

Article

Not peer-reviewed version

Convolution Neural Network Based Multi- Label Disease Detection using Tongue Images

[Vibha Bhatnagar](#) * and Prashant P Bansod

Posted Date: 8 April 2024

doi: 10.20944/preprints202404.0565.v1

Keywords: Key Words: Tongue Feature Extraction, Disease Diagnosis, segmentation, MobileNet V2 architecture, DenseNet 121 architecture, Convolutional Neural Network



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Convolution Neural Network Based Multi-Label Disease Detection using Tongue Images

Vibha Bhatnagar *, Prashant. P. Bansod

Department of Biomedical Engineering, Shri. G.S. Institute of Technology & Science. Indore 452003, India; ppbansod43@gmail.com

* Correspondence: bhatnagarvibha09@gmail

Abstract: Purpose: Tongue image analysis for disease diagnosis is an ancient traditional non-invasive diagnosis technique widely used by traditional medicine practitioners. Deep Learning based multi-label disease detection models offer tremendous potential to clinical decision support systems, by facilitating preliminary diagnosis. Methods: In this work, we propose a multi-label disease detection pipeline, where in observation and analysis of tongue images captured and received via smartphones assist in predicting the health status of an individual. All images are voluntarily given by subjects consulting collaborating physicians. Images thus acquired are first and foremost classified either into a diseased or a normal category by a 5-fold cross-validation algorithm using a convolution neural network (MobileNetV2) model for binary classification. Once the diseased label is predicted, the image is used to diagnose multiple diseases using the prediction algorithm based on DenseNet121. Results: Average accuracy of 93 % was achieved in classifying diseased from normal healthy tongue by detection model with MobileNetV2 architecture. Multilabel disease classification produced more than 90% accurate results for the seven class labels considered. Conclusion: AI based image analysis shows promising results and an extensive dataset could provide further improvements to this approach. Rather than employing high-cost sophisticated image capturing setup, experimenting with smartphone images opens opportunity to provide preliminary health status to individuals on smartphone prior to further line of treatment and diagnosis.

Keywords: Tongue Feature Extraction; Disease Diagnosis; segmentation; MobileNet V2 architecture; DenseNet 121 architecture; Convolutional Neural Network

1. Introduction

Technological breakthrough has made clinical investigations and test for disease diagnosis feasible to great extent. Despite of available high end diagnostic tools some common ailments can be diagnosed by observation of certain human indices such as pulse (nadi), eyes, nails and tongue as per ayurveda. Ancient traditional medicine made these parameters as means of diagnosis with no advanced technology tools at hand to predict many abnormalities related to the health status of internal body organs. The paper presents a novel method for multilabel disease classification from tongue analysis.

The use of decision support systems has been increasing over the past decades, modern deep learning models allow reliable classification and object detection of medical images showing remarkable accuracy comparable to that of physicians. An automated tongue diagnosis system is a means to bridge the gap between traditional method of diagnosis and modern western medicine. Subjective nature of the same related to the knowledge base and thorough practice of an expert traditional practitioner can be vanquished using quantitative analysis to make it accountable and acceptable. In this paper we have developed a multi disease classification algorithm which is able to predict multiple disease diagnosis with appreciable accuracy. Eight diseased states are considered based on the collected data samples. The Contributions of this paper are as follows:

- This paper presents a MobileNetV2 deep learning model to detect the disease tongue from a normal healthy tongue. Five-fold cross validation and transfer learning technique is used for this binary classification. Main achievement is sanguine results achieved with images taken by smartphone cameras rather than by standard equipment.
- Multilabel classification model is proposed for eight common categories of ailments. DenseNet 121 architecture is utilized for disease classification, satisfactory results achieved with the small dataset. Eight disease labels are diabetes, hypertension, acidic peptic disease, pyrexia, hepatitis, cold cough, gastritis and others.

Organization of this paper is as follows: ‘Related Work’ summarizes the tongue feature extraction techniques used. ‘Methodology’ section elaborates upon flow of the related work, dataset and training details along with experimentation and evaluation Metrics used, and finally performance analysis is reported in ‘Results and Discussion’ section. Conclusion and Future Work’ section finally concludes are research work.

2. Related Work

Quantifying the tongue diagnostic attributes is one of the main challenges in automation of tongue analysis for disease diagnosis. Study of tongue conditions based on ayurveda are presented by [1] related to health status of an individual. Summary different tongue attributes like colour coating, texture and geometric shape are analysed for predicting any particular diseased condition by [2] as practiced in oriental medicine practices. Automation of tongue analysis system essentially requires an image capturing device with high resolution images for accurate extraction of tongue features for disease predictions in agreement with the traditional medicine practitioners. Different imaging setups are explored using high end CCD cameras [3–9], Hyperspectral cameras [10–16] to smartphone cameras [17–24]. Tongue region needs to be segmented from the images captured which consist of teeth lips skin area before feature extraction and classification. Conventional approaches as well as ANN and deep learning algorithms have been explored over a period of time for tongue area segmentation and feature extraction. Most of the work done in this area targeted a particular disease and tongue features related to it, probably due to lack of digital dataset pertaining to all features possible for various diseases. Some common diseases such as diabetes, appendicitis, gastritis was targeted and relevant features and classification was done using statistical techniques [25–28]. Hybrid model with statistical methods for feature extraction and machine learning algorithms for classification were also developed [29–33].

