

Article

Palatal Rugoscopy for Biometric Identification based on ResNet50 Feature Extraction and Machine Learning Classifiers

Tamara Afif Anai^{*1}

1. Department of Basic Science (Computer Science), Dentistry, Tikrit University, Tikrit, Iraq

*Correspondence: tamsamka@tu.edu.iq

Abstract: The Palatal Rugoscopy have highly stable and unique morphological aspects and can be used in odontology as a ‘fingerprint’ of the mouth for reliable human identification in Forensic Dentistry. Though AI is increasingly being used in biometrics, the need for local datasets, especially for the Iraqi population is stark from College of Dentistry Teaching Hospital at the University of Tikrit. In addition, deep learning on newly collected small medical dataset results in significant overfitting and computational burden. Based on this, this study aims to present a lightweight hybrid ML-based system customized for a new manually curated dataset of palatal rugoscopy from Iraq population. The pipeline only retrains the parameters of the ResNet50 backbone for deep feature extraction, without retraining the whole model, which can be beneficial due to the number of parameters. High dimensional features are extracted and then classified using robust classifiers such as SVM, K-NN, ANN and the model generalization was guaranteed through targeted geometric and radiometric data augmentation. It is observed that SVM has the highest overall performance which yields 98.00% accuracy, 97.00% precision, 98.00% recall, and 97.33% F1-Score among the tested classifiers, even with a small number of initial samples. This first-of-its-kind study in the region introduces a highly efficient, computationally lightweight proof-of-concept for forensic identification which can be used to develop scalable, automated biometric systems in forensic Dentistry.

Keywords: Palatal Rugoscopy, Biometric, ResNet50, SVM, KNN, ANN

Citation: Anai, T. A. Palatal Rugoscopy for Biometric Identification based on ResNet50 Feature Extraction and Machine Learning Classifiers. Central Asian Journal of Mathematical Theory and Computer Sciences 2026, 7(4), 433-453.

Received: 27th Aug 2026

Revised: 15th Sep 2026

Accepted: 07th Oct 2026

Published: 27th Oct 2026



Copyright: © 2026 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>)

1. Introduction

One of the most important needs in Forensic Dentistry (FD) and medico of legal studies is the identification of individual. These are highly reliable and accurate methods of identification that are traditionally used, such as fingerprints, dental records, and DNA profiling. However, they cannot be always applied when the condition of the biological evidence is very serious such as severe trauma, fire, mass disaster or high-level postmortem decomposition, where the biological evidence may be lost. These limitations will lead to a search for other and additional biometric indicators which are stable, unique and not subject to post mortem changes. Thus, the automatic image analysis techniques are gaining more importance in clinical and forensics studies. Regulatory image acquisition protocols will include uniformity in image resolution, orientation, illumination, contrast, and scale, which will reduce variability caused by variations in imaging equipment. The automated of classification makes the diagnosis process reliable and reproducibly [1-3]. The palatal rugae are among the many anatomical features studied in the field of FD due to their unique, stability and resistance to post-mortem changes. Palatal rugoscopy are uneven, irregular ridges of mucosa that are situated in the anterior part of the hard palate, which extends laterally off of the incisive papilla. They are very

convenient to use in trauma cases, fire cases and cases of high degree of decomposition because they are safe to use intramuscularly and are not exposed to environmental or mechanical factors. Every individual has distinct morphological features of palatal rugae, and they do not change significantly in the course of life. The research results are preliminary and have been established to determine their forensic utility through demonstrating uniqueness and consistency over time. Besides forensic identification, palatal rugoscopy are also significant in clinical dentistry, and they help in the process of mastication, speech articulation, orthodontic evaluation, and prosthodontic planning. They are also fixed in position and can serve as a reference point for the evaluation of dental treatment. Palatal rugoscopy analysis has been demonstrated as a clinically and forensically important technique, but has traditionally relied on manual methods such as dental casts, photographs and visual tracing. Though effective, these techniques are time consuming and may be subjective and non-standardized by the observer. This has led to the need for structured, repeatable solutions that integrate different technical supports with dental know-how, to ensure precision and reliability [4-6].

The recommended this approach to reduce the amount of manual work and reduce observer bias and to enhance the accuracy, and reproducibility. The main dataset was limited to 200 original images because of its novel use in the analysis of palatal Rugoscopy collected from college of dentistry, University of Tikrit, Iraq. We used a very simple yet very efficient model to cope with this small amount of data and to minimize the risk of overfitting, carefully. Although it is simple in its architecture, the selected model utilizes a rich and novel methodological approach specific to extracting the unique features of palatal rugoscopy. Due to the complete absence of prior benchmark datasets or local comparative literature for this specific category, this work serves as a foundational proof-of-concept for future scalable expansions. A multi-stage processing and augmentation approach was used to preprocess the raw palatal rugae image for feature extraction and to create a robust model under limited sample sizes as following steps:

1. The first step is standardized preprocessing and normalization.

To ensure that all input RGB images had the same spatial dimension of $224 \times 224 \times 3$ pixels as required by the ResNet50 backbone architecture. The Standard Caffe-style preprocess_input transformation used for ResNet50 was applied to preprocessed images.

2. Synthetic Data Augmentation

To artificially increase intra-class variability, avoid over-reliance on deep feature orientation, and resist overfitting problems in down-stream classifiers based on the empirical data, a pipeline for data augmentation was systematically applied to the data set, consisting of five folds of data augmentations.

3. Deep Feature Representation Vector

After preprocessing and augmentation, the images were passed to the frozen pre-trained ResNet50 model for processing. Each processed image was converted to a small but very informative 2048-dimensional feature vector using the Global Average Pooling (GAP) layer and immediately fed into the next stage of classification (K-NN, SVM, and ANN) models.

Gap of Study

Even with the significant progress of AI in biometric identity verification and FD, there are still several key research areas that need to be addressed to enable widespread adoption of automated systems:

- a. Most current biometric databases are extremely biased towards certain populations on Earth. A significant lack of ethnically specific datasets has been detected for palatal rugoscopy in the Middle East especially in Iraq.
- b. Deep Convolutional Neural Networks (CNNs) have achieved state-of-the-art results on image classification while are extremely complex with hundreds of millions of parameters that make them highly prone to overfitting when trained from scratch on newly curated small-scale medical or forensic datasets.

- c. End-to-end deep learning pipelines generally need considerable computational power (high-end GPUs). This is a problem for deployment in standard forensic laboratories or clinical environments where rapid (lightweight and offline) inference is required.

Contributions of Study

To overcome these basic problems, a novel, lightweight hybrid machine learning framework is suggested in this study. This research has several specific contributions including providing a new benchmark for forensic biometrics for data-poor environments:

- Introducing the first carefully and manually curated dataset of palatal rugae for the Iraqi population. This database becomes a vital proof of concept and benchmark for future research in forensic odontology in the area.
- Hybrid Framework, We developed a very efficient, hybrid classification pipeline that separates feature extraction from decision making. We were able to successfully map complex morphological rugae patterns into dense vectors of 2048-dimensional without parameter-rich retraining of a pre-trained ResNet50 backbone, which is used only as a generic feature descriptor.
- Transfer Learning (frozen weights) combined with targeted geometric and radiometric data augmentation, followed by simple, robust classical classifiers (SVM, K-NN and ANN) is an effective overfitting mitigation strategy that is demonstrated empirically. This offers a scalable framework for implementing deep learning technology with very constrained medical datasets.
- The proposed framework will drastically reduce the training time and computational cost, as only shallow classifiers are used, hence avoiding deep convolutional layers. It demonstrates high Accuracy, Recall, Precision and F1 score, showing that high precision automated biometric identification can be performed on standard clinical computing systems.

The remainder of the article as follow: In section 2 show the related of work, Section 3 presents the Proposed-Architecture, section 4 display the result and discussion and Section 5 presented conclusion and future work.

Related Works

Palatal Rugoscopy analysis has seen a substantial advancement from manual FD to automated biometric identification. The palatal Rugoscopy have been recognized for their biological uniqueness, post-mortem stability and resistance to the environmental trauma and have acquired considerable importance in the fields of forensic identification and digital orthodontics [7-12]. Research has since broadened to automatic texture extraction and non-linear filtering methods [7, 9] to do away with manual landmark selection. In order to extract non-linear features from rugae 9 images, Wu et al. [7] proposed a Gabor filter-based framework which segmented the images into regional sectors. They had a 95.67% accuracy in comparing pre- and post-orthodontic treatment cases. Likewise, Principal Component Analysis (PCA) was used for dimensionality reduction and Adaptive Neuro-Fuzzy Inference System (ANFIS) for classification by Kusumaputri et al. [9]. The system obtained classification accuracy of 91.38% and a fast computation time of 0.2201 seconds, but its generalization ability was limited due to the small number of training data.

