scientific reports

Check for updates

OPEN

AI-driven tongue image analysis for diagnosing and predicting coronary artery disease

Alireza Hekmat-Ardakani[1], Hamid R. Rabiee[2]✉, Armin Behnamnia[2], Amir Hooman Kazemi[1], Amir Sobh Rakhshankhah[3], Ramtin Mahmoudi Kashani[4], Alireza Yargholi[5], Bahar Morshed-Behbahani[6], Abolfazl Malekahmadi[2], Hamidreza Hekmat[7]✉ & Mehrdad Karimi[1]✉

Coronary artery disease is a leading cause of mortality worldwide, with current diagnostic methods presenting challenges of invasiveness and accuracy. This study investigates artificial intelligence-powered tongue image analysis as a novel, non-invasive diagnostic tool for coronary artery disease. Drawing on principles from Traditional Iranian Medicine, we analyzed tongue features, including color, fissures, coating, moisture, and red spots, in 906 patients undergoing coronary angiography. A dataset of 12,000 tongue images was used to train artificial intelligence models, which were created using high-resolution images captured with a Visioface device. The Segment-Anything Model facilitated tongue segmentation, while DeepLabv3+/DINOv2 architectures were employed for feature extraction and stenosis classification. Results indicated that artificial intelligence predictions achieved up to 80% accuracy for left circumflex artery stenosis, highlighting tongue color and fissures as significant predictors of this condition. This study presents artificial intelligence-driven tongue analysis as a rapid and cost-effective adjunct for screening coronary artery disease, particularly in resource-limited settings, and warrants further validation in diverse populations.

Keywords Tongue image analysis, Coronary artery disease, Traditional iranian medicine, Artificial intelligence

Cardiovascular diseases are the leading global cause of death, with coronary artery disease (CAD) being the primary contributor[1]. Notably, coronary artery disease can be asymptomatic in a significant proportion of individuals[2]. Early and precise diagnosis of these disorders can significantly reduce their mortality and morbidity. Simple and accessible methods for diagnosing heart disease can help practitioners make more convenient and faster diagnoses, ultimately improving patients' health. While coronary angiography remains the gold standard for diagnosing CAD, its time-consuming nature and high costs, coupled with its invasive procedures, often deter physicians from ordering this test except in critical cases, and patients may be reluctant to undergo it[3]. Electrocardiography (ECG) is a widely used, noninvasive, and cost-effective diagnostic procedure that does not necessitate expensive equipment. While interpreting ECG results requires specialized training, this does not preclude many individuals from obtaining such training. Notably, the average accuracy of ECG in diagnosing CAD is approximately 57.8%, highlighting its limited diagnostic performance in assessing CAD[4]. Recently, a risk factor classifier has been used in conjunction with ECG to enhance its accuracy in diagnosing CAD. While this approach can improve ECG accuracy to as much as 94%, the classification and consideration of these risk factors necessitate a comprehensive understanding of the patient's medical history and records, which is often unfeasible in emergencies[5].

[1]Department of Traditional Medicine, School of Persian Medicine, Tehran University of Medical Sciences, Tehran, Iran. [2]Department of Computer Engineering, Sharif University of Technology, Tehran, Iran. [3]Department of Cardiology, School of Medicine, Tehran University of Medical Sciences, Sepehr heart center, Baharloo hospital, Tehran, Iran. [4]Department of Computer Science, Amirkabir University of Technology, Tehran, Iran. [5]AICT Research Institute, Sharif University of Technology, Tehran, Iran. [6]Department of Midwifery and Reproductive Health School of Nursing and Midwifery, Iran University of Medical Sciences, Tehran, Iran. [7]Department of Cardiology, School of Medicine, Ziaeian Hospital, Baharloo Hospital, Tehran University of Medical Sciences, Tehran, Iran. ✉email: rabiee@sharif.edu; hekmat.hamidreza@yahoo.com; bahmanhekmat@yahoo.com

The exercise tolerance test enhances the diagnostic accuracy for CAD to 77%; however, its implementation necessitates specialized facilities and equipment, limiting its feasibility in all healthcare settings[6]. Additionally, echocardiography serves as another non-invasive diagnostic modality for CAD, demonstrating a sensitivity of 87% in identifying the condition[7]. However, echocardiography requires specialized personnel and appropriate equipment, which are not consistently available across all clinical centers[8]. Non-invasive cardiac imaging techniques that utilize radiation, such as CT scans, CT angiography, and myocardial perfusion scans, can enhance diagnostic sensitivity and accuracy, achieving rates of up to 94%[9]. In these instances, in addition to being costly and necessitating dedicated space, specialized personnel, and specific tools and equipment, these methods are further limited by their reliance on radiation. Therefore, the availability of more non-invasive techniques increases the likelihood of swiftly achieving an accurate diagnosis.

The application of artificial intelligence (AI) has recently emerged as a transformative force in cardiovascular diagnostics, offering promising solutions to overcome the limitations of traditional methods. Beyond conventional approaches, AI is being successfully applied to various non-invasive modalities. For instance, AI-enabled electrocardiography (AI-ECG) models have shown remarkable accuracy in identifying a wide range of cardiac conditions, from predicting long-term atrial fibrillation risk to detecting left ventricular hypertrophy. Similarly, AI algorithms are being integrated with echocardiography and cardiac MRI to automate image analysis, improve measurement accuracy, and facilitate early detection of structural heart diseases. Furthermore, recent studies have demonstrated the potential of AI in analyzing chest X-rays and retinal images to predict cardiovascular risk, proving that subtle visual biomarkers can be leveraged to screen for systemic diseases like CAD[10,11].

Tongue assessment is a common practice in many holistic medicines, such as Traditional Iranian Medicine and Traditional Chinese Medicine, used by medical practitioners to evaluate a patient's condition in various diseases, including heart disease. As an anatomical structure, the tongue serves as a reflecting indicator of the body's physiology and pathology. The examination of the tongue is based on visual clues, with color, shape, movement, texture, and coating all playing important roles in the diagnostic process. Furthermore, the geometric arrangement of the tongue, which includes measurements such as thickness, size, fissures, and indentations, emerges as a contributing factor in evaluating an individual's health[12].

Recent studies suggest that the tongue can serve as a valuable diagnostic tool for detecting manifestations of internal and systemic diseases[13]. For example, iron deficiency anemia causes flattening of the tongue papillae. In the context of autoimmune diseases, the tongue manifestations are primarily due to the inflammation caused by the disease. For example, in Sjögren's syndrome, the tongue is dry and appears bald, with a tombstone-like texture, exhibiting visible cracks and fissures. In Kawasaki, the tongue becomes swollen and red, and due to swelling of the papillae, it becomes strawberry-shaped. In scleroderma, the tongue becomes smooth, and its mobility decreases. In amyloidosis, a large and swollen tongue (macroglossia) is often present, and in lupus, granulomatous ulcers may appear on the tongue. Endocrine disorders also present with distinct manifestations on the tongue. These include tongue enlargement and macroglossia in hypothyroidism, as well as patchy brown macular pigmentation on the tongue in Addison's disease. Tongue symptoms in diabetic patients include a dry tongue with large areas of thick, yellow coating[13–15]. Cardiovascular diseases may also present with a range of symptoms on the tongue, underscoring the notion that the tongue serves as a window to cardiac health[16]. The coloration of the tongue is a significant factor in predicting the severity of coronary artery stenosis[17,18]. In patients with stasis and coronary artery disease, the tongue often exhibits a bluish or reddish-purple coloration[19].

Despite the ease and noninvasiveness of tongue examination, obtaining an objective and standardized evaluation proved to be a difficult task. Recent scholarly efforts have focused on addressing these issues. In recent years, artificial intelligence has become increasingly popular, enhancing various diagnostic capabilities through objective and standardized observations[20].

Artificial intelligence and its associated technologies can analyze vast amounts of image data in medical science to more accurately determine the relationship between observed symptoms and signs in these images and the related diseases.

Within the medical field, several straightforward applications of artificial intelligence utilizing artificial neural networks can be highlighted. An exemplary application is assisting with medical diagnoses, such as those for melanoma. This is accomplished by using dermoscopic images of specific lesions obtained with a cell phone camera[21].

Artificial intelligence algorithms have achieved precision comparable to that of professional practitioners using this approach[22]. As a result, the concept of computer-assisted tongue diagnosis offers great promise for providing more consistent and objective health evaluations, particularly in fields such as heart disease assessment[23]. This diagnostic technique can play a crucial role in cases where the rapid identification and implementation of therapeutic interventions are critical for limiting the progression of the patient's condition, such as cardiovascular disorders.

Recognizing the high prevalence and critical importance of CAD, we have initiated a research project to investigate potential correlations between CAD and symptoms related to the tongue. This study aims to leverage advanced AI techniques to identify and analyze the relationship between tongue manifestations and CAD, thereby contributing to better diagnostic and preventive strategies in cardiovascular health.

