truth, it is carried out by taking random data taken. Then a classification process is carried out according to the method by comparing the image obtained with the GPI and DOPI classifications. As for the classification process, it will produce four types of classifications, namely with a level of DOPI 0 (no clinical pigmentation), DOPI 1 (no mild clinical pigmentation), DOPI 2 (moderate clinical pigmentation), and DOPI 3 (severe clinical pigmentation)[6].

B. *Image Processing*

In the image processing phase, cropping will be carried out first to focus on the pigmentation in the gingiva. The image then will be segmented to ensure that the pigmentation object is clearly visible, providing more accurate data According to M. Nasir's research, the initial process of image processing was carried out cropping, which aimed to focus the image on the pigmentation of the smoker melanosis in the smoker's gums [20][21].

C. *Greyscale*

Colour representation in the imagery is achieved by using a combination of one or more colour channels that are combined to form the colours used in the imagery. The representation used to store colours, determining the number and nature of colour channels, is commonly known as colour spaces. In general, the image has 3 colour channels which consist of red (R), Green (G), and Blue (B). Meanwhile, grayscale imagery uses a single-channel colour space that is limited to 2 – bit (binary) or intensity (grayscale) colour space [17].

In the gingiva image processing process, following the cropping process, a grayscale conversion is performed [18]. One common usage for digital image processing, particularly in industries like medical, is the conversion of colored images to grayscale. This conversion provides a number of advantages, including lower printing costs, sometimes better visual clarity, and other useful uses [22].

D. *Contrast Stretching*

The image histogram is employed for contrast stretching (also known as normalization), which expands the pixel intensity range of the input image to cover a broader dynamic range in the output image. To perform contrast stretching, the upper and lower limits of pixel values to be normalized, denoted as a and b, must first be identified. These are generally the upper and lower limits of the pixel quantization range used (i.e. for 8-bit images, a 255 and b 0) [17]. In its simplest form, the first part of the contrast stretch operation scans the input image to determine today's maximum and minimum pixel values, denoted by c and d, respectively. Then, based on these four values (a, b, c and d) [23].

E. *Median Filter*

Median filters are one of the noise reduction methods that are often used in image or signal processing. As it functions as a noise reducer, the median filter is used in the pre-processing process to improve the quality of the results to be the input of the following process (for example, edge detection in an image) [24]. This method effectively removes the 'salt and pepper type of noise, namely black and white spots noise. The median filter works by moving through pixel after pixel of the image, replacing the pixel value with the median value of the neighbouring pixel [17].

F. *Thresholding*

Thresholding is the binarisation process of an image [25]. In this process, the image that has gone through the filter process will be processed binary, whose value of T has been adjusted so that the result follows the desired target. For example, this study focused on blackish or brown pigmentation of the gums. If you have found the correct T value, then setting all the intensity of all pixels below the value of T will be worth 0, while for intensity with a value above T, it will be worth 255. This method can display an image with a binary representative [17].

G. *Contours and Masking*

Contours are lines connecting all points along an image's boundaries with the same intensity. Contours help analyze shapes, find desired sizes, and detect objects.[26] After obtaining the designated contour area, the image results are processed by the masking method to ensure that the contour area obtained is under the pigmentation, which is an essential point in this study. Where the masking process is the process of imagery inserting or covering an object so that the inserted image looks transparent and blends with the object covered [17].

H. *Image Segmentation*

Segmentation is the name given to a generic process in which an image is divided into regions or their constituent objects. Segmentation occupies a critical role in image processing because it is often a vital first step that must be successfully performed before subsequent tasks such as feature extraction, classification, description, etc. The primary purpose of segmentation is to partition the images into mutually exclusive territories to attach meaningful labels. Segmented objects are often called foregrounds and the rest of the image is the background. Accurate image segmentation depends mainly on the type of object or region we want to identify [27][28].

I. Dataset Preparation

After completing the image processing stage, the dataset preparation process will be carried out. Where for the preparation of this dataset will be divided into 2 parts of the dataset, including training data and testing data. If the data obtained experiences a case where the data obtained is unbalanced, a synthetic minority oversampling technique (SMOTE) process would be carried out. SMOTE is the most popularly used method of oversampling. SMOTE is done by adding synthetic data to minority classes. Over-sampling in the minor class is done by creating synthetic samples, which are new samples generated, rather than duplicating data. SMOTE is carried out by increasing the amount of data on minority classes by generating new data based on the k of the nearest neighbor. Data on minority classes are oversampled by taking data on minority classes and adding synthetic samples along the lines connecting one or all of the nearest neighbours of the minority class data. The number of neighbours k was randomly selected [29].