This problem was addressed by Shangguan et al. [8] who developed a direction vector-based correction algorithm that fits orientation lines by using the anatomical landmarks, including the incisive papilla and the midpoints of molars. Their approach was able to achieve a mean angular correction error of less than 1 and a variance of less than 0.5, which led to a strong image pre-processing for forensic identification pipelines.

In recent years, technologies have been added that utilise deep learning models and automated web management systems [10-12] with the advancement of AI. To overcome the limitation of 2D images, Bsat et al. [10] developed a hybrid deep learning system,

SegNet, for 3D dental motion tracking by incorporating spatial and channel attention mechanisms. The method they developed resulted in a segmentation accuracy of 98.69% and an IoU of 84.5% in 3D space after performing 2D semantic segmentation on multi-view projections and back-projecting the texture onto the 3D mesh models of the maxilla. Moreover, they were able to precisely determine the 3D translations and rotations of the teeth using the stable 3D rugae areas as reference points for an Iterative Closest Point (ICP) registration algorithm without requiring any manual registration markers.

Amos et al. [11] designed a hybrid dento-technical framework that utilizes Convolutional Neural Networks (CNNs) model for automatic image segmentation and traditional machine learning models (SVM and Random Forest) for quantification of morphometric changes after RPE. Concurrently, Morye et al. [12] designed an automatic web-based system based on OpenCV (Otsu thresholding and Gaussian blur), MD5 hash of images, and histogram matching to classify rugae according to the presence of structured 4+4 patterns for fast forensic database queries. Table 1 display the related work with method.

Table 1. The method related work

Method	Methodology & Algorithms	Dataset & Availability	Limitations & Weaknesses	Accuracy
Wu et al. [7]	Non-linear Gabor filter feature extraction, image sectoring, and MATLAB-based matching.	30 images (Private dataset)	Sensitive to camera tilt and lighting variations; restricted to 2D image analysis.	95.67 %
Shangguan et al. [8]	Orientation vector correction algorithm, line fitting, and angle-rotation adjustment.	120 images (Private dataset).	Relies heavily on intact anatomical landmarks (incisive papilla, molars); fails with severely damaged arches.	
Kusumaputri et al. [9]	Principal Component Analysis (PCA) for feature selection + Adaptive Neuro-Fuzzy Inference System (ANFIS). SegNet Encoder-Decoder with Spatial/Channel Attention + 3D	Intraoral rugae photos divided into training/testing sets (private dataset).	Small sample size; sensitive to pre-processing alignment errors; lacks 3D structural evaluation.	91.38%
Bsat et al. [10]	Back-Projection & Iterative Closest Point (ICP) registration.	86 images (Private clinical dataset).	Reduced segmentation accuracy on primary, unerupted, or impacted wisdom teeth.	84.5%
Amos et al. [11]	Hybrid framework: CNN for 2D image segmentation + Classical ML	16 images (Limited clinical dataset).	Small sample size; limited to 2D morphometrics without	

	(SVM / Random Forest) + Paired t-tests.		depth/volume tracking.
	OpenCV preprocessing (Otsu thresholding, Gaussian blur), MD5 hashing, histogram matching, Flask web framework.	Multi-source forensic database of palatine rugae images (Semi-public/private dataset).	Uses deterministic image hashing instead of deep spatial feature representations; sensitive to noise.
Morye et al. [12]			

2. Methodology

Dataset Description and Collection

In this research, the data were collected through palatal rugoscopy of 200 individuals. Considering the similarity to exploratory FD studies, which have focused on morphometric rugoscopy analysis. Data collection manually from the Teaching Hospital of the College of Dentistry, University of Tikrit, Iraq, after obtaining research ethics approval, and all samples were anonymized. Consistency was achieved by obtaining digital images of the palatal area with the aid of high-resolution intraoral photography with the help of standardized illumination, and orientation. The datasets are available from the corresponding author. Figure 1 shows the example of the digital palatal rugoscopy. The shape of palatal rugoscopy analysis based on the classification of the rugae pattern by Thomas and Kotze [13] as shown in Figure 2. The shape of the rugae pattern can be divided into straight, wavy, curvy, circular, diverging, and converging. Figure 3. Shows the shape of palatal rugoscopy.



Figure 1. The digital palatal rugoscopy images.

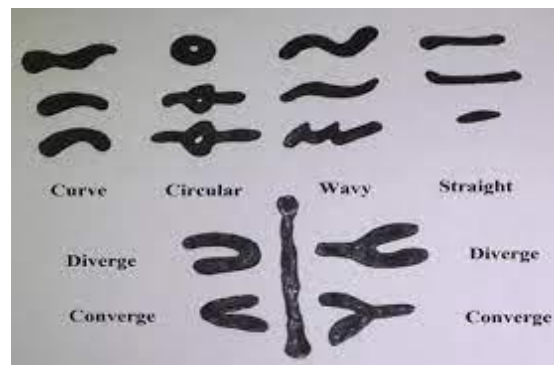


Figure 2. Thomas and Kotze classification of palatal rugosity pattern [13].

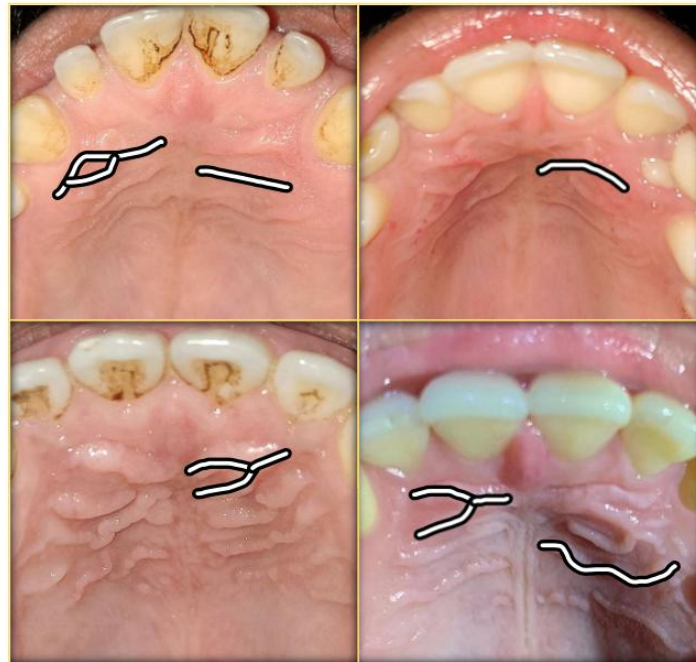


Figure 3. Palatal rugosity types.

Data Acquisition

Using the single reflector camera (Canon EOS 300D), which is set the same parameters, fixed position and fixed focal length, to obtain the high resolution and of reference value intraoral palatal rugosity images. The image processing software Photoshop is used to perform a preliminary cutting process. The camera has the following features: Specially customized equipment to make the image standard unified. All pictures of the palatal rugosity are used; the images are transmitted to the computer in JPEG format.

The Computational Framework

Training of the model and evaluation was done locally and an NVIDIA GeForce RTX 5060 OC (8GB GDDR7) GPU was used to accelerate the computation time. The initial the entire set of 200 high-resolution images after become 1000. In order to train the complex Ensemble network under the constraint of 8GB VRAM on the GPU, memory-management techniques like mixed precision training and gradient accumulation were used. This enabled the system to sustain an effective batch size that is large enough to guarantee stable convergence, while avoiding memory constraints.

Data Augmentation Techniques

High fidelity primary data was manually collected, curated and annotated. The original data set was 200 images. Four augmented versions of the image were created, each with some modification of the originals (e.g., rotation, linear contrast changes, and

Gaussian noise addition) were used to prevent the risk of overfitting for the model (1000 images (5×200)). Then the data was split into train and test sets with 80:20 ratio, which means there are 800 training and 200 test samples.

Proposed Architecture

The basic architecture proposes a lightweight yet robust end-to-end framework designed for the image classification of a novel dataset collected within the context of Iraq. Due to the complete absence of prior benchmark datasets or local comparative literature for this specific category, this work serves as a foundational proof-of-concept for future scalable expansions. The proposed architecture follows a structured four-stage execution pipeline as shown in Figure 4. Figure 5 shows the proposed Architecture.