Table 1. Summary of literature review of ML based models for specific tongue feature.

Tongue Features/ Disease targeted	Dataset	Method employed / Disease investigated	Reference
5 tongue body colour, 6 tongue coating colour	1080 subjects	k-means clustering algorithm applied to the images acquired using DSO1 state of the art acquisition system	[34]
Diabetes	732 subjects	TFDA -1 used to capture images and extract, texture, coating features along with an Auto - Encoder algorithm to extract tongue features then fusion of the two set of features done using k-means algorithm for classification	[35]
Tongue area detection calibration and	50 subjects	Tongue detection using faster -RCNN, Feature extraction models ResNet -50, VGG-16 and	[36]

constitution classification		Inception-V3, alongside LBP for texture features and Colour-Moment for colour feature. Model evaluated using classifiers SVM and Decision tree.	
Colour & Texture features	702 images,	Gray Level co-occurrence Matrix (GLCM), along with LEAD (Multilabel Learning Algorithm) with threshold determining algorithm for improved results over other existing techniques	[37]
Multifeature extraction	268 images,	GLA (Generalized Lloyd Algorithm) to extract colour and texture features from the tongue surface.	[38]
Seven categories fissured tongue, tooth-marked, statis, spotted, greasy, peeled and rotten coating	8676 images	faster R- CNN a region-based network achieved an accuracy of 90.67%	[39]
11 features on the tongue surface considered	482 images	ResNet-34 architecture, 86% accuracy for 11 features identified	[40]
Tooth marked tongue related to spleen defeciency	1548 images	ResNet-34 architecture, 90% accuracy	[41]
Gastritis	263 gastritis patients, 48 healthy.	features related to gastritis were extracted using constrained high dispersal neural network Ada Boost, SVM (support Vector Machine), MLP (Multilayer Perceptron Classifier)-	[42] [43]
11 disease categories plus healthy tongue images	936 images, 78 images for each of the 12 disease categories including healthy	extracted tongue features by VGG 19 network supported by Random Forest classifier achieved 93.7% accuracy.	[44]
12 disease categories including healthy	936 images, 78 images for each of the 12 disease categories	Designed IoT base Automated synergic deep learning tongue colour image analysis model giving 98.3% accuracy for disease diagnosis and classification.	[45]
Iron deficiency	95 images from Harvard dataset	Explored the possibility of monitoring health status by tongue images using CNN algorithm which could be deployed on android mobile app.	[46]

[47] Reviewed current trends in tongue diagnosis. They also trained classification model Random Forest and Support vector machine on tongue dataset with tongue region divided into 5 parts as per internal organs layout, and extracted seven colour spaces from the five extracted parts. Further they trained VGG and ResNet pretrained models on tongue constitution classification. ResNet 50 showed best performance with 64.52 % accuracy. Tremendous efforts have been put in to effectively utilize the full potential of the tongue diagnosis system for over two decades. Enhancement in tools and techniques for achieving the wholistic approach and to overcome the subjective nature of diagnosis is an ongoing process a need to set some standard protocols for it to be readily acceptable by the end users. Majority of the research has been targeted towards a particular disease and worked towards extracting tongue features for the same. Aim of this paper is a step towards developing a decision support system with use of deep architecture to predict a wholistic tongue diagnosis, not just confined to any particular tongue features or diseases.

3. Methodology

The implemented tongue image classification pipeline illustrated in Figure 1, after image acquisition preprocessing and tongue area segmentation, steps involved for classification are summarized as follows:

- Stratified 5-fold cross validation for diseased risk classification.
- Ensemble learning strategy: bagging, to reduce variance in data.
- Up sampling of dataset to set some minimum sample size in each class.
- Extensive real time data augmentation for training models.
- Class weighted focal loss, to tackle class imbalance.
- Individual training for multi disease classification and disease risk detection.

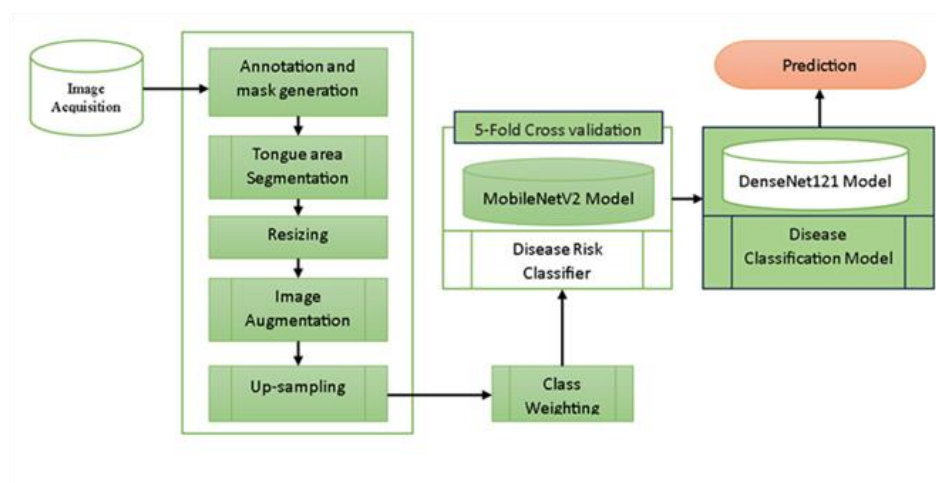


Figure.1. Proposed pipeline for multi-label disease detection using tongue images

3.1. Tongue Analysis Dataset

Total 1095 images of subjects suffering from one or more than one ailment is acquired using smartphone cameras with image resolution greater than equal to 8 mega pixels (Samsung A50, iPhone, one plus). Total 822 images for healthy individuals are also collected through willing individuals. All images are collected after necessary consent from individuals eager to be part of our study. Raw images are of different resolution and various sizes. For the proposed model images are annotated with eight conditions other than normal and disease risk categories as listed in Table.2. Label class others includes some uncommon ones like CAD (coronary artery disease), CKD (chronic kidney disease), COPD (Chronic Obstructive Pulmonary Disease), Epilepsy, Vertigo.