The primary aim of this study is to establish a precise relationship between tongue signs and CAD through artificial intelligence (AI) and tongue image analysis. This could lead to the development of software capable of analyzing coronary artery status based on tongue images. The specific objectives of the study are as follows:

- To train artificial intelligence to recognize pertinent findings on the tongue.
- To extract the coronary angiography results of the study participants and photograph their tongues under consistent conditions.

- To identify the correlation between tongue findings and coronary angiography results.

Within this rapidly expanding field, our study introduces a novel and non-invasive approach by leveraging traditional tongue diagnosis with advanced AI and computer vision techniques. While other AI-driven methods rely on complex and sometimes costly imaging, our method harnesses readily available visual information from the tongue. This makes it a unique and potentially scalable solution for mass screening, particularly in low-resource settings. By integrating the rich visual biomarkers of the tongue with a sophisticated deep learning framework, our research complements existing non-invasive diagnostic tools. It offers a new avenue for the early detection and risk stratification of CAD.

Methods

This was a cross-sectional study. From July 2021 to February 2022, tongue images were collected from 906 patients hospitalized in the angiography ward of Baharloo Hospital, affiliated with Tehran University of Medical Sciences, due to cardiac symptoms. The methodology of this study was approved by the Ethics Committee of Tehran University of Medical Sciences under code IR.TUMS.MEDICINE.REC.1400.675. All procedures conducted in this study were carried out in accordance with the guidelines of this committee.

Eligibility criteria

Inclusion criteria: Patients who were hospitalized in the angiography department and were hemodynamically stable were included in the study.
Exclusion criteria: The following criteria were used for exclusion from the study: non-cooperative patients; unconscious or hemodynamically unstable patients; patients with cognitive or behavioral impairments; those with active cancer or active oral infections; individuals experiencing symptoms related to oral cancer within the past two years; pregnant or lactating women; patients with congenital heart disease; and individuals with geographic tongue.

Data collection

All participants received detailed information regarding the research objectives and methodology, and written informed consent was obtained before their participation. The informed consent form includes consent for the publication of data related to the tongue images and test results in open-access journals. The informed consent form is provided as a Supplementary File. Their medical history and demographic information were recorded in the relevant forms. (Supplementary file) The imaging procedure took place in a room without windows or access to natural light. After rinsing the patients' mouths with a normal saline solution, their tongues were examined and recorded by a specialist in Iranian Traditional Medicine. The examination assessed the tongue's shape, color, presence of tooth marks, fissures, and the distribution, color, and thickness of its coating. Additionally, the moisture level of the tongue and the presence of red spots were evaluated. (Supplementary file)

We used the Visioface device, which features 210 LEDs for even illumination of the tongue, captures 14-megapixel photographs, and includes a head and chin support for stabilization during photography. The captured images are transmitted directly to a computer for storage and processing. (Fig. 1) Tongue photos



Fig. 1. The clinical imaging setup for patient tongue analysis. This figure illustrates the standardized procedure for capturing tongue images using a high-resolution Visio Face camera. The setup ensures consistent lighting and distance, which is crucial for minimizing environmental variability and guaranteeing the reproducibility of the visual data.

were taken from the front, right, and left sides. Each image was labeled with a unique code to maintain patient confidentiality.

In parallel with the imaging process, a subset of our tongue images was manually annotated by a board-certified expert in Traditional Medicine. These expert annotations served as the ground truth for training our automated feature extraction models. It is crucial to note that while this expert evaluation provided the necessary foundation, our primary deep learning model was not trained on these subjective ratings. Instead, the model was designed to extract objective, quantitative features directly from the images. For coating (thickness and area) and tongue color, we don't use expert annotations; instead, a fixed, automated image processing algorithm based on CIELAB color space is used, which doesn't contain any assumptions other than the definition of the CIELAB space.

For fissures, wetness, and other features, expert annotation was performed by A.H.A., a senior physician with 27 years of clinical experience and currently a sixth-year student of Iranian Traditional Medicine. The annotation process was supervised by M.K., Professor of the School of Iranian Medicine with over 20 years of experience, and A.H.K., Assistant Professor in Traditional Chinese Medicine trained at Peking University with more than 15 years of clinical and teaching experience. Three experts validated the area and existence of each feature on the tongue, as it can be challenging in many cases. For more details, we explain the visual criterion for each feature below,

1. Fissure: Tongue fissures are deep grooves or furrows that appear on the dorsal surface of the tongue.
2. Wetness: Moisture and wetness on the tongue caused by salivary secretions.
3. Red Spot: Red spots may appear on the tip, lateral borders, or root of the tongue. They are larger than a pinpoint but smaller than a mole.
4. Teeth marks: Teeth marks on the lateral borders of a large tongue are caused by pressure from adjacent teeth, resulting in alternating concave and convex patterns.

These complex morphological features (fissures, wetness, dental marks, and red spots) were quantified by a specialized computer vision algorithm. The robustness of this algorithm was established through training on an external database extracted from over 1,000 tongue images, which were annotated and validated by three experts, confirming its ability to accurately translate visual data into objective numerical inputs.

These structured annotations served one leading role in our pipeline:

Training the segmentation model DeepLabv3 + to identify and mask each feature accurately.
Furthermore, these masks were used as input features to fine-tune the DINOv2 classification head for predicting vessel stenosis and patient outcomes.

Features for half of the samples were generated by the trained DeepLabv3 + model, instead of direct expert annotations. By relying on a hybrid dataset of cross-validated expert annotations, AI-computed metrics, automated image processing algorithms (for coating and color features)—rather than purely subjective scoring—we reduced any potential bias in the annotation process. This hybrid approach leverages clinical expertise while maintaining quantitative objectivity throughout the training and evaluation stages.

All patients underwent coronary angiography. The summary of the study implementation steps is illustrated in Fig. 2.

In this study, the results of coronary angiography were evaluated in four parts. (Fig. 3)

1. Left main coronary artery (LMCA)
2. **Left anterior descending artery (LAD)**: The angiographic results of diagonal arteries 1 and 2, as well as the septal branch, which are minor branches of the LAD, are not considered.
3. **Left circumflex artery (LCX)** The angiographic results of the obtuse marginal arteries (OM1 and OM2) and ramus arteries are included in this part.
4. **Right coronary artery (RCA)** In this section, we considered the angiographic results of the posterior descending artery (PDA), posterior left ventricular (PLV) branch, and right ventricular (RV) branch.

In each part, the results of the worst segment will be considered. For example, the worst angiography result among the PLB, PDA, and RCA vessels will be used in the RCA part of the result. (Supplementary file)

Findings for each angiogram were scored using the modified Agatston score[24]. According to this grading system, the severity of stenosis is categorized as follows:

- **Mild**: 1–49% narrowing.
- **Moderate**: 50–69% narrowing.
- **Severe**: 70–99% narrowing.
- **Occluded**: 100% blockage.

We followed the patients until the results were determined. The outcomes are classified into four categories: drug treatment, PCI, CABG, and others (e.g., ablation for premature ventricular contractions, valve replacement).

Artificial intelligence

We have a dataset comprising facial images of 906 subjects captured from three angles: front and two sides, along with their angiography results, which are labeled with the degree of stenosis for different regions of the LCX, RCA, and LAD. Additionally, we have more than 12,000 images from the normal population without specific labels. Our current focus is on the frontal images of the subjects.