J. Classification Process

K-Nearest-Neighbors (KNN) is a non-parametric classification method which is simple but effective in most cases. For the data of the t record to be classified, the neighbouring k is taken, forming the environment t. The majority voting among data records in an environment is typically used to decide classifications for t with or without consideration of distance-based weighting. However, to implement kNN, we need to select the appropriate k value, and the success of the classification depends mainly on this 11value. In a sense, the kNN method is biased by k. There are many ways to choose the value of k, but the simplest is to run the algorithm many times with different values of k and choose the one with the best performance.

In this process, the features used are the Classification of Pigmentation Degrees, focusing on the extent of the image containing pigmentation and the Colour Degree of Oral Pigmentation Dummett, which focuses on the colour features in the pigmentation of smoker melanosis. By getting the value of each predetermined feature through digital image processing. In addition to paying attention to the features of the imagery used in this classification process using distance calculations using the Minkowski distance formula [30].

K. Data Testing Scenarios

Data testing is conducted in three stages: the image processing stage, dataset preparation stage, and image classification stage. This ensures a thorough evaluation of the entire system.

III. EXPERIMENTAL RESULT

This chapter will explain the testing of systems that have been formed according to the methodology, namely image processing, data sets, and classification, as well as the preparation of data sets using Hedin and DOPI classifications, which aim to determine the ground truth of pigmentation following previous studies. Meanwhile, all programming and algorithm experiments were carried out in this study using the python programming language with the Jupyter Notebook IDE.

There are three stages carried out in this section: (1) Explicit Imagery. In this stage, the results consist of oral images, as shown in Figure 5. During image capture, the distance between the camera and the object, the focus, and the use of a flash are critical. The distance significantly affects image cropping results, potentially distorting the cropped image. The focus influences the segmentation process; poor focus can prevent accurate extraction of the pigmentation area, as the distorted neighborhood values cause errors. Flash use affects light intensity, complicating digital processing. (2) Image Cropping. This step produces images focusing on the gingiva as seen in Figure 6. However, the resulting cropped images have varying pixel dimensions. (3) Ground Truth Creation. This stage produces DOPI segmentation levels of 0, 1, 2, and 3 as seen in Figure 7, where DOPI 0 indicates no clinical pigmentation, DOPI 1 represents mild brown pigmentation, DOPI 2 indicates moderate brown or mixed pink and brown pigmentation, and DOPI 3 indicates severe dark brown, blue, or black pigmentation.

Fig. 5. Image Retrieval Results

Fig. 6. Image Crop and Resize Results

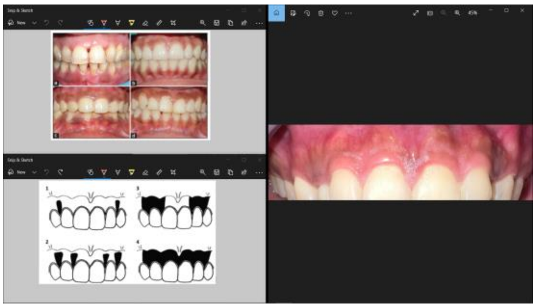

Fig. 7. Pigmentantion Type Determination

This image processing process is carried out to find features in the image before the classification process. The process begins with resizing the images, adjusting the width and length of the cropped images to ensure uniformity in size across all samples. This process aims to make all images of the same size to extract features for each image obtained more constant.

Furthermore, the image change process that begins with having 3 colour channels (Red, Green, and Blue) will be converted into grayscale imagery, which aims to simplify the imagery into 2 colour channels. Once the image becomes simpler, contrast stretching (also known as normalization) is performed, stretching the input image’s pixel intensity range to occupy a /more extensive dynamic range in the output image.