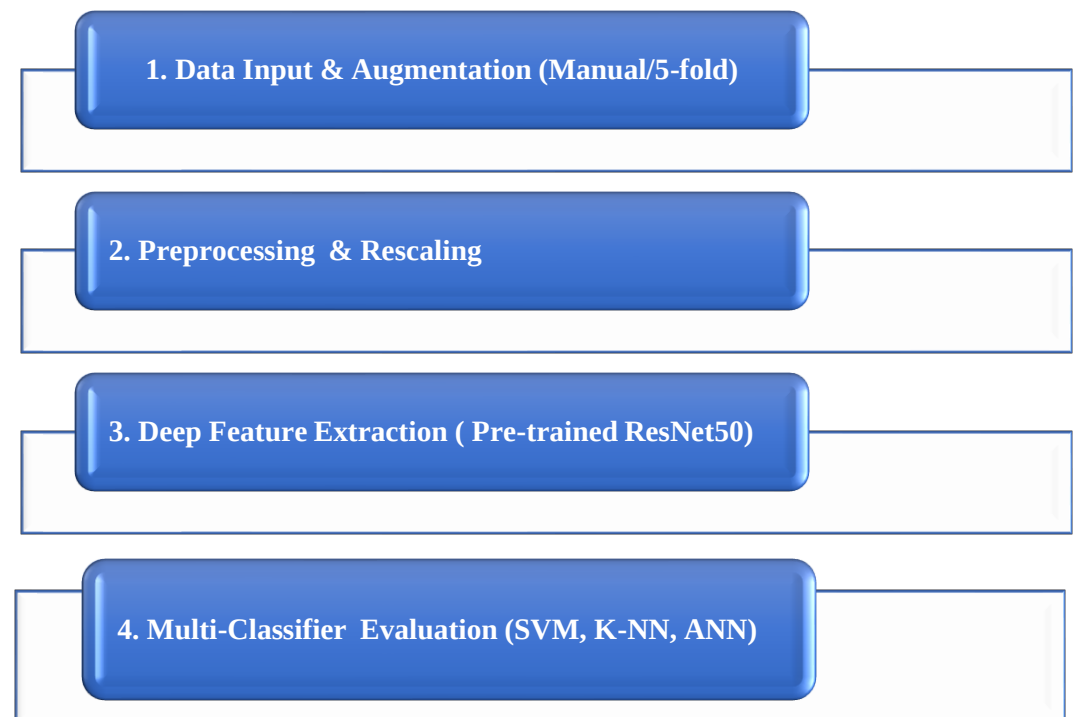


Figure 4. Four-stage execution pipeline for the proposed architecture

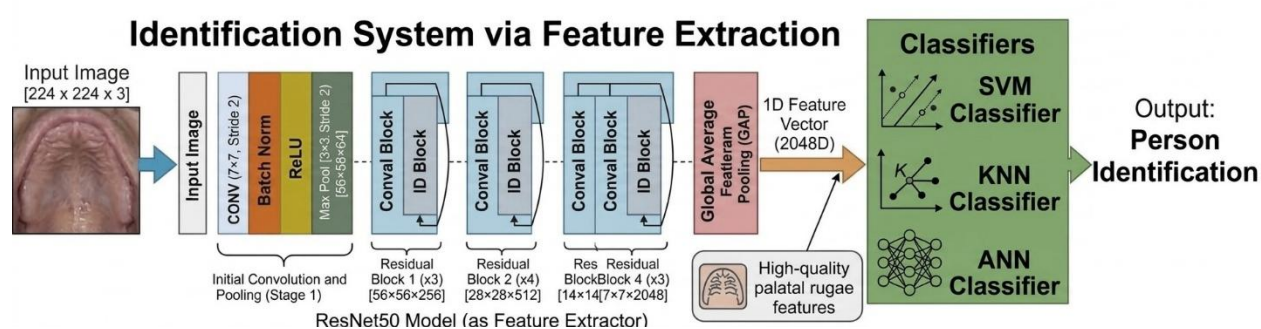


Figure 5. The proposed Architecture.

Methodological Pipeline

1. Data Acquisition and Augmentation

Primary data was manually collected, curated, and annotated to ensure high-fidelity labeling. To address sample scarcity and expand spatial variance, synthesized geometric and radiometric variations were introduced specifically, rotations ($\pm 3^\circ$), linear contrast adjustment ($\alpha = 1.2, \beta = 10$), and additive Gaussian noise ($\mu = 0, \sigma = 0.1$). Five versions are created for each original image (1 original + 4 versions) to increase the number of images and improve the efficiency of the model, as follows:

Version 1: The original image.

Version 2: Rotating the image by +3 degrees.

Version 3: Rotating the image by -3 degrees.

Version 4: Change contrast and brightness.

Version 5: Adding random Gaussian noise in a normal distribution.

2. Overfitting Strategies

To ensure model generalization and rule out empirical overfitting under small-sample regimes, several strict architectural constraints were enforced.

3. Preprocessing

All input images were standardized to a spatial resolution of $224 \times 224 \times 3$ pixels and normalized according to channel-wise zero-centered mean subtraction. The steps to prepare the images before feeding them into the deep learning model are:

- Read the images.
- Resizing: Making all the images the same size, which is 224×224 pixels for ResNet50.
- Color Conversion: Transform images to the RGB color space (and convert grayscale images to RGB, if there are grayscale images).
- Normalization: Using the preprocess function that is specific to ResNet50 to normalize pixel values and transform the range to be consistent with the pre-trained model ImageNet weights.

4. Fixed Feature Extractor

The convolutional weights of the ResNet50 backbone remained entirely frozen during execution, eliminating the parameter-heavy risk associated with fine-tuning deep architectures on limited data. Global Average Pooling (GAP) yielded a high-level 2048-dimensional representation vector per image without retraining the underlying convolutional layers. To extract the distinctive feature vector of each image, a pre-built deep learning model is used as follows:

- ResNet50 model is pre-trained ImageNet weights.
- Removing Top Layers (Feature Extractor Mode): False discards the original classification layers and only keeps the layers that perform feature extraction.
- Pooling: Average pooling was used to reduce the output of the layer to a 1D feature vector.
- Vector Generation: For each image, passing it through the model to get a deep feature vector of 2048 dimensions (2048-dimensional vector).

5. Low Architectural Complexity

After extracting features and splitting the data into 80% for training and 20% for testing. Classifiers were kept structurally simple K-Nearest Neighbors (K-NN, $k = 1$), Support Vector Machine (SVM with a Linear Kernel, $C = 1.0$), and a shallow Artificial Neural Network (MLP-ANN with a single 512-unit hidden layer, ReLU activation, and Adam optimization) to prevent the decision boundary from memorizing noise.

6. Regularization Mechanisms

Explicit L_2 weight decay penalty (10^{-4}) was incorporated into the MLP optimizer to penalize over-parameterization.

7. Strict Out-of-Sample Validation

Performance was evaluated using a strict 80/20 train-test split, ensuring that the test metrics reflect inference capability on completely unseen visual data.

8. Evaluation

Model accuracy (Accuracy), detailed performance reports (Classification Report such as Precision, Recall, F1-Score), plotting Confusion Matrix, and Precision-Recall Curve.

The main transformation and change are the Transfer Learning approach: Using a Pre-trained Model (Feature Extractor, not training from scratch) on your images with a highly deep and massive network (ResNet50). It has already a much better grasp on

understanding and deciphering image features, because it has been pre-trained on millions of images (ImageNet).

Separation between Feature Extraction and Classification: Deep learning is applied only to encode complex images in a numerical feature vector (2048 dimensions). These vectors are then fed to classic machine learning classifiers (SVM, K-NN and ANN) and passed for final classification. This method significantly cuts down on the amount of data required and the training time as compared to training an end-to-end deep neural network. Table 2 shows the parameters of proposed architecture.

