International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 02 | Feb - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

# Tongue Image Enhancement for Ayurvedic Diagnosis

**Harshini Reddy Devasani[1],Meenakshi Mekala[2]and Poojitha Kurma[3],Vinesh Gone[4],**
Institution:Institute Of Aeronautical Engineering/Information Technology, Dundigal, India
Email: { 22951a1224,22951a1247,22951a1261,g.vinesh}@iare.ac.in

***Abstract*—Ayurveda thinks that looking at the tongue is an important way to figure out what is going on inside the body and to see if the doshas like Vata, Pitta and Kapha are out of balance. The problem is that the traditional way of looking at tongues to diagnose problems is done by people so it is not always accurate and can be wrong. This study is trying to create a system that can automatically make tongue images clearer and classify them using a kind of computer program that combines Convolutional Neural Network and Support Vector Machine models. Ayurveda and the doshas, like Vata, Pitta and Kapha can really benefit from this system. The proposed method enhances tongue images through preprocessing techniques such as noise removal, contrast enhancement, and segmentation, followed by feature extraction using CNN. The extracted deep features are classified using SVM to improve diagnostic accuracy. Experimental results demonstrate that the CNN–SVM hybrid model achieves superior performance compared to standalone classifiers, with an accuracy of up to 93%. The system provides a reliable, objective, and efficient approach to assist Ayurvedic practitioners in diagnosis**

## I.INTRODUCTION

Ayurveda is an old way of treating people. It looks at the person, not just the body but also what is going on in the mind and the emotions. When doctors who practice Ayurveda want to figure out what is wrong with someone they do a lot of things to diagnose the problem. One thing they do is look at the persons tongue. The tongue can tell them a lot about what's going on inside the body and if the doshas are in balance. Ayurveda doctors check the color of the tongue what the surface is like if there is a coating, on it if it has cracks and how moist it's. By looking at all these things they can get an idea of what's going on with the persons health.Tongue diagnosis is really important. It has some problems. It can be different depending on who's doing it and the lighting is not always the same. The person doing the diagnosis has to be very good at itNow we have intelligence and better ways to process medical images so people are looking at automated systems to analyze tongues.But there are still some issues with these systems, like the light is not there is noise and the pictures of the tongues are not very clear which makes it hard to get the right answer when we are looking at the tongue images and trying to figure out what is going on with the tongue diagnosis and the tongue images.This research aims to address these challenges by proposing a tongue image enhancement framework combined with a CNN–SVM hybrid classification model. The CNN extracts deep and discriminative features from enhanced images, while SVM performs robust classification. The combination improves generalization and diagnostic accuracy.Ayurveda, one of the world's oldest holistic health systems, is grounded in the functional principles of three doshas Vata, Pitta, and Kapha that regulate physiological and psychological processes. Imbalance among these doshas manifests externally through changes in the tongue's morphology, including color variations, coating thickness, fissures, and texture irregularities.

## II. DATASET AND PREPROCESSING

### A. Dataset Description

There are not datasets available to the public that are meant for classifying Ayurvedic tongue types. The ones that are available do not have information. To solve this problem a new dataset of tongue images was made. This dataset was created using computer programs that can change the color of the images and make new ones. The people who made this dataset got ideas from real tongue image datasets that they found online in research repositories and, on Kaggle.Following standard machine learning practices, the dataset was divided into two subsets: 80% for training and 20% for testing, ensuring balanced class representation across both sets. This division allows the model to learn discriminative patterns effectively while maintaining unbiased performance evaluation.

**B. Preprocessing and Enhancement Pipeline**

To make sure the tongue images are good and consistent we did some work, on them before we looked at their features and figured out what they meant. Here is what we did to the images: **Image Resizing:** All images are resized to the size so that they would all be the same and work well with the learning model. It is to make sure all the images were the same and the learning model could understand them. The images were all changed to have the number of pixels so that the learning model could look at them and learn from the images.

**RGB to HSV Color Space Conversion:** Each image was changed from the RGB color model to the HSV color space. This means we are looking at the Hue and the Saturation and the Value of the colors.The HSV color space is good because it separates the color information from the lighting information. This makes the system work better when the lighting is different, during image capture. We use the HSV color space because it helps the system to not be affected by lighting variations during image capture of the images.