Table 2. Annotation Frequency for each class in the dataset.

Disease	Samples	Disease	Samples
Diabetes (DM)	112	Hepatitis	183
Blood Pressure (BP)	138	Cold Cough	150
Acid Peptic Disease (APD)	156	Gastritis	189
Pyrexia	98	Others	429

3.2. Preprocessing and Image Augmentation

Images captured are of different sizes and formats primary processing is done by converting all images to jpeg of 256 x 256 pixels. Next step involves segmentation of the tongue area of interest by Double U-Net architecture [48]. In order to increase data variability further preprocessing methods used such as image augmentation for up sampling to balance class distribution and real time augmentation during training to obtain unique and novel images in each epoch thus improving the model’s performance. Up sampling is done to ensure each label occurred at least 100 times in the dataset which increased the total diseased tongue images to 2729. Rotation, flipping and altering brightness saturation and hue is used for real time augmentation.

3.3. Deep Learning Models

Our Pipeline combines two different types of image classification methods, a binary classification for normal/diseased tongues and a disease label classifier for multilabel annotated images. We have used the AUCMEDI [49] platform to develop our pipeline in both cases we are using pretrained models to reduce time and cost of training a fresh model. Transfer learning is applied with frozen layers except the classification head. After 10 epochs of training the freezing is undone, so that the weights can be adapted to the new task.

The two architectures used are chosen to be compatible to the low resource requirement and possible to be deployed on android platform.

Diseased Tongue Detector

MobileNetV2 is used for binary classification to categorize between normal and unhealthy tongue images.. We used 5-fold cross validation with bagging method for ensemble with Random Forest and Soft Majority Voting. Dataset with two classes is split into three parts as train, test model and test ensemble set. Ensemble dataset is used for bagging method by Random Forest classifier. The train set is further divided into train and validation set for 5-fold cross validation wherein, training-set is split into 5 parts and training is performed 5 times. Each time one of the five parts, a different one, is used for validation and the other four parts are used for training. Due to imbalance in images in the two categories class weights is computed and categorical focal loss is considered as the loss function. Testing of model performance done on the test data and evaluation metrics for training each fold are computed. Ensemble by 5-fold bagging is done by Random Forest and Soft Majority vote methods. Each model is trained for 50 epochs and dynamic learning rate set to maximum decrease to 1e-7 with factor of 0.1 in case of no improvement is monitored for validation loss, with patience of 5 epochs. Early stopping and Model checkpoint method is also used with patience level set as 12 epochs with no improvements in validation loss. Finally, all the individual folds F1 Scores along with ensemble models is compared to indicate the best performance of the models.

Disease Label Classifier

This is a multilabel classifier using Densenet121 architecture for the eight disease classes. Multilabel focal loss is used with categorical accuracy. Model is trained for 100 epochs with dynamic learning rate set to maximum decrease to 1e-7 with factor of 0.1 in case of no improvement is

monitored for validation loss, with patience of 5 epochs. Early stopping and Model checkpoint method is also used stopping after 12 epochs with no improvements monitored.

The experiments were performed with hp Pavilion laptop with a 1.60 GHz intel i5 8th generation processor and 8 GB of Ram. Training of the model was done on Google Colab Python 3 Google Compute Engine backend GPU.

3.4. Evaluation Metrics

Metrics considered for quantitative evaluation of the model on test data are Precision, Recall and F1 – Score. *Precision* is number of true positives divided by the number of total positive predictions

$$Precision = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Positive\ (FP)}$$

Recall measures the model's ability to predict the positives, it is true positive divided by the true positive and false negative.

$$Recall = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Negative\ (FN)}$$

F1 score is the harmonic mean of Precision and recall given by

$$F1Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Categorical Accuracy calculates the percentage of predicted values that match the actual truth values. Calculation for the same is done as follows:

- First identify the index at which the maximum value occurs using `argmax()`.
- If it is the same for both predicted and true value, it is considered accurate.

Here since maximum value index is observed predicted value can be logit or probability function.

$$Categorical\ accuracy = \frac{accurately\ predicted\ records}{total\ number\ of\ records}$$

Receiver Operating Characteristics Curve

Receiver Operating characteristic (ROC) curve based on the calculation of average ROC for the 5-folds. Practically for each fold different False- and True-Positive-Rates exists and mean cannot be calculated hence, first aggregation of all False-Positive-Rates (FPR) into one vector is done serving as the x axis of average ROC curve. The True-Positive-Rate (TPR) are to be interpolated. The TPR is the corresponding y-points to the x-points that are already collected.