Table 2. The parameter of proposed architecture

Steps	Parameter	Value
Input Resolution	RGB	224 x 224 pixels
Data Splitting	80% Train, 20% Test	(4 images/subject) (1 image/subject)
Normalization Strategy	ResNet50 preprocess_input	(Caffe-style: zero-centered per-channel)
Data Augmentation	Rotation	+3° & -3°
	Contrast Adjustment	alpha = 1.2, beta = 10
	Gaussian Noise	mean = 0, std = 0.1
ResNet50Feature Extractor	Pre-trained Weights	ImageNet
	Include Top Layers	False
	Pooling Method	Average Pooling
	Input Image Size	224 x 224 pixels
	Caffe-style	zero-centered per-channel
(K-NN)	Number of Neighbors (k)	1
	Distance Metric	Euclidean
(SVM)	Kernel Type	Linear
	Regularization Parameter (C)	1.0
(ANN)	Hidden Layers Structure	(512,) (1 hidden layer with 512 nodes)
	Maximum Iterations	500
	Activation	ReLU
	Optimizer	Adam
	Base Learning Rate	0.001
	Batch Size	Minibatch (200)
	Dropout Rate	None
	Weight Decay (L2 penalty)	0.0001
	Random Seed	42

ResNet50 Architecture

The ResNet50 model was chosen because it is well known for its high-performance utilizing ImageNet pre-trained weights to extract high-order and complex textural features [14,15]. ResNet50 is based on deep residual learning, which involves the use of skip connections to bypass certain layers, mitigating the vanishing gradient problem and allowing for the extraction of deep patterns as shown in Figure 6 [14]. This allows the model to identify fine details in medical images, making it particularly effective for distinguishing between various types of tissue abnormalities. Its structure and feature extraction steps are organized as follows:

- a. Initial Convolution and Pooling: A convolution with 64 filters and a stride of 2 is followed by a max pooling layer to reduce dimensionality.

- b. Residual Blocks: Four blocks in sequence gradually deepen the feature extraction, resulting in a complex hierarchy of features with a depth of up to 2048 channels.
- c. Remove Top Layers (Feature Extractor Mode): The original fully connected classification layers are removed from the network; only the base network is kept to serve exclusively as a feature extractor.
- d. Global Average Pooling (GAP): Average pooling is used to down-sample the final spatial output into a 1D feature vector.
- e. Vector Generation: For each fed image, a unique deep feature vector of 2048 dimensions is generated, capturing the image's essential features.

Specifically, the proposed feature extraction methodology begins with an input layer that processes images sized $224 \times 224 \times 3$, where the three channels represent the RGB color values. This standardized input format ensures compatibility and allows for effective feature extraction. The first convolutional layer uses a 7×7 kernel with a stride of 2, shrinking the spatial resolution to $112 \times 112 \times 64$. Low-level features (such as edges and textures) are captured by this layer, while Batch Normalization (BN) stabilizes the learning process, and the ReLU activation function introduces non-linearity to aid gradient propagation. A 3×3 MaxPooling layer with a stride of 2 further reduces the dimensions to $56 \times 56 \times 64$, discarding non-essential patterns.

To extract features deeply while preventing the vanishing gradient problem, the model utilizes four residual blocks featuring skip connections, which preserve critical information from one layer to the next. Residual Block 1 preserves spatial sizes at $56 \times 56 \times 256$ while enhancing feature depth. Residual Block 2 decreases the spatial size to $28 \times 28 \times 512$, enabling the network to learn more intricate features. Residual Block 3 further reduces the spatial dimensions to $14 \times 14 \times 1024$. Finally, Residual Block 4 fine-tunes the high-level feature representations, reducing the dimensions to $7 \times 7 \times 2048$.

To preprocess the extracted features, the model applies (GAP). This operation collapses the spatial dimensions, eliminates spatial redundancy, and outputs a definitive 1D deep feature vector of 2048 dimensions per image. This dense representation retains the most important and salient textural information.

Ultimately, by utilizing the pre-trained ResNet50 model strictly as a feature extractor, the proposed method leverages deep residual connections to obtain high-quality representations. This approach effectively models complex structural patterns, enabling powerful analysis and generalization for the purpose of person identification in palatal rugoscopy. Table 3 display the different between ResNet50 model and Proposed ResNet50 Architecture.

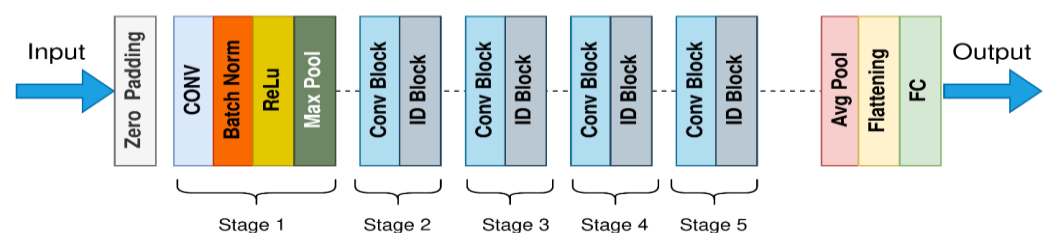


Figure 6. ResNet50 Architecture [14].

Table 3. Comparison between ResNet50 Architecture and Proposed ResNet50 Architecture.

Component	ResNet50 Architecture	Proposed ResNet50 Architecture
Dataset	General Image	Specific Dimensions Image
Stage1	CONV + BN + ReLU + MaxPool	CONV(Stride2) + BN + ReLU + MaxPool(Stride2)

Residual Blocks (Stage 2..5)	Conv Block / ID Block	Added precise dimensions for each stage
Output-Head	AvgPool + Flattening + FCLayer	Global Average Pooling (GAP) only
Output	Class Probabilities	(Removed FC and Top Layers) 2048-D-Feature Vector

Classification

Classification of the approach that identifies to what category a new item or object belongs based on the training data set that contains items or objects whose category label is known. There are many Classification techniques such as K-Nearest Neighbors(K-NN), Support Vector Machine(SVM), Neural Networks(ANN), Decision Trees(DT) so on. In this study used to SVM, K-NN, and ANN to identification the individual based on the Palatal Rugoscopy images as follows:

1. K-Nearest Neighbors(K-NN)

K-NN is one of the classification algorithms that is simple to understand and easy to implement. K-NN was proposed in the early 1950's. It finds a wide range of applications in pattern recognition, data mining, so on. KNN classifier works on 'k' closest points, or the nearest neighbors, to perform the classification. K-NN have three parameters. Firstly, the set of stored records, also known as the training set; secondly, the distance measure to calculate the distance between the records; and lastly, the value of 'k'[16].

To calculate the distance between the training dataset and the unknown test can be used the Euclidean distance. Eq. (1) presents the Euclidean distance.

$$d_{(p,q)} = \sqrt{\sum_{i=1}^k (p_i - q_i)^2} \quad (1)$$

Where $i = \text{from } 1 \dots k$, p_i & q_i are the i^{th} attributes of data p & q

An important problem that one has to address when using the K-NN classifier is the selection of the value of 'k'. In case 'k' value is selected which is too small, there is a possibility of noise points being included. The 'k' value selected could lead to neighbouring points belonging to other existing classes being included in the neighborhood. The 'k' value chosen is $k = 1$. 80% of the dataset is taken for training and 20% are picked at random to be used for testing.

2. Support Vector Machine (SVM)

SVM is a machine learning method that can be used for regression and classification problems. The training data categorized into two classes, could be classified to one of the two classes easily. For this reason, SVM was employed for the purpose of classification mainly into two classes, making it a non-probabilistic binary linear classifier [17]. Can be expanded and used in multiclass classification as well. According to the formal definition of SVM, it will build a hyperplane or a collection of hyperplanes in an infinite or high dimensional space, which can be used for classification problems. The separating hyperplane can be written as in Eq. (2)

$$W \cdot X + b = 0 \quad (2)$$

Where $W = w_1 \dots w_n$ it is weight vector for n attributes,

b is the bias and the training of tuple $- \rightarrow X$.

The best hyperplane is the one which has the largest margin separation between both the classes. The maximum-margin can be mathematically expressed as in Eq. (3)

$$\text{Maximum - Margin} = \frac{2}{\|W\|} \quad (3)$$

Where $\|W\|$ is 'Euclidean', W represented as $\sqrt{W \cdot W}$

There are various kernel functions that can be used in SVM, such as polynomial, Gaussian, and sigmoidal kernels. The main purpose of the kernel is to accept the data and convert it into the appropriate format.

SVM is used to classify the unknown Palatal Rugoscopy and identify the person using Linear kernel function as in Eq. (4), 80% of the dataset for training the classifier, and 20% of the dataset for testing randomly.

$$K(x, x') = x^T x' \quad (4)$$

Where:

x, x' : Feature vectors

$x^T x'$: Similarity

3. Artificial Neural Network (ANN)

The ANN has a number of layers. These layers comprise basic processing elements termed neurons that are intertwined to form a composite communication network. Every single neuron is related to a transfer function and has an associated weight, which can be either + value or – value. The neurons are organized in the Input layer, Hidden layers, and Output layer(s) [18] are based on the interconnectedness of the neurons. The architecture of a neural network can be shown as in Figure 7. Most of the neural networks consist of a learning rule which modifies the weighted connections according to the input features. There are numerous learning rules. But the most commonly used is the delta rule. Whenever an input feature is fed into the neural network's input layer, it makes a random guess as to what it is. It then sees how far the answer is from the actual answer and makes required changes to the connection weights.