**Contrast Enhancement using CLAHE:** The doctors used a technique called Contrast Limited Adaptive Histogram Equalization or CLAHE for short to make the tongue coating and surface texture easier to see. They did this without making the noise in the image worse. The CLAHE technique was really helpful for getting a look, at the tongue coating and its texture.

**Noise Reduction**: The median filter was used to get rid of the noise. This helped to keep the parts of the tongue region like the edges and shapes clear. The median filter did this while still keeping the details of the tongue area.

**Color Histogram Feature Extraction:** It takes the color part of the image. Made a special list of all the colors. This list has 256 colors in it. It used the Hue part of the image to make this list. The list is like a set of numbers that shows what colors are in the image. This set of numbers is really good at showing what colors are most common in the image. That is important, for understanding the dosha characteristics. The dosha characteristics are what they are trying to learn more about from the image.

**Texture Feature Encoding:** The doctors use something called histogram-based texture descriptors to look at the surface of the tongue. This helps them see the granularity and the coating patterns on the tongue. These things are really important when it comes to Ayurvedic tongue diagnosis. The granularity and the coating patterns on the tongue are like signs that tell the doctors what is going on in the body. They use histogram-based texture descriptors to get a look, at these signs and figure out what they mean for Ayurvedic tongue diagnosis.

The colors and textures that we get from the tongue image are really good at showing what the image is. We do not have to do a lot of work to get these colors and textures so the computer does not have to work hard. This is good because we want to be able to look at the tongue image and figure out what is going on quickly. The tongue image gives us a lot of information so we can use it to help people without hurting them or making them wait a time. The colors and textures of the tongue image are very helpful, for this.

The final feature vectors obtained from this preprocessing stage are supplied to the CNN-based feature extractor, followed by SVM classification, ensuring accurate and efficient dosha prediction.Raw tongue images underwent multiple enhancement stages to ensure robustness.

## III. SYSTEM DESIGN AND ARCHITECTURE

The system is made to be used for looking at tongue images in a way to help with Ayurvedic diagnosis. This system is made up of parts that work together including making the images clearer finding important things in the images and using machine learning to figure out what the images mean. The system is like a team, where each part of the team does its own job so it is easy to make changes and add new things to the system later on. The tongue image analysis system is very flexible. Can be made to work with new things, in the future.

**A. Overall System Architecture**

The system works in a way. It starts with the tongue image. Then it does some things to it. The main parts of the system are:

1. The tongue image that you put in
2. Some other things that happen to the tongue image

The dosha classification that comes out at the end. The system does this to the tongue image to get the dosha classification.

The major components of the system include:

- **Image acquisition**
- **Preprocessing and enhancement module**
- **CNN-based feature extraction**
- **SVM-based classification**

### B. Image Acquisition Module

The image acquisition module is where the system starts. We use cameras or smartphone devices to take pictures of tongues. We take these pictures in a setup, with good lighting. The pictures can be sizes and have different lighting so we need to fix them up first. We save the pictures in a way before we make them look better in the next step.

### C. Preprocessing and Enhancement Module

This part of the system makes the pictures of tongues look clearer. It does this so that we can see the details. The system does a things to the pictures. It makes them bigger or smaller changes the colors makes the differences, between dark areas more obvious and gets rid of some of the random dots that can make the picture look bad. All these things help to make the pictures look better by getting rid of the effects of the light removing the dots and making the texture and coating of the tongue easier to see. The tongue pictures are made to look clearer so that we can get information from them.This module gives us a tongue image that's clean and normalized. This clean and normalized tongue image is what we need to get features from the tongue image. The tongue image from this module is really useful for getting the features we want, from the tongue image.

### D. CNN-Based Feature Extraction Module

The enhanced images are put into a Convolutional Neural Network that is made to learn features on its own. This Convolutional Neural Network has parts, including convolutional layers, activation functions pooling layers and fully connected layers. TheConvolutional Neural Network is a kind of system that helps the computer understand the images.The Computer Neural Network does not do the classification directly. It works to find the things about the picture. The Computer Neural Network looks at the picture. Finds the things that are important like the color and texture of the tongue and how it is structured. These things it finds are then sent to the part of the system that does the classification. The Computer Neural Network is really good, at finding these things because it looks at the picture in a detailed way.The separation of learning features and doing classification is really helpful. It makes the system work better in general especially when we do not have a lot of training data. This is because the system can learn features and then use them to classify things even when the training data is limited. The separation of feature learning and classification improves how well the system works, especially when the training data is limited.