4. Results and Discussions

Disease risk detection model is trained on the dataset with healthy and unhealthy tongue images. Up sampling of the unhealthy image data is done to ensure at least 150 images in each of the 8 categories. Real time Image-augmentation used to increase the image-set artificially by adding small transformations to the original images such as rotations or changes of the contrast or saturation. With online image-augmentation the transformations are applied to each image when loaded with the data generator and need not be saved to disk. For multilabel disease classification evaluation one hot encoded file serves as interface defining true labels for the model.

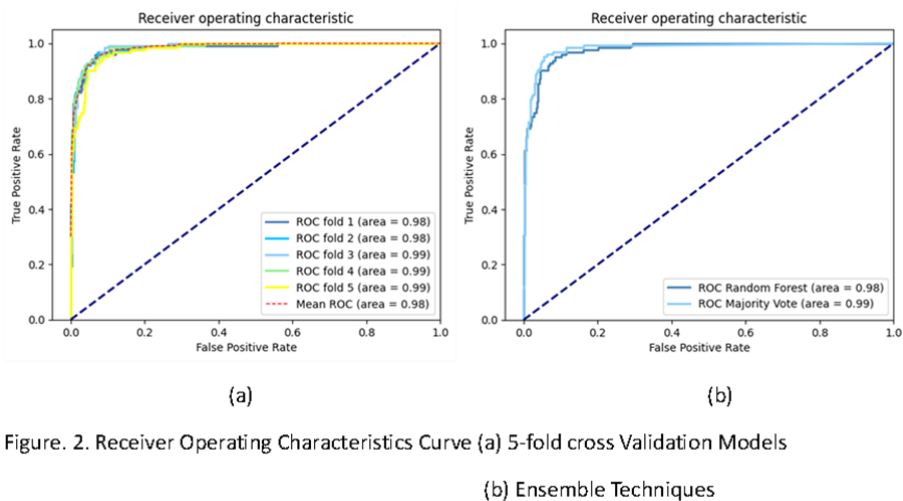
4.1. Performance Analysis of Disease Risk Detector

The sequential training of the disease risk detector with 5-fold cross validation with MobileNet V2 architecture approximately took 4.5 hrs with 45 epochs for each fold on an average, when trained

on shared GPU from COLAB. Disease classification model with DenseNet121 architecture took over 18 hrs to train. Performance results of the Disease Risk model w.r.t metrics considered on test model dataset of 485 images in all consisting of 361 unhealthy tongue images and 124 normal healthy tongue images are as shown in Table 3. Concise ROC Curve Figure 2, for the 5-folds along with ensemble techniques with mean value indicate appreciable performance of the model for binary classification.

Table 3. Performance Metrics for the 5 folds of cross validation of Disease Risk Model.

		Precision	Recall	F1-Score	Accuracy
Fold -1	Diseased	0.97	0.96	0.96	0.95
	Normal	0.88	0.91	0.90	
Fold -2	Diseased	0.99	0.91	0.95	0.92
	Normal	0.78	0.97	0.87	
Fold -3	Diseased	0.99	0.92	0.95	0.92
	Normal	0.81	0.96	0.88	
Fold -4	Diseased	0.97	0.96	0.96	0.95
	Normal	0.88	0.92	0.90	
Fold -5	Diseased	0.97	0.93	0.95	0.93
	Normal	0.82	0.92	0.97	



Best model is saved for each fold of cross-validation. Predictions of each fold on a test ensemble dataset with 485 images are combined to form an ensemble. Implementation of the Random Forest ensemble done by picking random samples and features to build many decision trees on bootstrapped dataset by a repeated process. Prediction of new data is accomplished by bagging operation wherein final decision is the most common decision amongst all decision trees.

Aggregation done by Soft Majority Voting ensemble to create a new predictions matrix from the generated predictions. For each sample the category that has the max. value for this (ensemble) predictions-matrix is taken as the final prediction. Ensemble of the 5-fold predictions in this particular case with two classes does not show enhanced performance, F1 Scores for Random Forest and Soft Majority Voting are comparable to the individual folds F1 score as depicted in Figure 3.

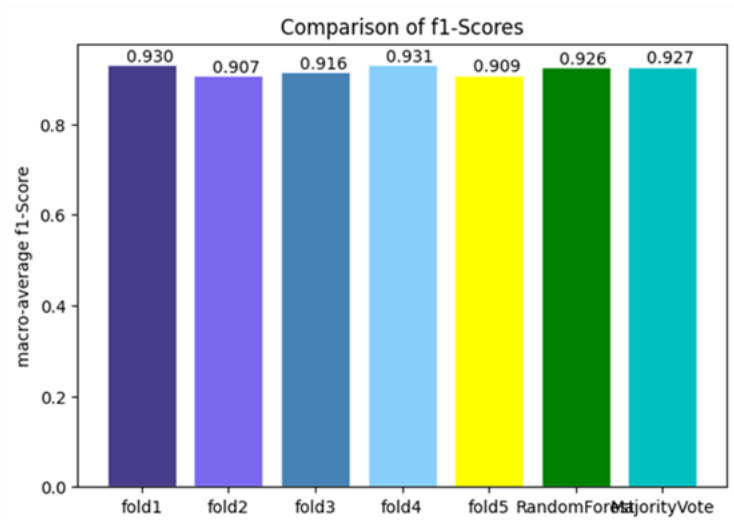


Figure. 3. Comparison Bar graph for five folds and ensemble techniques

4.1. Performance Analysis of Multi-Label Disease Classification Model

DenseNet 121 model with four dense blocks with [6,12,16,24] number of layers in each block is utilized for classification of diseases. Dataset consisted of 7 specific disease labels and one with others label under which all diseases which had less than 10 subjects falling in that particular category are placed. Model training stopped at 41 epochs. The performance parameters for the eight labels considered on a test data of 474 images is as in Table 4. ROC Curve in Figure 4 gives an idea of true positive predictions for each category.