In this work a two layered feed forward neural network is used. Softmax transfer function is mainly used for the multiclass classification. Back propagation algorithm is used as the training algorithm. A Multi-Layer Perceptron (MLP) classifier with a broad structure having one hidden layer of 512 neurons was implemented for the classification stage. This setup allows the network to effectively represent non-linear separability in the high dimensional feature space obtained by the backbone model. To achieve full convergence of the model, it was trained with a maximum of 500 iterations and with standard optimization settings. The random seed was fixed at 42 in order to maintain experimental reproducibility.

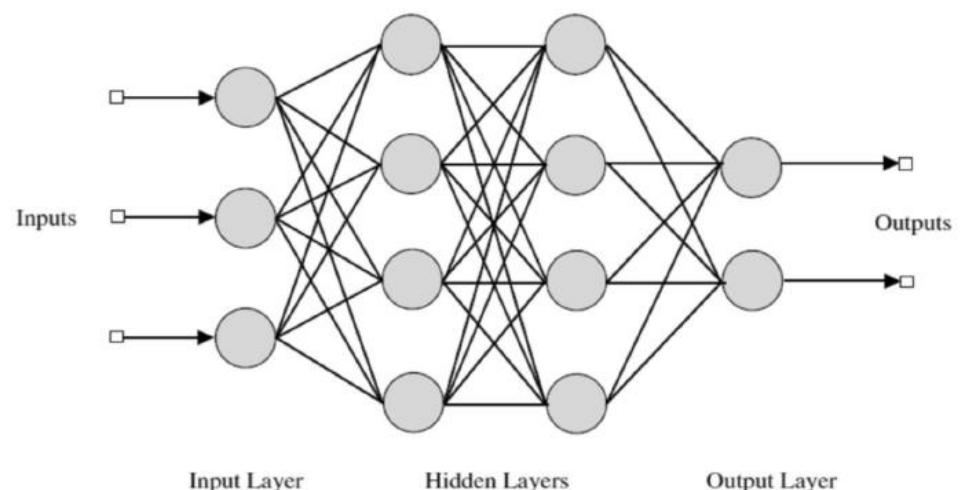


Figure 7. Display The ANN structure.

3. Results and Discussion

The experimentation was conducted on 1000 images. The data were collected through palatal rugoscopy of 200 individuals. This approach is similar to exploratory forensic dentistry studies that have focused on morphometric rugae analysis. The resulting feature vector was used in the classification process. The original data set was 200 images. Four augmented versions of the image were created, each with some

modification of the originals (e.g., rotation, linear contrast changes, and Gaussian noise addition) were used to compensate for the limited dataset size in the computational phase to prevent the risk of model overfitting for a total of 1000 images (5×200). Table 4 shows the Starting Data Augmentation and Deep Feature Extraction, and Figure 8 displays the augmented versions of the image.

Table 4. The Starting Data Augmentation and Deep Feature Extraction
Data Augmentation and Deep Feature Extraction (ResNet50)

Processed 20 persons (100 images generated)...

Processed 40 persons (200 images generated)...

.....

.....

.....

.....

Processed 200 persons (1000 images generated)...

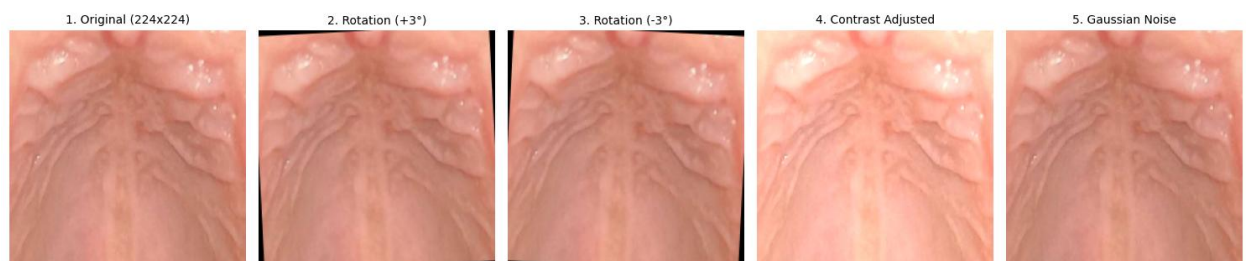


Figure 8. Data Augmentation Visual Example.

Classification of Performance

The classification of performance using (Accuracy, Precision, Recall, and F1- Score) given by Eqs. (5), (6), (7) and (8) [19]. Figures (9,10, and 11) show Performance Analysis of KNN, SVM, and ANN respectively. Table 5 displays the performance for different classifiers.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (6)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (7)$$

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (8)$$

Where:

True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN)

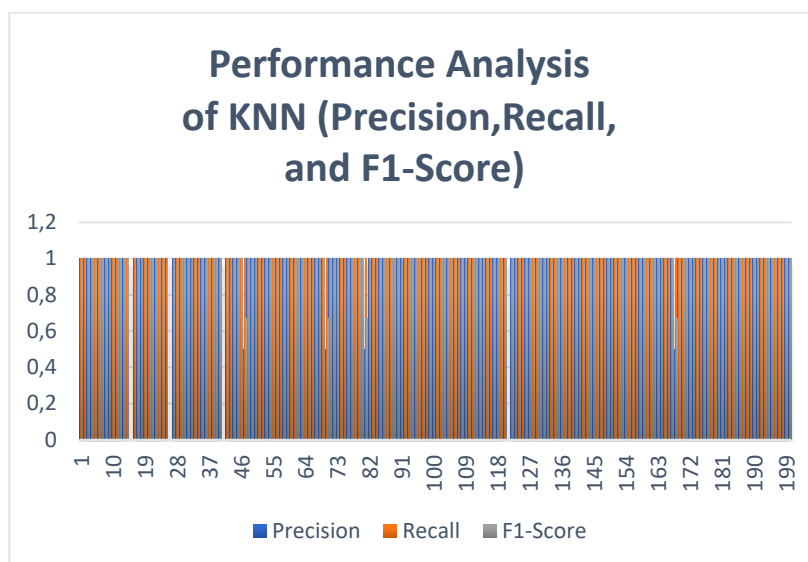


Figure 9. Performance Analysis of KNN (Precision, Recall, and F1-Score)

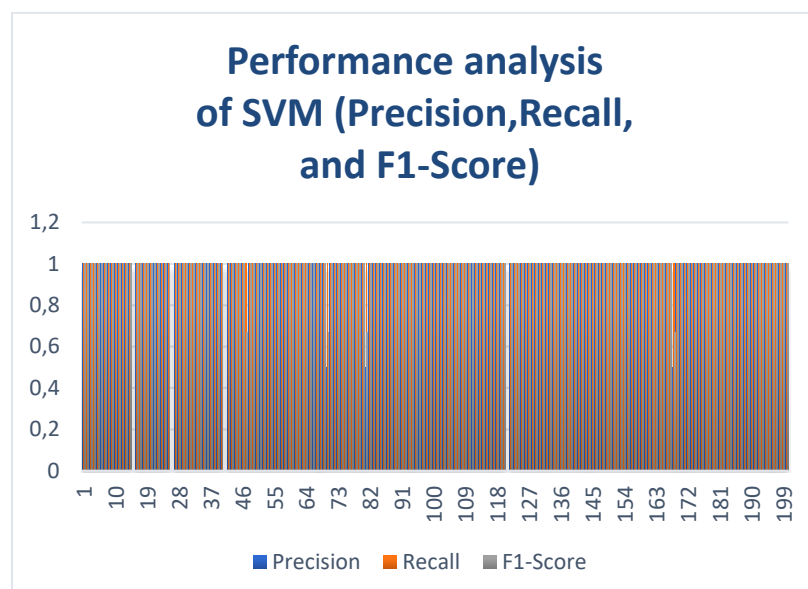


Figure 10. Performance Analysis of SVM (Precision, Recall, and F1-Score)