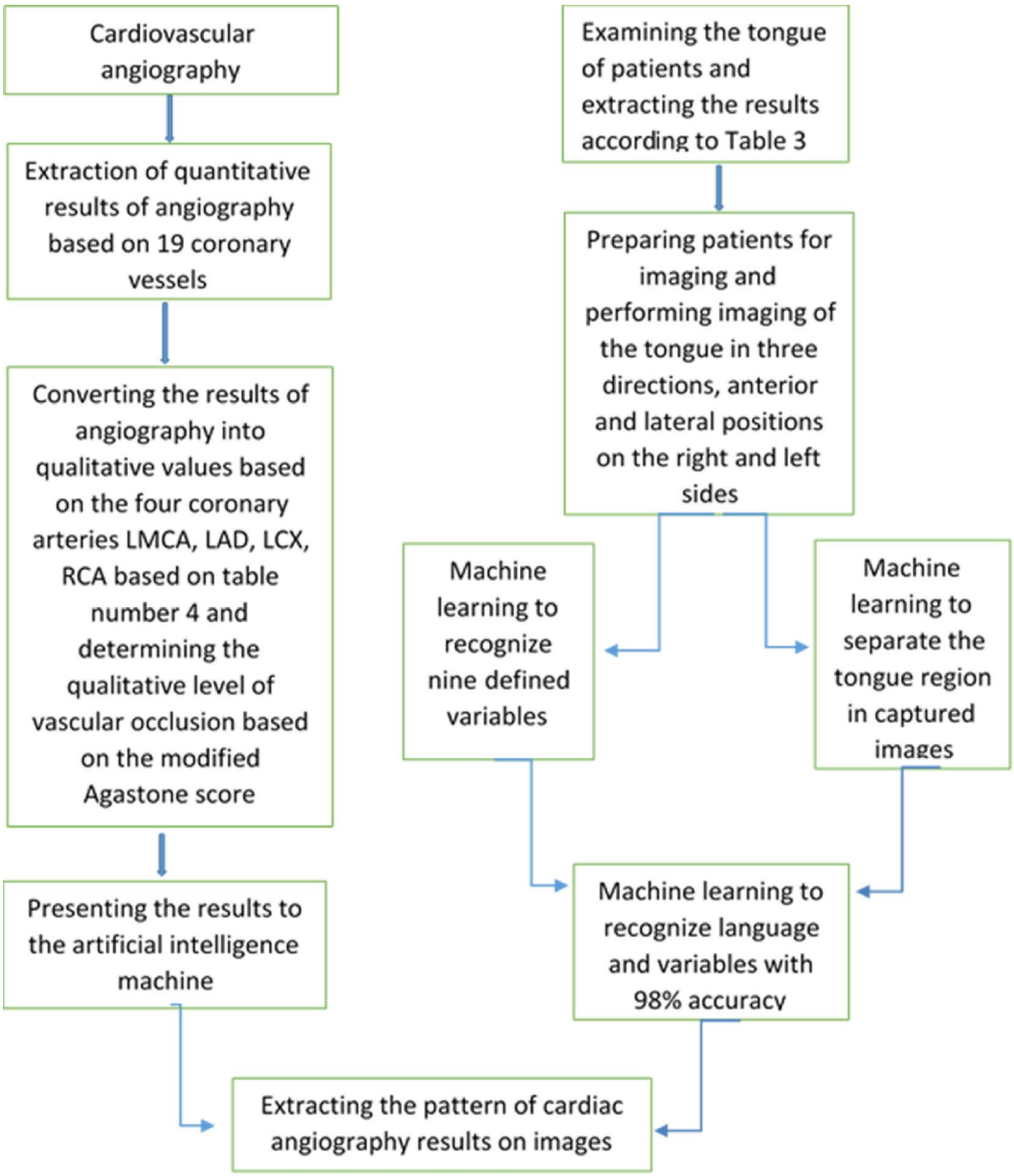


Fig. 2. The complete study pipeline from data acquisition to CAD prediction. This flowchart details the multi-stage process of our research. It begins with the initial collection of tongue images, followed by advanced image preprocessing, computational feature extraction, and finally, the AI-driven classification to predict the likelihood of Coronary Artery Disease (CAD). Each step is designed to transform raw visual data into a robust, predictive biomarker.

To overcome the limitations of a relatively small, labeled clinical dataset, our methodology incorporates a two-stage training process. For feature extraction, our deep learning models were first pre-trained on a large, auxiliary dataset of 12,000 unannotated tongue images from the general population. This self-supervised pre-training was specifically critical for learning to identify and segment features with low prevalence, such as teeth marks and red spots, and for generalizing to highly variable features like moisture and fissures, which require a large number of images for robust training. This approach is crucial as it leverages a vast amount of available unlabeled data to establish a strong foundational knowledge for the model. Following this, the pre-trained models were fine-tuned on our primary, labeled dataset of 906 CAD patients for the specific task of disease prediction. This transfer learning approach significantly enhances the models' performance and reduces the risk of overfitting to the limited clinical data.

Our proposed method consists of three main steps: (1) extraction of the tongue from the facial image, (2) extraction of tongue characteristics from the tongue image, and (3) estimation of the degree of coronary artery occlusion based on the extracted characteristics and tongue image. The whole pipeline is illustrated in Fig. 4. In

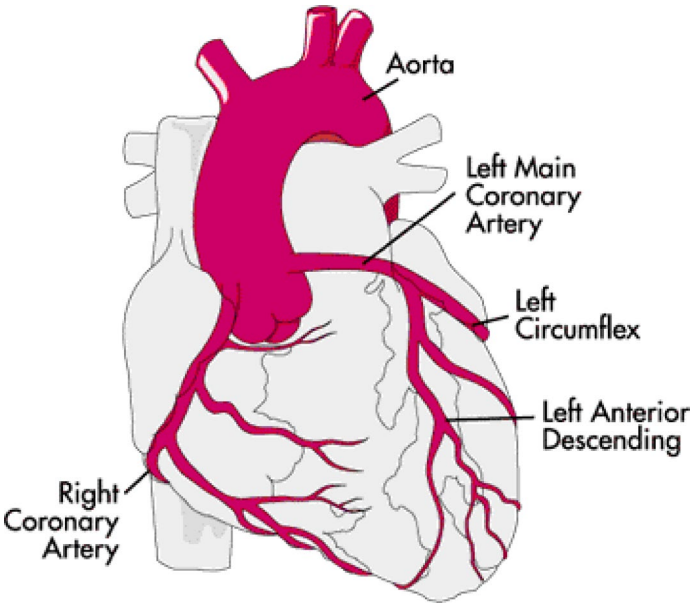


Fig. 3. Anatomical reference for the major coronary arteries. This diagram provides a visual guide to the three main coronary arteries: the Left Anterior Descending (LAD), Left Circumflex (LCX), and Right Coronary Artery (RCA). This figure serves as a crucial reference for understanding the spatial relationship between the model's predictions and the specific arteries discussed in our results.Taken from the link below:http://anatomyof.weebly.com/great-vessels-of-the-heart.html.

the following sections, each of these three steps will be explained in detail. Here are the implementation details and setup for the main CAD detector models.

Pretrained CAD model (DINOv2) setup

Model: Vision Transformer Large (ViT-Large).
Framework: DINOv2.
Pretraining method: Self-supervised learning.
Weights: vit_large_patch14_dinov2.lvd142m from timm library.
Pretraining dataset: LVD-142 M.

Transfer learning & fine-tuning strategy

Parameter freezing: None (Full fine-tuning).
Backbone (ViT): All parameters in the transformer blocks were trainable (requires_grad = True).
Input layer: The initial Conv2d patch projection layer was replaced to handle custom input channels; this new layer was trained from scratch.
Classifier head: A custom MLP head was added and trained from scratch.
Structure: Linear(1024 -> 512), ReLU, Dropout($p = 0.3$), Linear(512 -> 3).

Training hyperparameters

Optimizer: AdamW.
Initial learning rate: 1e-3.
Weight decay: 1e-4.
LR scheduler: OneCycleLR.
Max learning rate: 1e-4.
Loss function: Weighted Cross-Entropy (weights based on inverse class frequency of training set).
Training duration: 150 epochs.

Data handling

Batch size: 8.
Input image size: 518×518.
Data split: 80% Train / 20% Validation (Stratified).