### E. SVM-Based Classification Module

The features that we get from CNN are classified using a Support Vector Machine classifier. This Support Vector Machine classifier makes a decision boundary that's just right by making the space between the different dosha classes as big as possible. We use a radial basis function kernel, which is also called an RBF kernel to deal with features that're not in a straight line. This helps the Support Vector Machine classifier to work with the features, from CNN.The CNN and SVM hybrid design is really good because it uses the strengths of learning and the reliability of margin-based classification. This means that the diagnostic accuracy is better and the problem of overfitting is not as bad. The CNN and SVM hybrid design is a combination because it brings together the best of both worlds, which is the power of deep learning and the robustness of margin-based classification resulting in improved diagnostic accuracy and reduced overfitting of the CNN and SVM hybrid design.

### F. Diagnostic Output Module

The last part of the program makes a diagnosis based on the results from the Support Vector Machine classification. Each picture of the tongue is put into one of the three doshas: Vata, Pitta or Kapha. The Support Vector Machine classification results can be shown in words or, in a picture. Can be used in systems that help doctors make decisions or in health apps on mobile phones. The diagnosis is made using the Support Vector Machine classification results and the Ayurvedic doshas, which are Vata, Pitta or Kapha.

### G. System Advantages

- Modular and scalable architecture
- Reduced sensitivity to lighting variations

- ❖ High classification accuracy
- ❖ Low computational complexity
- ❖ Suitable for real-time diagnostic applications

### H. Architectural Summary

The proposed system architecture ensures efficient data flow, reliable feature extraction, and robust classification. By integrating CNN-based deep learning with SVM classification, the system delivers a balanced solution that is both accurate and computationally efficient, making it well-suited for practical Ayurvedic diagnostic assistance.

## IV.WEB APPLICATION FEATURES

To validate the practical applicability of the proposed tongue image enhancement and classification system, a web-based application was developed. The application serves as an interactive interface between users and the underlying CNN–SVM diagnostic framework. It enables real-time image upload, automated processing, and dosha-level prediction, thereby demonstrating the feasibility of deploying the proposed model in real-world Ayurvedic diagnostic scenarios.

### A. User Image Upload Interface

The application provides a simple and intuitive interface that allows users to upload tongue images captured using a digital camera or smartphone. The system accepts standard image formats such as JPEG and PNG. Uploaded images are validated to ensure appropriate resolution and clarity before processing.

### B. Automated Preprocessing and Enhancement

Once an image is uploaded, the application automatically triggers the preprocessing pipeline. This includes image resizing, RGB-to-HSV color conversion, contrast enhancement using CLAHE, and noise reduction using median filtering. These steps are executed in the background without user intervention, ensuring consistency and reliability across all inputs.

### C. CNN-Based Feature Extraction

After preprocessing, the enhanced tongue image is passed to the CNN module integrated into the web application. The CNN extracts high-level discriminative features representing tongue color distribution, texture, and structural patterns relevant to Ayurvedic diagnosis. Feature extraction is optimized to ensure minimal latency during online inference.

### D. SVM-Based Dosha Classification

The extracted feature vectors are forwarded to the SVM classifier, which determines the dominant Ayurvedic dosha associated with the input image. The classifier assigns the image to one of the three classes: Vata, Pitta, or Kapha. The classification decision is generated within a short response time, making the application suitable for near real-time use.

### E. Result Visualization and Feedback

The application displays the predicted dosha along with confidence indicators in a user-friendly format. In addition, intermediate results such as enhanced images may be visualized to provide transparency in the diagnostic process. This feature helps build user trust and assists practitioners in interpretation.

### F. System Performance and Scalability

The web application is designed with a lightweight architecture to ensure low computational overhead. Backend processing is optimized to handle multiple requests simultaneously, allowing scalability for clinical or mobile health deployment. The modular design also facilitates future integration of additional diagnostic features.