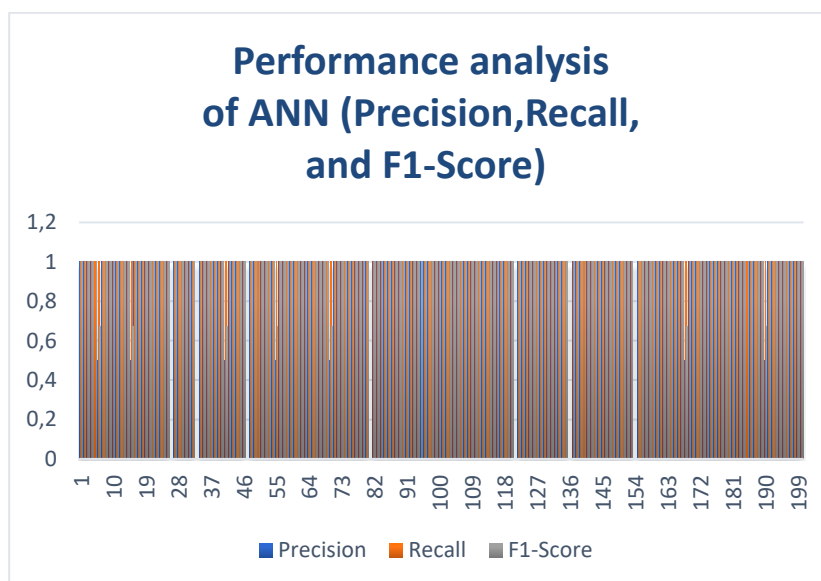


Figure 11. Performance Analysis of ANN (Precision, Recall, and F1-Score)

Table 5. The performance for classifier.

Classifier	Accuracy	Precision	Recall	F1-Score	Time
K-NN	96.50%	94.75%	96.50%	95.33%	0.07 seconds
SVM	98.00%	97.00%	98.00%	97.33%	1.15 seconds
ANN	96.50%	94.75%	96.50%	95.3%	10.74 seconds

Confusion Matrix

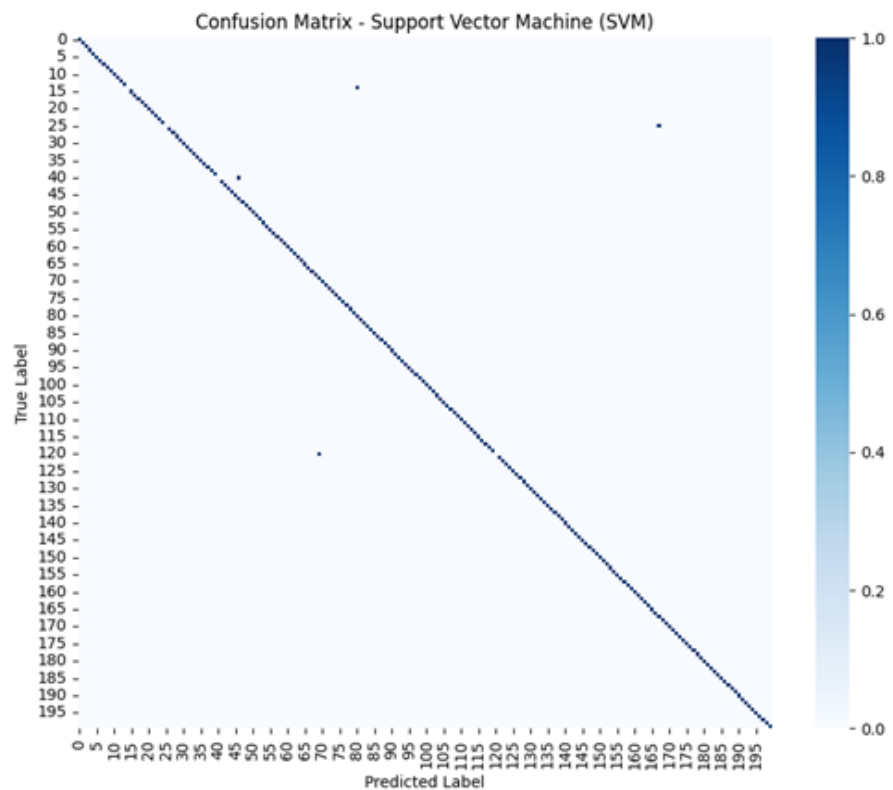
Confusion matrices were built for (SVM), (K-NN) and (ANN) models to assess the discriminative capability and stability of the extracted ResNet50 deep features within all 200 individual identities. As shown in Figures (12,13, and 14).

Each row is a correct identity label of the target image, while each column is the predicted identity label output by each model. All three matrices show a very strong trend line running from the upper left corner to the lower right corner. This strong diagonal concentration indicates that a vast majority of the test samples were correctly mapped to their true identities with an exceptionally high identification accuracy and a good separation of the features in the 2048-dimensional feature space.

SVM has the most diagonal characteristics among the models, with almost no off-diagonal components, indicating that it has a good generalization ability and is robust in high-dimensional deep feature representations.

ANN & K-NN : Both models also have good identification power, having a clear main diagonal line, with few, isolated misclassifications (sparse off-diagonal points).

In general, the very few off-diagonal predictions indicate that in the palatal rugae images of 200 subjects there is little confusion among classes, which shows that the ResNet50 framework is able to capture interesting and highly discriminative patterns in morphology and texture. Figure 15 Showing five Random Test Images with Predictions from SVM, K-NN, and ANN.

**Figure 12.** The Confusion-Matrix for SVM.

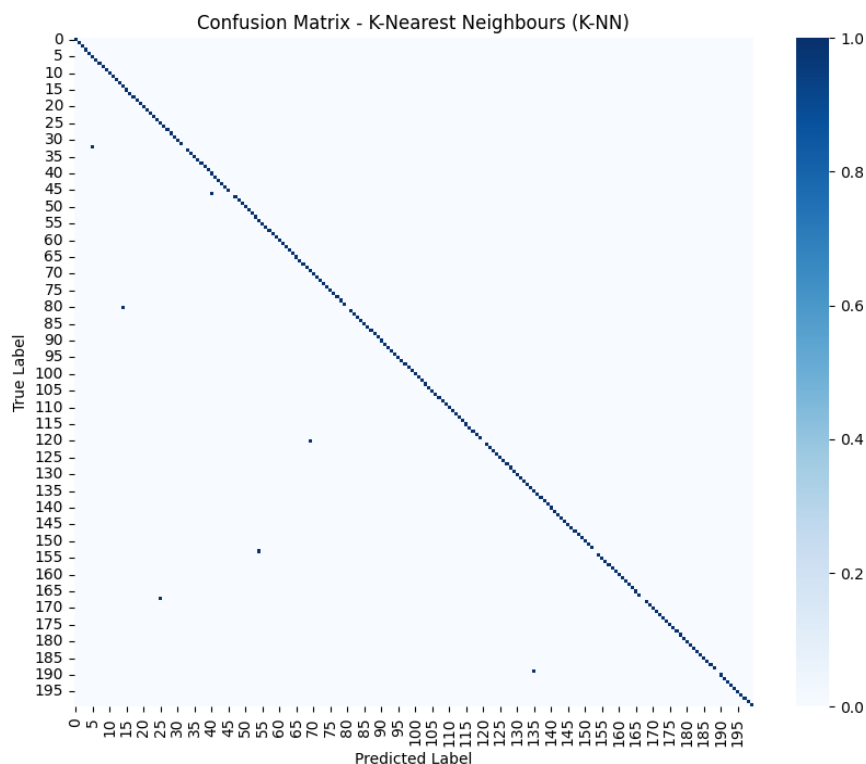


Figure 13. The Confusion-Matrix for K-NN.