Table 4. Performance Parameters for Disease Classification Model.

Disease	Precision	F1-Score	Accuracy
DM	0.9722	0.8203	0.9148
BP	0.9803	0.8658	0.9425
APD	0.9130	0.8038	0.9240
Pyrexia	0.9473	0.9183	0.9703
Hepatitis	0.9885	0.8958	0.9629
Cold Cough	0.9878	0.8901	0.9629
Gastritis	0.9798	0.9652	0.9870
Others	0.9034	0.7553	0.8093

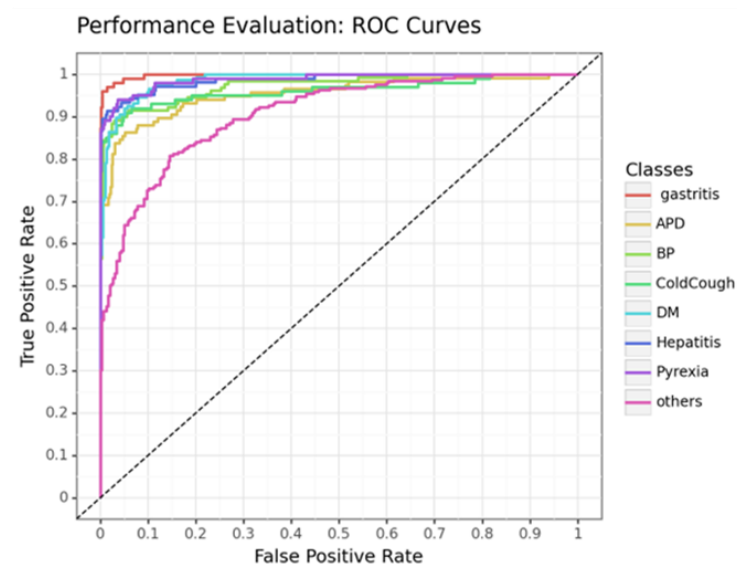
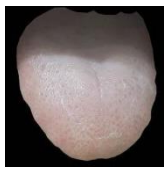
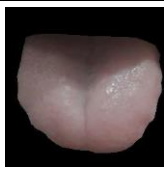
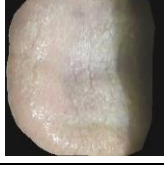


Figure .4. ROC curve of DenseNet 121 model on test dataset

It is observed that an average accuracy over 90% is achieved for each of the diseased class label. Some Sample results presented in Table 5 show the success of the DenseNet 121 model for disease classification. Highlighted rows in the table indicate inaccurate classification for 2 sample images. Class ‘others’ showed slightly lesser accurate results as also evident from ROC Curve and Performance indices.

Table 5. Sample Test Images with prediction probability and truth data.

Test Image Samples	Prediction Probability								Truth
	DM	BP	APD	PYR	HEP	CC	GAS	others	
	0.0456	0.0190	0.0400	0.0131	0.7392	0.0207	0.0256	0.7650	Hepatitis & others
	0.9177	0.0405	0.0055	0.0001	0.0033	0.9872	0.0071	0.9968	Diabetes cold cough, others
	0.0129	0.0198	0.0118	0.4670	0.0729	0.1006	0.0051	0.3405	Pyrexia
	0.4555	0.6903	0.0591	0.0014	0.8166	0.0654	0.0259	0.3004	Diabetes hypertension, Hepatitis

	0.0059	0.0235	0.7697	0.0071	0.0074	0.0315	0.9693	0.8268	APD, gastritis Others
	0.0004	0.0103	0.9436	0.0413	0.0029	0.0142	0.9705	0.7960	APD, gastritis Others
	0.0335	0.0188	0.0108	0.0987	0.4677	0.0124	0.7906	0.0795	Hepatitis, Gastritis
	0.1365	0.0628	0.2372	0.0261	0.0131	0.2798	0.0423	0.5847	Pyrexia
	0.0264	0.0036	0.1321	0.1261	0.0083	0.4880	0.1867	0.3476	Cold cough
	0.9768	0.8526	0.0079	0.0013	0.0009	0.9293	0.0010	0.9519	DM, hyper tension, Cold cough, others

Proposed model for Multi Disease Classification shows satisfactory results. The complete pipeline for disease risk detection and classification of multiple diseases can further be improved upon by adding more classification labels and using multiple models’ ensemble to achieve highest accuracy for all class labels.

Limitation of this work is small dataset with less variety of disease classes. Since only a single dense model is considered its performance accuracy for all class labels could not be uniform, this can be enhanced by incorporating ensemble learning method.