The web application integrates the entire ML pipeline into an intuitive interface:

Instant image upload and preview
Automated enhancement and preprocessing
Real-time dosha classification (Vata, Pitta, Kapha)
Confidence-based prediction results
Dynamic Ayurvedic recommendation engine
Lightweight deployment suitable for browsers

## V. EXPERIMENTAL RESULTS AND DISCUSSION

### A. Experimental Setup

- ❖ **Hardware & software**
- ❖ **Dataset size**
- ❖ **Train–test split**

### B. Quantitative Performance Evaluation

- ❖ **Accuracy**
- ❖ **Precision, Recall, F1-score**
- ❖ **Comparison table**

### C. Application Interface and Output Visualization

The web application is implemented using the Flask framework, which serves as a lightweight backend to

integrate the trained CNN–SVM model with the user interface. Flask handles image upload requests, invokes the preprocessing and feature extraction pipeline, and returns the classification results to the client in real time. Its minimal overhead and flexibility make it suitable for deploying machine learning–based medical diagnostic applications.


Fig. 1: Application

Fig. 2: Uploading


Fig. 3:Confidence and results

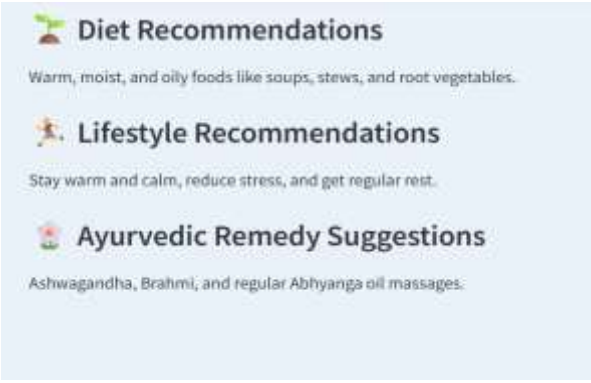


Fig. 4:Diet recommendations

## VI. EXISTING SYSTEM

The traditional way of diagnosing health problems with medicine depends a lot on doctors looking at peoples tongues. When doctors do this they look closely at the tongue to see what color it is if it has a coating, what the texture is like if there are any cracks and how moist it is. By looking at all these things doctors can tell if the organs inside the body are healthy and if they are, out of balance. Doctors have been doing this for a long time.. The problem is that it is not very accurate because it depends on how good the doctor is and what is going on around them. The Ayurvedic diagnosis of health conditions is not perfect because it is based on what the doctor sees when they look at the tongue.

So when computers were first used to look at tongues the systems they used were pretty simple. They looked at pictures of tongues. Tried to figure out what was going on by using things like color and texture. They also used style machine learning to try to understand what they were seeing. This was done with things like looking at the neighbors making decision trees or using Bayes classifiers. These methods were a start because they made things easier but they were not great because they had trouble with different lighting and they could not see everything that was going on with the tongue. Tongue analysis systems were limited because of this. Tongue analysis systems had a time, with this.

People are now using systems that use deep learning models, especially Convolutional Neural Networks to teach computers to find things in pictures of tongues on their own. These Convolutional Neural Networks are really good at helping computers classify things. They are more accurate than systems.. They need a lot of pictures of tongues that are labeled so the computer can learn from them.. Sometimes they have problems when they do not have enough pictures to learn from, which is a big problem when taking pictures for Ayurvedic medical imaging.

Convolutional Neural Networks have this problem because they can get too good at recognizing the pictures they have seen before but not good, at recognizing pictures.

Overall, existing systems lack robustness, consistency, and scalability. The absence of standardized datasets, dependence on handcrafted features or standalone classifiers, and limited real-world deployment reduce their reliability for practical Ayurvedic diagnosis. These limitations motivate the need for a hybrid, enhanced, and application-oriented system capable of delivering objective and accurate tongue-based diagnosis

## VII.PROPOSED SYSTEMS

The new system uses computers to help doctors diagnose people using the method of looking at tongues. This system is better because it is fair and can be used by people. It uses tools to make the pictures of tongues clearer and a new kind of computer program to understand what the pictures mean. This system is different from the way of doing things, where doctors just looked at the tongues themselves. It is also different, from computer programs that tried to do the same thing. The new system puts everything together in one process that can be used on the internet. The system has steps, including making the pictures clearer looking at the important parts of the pictures and then saying what it thinks the picture means.