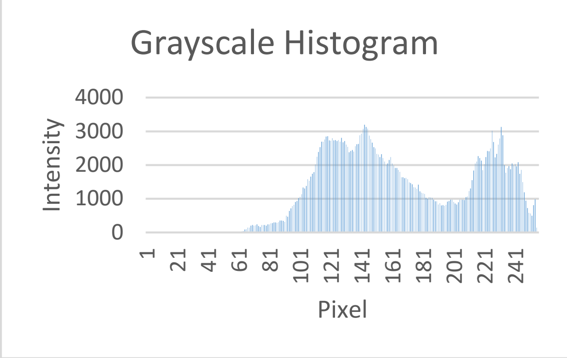

Fig. 8. Grayscale histogram

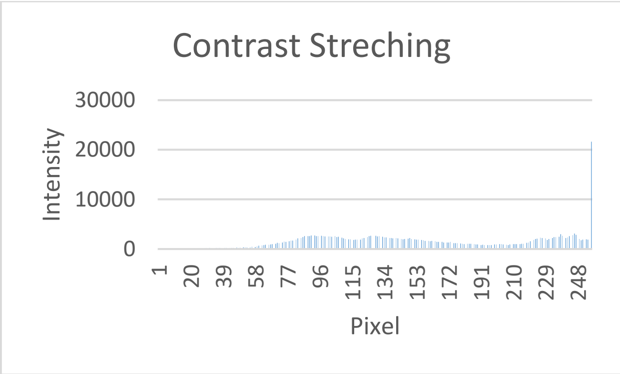

Fig. 9. Histogram Contrast Stretching

Median filters are one of the noise reduction methods that are often used in image or signal processing. As it functions as a noise reducer, a median filter is used in the pre-processing process to improve the quality of the results to be the input of the following process (for example, edge detection in an image). This method effectively removes the ‘salt and pepper type of noise, namely black and white spots noise. This can be observed in Figure 10, which illustrates the effect of the median filter. The filter operates by traversing the image pixel by pixel, replacing the value of each pixel with the median value of its neighboring pixels.

Fig. 10. Median Filter

Thresholding is a technique used to binarize an image. In this step, the image, having undergone filtering, is converted into a binary form, where the threshold value (T) is carefully adjusted to achieve the desired output. For instance, this study focuses on black or brown pigmentation on gums. A binary representation is produced once a suitable T value is established, with all pixel intensities below this value set to 0 and those above T given a value of 255. Figure 11 demonstrates the result of this thresholding process. Contours, which connect points along the boundaries of an image with the same intensity, are helpful for analyzing shapes, determining sizes, and detecting objects.

Fig. 11. Thresholding

After obtaining the designated contour area, the results of the image are processed by the masking method to ensure that the contour area obtained is following the pigmentation, which is an essential point in this study, where the masking process is the process of inserting or covering an object so that the inserted image looks transparent and blends with the object being covered.The result of this masking process is presented in Figure 12.

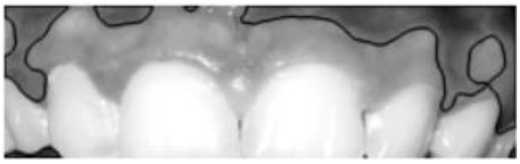

Fig. 12. Masking

Images with black gum coloring are the subject of this segmentation process. where the image was masked and the result was obtained. After that, bitwise OR surgery is carried out using the original image in comparison to image masking, which tries to produce a new image that highlights the smoker's melanosis pigmentation and removes characteristics like teeth and gums that lack pigmentation.

In this process, several features will be used from the imagery obtained through the segmentation process. For the first feature, look for the average value of each colour channel. In this process, calculations are carried out by looking for the average value of each colour pixel row, and then a process is carried out to calculate the average value of the colour pixel row. After getting the features of each colour, calculating the area of the segmentation area is carried out by calculating the area of the entire pixel area of the image area and then looking for pixels from the segmented image.The results of the segmentation process for DOPI 0 and DOPI 3 are presented in Figures 13 and 14.

Fig. 13. Pigmentation Segmentation Smoker Melanosis DOPI 0

Fig. 14. Pigmentation Segmentation Smoker Melanosis DOPI 3

Based on the DOPI classification, the feature extraction results exhibit skewed data distribution, with DOPI 2 having a significantly larger proportion of the data compared to DOPI 0, DOPI 1, and DOPI 3 as shown in Figure 15. This imbalance may impact the accuracy of the classification process.