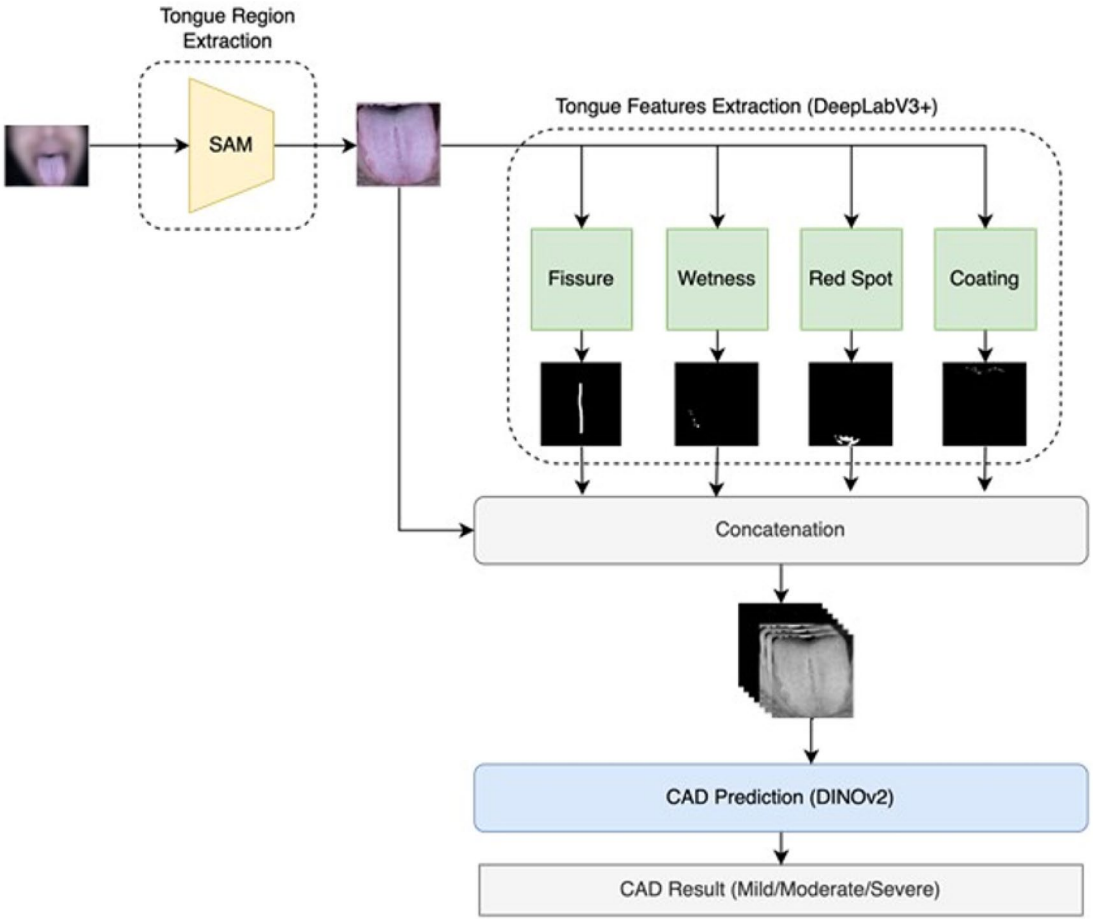


Fig. 4. The complete AI-driven pipeline for tongue image analysis and CAD classification. This detailed flowchart illustrates the entire process, starting with the patient's tongue image captured by the camera. The image is then passed through a series of automated AI-based preprocessing steps, including segmentation, fissure and red spot detection, and wetness and coating analysis. The computationally extracted features are then fed into the main AI module for the final prediction of CAD severity.

Data augmentation (training)

Geometric: HorizontalFlip($p = 0.3$), ShiftScaleRotate($p = 0.5$).
Normalization: ImageNet standard mean/std, VerticalFlip($p = 0.3$), RandomRotate90($p = 0.3$).

Tongue region extraction

The initial step involves processing the facial image for analysis. We employed the Segment-Anything Model (SAM)[25], an image segmentation model based on visual transformers, to leverage its pre-trained capabilities in extracting the tongue region from the facial context. Using three fixed prompts, one for the foreground and two for the background, as shown in Fig. 5, we consistently segmented the tongues from the facial images. This method establishes a reliable foundation for the subsequent analysis.

Tongue characteristics segmentation

In alignment with the principles of Iranian Traditional Medicine and other holistic medicines, such as TCM, we extract and note the suggested features that are critical for CAD. These include the coating, fissures, wetness, red spots, and tooth marks. Coatings are obtained differently from other features.

Detection of coating

To detect the coating, we process the image in the CIELAB color space. We focus on the 'A' channel, which represents the redness of the pixels (the intensity of red minus the intensity of green). This channel is denoised using a shifted truncation (ST) module defined as:

$$ST(x;\ a,\ b) = \begin{cases} b & x < a \\ x & otherwise \end{cases}$$

With parameters (a, b) = (145,119).

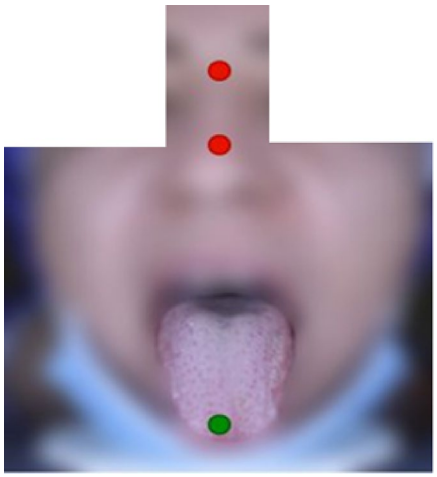


Fig. 5. Using SAM for facial image segmentation to extract the tongue region. This figure illustrates the function of the Segment Anything Model (SAM) in the initial stage of our image processing pipeline. Given an input prompt from the tongue and two different prompts from outside the tongue, the SAM model automatically and accurately identifies the "foreground" region, which is the tongue, and separates it from the "background" (the rest of the facial image). This method provides an efficient and accurate solution for the automated extraction of the tongue from the full image.

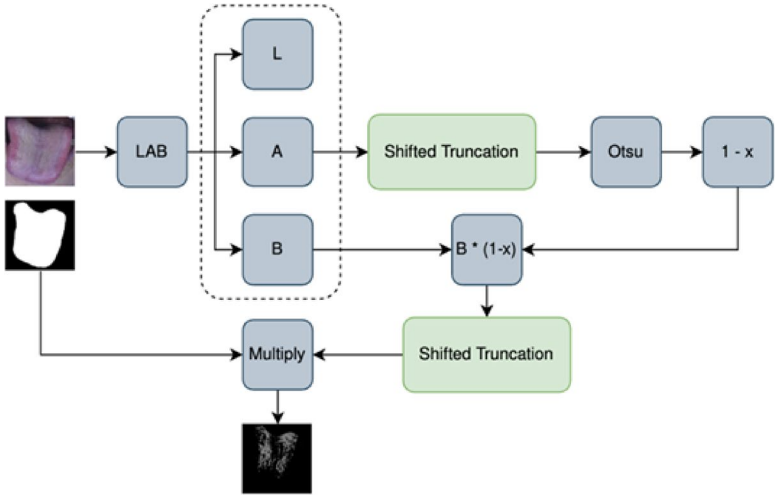


Fig. 6. Procedure for the quantitative analysis of tongue coating. This figure illustrates the computational process of extracting tongue coating. Regions of the tongue with low red color and high yellow tones, as measured in the CIELAB color space, are algorithmically identified as coating. The intensity and thickness of the coating are quantified based on the contrast of these pixels relative to the overall tongue color, providing an objective metric for this feature.

Otsu thresholding is then applied to isolate dominant areas. The inverse of this mask highlights regions with low red color intensity, which are then multiplied by the 'B' channel to identify regions exhibiting high yellow intensity and low red intensity. This result is denoised again using ST ($_x$; 120, 0). Finally, the output is filtered by the tongue mask to detect only the areas on the surface of the tongue, as illustrated in Fig. 6.

Detection of other tongue features

For the other characteristics, we utilize machine-learning techniques to develop models capable of performing pixel-wise segmentation of tongue images. To accomplish this, we require a training set of annotated tongue images. We organized mask annotation projects using the Label Studio library for each feature, with the annotations performed by a medical expert. A total of 450 tongue images were annotated and used to train our feature extraction models. Additionally, a dataset consisting of 12,000 tongue images from the general population was used to train the AI model to identify specific variables, including teeth marks, red spots, fissures, and moisture levels. These images were carefully reviewed by an Iranian medicine specialist, resulting in the selection and labeling of at least 200 representative images for each variable to ensure precision in the training process.

For segmentation, we implement the DeepLabV3 + model[26]. We train a separate model for each of the four characteristics. The training process employs focal binary cross-entropy loss as the objective function, defined as:

$$\sum {}_i\left(1-f(x)_i\right)^{\gamma} y_i log\left(f(x)_i\right)+f(x)_i^{\gamma}\ (1-y_i)\, log\left(1-f(x)_i\right)$$

Figure 7 illustrates examples of the extracted features.

Our ultimate goal is to develop a model that accurately detects the degree of occlusion in the LAD, LCX, and RCA arteries. The occlusion is categorized into three levels: "Nothing/Mild" (low), "Moderate" (medium), and "Severe/Occluded" (high). For each artery, we train a separate model. We use the DINOv2[27] model architecture, a ViT-based deep neural network, which is used as a foundation model in Computer Vision. We replace the classification head of the model with a two-layer fully connected neural network, where the size of the hidden layer is 512, and fine-tune the entire model on our collected and annotated dataset. We conducted multiple experiments with different multichannel configurations, incorporating various combinations of extracted tongue characteristics and the RGB tongue image. These experiments involved feeding the modified DINOv2 model with different multichannel inputs, as shown in Fig. 4. For each of the arteries, we train a model in a 3-class classification task. Weighted cross-entropy loss is used to train the model. Weighted cross-entropy is used when the distribution of data samples in different classes is uneven, resulting in class imbalance in the dataset. The binary classification output of the model is designed to discern between two categories of heart vessel occlusions as observed in patient examinations: those indicative of severe occlusion and those without severe occlusion. For training the model, we put aside 150 images for testing, 100 images for validation, and 655 others for training. We also use focal for S-HVO detection:

$$l\ \left(f\left(x_i\right),\ y\right)=\ \ \left(1-f\left(x\right)_i\right)^{\gamma}\ y_i \log\left(f\left(x\right)_i\right)+f\left(x\right)_i^{\gamma}\ \left(1-y_i\right)\log\left(1-f\left(x\right)_i\right)$$