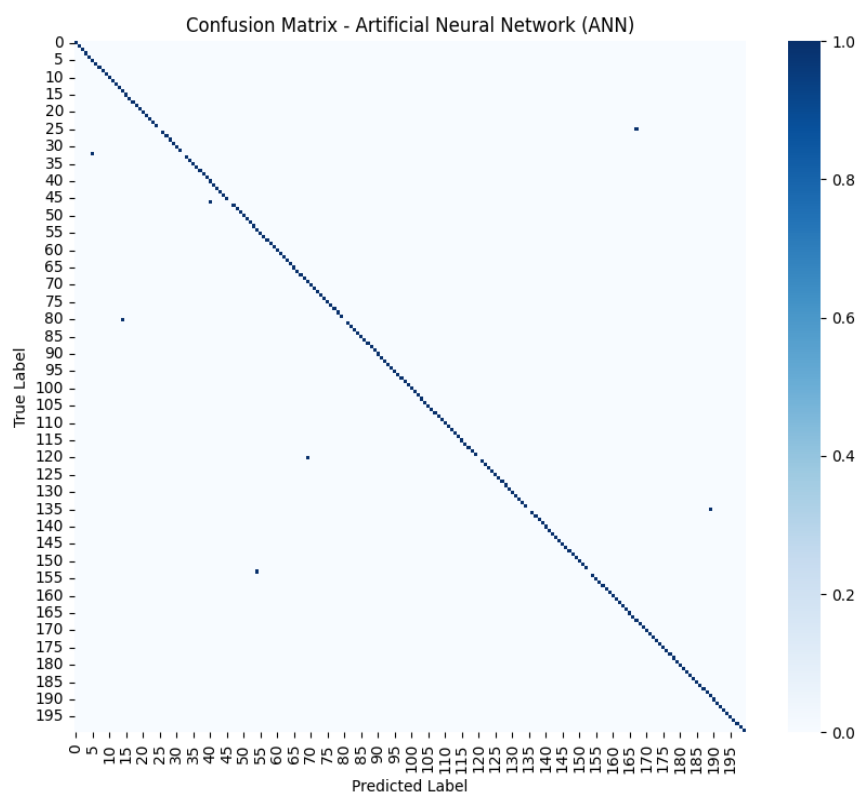


Figure 14. The Confusion-Matrix for ANN.



Figure 15. Random Test-Images.

Precision-Recall Curve

In order to thoroughly assess the multi-class classification effectiveness of the three machine learning classifiers, (K-NN), (SVM) and (ANN), micro-averaged Precision-Recall (PR) curves were plotted and compared across the 200 identity classes shown in the Figure 16.

Micro-averaging the PR curves shows remarkable performance for all models, with the top-right corner (Precision = 1.0, Recall = 1.0) being maintained. This is a very good trade-off between precision and recall, and shows that the 2048-D deep feature representation is obtained by means of a very robust ResNet50 architecture.

The ANN classifier (curve shown in green) has the most stable and best performance with a high level of precision over almost all ranges of the recall value with a sudden drop near the maximum boundary.

K-NN model (blue curve) has good performance and can be close to high-performance thresholds, but it slightly drops in high recall thresholds.

SVM model (orange curve) performs well with some minor fluctuations in the lower recall regions and quickly recovers in the regions where recall is higher to match the high precision ones.

Overall, the results of all three classifiers with high Area Under the PR Curve (AUC-PR) confirms the effectiveness and discriminative power of the proposed feature extraction methodology for identity recognition in palatal rugoscopy.

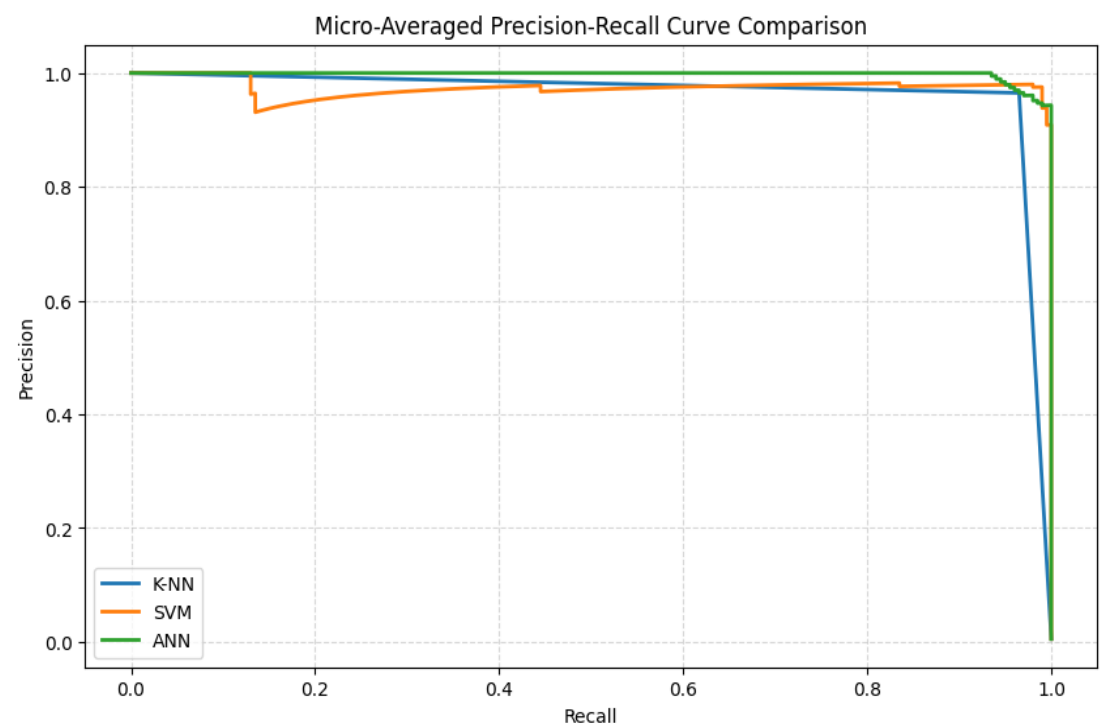


Figure 16. Precision-Recall Curve

Performance Comparison with Other the Methods

The patterns of rugae on tongue vary among individuals and these patterns can be used to identify individuals [20]. These results also were supported by population-based studies. The morphology of rugae was found to be different among Australian Aborigines and Caucasians by Kapali et al. [21] which showed that ethnicity affected the morphology of rugae. Hermosa et al. also observed high variability of the palatal rugae in a population from Peru that the palatal rugaescopy used to specific FD [22]. The results obtained by Indira et al. showed that there is a stable in the patterns of rugoscopy in the Indian population even after a time has passed [23]. Shukla et al. also showed the stability of palatal rugae patterns as they mentioned that these rugae patterns are age and treatment

independent [4]. from the following studies in FD have focused on computerized and three-dimensional methods of better analysis and reporting. The use of studies using digital three-dimensional imaging and craniofacial measurements have been proven to be more precise and more objective than conventional three-dimensional studies [1,2,24]. Also, there have been possibilities of automated pattern recognition and identification with the increasing use of artificial intelligence and machine learning in FD. Table 6 display a comparison of the accuracy between proposed architecture and other methods.

Table 6. Comparison the accuracy between proposed architecture and other methods

Method /year	Accuracy
Wu et al. (2017) [7]	95.67 %
Kusumaputri et al. (2020) [9]	91.38%
Bsat et al. (2025) [10]	84.5%
Amos et al. (2026) [11]	/
Morye et al. (2026) [12]	/
Proposed Architecture	
K-NN	96.50%
ANN	96.50%
SVM	98.00%

Challenges and Limitations

- A major constraint of this study is the small number of publically available, standardized benchmark datasets for Palatal Rugoscopy.
- Compared with the traditional biometric traits like fingerprints and facial images, the number of rugae datasets available in public repositories is almost none. Thus, a primary dataset needed to be built up with careful clinical collection and annotation, which limited the total sample size in the primary dataset.
- Although palatal rugae patterns have been studied in both forensic and clinical settings, the combination of these with sophisticated computer vision and deep learning systems is very little represented in the literature. This lack of comparative methodological baselines, which is a major challenge, was a hindrance in benchmarking computer-aided extraction models with previous literature.
- To our best knowledge, this is the first automatic and computer assisted system for personal identification based on palatal rugae patterns in Iraq. With no existing digital archiving systems in place, and operating as a pioneering study in the area, it was not without logistical and operational challenges to collect good quality clinical intraoral images in local dental facilities.
- To address these limitations in the data available, the proposed method utilized a powerful feature extraction paradigm that utilizes pre-trained ResNet50 weights as well as advanced rotation, contrast modification, and noise augmentation techniques. This led to effective model generalization, avoiding overfitting and successfully producing highly-discriminative 2048-dimensional feature representations.