Table 1 provides a summary of the data distribution among the DOPI classes. The Synthetic Minority Over-sampling Technique (SMOTE) will be used to balance the dataset as a result of this unequal distribution. As shown in Figure 16, the amount of synthetic data produced for the SMOTE process will be modified to correspond with the amount of data in the DOPI 2 class.

TABLE I. DATA DISTRIBUTION RESULTS FOR EACH CLASS

	DOPI 0	DOPI 1	DOPI 2	DOPI3
TOTAL	4	4	20	1

The K-Nearest Neighbors (kNN) technique was used to classify the data after feature extraction was finished. Using the Minkowski distance formula, the kNN classification divided the dataset into 30% for testing and 70% for training. Furthermore, parameters like the k-value (the number of nearest neighbors) and the p-value for the Minkowski formula were tuned. As seen in Table 2, the p-value was set to 3 and the k-value to 2.

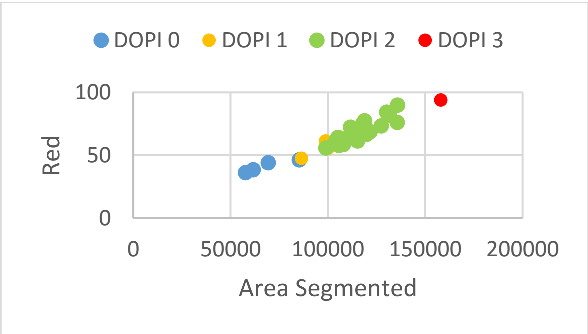

Fig. 15. Results of Data Distribution of Pigmentation Features of Smoker Melanosis

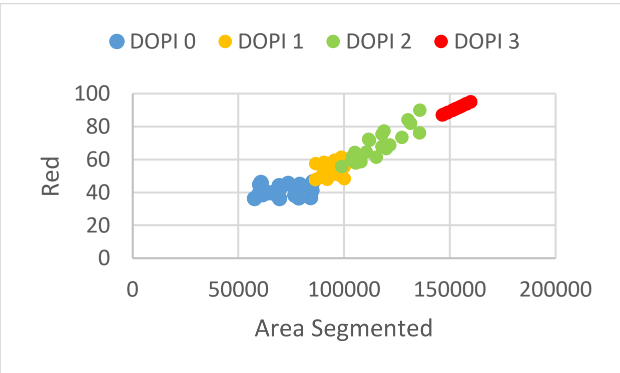

Fig. 16. Results of Data Distribution of Smoker Melanosis Pigmentation Feature (SMOTE)

TABLE II. CLASSIFICATION ACCURACY RESULTS

No	Class	Precision	Recall	F1 - Score
1	DOPI 0	1.0	1.0	1.0
2	DOPI 1	1.0	1.0	1.0
3	DOPI 2	1.0	1.0	1.0
4	DOPI 3	1.0	1.0	1.0

IV. CONCLUSION

This study created a methodology to categorize the degree of pigmentation in smoker's melanosis utilizing photographs of the gingiva using two classification techniques: the Dummett oral color pigmentation classification and the degree of pigmentation based on the pigmentation area. The testing phase yielded several important conclusions. First, when creating datasets and taking pictures with a smartphone, several factors need to be taken into account. Image quality was greatly affected by variables like focus, flash usage, and the distance of the camera from the item. Image processing was complicated by improper focus or distance, especially after cropping when the image was compressed. While the use of flash helped maintain consistent lighting conditions and improved pigmentation visibility, blurry photos decreased the accuracy of pigmentation information. Additionally, the cropping operation produced diverse image shapes because of the different diameters of the gum. Second, meticulous picture processing, including bitwise operations, was necessary to preserve the image's RGB output.

The parameters tested in this study showed that the system effectively processed images of smoker's melanosis pigmentation with high accuracy. Finally, the K-Nearest Neighbors (kNN) classification method, applied to the extracted features (mean values of color channels and segmentation area), showed promising results. The data distribution across DOPI 0, DOPI 1, DOPI 2, and DOPI 3 classes allowed for clear grouping in the feature extraction plots. Using kNN, the system achieved an accuracy of 1.0, with an F1-score of 1.0 for all DOPI classes (0-3). This demonstrates that the smoker's melanosis pigmentation can be successfully classified using kNN, provided that the extracted features align with pigmentation degree classifications and that consistent image capture conditions are maintained. Future work should focus on developing faster image processing methods that retain segmentation accuracy, as well as exploring more advanced classification techniques such as Convolutional Neural Networks (CNN) to further improve classification performance.