DINOv2 is a Deep Neural Network trained based on self-supervised learning, in which models are trained on vast amounts of unsupervised image data. This enhances the representation power of the DINO model, making it a powerful tool for training on small image datasets. The patch-based global representations of the ViT architecture in DINOv2 make it a strong choice for complex structures and patterns in small-scale datasets. Moreover, DINOv2 demonstrates strong domain transfer capability and is well-suited for our scenario, where images of the human tongue are relatively rare in public image datasets. DeepLabV3 + is also especially good for fine-tuning on small datasets. The precise boundary recovery and strong capacity of this model make it suitable for complex segmentation tasks. DeepLabV3 + not only maintains spatial resolution but also utilizes a vast receptive field using atrous (dilated) convolutions, while controlling memory usage. The multi-scale feature fusion in this model enables it to recognize patterns and entities across different scales. SAM is a foundation model for image segmentation, typically much larger than CNN models used for training on most segmentation datasets. A model of such a large scale, efficiently trained on vast datasets (e.g., SA-1B with 1 billion segmentation masks and 11 billion images), is ideal for zero-shot inference. We utilize the pre-trained SAM model to directly

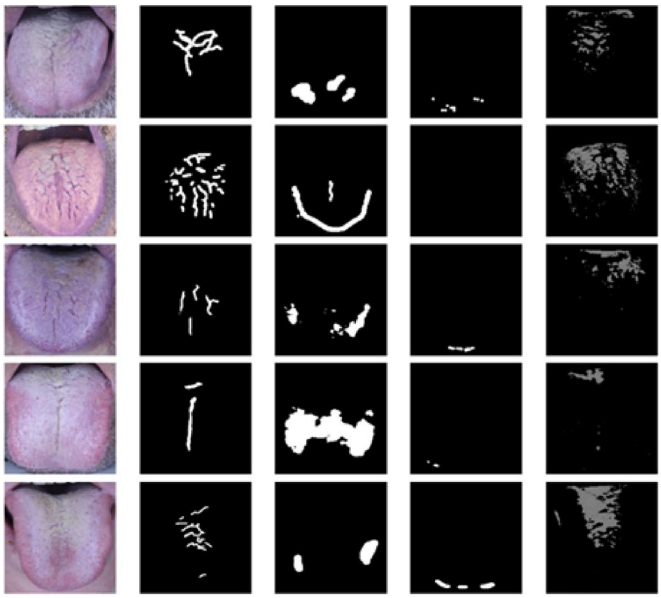


Fig. 7. Visualization of computationally extracted tongue characteristics. This figure provides a visual representation of the key features extracted by our AI pipeline from four example tongue images. In each row, the panels display (from left to right): (1) the original tongue image, (2) the computationally detected fissures, (3) wetness intensity, (4) the identified red spots, and (5) a visual map of the tongue coating. This visualization demonstrates the ability of our model to quantify and map key biomarkers objectively.

Characteristic	Average ± range
Sex	Male: 58% < br> Female: 42%
Age (years)	57.9 (Min: 29, Max: 86)
BMI (Kg/m^2)	27.85 (Min: 14.33, Max: 47.66)
Pulse rate (pulse/minute)	75 (Min: 18, Max: 118)
Systolic pressure (mmHg)	126.6 (Min: 80, Max: 194)
Diastolic pressure (mmHg)	79.4 (Min: 50, Max: 118)

Table 1. General characteristics of Participants.

Health Condition	Prevalence (%)
Hypertension (High blood pressure)	53.8%
Dyslipidemia	46.0%
Type 2 Diabetes Mellitus (DM type two)	34.1%
Smoking (1 to 100 pack/year)	30.8%
Opium use	18.8%
Constipation (more than two days between defecations)	26.6%
Hypothyroidism	8.6%
Hyperthyroidism	1.1%
Passive smoker (more than ten years)	3.3%

Table 2. Prevalence of health conditions among participants. Patients taking medication for their condition for more than a year are included in relevant categories.

and accurately identify the region of the tongue in the subject's face, eliminating the need for training or additional datasets. The prompt-based structure of SAM enables us to specialize the task of segmentation for our own context and application. By using appropriate prompts, the accurate detection of the tongue's region in the face was made possible due to SAM's strong out-of-distribution segmentation capability. Also, SAM is object-agnostic and segments primarily based on visual coherence rather than semantical entities in the image.

Results

Tongue images of all patients participating in the study were captured under uniform conditions following the acquisition of informed consent. Additionally, their coronary angiography results were obtained. The general characteristics of the participants are detailed in Table 1.

This study identified common cardiac symptoms, including chest pain (72.9%), pain in the left shoulder/back (73.7%), left arm numbness or pain (18.1%), epigastric pain (7.9%), and jaw/neck pain (2.3%). Multiple pain sites and additional symptoms such as pain in both arms, tingling at rest, chest tightness, and sensations like tearing or freezing were also reported.

The primary pain described was a sensation of pressure or burning in the chest, followed by shooting pain, heaviness, stabbing sensations, and electric shock-like feelings. Additionally, 57.8% experienced exertional dyspnea, 25.9% had shortness of breath with chest pain, and 5% experienced dyspnea at rest.

A significant proportion of patients possessed risk factors associated with cardiovascular disease. The prevalence of metabolic diseases and other risk factors among the patients is summarized in Table 2.

Angiographic results and outcome in study participants

The angiographic findings from the participants are summarized as follows:

- **Coronary artery status (CAS)**: A total of 120 patients (13.2%) exhibited completely open coronary arteries, with no evidence of even mild stenosis.
- **Left main coronary artery (LMCA) stenosis**: stenosis was observed in the LMCA of 48 patients, including 30 with mild stenosis, 9 with moderate stenosis, and 9 with severe stenosis.
- **Stenosis in major coronary arteries**: Among the three primary coronary arteries, the highest prevalence of stenosis was noted in the Left Anterior Descending Artery (LAD), followed by the Right Coronary Artery (RCA), and then the Left Circumflex Artery (LCX), respectively.

The severity of stenosis in the coronary arteries is presented in Table 3.

The most severely affected segments of the coronary arteries were primarily located in the middle portion of each vessel.

In terms of treatment outcomes, the most common intervention was PCI, administered to 51.2% of the patients. This was followed by medical therapy, which was used in 40.8% of cases, and coronary artery bypass grafting (CABG) surgery, performed in 6.5% of patients.

Coronary artery	Severe stenosis or occlusion (%)	Moderate stenosis (%)	Mild stenosis or no stenosis (%)
LAD	44.1%	8.2%	47.7%
RCA	31.4%	6.6%	62.0%
LCX	27.9%	6.6%	65.5%

Table 3. The severity of coronary artery stenosis. LAD: Left Anterior Descending Artery, RCA: Right Coronary Artery, LCX: Left Circumflex Artery. The percentages indicate the prevalence of each severity category for stenosis among the respective coronary arteries.

Feature set	Average accuracy	RCA accuracy	LCX accuracy	LAD accuracy
RGB + F	0.70	0.66	0.80	0.64
RGB + F + W	0.66	0.67	0.70	0.60
RGB + F + W + RS	0.64	0.65	0.70	0.57
C + F	0.63	0.68	0.68	0.52
C + F + W	0.64	0.70	0.70	0.52
C + F + W + RS	0.66	0.74	0.70	0.54

Table 4. Accuracy of AI predictions for stenosis in each vessel. LAD: Left Anterior Descending Artery, RCA: Right Coronary Artery, LCX: Left Circumflex Artery, RGB: Tongue color, F: Tongue fissure, W: Tongue wetness, RS: Tongue red spot, C: Tongue coating. The table summarizes the accuracy of AI predictions for stenosis in the RCA, LCX, and LAD, based on varying feature sets derived from tongue examination.

We compared the specialist's visual assessments of tongue shape, color, dental markings, fissures, coating, and moisture level against angiographically measured coronary artery occlusion. No statistically significant associations were observed ($p > 0.05$).

Results from artificial intelligence analysis

We utilized images of patients' faces and tongues as input for the AI analysis. The objective was for the AI device to predict the degree of stenosis in each of the three principal coronary arteries: the LAD, the RCA, and the LCX, categorizing stenosis into three degrees:

1. No Occlusion or Mild stenosis (0% to 49%).
2. Moderate Stenosis: (50% to 69%)
3. Severe Stenosis (70% to 99%) or complete occlusion.

We obtained the following segmentation performance metrics on 450 expert-annotated tongue fissure images: a Dice coefficient of 0.84, an Intersection over Union (IoU) score of 0.70, a precision of 0.81, and a recall of 0.83. These values indicate high segmentation accuracy and reliability for fissure extraction.

The accuracy of the AI predictions for each artery, considering the assessed tongue features, is presented in Table 4.

In our analysis, tongue coating is treated as a distinct variable, separate from tongue color. The AI model did not require traditional machine learning methods to recognize this variable; instead, it was trained using advanced image processing techniques.