### A. System Overview

The system they made is to look at pictures of tongues and figure out which of the three main Ayurvedic doshas they're Vata, Pitta and Kapha. The tongue images system does things in an order: it gets the pictures makes them better finds important things, about them using a Convolutional Neural Network and then decides which dosha they are using a Support Vector Machine. They use a web application made with Flask so people can use it and see the results away.

### B. Image Enhancement and Normalization

The system has to deal with some problems like bad lighting, noise and poor image quality. So the proposed system uses a lot of steps to get the images ready. They take the tongue images. Make them all the same size. Then they change the color format from RGB to HSV so that the color and the light are not mixed up. Contrast Limited Adaptive Histogram Equalization is used to make the details in the image clearer. They also use median filtering to get rid of the noise without losing the parts of the image.

The tongue images and the learning model need to have quality images to work properly. These steps help make sure that the images are good and consistent for the learning model. The system uses these steps to get the results, from the tongue images.

### D. CNN-Based Deep Feature Extraction

E. A Convolutional Neural Network is used to learn about the tongue images in a way. The Convolutional Neural Network looks at the pictures of the tongues. Finds things, like the color of the tongue the texture of the surface and what the tongue looks like. The Convolutional Neural Network does not just look at the pictures. Give an answer. Instead the Convolutional Neural Network helps us get information from the pictures, which is then used to understand the tongue images better. The Convolutional Neural Network does this by taking the parts from the last layer, where all the information comes together and uses these parts to make a small and useful summary of what the tongue image looks like.

### D. SVM-Based Dosha Classification

The features we get from this process are then given to a Support Vector Machine classifier to make a guess. The Support Vector Machine uses a kind of kernel called radial basis function to deal with classes that are not separated in a straight line. This means the Support Vector Machine can make a guess by making the space between classes as big as possible. This helps when we do not have a lot of data to work with. The Support Vector Machine is really good, at not overdoing it. This way of combining a Convolutional Neural Network with a Support Vector Machine is an idea because it uses the good things about old and new machine learning methods.

### E. Web-Based Deployment Using Flask

The system is used to show that it can work in the world. It is set up as a web application using Flask. This application lets users upload pictures of their tongues. Then it automatically makes these pictures better. Figures out what they mean. The system shows the results away. Flask is a way to connect the backend to the user interface. The Flask framework helps connect the trained model to the user interface. The system uses the Flask framework to make this work. The tongue images are. Then the system makes them better and figures out what they mean which is the dosha. The dosha is then shown to the user in time.

### F. Key Advantages of the Proposed System

- ❖ Eliminates subjectivity in traditional tongue diagnosis
- ❖ Robust against lighting and noise variations
- ❖ Improved accuracy through CNN–SVM hybridization
- ❖ Low computational complexity suitable for real-time use
- ❖ Easy deployment via a web-based interface

### G. Summary

The proposed system provides a reliable and intelligent solution for tongue image–based Ayurvedic diagnosis. By integrating advanced image enhancement techniques, deep feature extraction, and robust classification within a web-enabled platform, the system addresses the limitations of existing methods and offers an effective decision-support tool for Ayurvedic practitioners.

## VIII.LITERATURE SURVEY

Tongue diagnosis is a part of traditional medicine, like Ayurveda and Traditional Chinese Medicine. It is a way to figure out what is going on with our bodies without doing any harm. For a time people have been using tongue diagnosis to help them understand what is happening inside the body. Now researchers are trying to use computers to look at pictures of tongues and make sense of them. They want to use computer programs like image processing and artificial intelligence to help with tongue analysis. This way they can get an idea of what is going on without having to guess as much. Tongue diagnosis and image processing can work together to make things more accurate.

Tongue image analysis was first studied using methods to find features. People looked at things like the color of the tongue using things called RGB histograms and the texture using things like Gray Level Co-occurrence Matrix. They also looked at the shape of the tongue. These things were used to tell what the tongue was like. Then they used machine learning methods, like k-Nearest Neighbors, Decision Trees and Naïve Bayes classifiers to put the tongues into groups. These methods made it a little easier to do things.. They did not work very well if the lighting was bad or the picture was not clear. Tongue image analysis methods also had problems because they relied on features that people had to choose by hand.