4. Conclusions

The study accomplishes a highly accurate and automatic system for human Palatal Rugoscopy patterns' biometric identification, which is a pioneering study in this field in Iraq. Intraoral images $224 \times 224 \times 3$ were first fed into a fine-tuned ResNet50 model with all fully connected classification layers removed, and then into (GAP) to output a dense feature vector of length 2048 that was highly discriminative. The performance of identity mapping was tested on 200 individuals using three machine learning algorithms: (SVM), (K-NN), and (ANN). Experimental Results: It is observed that SVM has the highest overall performance which yields 98.00% accuracy, 97.00% precision, 98.00% recall, and 97.33% F1-Score among the tested classifiers. It reveals the excellent margin separation

performance of SVM with high order deep residual representations. The results of this study indicate that K-NN is an efficient algorithm, having a very small execution time of just 0.07 seconds, and simultaneously producing good recognition metrics of 96.50% Accuracy and 95.33% F1-Score. However, ANN was able to obtain the same accuracy (96.50%), but needed significantly longer computational time of 10.74 seconds. The normalized Confusion Matrices for all 200 target classes show very low levels of off-diagonal misclassifications (i.e., low inter-class interference) and sharp diagonal trends, further confirming this. In addition, Averaged Precision-Recall curves of all models stayed well near the top right corner - near the optimal curve. To recap, the ResNet50 feature embedding and SVM classifier is a robust, non-invasive, and automated approach for forensic dentistry and biometric identification that addresses the challenges of data scarcity and primary dataset limitations.

Recommendations for further research are as follows:

- a. Extending the primary data with a larger and more diverse population from various regional health institutions in Iraq for assessing the long-term stability and invariance of palatal rugae patterns throughout aging, orthodontic therapy and palatal trauma.
- b. Comprehending lightweight deep learning architectures (e.g., MobileNetV3, EfficientNet) and Vision Transformers (ViTs) for optimizing inference latency, facilitating real-time use on portable dental intraoral scanners and mobile devices.
- c. Designing a multimodal system that fuses together the parameters of palatal rugoscopy with other oral biometrics such as dental arch geometry, tongue prints or lip prints at the feature-level or score-level to reach absolute zero-error recognition performance in high-stakes forensic identification applications.

REFERENCES

- [1] M. Eason et al., "Three-dimensional assessment of tooth movements in clear aligner treatment of moderate and severe crowding," *Clinical Oral Investigations*, vol. 30, no. 1, p. 6, 2026. doi: 10.1007/s00784-025-06657-y
- [2] E. Dayan, M. Kurnaz, Ö. Ö. Kuşçu, S. Biren, and E. Sepet, "From Bite Marks to Vertical Facial Phenotype: Investigating the Relationship Between Craniofacial and Dental Arch Parameters Using Conventional Wax and Digital Models," *Journal of Forensic and Legal Medicine*, p. 103078, 2026. doi: 10.1016/j.jflm.2026.103078
- [3] A. Dubey et al., "From Skull to Face: MRI vs. CBCT-Based Soft Tissue Depths in Forensic 3D Craniofacial Reconstruction," *Journal of Forensic and Legal Medicine*, p. 103073, 2026. doi: 10.1016/j.jflm.2026.103073
- [4] D. Shukla, A. Chowdhry, D. Bablani, P. Jain, and R. Thapar, "Establishing the reliability of palatal rugae pattern in individual identification," *Journal of Forensic Dental Sciences*, vol. 3, no. 1, pp. 20–25, 2011. doi: 10.4103/0975-1475.81286
- [5] L. Krenmayr, R. von Schwerin, D. Schaudt, P. Riedel, and A. Hafner, "DilatedToothSegNet: Tooth Segmentation Network on 3D Dental Meshes Through Increasing Receptive Vision," *J. Imaging Inform. Med.*, vol. 37, pp. 1846–1862, 2024. doi: 10.1007/s10278-024-01061-6
- [6] S. Wang et al., "STSN-Net: Simultaneous Tooth Segmentation and Numbering Method in Crowded Environments with Deep Learning," *Diagnostics*, vol. 14, p. 497, 2024. doi: 10.3390/diagnostics14050497
- [7] X.-P. Wu, J.-N. Han, F. Pan, Y.-J. Wan, and B. Li, "Analysis of palatal rugae morphology before and after orthodontic treatment by a digital image recognition system," *International Journal of Morphology*, vol. 35, no. 2, pp. 420–424, 2017. doi: 10.4067/S0717-95022017000200005
- [8] H. Shanguan et al., "Research on correction method of palatal rugae morphology in forensic digital identification system," *International Journal of Morphology*, vol. 37, no. 1, pp. 324–330, 2019. doi: 10.4067/S0717-95022019000100324
- [9] F. H. Kusumaputri, S. A. Wibowo, and Y. Malinda, "Identifikasi individu berdasarkan pola citra rugae palatina menggunakan Adaptive Neuro Fuzzy Inference System (ANFIS)," *Jurnal Ilmiah FIFO*, vol. 12, no. 2, pp. 156–167, 2020. doi: 10.22441/fifo.2020.v12i2.005
- [10] A. R. El Bsat et al., "Three-dimensional semantic segmentation of palatal rugae and maxillary teeth and motion evaluation of orthodontically treated teeth using convolutional neural networks," *Diagnostics*, vol. 15, no. 11, Art. no. 1415, 2025. doi: 10.3390/diagnostics15111415

- [11] R. I. Amos, R. Kamalraj, and R. Pande M S, "A dento-technical model for the analysis and classification of palatal rugae," *International Research Journal on Advanced Engineering and Management*, vol. 4, no. 2, pp. 157–164, 2026. doi: 10.47392/IRJAEM.2026.0022
- [12] B. T. Morye, R. D. Tayade, and S. D. Thorat, "A comprehensive survey of automated palatine rugae recognition for forensic applications," *International Journal of Research Publication and Reviews*, vol. 7, no. 3, pp. 2683–2691, 2026. doi: 10.55248/gengpi.07.0326.0513
- [13] C. J. Thomas and T. J. Kotze, "The Palatal rugae pattern: a new classification," *J. Dent. Assoc. S. Afr.*, vol. 38, no. 3, pp. 153–157, 1983.
- [14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, pp. 770–778, 2016. doi: 10.1109/CVPR.2016.90
- [15] A. Zhang, Z. C. Lipton, M. Li, and A. J. Smola, "Dive into Deep Learning," *arXiv preprint arXiv:2106.11342*, 2021. doi: 10.48550/arXiv.2106.11342
- [16] P. K. Syriopoulos et al., "k NN Classification: a review," *Annals of Mathematics and Artificial Intelligence*, vol. 93, no. 1, pp. 43–75, 2025. doi: 10.1007/s10472-023-09923-2
- [17] C. E. Mbonu, K. Anigbogu, D. Asogwa, and T. Belonwu, "An explorative analysis of svm classifier and resnet50 architecture on african food classification," *arXiv preprint arXiv:2505.13923*, 2025. doi: 10.48550/arXiv.2505.13923
- [18] S. Sapna and A. Renuka, "Computer-aided system for leukocyte nucleus segmentation and leukocyte classification based on nucleus characteristics," *Int. J. Comput. Appl.*, vol. 42, no. 6, pp. 1–12, 2020. doi: 10.1080/1206212X.2018.1469375
- [19] T. A. Anai, M. A. Anaee, and S. Mohamed, "Optimized Deep Learning for Feature Extraction and Detection of Oral Cancer from Histopathological Images," *Wasit Journal of Computer and Mathematics Science*, vol. 5, no. 3, pp. 26–41, 2026. doi: 10.31185/wjcm.145
- [20] W. R. English et al., "Individuality of human palatal rugae," *Journal of Forensic Sciences*, vol. 33, no. 3, pp. 718–726, 1988. doi: 10.1520/JFS12497J
- [21] S. Kapali, G. Townsend, L. Richards, and T. Parish, "Palatal rugae patterns in Australian Aborigines and Caucasians," *Australian Dental Journal*, vol. 42, no. 2, pp. 129–133, 1997. doi: 10.1111/j.1834-7819.1997.tb00107.x
- [22] R. Hermosa, M. Garcia, and J. Castillo, "Study of the morphologic diversity of palatal rugae in a Peruvian population," *Journal of Forensic Odonto-Stomatology*, vol. 27, no. 2, pp. 39–46, 2009.
- [23] A. P. Indira, M. Gupta, and M. P. David, "Usefulness of palatal rugae patterns in establishing identity: Preliminary results from Bengaluru city, India," *Journal of Forensic Dental Sciences*, vol. 4, no. 1, pp. 2–5, 2012. doi: 10.4103/0975-1475.99149
- [24] A. F. de Oliveira, B. P. Zanollo, D. D. Moreira, and A. Franco, "External validation of a country-specific model for radiographic dental age assessment of Brazilian children and adolescents," *Morphologie*, vol. 110, no. 368, p. 101099, 2026. doi: 10.1016/j.morpho.2025.101099