For the evaluation of tongue coating, three parameters were considered: area, color, and thickness. The final measurement of coating was calculated as a product of these three factors, thereby accounting for the influence of each aspect of the coating on the overall assessment.

To identify which regions of the tongue contribute most significantly to predicting coronary artery stenosis, we divided the surface of the tongue into nine distinct segments. Utilizing the RGB + F indices, which demonstrated the highest predictive accuracy for stenosis in the LCX, we evaluated the impact of each segment on the prediction model by systematically removing one segment at a time. Figure 8.

Our findings indicate that removing any tongue segment led to a decline in prediction accuracy across all coronary branches. The most pronounced decrease for LCX occurred when the right third of the tongue was excluded, underscoring this region's critical contribution to model performance. Similarly, RCA accuracy fell most sharply upon removal of the right third, whereas LAD accuracy experienced its most significant drop when the middle third was omitted. Table 5 presents the quantitative accuracy reductions for LCX as an illustrative example.

Table 6 demonstrates the specific correlations between tongue features, their corresponding regions, and the prediction of coronary artery stenosis.

To assess clinical utility, we stratified model performance by stenosis severity—mild, moderate, and severe/occlusion—across the LAD, LCX, and RCA, generating confusion matrices (Fig. 9) for a granular sensitivity analysis at each severity level.

Average and the Gaussian confidence interval method for each metric reported in Table 7.

The model's sensitivity for mild/moderate/severe stenosis was:

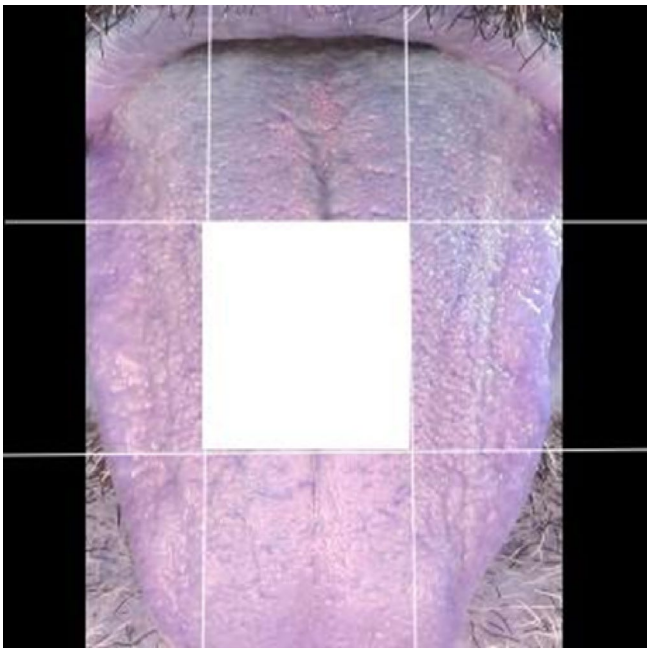


Fig. 8. Importance of the tongue region for CAD prediction. This figure illustrates the methodology used to identify the most influential anatomical regions of the tongue for Coronary Artery Disease (CAD) prediction. To assess the importance of each area, the tongue was systematically divided into nine distinct segments. The model's performance was then evaluated by sequentially occluding (i.e., removing) each of these segments and measuring the resulting drop in predictive accuracy. The segment whose removal caused the most significant decrease in model performance was identified as the region with the highest predictive power for CAD. The nine segmented regions, shown here, correspond to specific anatomical zones of the tongue, enabling a granular analysis of their contribution to the final diagnosis.

Removed segment	Column 1	Column 2	Column 3
Row 1	0.65	0.72	0.72
Row 2	0.65	0.67	0.71
Row 3	0.65	0.68	0.71

Table 5. Prediction accuracy after removal of tongue segments. The table presents the prediction accuracy for LCX after sequential removal of specific tongue segments.

Coronary artery branch	Most effective features	Most effective tongue region	Prediction accuracy
LCX	F + RGB	Right Third of the Tongue	80%
LAD	F + RGB	Middle Third of the Tongue	64%
RCA	F + C + W + RS	Right Third of the Tongue	74%

Table 6. Correlation of most effective tongue features and regions with coronary artery Stenosis. LAD: Left Anterior Descending Artery, RCA: Right Coronary Artery, LCX: Left Circumflex Artery, RGB: Tongue color, F: Tongue fissure, W: Tongue wetness, RS: Tongue red spot, C: Tongue coating.

- LCX: 84.2% / 75.0% / 71.0%
- RCA: 71.3% / 68.9% / 81.0%
- LAD: 69.4% / 68.0% / 57.7%

This confirms the model's ability to stratify disease severity, highlighting its value as a screening tool.

Prediction of patient outcomes based on tongue variables

The prediction of patient outcomes was conducted by analyzing tongue variables using an artificial intelligence device. Given that 98.8% of patient outcomes fell into three categories: medical therapy, percutaneous coronary intervention (PCI), and coronary artery bypass graft (CABG). We calculated the results specifically for these

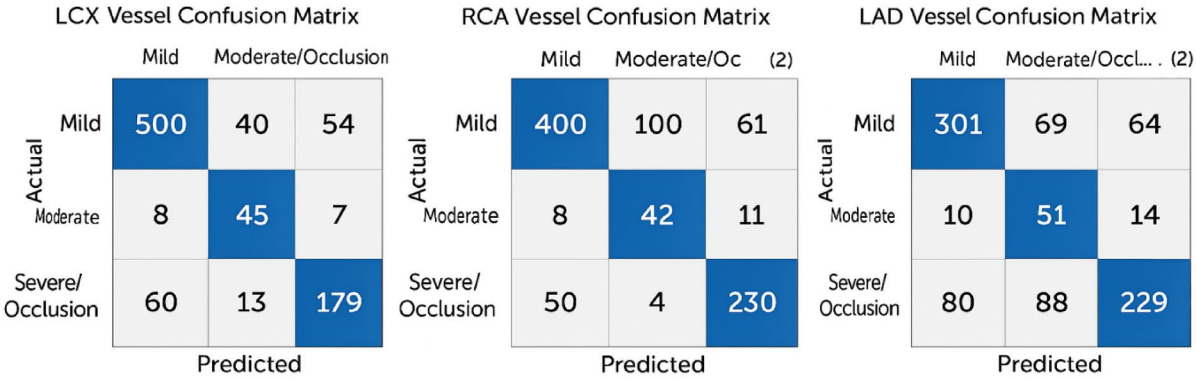


Fig. 9. Confusion matrices for coronary artery disease severity classificationthese visual confusion matrices present the classification performance of the model for each of the three main coronary arteries (RCA, LAD, and LCX). Each cell contains the count of samples and the corresponding percentage. The vertical axis (Y-axis) represents the Actual Class, and the horizontal axis (X-axis) represents the Predicted Class. High values along the main diagonal indicate correct predictions.

Vessel	Metric	Mean	95% CI lower	95% CI upper	Format
RCA	F1 score	0.6091	0.5407	0.6775	0.609 ± 0.068
	Accuracy	0.7749	0.7214	0.8284	0.775 ± 0.054
	AUC	0.5783	0.5275	0.6291	0.578 ± 0.051
LAD	F1 score	0.4520	0.3777	0.5263	0.452 ± 0.074
	Accuracy	0.5486	0.4773	0.6199	0.549 ± 0.071
	AUC	0.6216	0.5483	0.6949	0.622 ± 0.073
LCX	F1 score	0.6052	0.5256	0.6848	0.605 ± 0.080
	Accuracy	0.7714	0.7095	0.8333	0.771 ± 0.062
	AUC	0.5388	0.4759	0.6017	0.539 ± 0.063

Table 7. Performance metrics (F1 Score, Accuracy, AUC) with 95% CI for Vessel-Specific Models. LAD: Left Anterior Descending Artery, RCA: Right Coronary Artery, LCX: Left Circumflex Artery.

categories. Table 8 outlines the prediction accuracy of patient outcomes based on different combinations of tongue features analyzed by the artificial intelligence device.

Discussion

The findings of this study underscore the potential of tongue examination as a non-invasive and accessible diagnostic tool for assessing CVD. Utilizing artificial intelligence to analyze various tongue features such as size, color, fissures, coating, and moisture levels provides valuable insights into patients' underlying health, particularly for individuals at risk of CAD. Our results suggest that specific tongue features may serve as reliable indicators of CVD, supporting prior research that has documented associations between tongue morphology and systemic diseases[12–19]. These findings support further investigation into the integration of tongue examination into routine clinical practice as a complementary, non-invasive approach for cardiovascular assessment.