REFERENCES

[1] A. Drovandi, P. A. Teague, B. Glass, and B. Malau-Aduli, "A systematic review of smoker and non-smoker perceptions of visually unappealing cigarette sticks," *Tob. Induc. Dis.*, vol. 16, no. January, pp. 1–11, 2018, doi: 10.18332/tid/82191.

[2] S. M. Noar, D. B. Francis, C. Bridges, J. M. Sontag, K. M. Ribisl, and N. T. Brewer, "The impact of strengthening cigarette pack warnings: Systematic review of longitudinal observational studies," *Soc. Sci. Med.*, vol. 164, pp. 118–129, 2016, doi: 10.1016/j.socscimed.2016.06.011.

[3] S. M. Noar, M. G. Hall, D. B. Francis, K. M. Ribis, J. K. Pepper, and N. T. Brewer, "Pictorial cigarette pack warnings: A meta-analysis of experimental studies," *Tob. Control*, vol. 25, no. 3, pp. 341–354, 2016, doi: 10.1136/tobaccocontrol-2014-051978.

[4] S. Vellappally, Z. Fiala, J. Smejkalová, V. Jacob, and R. Somanathan, "Smoking related systemic and oral diseases," *Acta Medica (Hradec Kralove)*, vol. 50, no. 3, pp. 161–166, 2007, doi: 10.14712/18059694.2017.76.

[5] S. Araki, K. Murata, K. Ushio, and R. Sakai, "Dose-response relationship between tobacco consumption and melanin pigmentation in the attached gingiva," *Arch. Environ. Health*, vol. 38, no. 6, pp. 375–378, 1983, doi: 10.1080/00039896.1983.10545823.

[6] C. A. Hedin, "Smokers' Melanosis: Occurrence and Localization in the Attached Gingiva," *Arch. Dermatol.*, vol. 113, no. 11, pp. 1533–1538, 1977, doi: 10.1001/archderm.1977.01640110053007.

[7] T. Kato, S. Mizutani, H. Takiuchi, S. Sugiyama, T. Hanioka, and T. Naito, “Gingiva pigmentation affected by smoking among different age groups: A quantitative analysis of gingiva pigmentation using clinical oral photographs,” *Int. J. Environ. Res. Public Health*, vol. 14, no. 8, Aug. 2017, doi: 10.3390/ijerph14080880.

[8] M. A. Peres *et al.*, “Oral diseases: a global public health challenge,” *Lancet*, vol. 394, no. 10194, pp. 249–260, 2019, doi: 10.1016/S0140-6736(19)31146-8.

[9] M. Greenberg and M. Glick, *Oral Medicine Diagnosis & Treatment*, 2003.

[10] A. Nuraini, A. Harmoko, and B. Kiswanjaya, “Improvement of smoker prediction system based on hyperspectral image with hybrid deep learning model,” in *Int. Electron. Symp. 2021 Wireless Technol. Intell. Syst. Better Hum. Lives, IES 2021 - Proc.*, Sep. 2021, pp. 417–422, doi: 10.1109/IES53407.2021.9594025.

[11] F. I. Hawari *et al.*, “Respiratory health and quality of life in young exclusive, habitual smokers—a comparison of waterpipe smokers, cigarette smokers and non-smokers,” *Int. J. COPD*, vol. 14, pp. 1813–1824, 2019, doi: 10.2147/COPD.S205050.

[12] K. R. Rosalin Hongsathavij and Y. Kuphasuk, “Correlation of environmental tobacco smoke to gingiva pigmentation and salivary alpha amylase in young adults,” *Eur. J. Dent.*, vol. 11, no. 4, pp. 192–195, 2017, doi: 10.4103/ejd.ejd.

[13] R. N. Smith, D. L. Lath, A. Rawlinson, M. Karmo, and A. H. Brook, “Gingiva inflammation assessment by image analysis: measurement and validation,” *Int. J. Dent. Hyg.*, vol. 6, no. 2, pp. 137–142, 2008, doi: 10.1111/j.1601-5037.2008.00294.x.