5. Conclusion and Future Work

Main challenging aspect of automating tongue analysis for disease diagnosis is quantifying the diagnostic features from the tongue image at par with the expert practitioner’s observations. In this study we introduced a powerful multi disease detection pipeline for tongue image analysis, which exploits ensemble learning to combine the predictions of 5 individually trained models for better performance. It utilizes strategies such as transfer learning, class weighting as imbalance in diff classes, extensive real time data augmentation, focal loss utilization and used two different deep learning models to predict the diseased conditions with multiple class labels. The most significant accomplishment from this study is that we could ascertain that mobile phone images can be successfully used for tongue image analysis, a step towards reducing the cost and expertise of a high-end sophisticated image capturing device. This opens opportunity to explore more enhanced models with a larger dataset for performance improvement. In our future work we aim to build more robust ensemble model capable to classify all possible disease label classification feasible using tongue analysis, this inherently includes within it need to compile an exhaustive tongue image dataset for research.

Author Contributions: I Mrs. Vibha Bhatnagar a Ph. D. scholar has compiled the work after literature survey and developing algorithm which is then trained on the dataset and rigorously analyzed by Prof. P. P. Bansod my guide and mentor. Rough draft prepared by first author was revised and corrected critically by Prof. Bansod for intellectual content and final approval for submitted version.

Funding: This research received no external funding.

Institutional Review Board Statement: This is an observational study and hence no ethical approval is required, only voluntary agreement of all participants is essential. Individual consent was taken from all subjects and recorded with clinicians' diagnosis. Data collected only used for research work only without disclosing identity of any subject.

Informed Consent Statement: Informed consent was obtained from all individual participants included in the study is taken. They were clearly informed that the images given by them would be only used for authors research work and will not disclose any personal information related to the individual on any platform whatsoever.

Data Availability Statement: All the images employed for the experiment are collected from clinics of doctors who willingly agreed to contribute for the purpose by Mrs. Vibha Bhatnagar and team of undergraduate students.

Acknowledgments: The authors acknowledge the support of Medical professionals who contributed in enabling collection of dataset and annotation and labelling of the same, also we are thankful to student volunteers of the department of Biomedical Engineering, SGSITS, Indore.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Patil, M.K. Anatomical Study of Jhiva W.R.T Sam and Niram Prakriti Pariksha". International Ayurvedic Medical Journal, IAMJ. 1(2),151-159 (2017).
2. Vocaturo, E., Zumpano, E and Veltri, P. 2019. On discovering relevant features for tongue colored image analysis. In Proceedings of the 23rd International Database Applications & Engineering Symposium (IDEAS '19). Association for Computing Machinery, New York, NY, USA, Article 12, 1–8. <https://doi.org/10.1145/3331076.3331124>.(2019).
3. Chiu, C.C. A novel approach based on computerized image analysis for traditional Chinese medical diagnosis of the tongue. Proceedings of the Computer methods & Programs of Biomedicine Conference; (2000) Feb; Elsevier; 2000:61(2). p. 77-89.
4. Wang, Y., Zhou, Y., Yang, J., Xu, Q. An Image Analysis System for Tongue Diagnosis in Traditional Chinese Medicine. Proceedings of the International Computational & Information Science Conference; 2004; Springer – Verlag Berlin, Heidelberg; (2004). p.1181-1186.
5. Zhang, H., Wang, K., Zhang, D., Pang, B., Huang, B. Computer aided tongue diagnosis system. Proceedings of the 27th Annual Conference on Engineering in Medicine & Biology Society; (2005) Jan 17-18; IEEE; 2006. p. 6754–6757.
6. Jiang, L., Xu, W., Chen, J. Digital imaging system for physiological analysis by tongue color inspection. Proceedings of the 3rd Innovative Engineering & Applications Conference; 2008 June3-5; IEEE; 2008. p.1833–1836
7. Xu, J., Tu, L., Ren, H., Zhang, Z. A Diagnostic Method Based on Tongue Imaging Morphology. Proceedings of the 2nd international Conference on Bioinformatics & Biomedical Engineering; (2008) May16-18; IEEE. 2008. p. 2613-2616.
8. Wang, X., Zhang, B., Yang, Z., Wang, H., Zhang, D. Statistical Analysis of Tongue Images for Feature Extraction & Diagnostics. IEEE transactions on Image Processing. (2013); 22(12): 5336 – 47.
9. Wang, X., Zhang, D. A High-Quality Color Imaging System for Computerized tongue Image Analysis. Journal of Expert Systems and Applications. Elsevier; (2013) November 1; 40(15): 5854-66.
10. Liu, Z., Zhang, D., Yan, J.Q., Li, Q.L., Tang, Q.L. Classification of hyperspectral medical tongue images for tongue diagnosis. Computerized Medical Imaging and Graphics. (2007); 31(8):672-678.
11. Li, Q., Liu, J., Xiao, G., Xue, Y. Hyperspectral tongue imaging system used in tongue diagnosis. Proceedings of 2nd International Conference on Bioinformatics & Biomedical Engineering; (2008) May 16-18; IEEE; 2008. p. 2579–2581.
12. Li, Q., Wang, Y., Liu, H., Sun, Z., Liu, Z. Tongue fissure extraction and classification using hyperspectral imaging technology. Applied Optics; (2010);49(11): 2006-2013.
13. Li, Q., Lui, Z. Tongue color analysis and discrimination based on hyper spectral images. Computerized Medical Imaging & Graphics;(2009); 33:217–221.