This study also investigated the relationship between tongue symptoms and stenosis of the three main coronary arteries, LAD, LCX, and RCA. Leveraging AI, we developed a rapid, non-invasive method to accurately

Tongue feature set	Prediction accuracy
RGB + F	0.61
RGB + F + W	0.68
RGB + F + W + RS	0.68

Table 8. Prediction of patient outcomes based on tongue features.

predict stenosis in these arteries, providing a novel quantitative assessment of stenosis severity. In a preliminary comparison, a single traditional medicine expert's visual assessments of tongue features showed no statistically significant correlation with angiographic stenosis in any coronary artery or with the deep learning model's predictions ($p > 0.05$). Traditional tongue diagnosis suffers from subjectivity and poor reproducibility, as expert-evaluated features showed no significant correlation with CAD. Our AI-driven approach overcomes this by extracting objective, quantifiable metrics such as CIELAB color values and fissure width measurements directly from images. This shift to computational analysis represents a significant advance in non-invasive diagnostics.

We conducted a severity-stratified evaluation of our AI-driven tongue image model across the three main coronary arteries, using confusion matrices and sensitivity metrics. LCX predictions were most sensitive for mild (84.2%) and moderate (75.0%) stenosis, underscoring its value as an early screening tool. RCA sensitivity peaked at 81.0% for severe/occlusive lesions, highlighting its potential for urgent triage. LAD performance was highest for moderate stenosis (68.0%), aiding in risk stratification. This granular analysis—beyond overall accuracy—demonstrates the model's differential strengths and supports tongue image analysis as a noninvasive method for both detecting CAD and assessing its severity.

A key finding of our study is the variation in the model's predictive performance across the three major coronary artery branches. While the model showed strong accuracy for LCX stenosis, its performance for LAD and RCA was comparatively lower. We hypothesize that this discrepancy may be attributed to the anatomical and pathological variations that manifest on the tongue. The visual biomarkers associated with LCX stenosis, such as changes in the central tongue body and prominent fissures, may be more pronounced or consistent across the patient cohort, making them easier for the AI to learn. In contrast, the features associated with LAD and RCA may be more subtle or complex, requiring a larger and more diverse dataset to be fully captured. This observation underscores the need for further research to better understand the intricate relationships between the tongue and the arteries.

Model accuracy varied by artery, with LCX and LAD stenosis being predicted most accurately using fissures (F) and tongue color (RGB), highlighting these as key predictors. In contrast, RCA stenosis required a broader feature set—fissures, coating, red spots, and wetness—suggesting its association with a wider range of tongue abnormalities.

Our study uniquely mapped tongue regions to predict artery-specific stenosis. Features in the tongue's right third most strongly predicted LCX and RCA stenosis, while the middle third was the primary predictor for LAD stenosis. This regional correlation provides a more localized clinical interpretation of diagnostic findings.

This study demonstrates a significant statistical association between specific tongue regions and coronary artery stenosis, resonating with mappings long asserted by Traditional Chinese and Iranian Medicine. Focusing on distinct thirds of the tongue enhances the accuracy of CAD diagnostic models compared to whole-tongue analysis. For future research, exploring the underlying physiological mechanisms through imaging and histological studies could establish causality. Integrating hyperspectral tongue imaging and validating the findings in larger, more diverse cohorts will enhance the clinical utility of these region-specific diagnostic markers.

Identifying the specific affected vessel and the extent of the occlusion is crucial for timely intervention. Notably, this is the first application of AI-enhanced tongue examination to predict patient outcomes, highlighting its potential to guide urgent clinical decisions. To our knowledge, this study is the first to examine CAD diagnosis through a traditional diagnostic framework, for which no prior consensus scales or criteria exist. Accordingly, we report our AI model's performance independently and emphasize the necessity for future work to develop standardized expert scoring guidelines and multi-center datasets for robust comparative validation.

A recent study has explored the use of tongue images as predictive indicators for CAD in individuals with hypertension. Their results indicated that tongue features could serve as valuable biological markers for CAD diagnosis, achieving a predictive accuracy of 76.0% among patients with hypertension-related risk factors. This aligns with the growing body of evidence supporting the potential of tongue analysis as a non-invasive diagnostic tool in cardiovascular risk assessment[28]. Nevertheless, our study not only predicted the extent and severity of stenosis in each of the three main coronary arteries but also forecasted patient outcomes, achieving higher accuracy compared to the referenced studies.

In our application of artificial intelligence, we incorporated advanced image processing techniques alongside deep learning methodologies. Image processing was employed to detect and quantify specific tongue features, including coating area, thickness, and color. Simultaneously, deep learning algorithms analyzed various characteristics, including color variations, fissures, moisture levels, teeth marks, and red spots, providing a comprehensive assessment of tongue morphology. Our findings indicate that tongue color and the presence of tongue fissures were the most significant predictors of coronary artery stenosis.

Several studies have confirmed the significance of tongue color in predicting the severity of coronary artery stenosis. These studies have demonstrated that tongue color can serve as a reliable indicator for assessing the degree and severity of vascular obstruction[17]. Additionally, the critical role of tongue color has been highlighted

in diagnosing and evaluating the progression of cardiovascular disease[29]. These findings are consistent with the results of our own study and underscore the importance of observable tongue features in predicting and assessing vascular health.

Fissured tongue, predominantly observed in adults and rarely in children under 10, is linked to underlying inflammation and leukocyte infiltration beneath the epithelium. It serves as a potential indicator of inflammatory status[30,31]. In traditional Chinese medicine, this condition is associated with excessive heart activity and blood exhaustion[32]. The fissured tongue is closely connected to inflammatory processes and vascular abnormalities, which are significant factors in the development of coronary artery disease.

Our study's strengths include the ability to predict the severity of stenosis in each of the three major coronary arteries and to forecast patient outcomes based on artificial intelligence analysis of tongue images. Additionally, our findings indicate that symptoms present on the right one-third of the tongue have a stronger correlation with coronary artery stenosis. To the best of our knowledge, this has not been previously reported.

One limitation of our study is that the mean age of participants was 57.9 years, and younger individuals were omitted. This may restrict the generalizability of the findings to a broader age group. Furthermore, the images used in this study were captured in a controlled environment without natural light, utilizing high-performance equipment. Future research should aim to analyze images taken with standard cameras in diverse settings and lighting conditions to establish acceptable imaging standards. Despite leveraging a large auxiliary dataset to alleviate data constraints, our primary clinical cohort ($n = 906$) remains limited. Future studies should validate model performance on larger, diverse, multi-center datasets to ensure generalizability. Additionally, although self-supervised pretraining addresses data scarcity, its dependence on general-population images may introduce biases that warrant further investigation.

While our model shows robust performance, particularly for LCX stenosis, we recognize the inherent limitations of a single-center study. Therefore, we emphasize that external validation on a large, independent, and geographically diverse patient cohort is the critical next step to confirm our findings and to pave the way for the clinical adoption of this technology.

While our study represents a novel application of AI to tongue imaging, its significance is best understood when contextualized within the broader landscape of non-traditional cardiovascular diagnostics. Recent advancements have demonstrated the remarkable potential of AI in analyzing various non-invasive data sources for assessing CAD risk. For example, AI algorithms have shown high accuracy in interpreting electrocardiograms (ECGs) to predict cardiac conditions. In contrast, others have successfully leveraged chest X-rays and retinal images to estimate cardiovascular risk. These approaches, by exploiting subtle patterns invisible to the human eye, are revolutionizing the field of early detection[33,34].

Despite the promising nature of these methods, our tongue-based approach offers several unique advantages. Unlike AI-ECG and other digital health tools, our method relies on visual biomarkers that are not susceptible to electrical noise or patient movements. It is also more accessible and cost-effective than chest X-rays or retinal imaging, making it a highly scalable solution for mass screening, particularly in primary care and remote areas. The tongue's holistic representation of systemic health provides a complementary set of biomarkers that are distinct from other modalities, offering a novel and promising avenue for early detection and risk stratification of CAD.

While this study investigates the potential benefits of tongue diagnosis for CAD, future research is necessary to validate and refine our diagnostic model in a larger, more diverse population. We anticipate that further research in this area will lead to the development of software capable of analyzing tongue images captured by digital cameras in various settings. This software aims to provide objective, image-based biomarkers for the early detection and monitoring of diseases affecting internal organs.

Data availability

The data will be available on a reasonable request from the corresponding author in due course.