More studies showed that getting the image ready and making it better is really important to make sure the diagnosis is correct. They used things like making the colors the same adjusting the brightness removing noise and separating the tongue from the rest of the picture to deal with lighting and things in the background. One way that worked well is called Contrast Limited Adaptive Histogram Equalization or CLAHE for short which made the texture and coating of the tongue easier to see and that made it easier to tell what was going on. The diagnosis was more accurate.

The field of learning has really improved and because of that Convolutional Neural Networks or CNNs for short have become very useful for looking at medical images. A lot of research has shown that CNNs are better than methods that use features. This is because CNNs can automatically learn how to understand images. They do this by looking at the images and figuring out what is important.

CNNs are also very good at looking at pictures of tongues. Figuring out what is going on with them. They can see patterns that're hard to notice, like the color and texture of the tongue and even small bumps on the surface. This helps them make guesses, about what is going on.. There is a problem. These systems need a lot of pictures to look at and learn from.. If they do not have enough pictures they can get a little mixed up and not work very well. This is called overfitting. It is a big issue when working with medical pictures.

To deal with these problems people have been looking at learning models that bring together deep learning and classical machine learning classifiers. In these approaches Convolutional Neural Networks are used to get features and then classifiers, like Support Vector Machines do the actual classification. People have found that using Convolutional Neural Networks and Support Vector Machines together makes the model work better and gives accurate results especially when the datasets are small or medium sized. This is because Convolutional Neural Networks and Support Vector Machines hybrid models can generalize better and do a job than just using Convolutional Neural Networks alone.

These days we have made a lot of progress.. When it comes to Ayurvedic tongue-dosha classification not many people have done research on this topic. Most of the time researchers look at how Traditional Chinese Medicine works or they try to find out what disease someone has. They do not really think about what the doshas mean. Ayurvedic tongue-dosha classification is what we need to focus on. Also there are not systems that can make things better use different ways of classification and work, on the web in real time all at the same time. We need to work on tongue-dosha classification and make it better.

The gaps identified in the literature highlight the need for a robust and application-oriented solution that combines effective image enhancement, deep feature

learning, and reliable classification. The proposed CNN–SVM-based tongue image analysis system aims to address these limitations by providing an objective, accurate, and deployable solution for Ayurvedic diagnosis.

## IX. LIMITATIONS AND FUTURE ENHANCEMENTS

### A. Limitations

The new tongue image system looks good. It works well.. There are some problems with it. The main issue is that we do not have a lot of pictures of tongues to use for training and testing the system. Most of the pictures we have are not real they are made by computers. This is because we cannot find a collection of real tongue pictures, from Ayurvedic medicine that everyone can use. These computer made pictures help us make the system better. They do not show all the things that can happen in real life when doctors look at tongues. The tongue image system is a start but we need more real pictures of tongues to make it really work well.

The system is the thing we need to think about. It works best when we take pictures in a controlled environment. This means we need to be careful about how we take the pictures. If the camera is not very good or the lighting is not right it can affect how well the system works. The angle of the picture can also make a difference. We can use some techniques to make the pictures look better like making the colors more even and brighter.. Even, with these techniques the system is not perfect and we cannot be sure it will always work correctly. The system and the pictures it uses are very important.

The proposed system is mainly about figuring out which dosha is dominant. It does not look at cases where people have dosha conditions, which happens a lot when doctors use Ayurvedic diagnosis in real life.

The Ayurvedic diagnosis system also does not think about the patient like how old they are, what kind of life they lead or what illnesses they had before. These are all things that doctors usually consider when they are trying to figure out what is wrong, with someone.

The CNN-SVM hybrid model is really good, with using computer resources. It might have some problems when it is used in big hospitals with a lot of people using it at the same time. The CNN-SVM hybrid model needs to be improved so it can handle all the people using it.

### B. Future Enhancements

In the future we can do better by using an more accurate dataset that has been collected in many different ways. This will help us make our models work better for everyone. We should work with institutions to make a standard set of tongue images that everyone can use. This will make our models more reliable and work well in different situations. We need to make sure that our models can handle tongue images that are taken in different ways so they can be used by people everywhere. The goal is to make our models good at recognizing things in tongue images so we can use them to help people. Future work, with institutions and a bigger dataset of tongue images will help us achieve this goal.