14. Yamamoto, S., Tsumura, N., Nakaguchi, T., Namiki, T., Kasahara, Y., Terasawa, K., Miyake, Y. Early Detection of Disease Oriented State from Hyperspectral Tongue Images with Principal Component Analysis and Vector Rotation. *Proceedings of Annual International Conference on Engineering in Medicine & Biology Society*; (2010); IEEE ;2010. p. 3025 -3028.
15. Li, Q., Wang, Y., Liu, H., Sun, Z. AOTF based Hyperspectral Tongue Imaging System and Its Applications in Computer-aided Tongue Disease Diagnosis. *Proceedings of 3rd International Conference on Biomedical Engineering and Informatics*; (2010); IEEE; 2010. p. 1424-27.
16. Liu, Z., Wang, H.J., Li, Q. Tongue Tumour Detection in Medical Hyperspectral Images. *Sensors*; MDPI; (2012); 12: 162-174.
17. Ryu, I., Itiro, S. A tongue diagnosis system for personal healthcare on smartphone. *Proceedings of the 5th Augmented Human International Conference*; 2014 March; Waseda University Press; (2014).
18. Duan, Y., Xu, D. (Interdisciplinary Innovations Fund (IIF) 2012/2013 Awards (MU)). I Tongue: An iPhone App for Personal Health Monitoring Based on Tongue Image. Final Report; University of Missouri, Columbia. (2014).
19. Hu, M.C., Cheng, M.H., Lan, K. C. Color Correction Parameter Estimation on the Smartphone and its Application to Automatic Tongue Diagnosis. *Journal of Medical Systems*. (2016); 40:18.
20. Tania, M.H., Lwin, K.T., Hossain, M.A. Computational Complexity of Image Processing Algorithms for an Intelligent Mobile Enabled Tongue Diagnosis Scheme. *Proceedings of 10th International Conference on Software, Knowledge, Information Management & Applications*; (2016) Dec 15-16; IEEE; 2016. p.29 - 36.
21. Kanawong, R., Obafemi-Ajayi, T., Liu, D., Zhang, M., Xu, D., Duan, Y. Tongue Image Analysis and Its Mobile App Development for Health Diagnosis. *Adv Exp Med Biol*.(2017):99-121.
22. Hu, M.C., Lan, K.C., Fang, W.C., Huang, Y.C., Ho, T.J., Lin, C.P., Yeh, M.H., Raknim, P., Lin, Y.H., Cheng, M.H., He, Y.T., Tseng, K.C. Automated Tongue Diagnosis on the Smartphone and its Applications. *Computer Methods and Programs in Biomedicine*. (2019); 174(C):51-64.
23. Zhou, Z., Peng, D., Gao, F., Leng, L. Medical Diagnosis Algorithm Based on Tongue Image on Mobile Device. *Journal of Multimedia Information System*. (2019) June; 6(2):99-106.
24. Smith, Z.J., Chu, K., Espenson, A.R., Rahimzadeh, M., Gryshuk, A., Molinaro, M., Dwyre, D.M., Lane, S.M., Matthews, D., Wachsmann-Hogiu, S. Cell-Phone-Based Platform for Biomedical Device Development and Education Applications. *Article in PLoS One*. (2011) March; 6(3): e17150.
25. Haralick, R.M., Shanmugan, K., Dinstein, I. Textural features for Image Classification. *Journal of IEEE Transactions on Systems, Man, and Cybernetics*. (1973); SMC-3(6): 610-21.
26. Pang, B., Zhang, D. Computerized tongue diagnosis based on Bayesian networks. *IEEE Trans. Biomed. Eng*; (2004); 51:1803–1810.
27. Pang, B., Zhang, D., Wang, K. Tongue Image analysis for appendicitis diagnosis. *Information Sciences*; Elsevier; (2005); 175(3):160-176.
28. Kim, J., Son, J., Jang, S., Nam, D.H., Han, G., Yeo, I., Ko, S.J., Park, J.W., Ryu, B., Kim, J. Availability of Tongue diagnosis system for Assessing Tongue Coating thickness in Patients with Functional Dyspepsia. *Evid based Complementary & Alternative Medicine: Article ID348272*.(2013).
29. Jayanti, S.K., Shanmugapriyanga, B. Detecting Diabetes Mellitus Gradient Vector Flow Snake Segmented Technique. *International Research Journal of Engineering and Technology*; (2017); 4(4):1238-1244.
30. Zhang, D., Zhang, H., Zhang, B. Detecting Diabetes Mellitus and No proliferative Diabetic Retinopathy Using CTD. *Tongue Image Analysis*; Springer Singapore; (2017), March. p.303-325.
31. Sandhya, N., Rajasekar, M. Tongue Image Analysis for Hepatitis Detection Using GA-SVM. *Indian Journal of Computer Science and Engineering*. (2017); 8(4) :526-534.
32. Mrilaya, D., Pervetaneni, P., Aleperi, G. An Approach for Tongue Diagnosing with Sequential Image Processing Method. *International Journal of Computer Theory and Engineering*. (2012); 4(3): 322-328.
33. Dhanalakshmi, M., Pervetaneni, P., Aleperi, G. Applying Wavelet Transforms and Statistical Feature Analysis for Digital Tongue Image. *IOSR-JCE, IOSR Journal of Computer Engineering*, p-ISSN 2278-8727: Feb 2014;16(IX):95-102. (2014).
34. Kawanabe, T., Kamarudin, N.D., Ooi,C.Y., Kobayashi,F., Xiaoyu, M., Sekine,M., Wakasugi,A., Odaguchi, H., Hanawa, T. Quantification of tongue colour using machine learning in Kampo medicine, *European Journal of Integrative Medicine*, Volume 8, Issue 6,2016,Pages 932-941, ISSN 1876- 3820. <https://doi.org/10.1016/j.eujim.2016.04.002>.(2016).
35. Li, Jun and Hu, Xiaojuan and Tu, Liping and Cui, Longtao and Jiang, Tao and Cui, Ji and Ma, Xuxiang and Yao, Xinghua and Shi, Yulin and Wang, Sihan and Liu, Jiayi and Wang, Yu and Li, Yongzhi and Zhou, Changle and Huang, Jingbin and Xu, Jiatuo, *Diabetes Tongue Image Classification Using Machine Learning and Deep Learning*. <https://doi.org/10.2139/ssrn.3944579>.
36. Ma, J., Wen, G., Hu, Y., Chang, T., Zeng, H., Jiang, L., Qin, J. Tongue image constitution recognition based on Complexity Perception method.(2018). <https://doi.org/10.48550/arXiv.1803.00219>.