[14] H. Seshan and M. Shwetha, “Gingiva inflammation assessment: Image analysis,” *J. Indian Soc. Periodontol.*, vol. 16, no. 2, pp. 231–234, 2012, doi: 10.4103/0972-124X.99267.

[15] T. Kato *et al.*, “Measurement of reduced gingiva melanosis after smoking cessation: A novel analysis of gingiva pigmentation using clinical oral photographs,” *Int. J. Environ. Res. Public Health*, vol. 13, no. 6, 2016, doi: 10.3390/ijerph13060598.

[16] H. Kaur, S. Jain, and R. Sharma, “Duration of reappearance of gingiva melanin pigmentation after surgical removal—A clinical study,” *J. Indian Soc. Periodontol.*, vol. 14, no. 2, p. 101, 2010, doi: 10.4103/0972-124X.70828.

[17] A. Vyas, S. Yu, and J. Paik, *Fundamentals of Digital Image Processing*, 2018.

[18] S. Shawal, M. Shoyab, and S. Begum, “Fundamentals of Digital Image Processing and Basic Concept of Classification,” *Int. J. Chem. Process Eng. Res.*, vol. 1, no. 6, pp. 98–108, 2014, doi: 10.18488/journal.65/2014.1.6/65.6.98.108.

[19] T. Eckhard, E. M. Valero, J. L. Nieves, and F. Mesa, “The Labial Teeth and Gingiva Photographic Image Database LTG-IDB,” [online resource].

[20] Q. Wang and Y. Yuan, “Learning to resize image,” *Neurocomputing*, vol. 131, pp. 357–367, 2014, doi: 10.1016/j.neucom.2013.10.007.

[21] G. Iwasokun, "Image Enhancement Methods: A Review," *British Journal of Mathematics & Computer Science*, vol. 4, no. 16, pp. 2251–2277, 2014, doi: 10.9734/bjmcs/2014/10332.

[22] C. Saravanan, "Color image to grayscale image conversion," in *2010 2nd International Conference on Computer Engineering and Applications (ICCEA)*, 2010, vol. 2, pp. 196–199, doi: 10.1109/ICCEA.2010.192.

[23] M. Nasir *et al.*, "New method using image analysis to measure gingiva color," *Journal of Osaka Dental University*, vol. 38, no. 1, pp. 37–40, 2004, doi: 10.18905/jodu.38.1_37.

[24] G. Gupta, "Algorithm for Image Processing Using Improved Median Filter and Comparison of Mean, Median and Improved Median Filter," *International Journal of Soft Computing*, no. 5, pp. 304–311, 2011. [Online]. Available: http://ijsce.org/attachments/File/Vol-1_Issue-5/E0234101511.pdf

[25] P. Roy, S. Dutta, N. Dey, G. Dey, S. Chakraborty, and R. Ray, "Adaptive thresholding: A comparative study," in *2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*, 2014, pp. 1182–1186, doi: 10.1109/ICCICCT.2014.6993140.

[26] S. Dey, R. Sarkar, K. Chatterjee, P. Datta, A. Barui, and S. P. Maity, "Pre-cancer risk assessment in habitual smokers from DIC images of oral exfoliative cells using active contour and SVM analysis," *Tissue & Cell*, vol. 49, no. 2, pp. 296–306, 2017, doi: 10.1016/j.tice.2017.01.009.

[27] A. Rana, G. Yauney, L. C. Wong, O. Gupta, A. Muftu, and P. Shah, "Automated segmentation of gingiva diseases from oral images," in *2017 IEEE Healthcare Innovation Point-of-Care Technologies (HI-POCT)*, 2017, vol. 2017-Dec., pp. 144–147, doi: 10.1109/HIC.2017.8227605.

[28] T. Eckhard, E. M. Valero, and J. L. Nieves, "Labial teeth and gingiva color image segmentation for gingiva health-state assessment," in *6th European Conference on Colour in Graphics, Imaging, and Vision (CGIV)*, 2012, pp. 102–107.

[29] C. Nitesh V, B. Kevin W, H. Lawrence O, and P. K. W, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002, doi: 10.1002/eap.2043.

[30] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, "KNN model-based approach in classification," in *Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 2888, pp. 986–996, 2003, doi: 10.1007/978-3-540-39964-3_62