Received: 4 June 2025; Accepted: 11 November 2025
Published online: 29 December 2025

References

1. Tsao, C. W. et al. Heart disease and stroke Statistics-2023 update: A report from the American heart association. *Circulation* **147**, e93–e621. https://doi.org/10.1161/cir.0000000000001123 (2023).
2. Talaei, M. et al. Incidence of cardiovascular diseases in an Iranian population: the Isfahan cohort study. *Arch. Iran. Med.* **16**, 138–144 (2013).
3. Bertrand, T., Bartlett-Esquilant, G., Fischer, K. & Friedrich, M. G. Patient and physician preferences for non-invasive diagnostic cardiovascular imaging technologies: a discrete choice experiment. **6**, 15, https://doi.org/10.1186/s41687-022-00419-0 (2022).
4. Mahmoodzadeh, S., Moazenzadeh, M., Rashidinejad, H. & Sheikhvatan, M. Diagnostic performance of electrocardiography in the assessment of significant coronary artery disease and its anatomical size in comparison with coronary angiography. *J. Res. Med. Sciences: Official J. Isfahan Univ. Med. Sci.* **16**, 750–755 (2011).
5. Jahnavi, D. et al. Non-invasive Coronary Artery Disease Screening Based on Electrocardiogram Characteristics and Clinical Risk Factors. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference* 1–4, (2023). https://doi.org/10.1109/embc40787.2023.10340892 (2023).
6. Attar, A. et al. Accuracy of exercise tolerance test in the diagnosis of coronary artery disease in patients with left dominant coronary circulation. *Indian Heart J.* **69**, 624–627. https://doi.org/10.1016/j.ihj.2017.02.009 (2017).
7. Peteiro, J. & Bouzas-Mosquera, A. Exercise echocardiography. *World J. Cardiol.* **2**, 223–232. https://doi.org/10.4330/wjc.v2.i8.223 (2010).
8. Marwick, T. H., Nemec, J. J., Pashkow, F. J., Stewart, W. J. & Salcedo, E. E. Accuracy and limitations of exercise echocardiography in a routine clinical setting. *J. Am. Coll. Cardiol.* **19**, 74–81. https://doi.org/10.1016/0735-1097(92)90054-q (1992).

9. Hagar, M. T. et al. Accuracy of Ultrahigh-Resolution Photon-counting CT for detecting coronary artery disease in a High-Risk population. **307**, e223305, https://doi.org/10.1148/radiol.223305 (2023).
10. Hong, Y. et al. Predicting categories of coronary artery calcium scores from chest X-ray images using deep learning. *J. Cardiovasc. Comput. Tomogr.* **19**, 331–339. https://doi.org/10.1016/j.jcct.2025.03.010 (2025).
11. Hu, W. et al. Real-world feasibility, accuracy and acceptability of automated retinal photography and AI-based cardiovascular disease risk assessment in Australian primary care settings: a pragmatic trial. *NPJ Digit. Med.* **8** https://doi.org/10.1038/s41746-025-01436-1 (2025).
12. Jung, C. J., Jeon, Y. J., Kim, J. Y. & Kim, K. H. Review on the current trends in tongue diagnosis systems. *Integr. Med. Res.* **1**, 13–20. https://doi.org/10.1016/j.imr.2012.09.001 (2012).
13. Islam, N. M., Bhattacharyya, I. & Cohen, D. M. Common oral manifestations of systemic disease. *Otolaryngol Clin North Am.* **44**, 161–182. https://doi.org/10.1016/j.otc.2010.09.006 (2011). vi.
14. Bugălă, N. M., Carsote, M., Stoica, L. E., Albulescu, D. M. & Ţuculină, M. J. New Approach to Addison Disease: Oral Manifestations Due to Endocrine Dysfunction and Comorbidity Burden. 12,https://doi.org/10.3390/diagnostics12092080 (2022).
15. Hsu, P. C. et al. The tongue features associated with type 2 diabetes mellitus. *Medicine* **98**, e15567. https://doi.org/10.1097/md.0000000000015567 (2019).
16. Cheng, T. O. Tongue as the window to the heart. *Chin. J. Integr. Med.* **17** https://doi.org/10.1007/s11655-011-0824-x (2011).
17. Li, J. et al. Tongue color parameters in predicting the degree of coronary stenosis: a retrospective cohort study of 282 patients with coronary angiography. *Front. Cardiovasc. Med.* **11-2024** https://doi.org/10.3389/fcvm.2024.1436278 (2024).
18. Xia, Y. et al. Tongue diagnosis based on hue-saturation value color space: controlled study of tongue appearance in patients treated with percutaneous coronary intervention for coronary heart disease. *Intell. Med.* **3**, 252–257. https://doi.org/10.1016/j.imed.2022.09.002 (2023).
19. Feng, Y. et al. Study on the tongue manifestations for the blood-stasis and toxin syndrome in the stable patients of coronary heart disease. *Chin. J. Integr. Med.* **17**, 333–338. https://doi.org/10.1007/s11655-011-0615-4 (2011).
20. Segawa, M. et al. Objective evaluation of tongue diagnosis ability using a tongue diagnosis e-learning/e-assessment system based on a standardized tongue image database. *Front. Med. Technol.* **5**, 1050909. https://doi.org/10.3389/fmedt.2023.1050909 (2023).
21. Zouridakis, G. et al. Melanoma and other skin lesion detection using smart handheld devices. *Methods Mol. Biology (Clifton N J).* **1256**, 459–496. https://doi.org/10.1007/978-1-4939-2172-0_30 (2015).
22. Phillips, M. et al. Assessment of accuracy of an artificial intelligence algorithm to detect melanoma in images of skin lesions. *JAMA Netw. open.* **2**, e1913436. https://doi.org/10.1001/jamanetworkopen.2019.13436 (2019).
23. Xie, J. et al. Digital tongue image analyses for health assessment. 1, 172–198, https://doi.org/10.1515/mr-2021-0018 (2021).
24. Hanifehpour, R., Motevalli, M., Ghanaati, H., Shahriari, M. & Aliyari Ghasabeh, M. Diagnostic accuracy of coronary calcium score less than 100 in excluding coronary artery disease. *Iran. J. Radiology: Q. J. Published Iran. Radiological Soc.* **13**, e16705. https://doi.org/10.5812/iranjradiol.16705 (2016).
25. Ronneberger, O., Fischer, P. & Brox, T. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015.* (eds Nassir Navab, Joachim Hornegger, William M. Wells, & Alejandro F. Frangi) 234–241 (Springer International Publishing).
26. Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F. & Adam, H. In *Computer Vision – ECCV 2018.* (eds Vittorio Ferrari, Martial Hebert, Cristian Sminchisescu, & Yair Weiss) 833–851 (Springer International Publishing).
27. Oquab, M. et al. DINOv2: Learning Robust Visual Features without Supervision. (2023).
28. Duan, M. et al. Feasibility of tongue image detection for coronary artery disease: based on deep learning. *Front. Cardiovasc. Med.* **11**, 1384977. https://doi.org/10.3389/fcvm.2024.1384977 (2024).
29. Hassoon, A. R., Al-Naji, A., Khalid, G. A. & Chahl, J. Tongue disease prediction based on machine learning algorithms. *Technologies* **12**, 97 (2024).
30. Järvinen, J., Mikkonen, J. J. W. & Kullaa, A. M. Fissured tongue: A sign of tongue edema? *Med. Hypotheses.* **82**, 709–712. https://doi.org/10.1016/j.mehy.2014.03.010 (2014).
31. Ren, Q. et al. A quantitative diagnostic method for phlegm and blood stasis syndrome in coronary heart disease using Tongue, Face, and pulse indexes: an exploratory pilot study. *J. Altern. Complement. Med.* **26**, 729–737. https://doi.org/10.1089/acm.2020.0008 (2020).
32. Maciocia, G. *Tongue Diagnosis in Chinese Medicine* Vol. 3, 76–77 (Eastland, 1995).
33. Udoy, I. A. & Hassan, O. AI-Driven technology in heart failure detection and diagnosis: A review of the advancement in personalized healthcare. *Symmetry* **17**, 469 (2025).
34. Martínez-Sellés, M. & Marina-Breysse, M. Current and future use of artificial intelligence in electrocardiography. *J. Cardiovasc. Dev. Disease.* **10** https://doi.org/10.3390/jcdd10040175 (2023).

Acknowledgements

We gratefully acknowledge Dr. Boroujerdi, the esteemed head of the Sepehr Angiography Department at Baharloo Hospital, and the dedicated staff of the department for their exceptional cooperation.We also extend our gratitude to Ms. Yasaman Hekmatardakani for her valuable assistance in data entry related to the patients' demographic, angiographic, and tongue examination information. This study forms part of the PhD thesis of the primary investigator at the Department of Iranian Medicine, Tehran University of Medical Sciences, Iran.

Author contributions

AHA was the principal investigator, designed the study, and wrote the first draft. MK and HH supervised the study, provided key information, and contributed to the writing process. HRR, AHK, and AR were the study advisors, who provided background information and helped with the methodology. HRR, AB, AY, AM, and RM contributed substantially to the design and drafting of the work. BMB made substantial contributions to the conception of the work and drafted the work. All authors have read and approved the final manuscript and agree to be accountable for all aspects of the work, such that they will answer questions related to the accuracy or integrity of any part of the work.

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval and consent to participate

The ethics committee of Tehran University of Medical Science (TUMS) approved the study. (Ethic code: IR.TUMS.MEDICINE.REC.1400.675).

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-28417-4.

Correspondence and requests for materials should be addressed to H.R.R., H.H. or M.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025