37. Zhang, X.F., Zhang, J., Hu, G.Q., and Wang, Y.Z. Preliminary Study of Tongue Image Classification Based on Multi-Label Learning. Springer; ICIC 2015, Part III LNAI9227. (2015). https://doi.org/10.1007/978-3-319-22053-6_23:2080220.
38. Chen, L., Wang, B., Zhang, Z., Lin, F., and Ma, Y. (2017). Research on Techniques of Multifeatures Extraction for Tongue Image and Its Application in Retrieval. Hindawi, Computational and Mathematical Methods in Medicine; Volume 2017, Article ID 8064743, 11 pages. <https://doi.org/10.1155/2017/8064743>, (2017).
39. Jiang, T., Lu, Z., Hu, X., Zeng, L., Ma, X., Huang, J., Cui, J., Liping, T., Zhou, C., Yao, X., Xu, J. Deep Learning Multi-label Tongue Image Analysis and Its Application in a Population Undergoing Routine Medical Checkup. Evidence-Based Complementary and Alternative Medicine. 1-12. (2022). <https://doi.org/10.1155/2022/3384209>.
40. Li, J., Zhang, Z., Zhu, X., Zhao, Y., Ma, Y., Zang, J., Li, B., Cao, X., Xue, C. Automatic Classification Framework of Tongue Feature Based on Convolutional Neural Networks. Micromachines (Basel). (2022) Mar 24;13(4):501. PMID: 35457806; PMCID: PMC9025353. <https://doi.org/10.3390/mi13040501>.
41. Wang, X., Liu, J., Wu, C., Liu, J., Li, Q., Chen, Y., Wang, X., Chen, X., Pang, X., Chang, B., Lin, J., Zhao, S., Li, Z., Deng, V., Yi, Lu., Zhao, D., Chen, J. Artificial intelligence in tongue diagnosis: using deep convolutional neural network for recognizing unhealthy tongue with tooth-mark. Elsevier, Computational and Structural Biotechnology Journal. 18(2020).973-980. <https://doi.org/10.1016/j.csbj.2020.04.002>.
42. Meng D, Cao G, Duan Y, Zhu M, Tu L, Xu D, Xu J. Tongue Images Classification Based on Constrained High Dispersal Network. Evid Based Complement Alternat Med. 2017; 2017:7452427. <https://doi.org/10.1155/2017/7452427>. (2017) Mar 30. PMID: 28465706; PMCID: PMC5390589.
43. Kanawong, R., Ajayi, TO., Ma, T., Xu, D., Li, S., and Duan, Y. Automated Tongue Feature Extraction for Zheng Classification in Traditional Chinese Medicine. Hindawi Publications corporation. (2012). <https://doi.org/10.1155/2012/912852>.
44. Rajakumaran, S., & Sashikala, J. An Automated Tongue Color Image Analysis for Disease Diagnosis and Classification Using Deep Learning Techniques. European Journal of Molecular & Clinical Medicine, 7, 4779-4796. (2021).
45. Mansour, R.F., Althobaiti, M.M., and Ashour, A.A. "Internet of Things and Synergic Deep Learning Based Biomedical Tongue Color Image Analysis for Disease Diagnosis and Classification," in IEEE Access, vol. 9, pp. 94769-94779, (2021). <https://doi.org/10.1109/ACCESS.2021.3094226>.
46. Soma, P., Saradha, K.R., Jothika, S. and Dharshini, S. Tongue Diagnosis using CNN for Disease Detection. International Journal of Electrical and Electronics Research (IJEER), Volume 10, Issue 4 | Pages 817-821, e-ISSN: 2347-470X.(2022).
47. Xie, J., Jing, C., Zhang, Z., Xu, J., Duan, Y., Xu, D. Digital tongue image analyses for health assessment. Med Rev (Berl). (2022) Feb 14;1(2):172-198. <https://doi.org/10.1515/mr-2021-0018>. PMID: 37724302; PMCID: PMC10388765. (2021).
48. Mayer, S., Müller, D and Kramer, F. Standardized Medical Image Classification across Medical Disciplines. October (2022). <https://doi.org/10.48550/arXiv.2210.11091>.
49. Bhatnagar, V., Bansod, P.P. Double U-Net a Deep Convolution Neural Network for Tongue Body Segmentation for Diseases Diagnosis. Proceedings of International Conference on Communication and Computational Technologies. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-19-3951-8_23:293-303. (2023).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.