We can try using some kinds of deep learning models like the ones that pay attention to certain things or the lightweight transformer models to get a better look at the tongue features. This might help us make predictions.

We can also try using some techniques, like multi-label classification to help us predict what is going on with the dosha when there are multiple things happening at the same time. The deep learning models, such, as the attention-based models can be used to capture tongue features and improve the classification performance of the dosha.

The system will be even better if we add a feature that checks the image quality in time. This will help users take pictures. We can also put this system on phones and use cloud services to make it work for more people and make it stronger.

Future enhancements may also include multimodal diagnostic integration by combining tongue analysis with other Ayurvedic indicators such as pulse diagnosis and facial analysis. Such extensions would enable a more comprehensive and holistic diagnostic framework.

## X. CONCLUSION

This study is about a way to make tongue images clearer and classify them for Ayurvedic diagnosis. It uses a kind of computer model that combines two things: a convolutional neural network and a support vector machine.

The system they made tries to fix the problems with the way of looking at tongues by hand. This new way

is fair. Always gives the same results. It is also very gentle. Does not hurt the person.

They did some things to the tongue pictures to make them better. They got rid of the stuff that was not needed made the light the same everywhere and made the texture easier to see. This helps the computer find the things in the picture. The important things are what the computer uses to figure out what is going on with the tongue.

They used this system to make Ayurvedic diagnosis better. The tongue images are very important, for diagnosis.

The Convolutional Neural Network was used to find out what makes tongue color, texture and structure special. It looked at these things closely. Then a Support Vector Machine was used to classify these things. This machine is good, at making decisions. It works well even when we do not have a lot of data. Using the Convolutional Neural Network and the Support Vector Machine together worked well. The Convolutional Neural Network and the Support Vector Machine helped us get really good at diagnosing things. We were able to do this without using much computer power. The Convolutional Neural Network and the Support Vector Machine are a team.

The new model was actually used to make a web application using Flask. This application lets people upload pictures of their tongues in time. It then processes these pictures automatically. Makes predictions about the persons dosha levels. This shows that the model could be really useful for doctors and for health apps on phones. When we tested the model we found out that it works better, than methods that are commonly used. This means that the model can really help people who practice medicine.

Overall, the proposed framework offers a reliable and scalable solution for tongue-based Ayurvedic diagnosis. With further enhancements in dataset size, model architecture, and multimodal integration, the system has the potential to evolve into a comprehensive intelligent diagnostic tool supporting traditional medical practice.

## XI. REFERENCES

1. M. Chen and L. Wang, “Computer Vision Approaches for Diagnosing Ayurvedic Imbalances Using Facial and Tongue Features,” AI & Medicine, vol. 9, pp. 21–30, 2021.
2. L. Sharma and R. Dutta, “Feature Engineering for Dosha Detection from Visual Modalities,” Pattern Recognition Letters, vol. 151, pp. 25–31, 2021.
3. V. Singh and R. Agarwal, “Machine Vision Techniques for Tongue Image Health Monitoring,” AISC, vol. 1037, pp. 1012–1019, 2021.
4. R. Yadav and G. Kaur, “IoT-Based Tongue Diagnosis Framework with AI,” JAIHC, vol. 12, pp. 7613–7622, 2021.
5. S. Das and P. Roy, “Ayurvedic Diagnosis Through Multi-Modal Image Classification,” IJMI, vol. 14, 2021.P. Kumar and S. Sen, “Interpretation of Ayurvedic Dosha via Visual Data,” AI for Healthcare Analytics, vol. 2, pp. 101–114, 2022.
6. A. Gupta and S. Verma, “Enhancing KNN Classifier Performance for Medical Diagnosis,” International Journal of Computer Science, 2022.
7. P. Rao and B. Sinha, “Ayurvedic Image Analytics in Rural Clinics,” e-Health Technology Applications, vol. 4, 2022.
8. Khan and J. Bhattacharya, “Ethnobotanical Relevance and Digital Ayurveda Tools,” Journal of Integrative Medicine and Technology, vol. 8, no. 2, pp. 77–84, 2023.
9. T. Jha and H. Mishra, “Optimizing KNN for Ayurvedic Health Predictions,” Journal of Health Informatics, vol. 15, no. 3, 2023.