ISTP

RESEARCH ARTICLE

People Recognition via Tongue Print Using Deep and Machine Learning

Ahmed Shallal Obaid, Mohammed Y. Kamil, and Basaad Hadi Hamza

College of Science, Mustansiriyah University, Baghdad, Iraq

Abstract: The tongue is a unique organ that is well protected inside the mouth and not affected by external factors; it is also difficult to forge. Several biometric systems are widely used for authentication and recognition, such as fingerprints, faces, iris, sound, and retina. Traditional biometrics represent a challenge and an obstacle as they can be falsified, duplicates can be made (e.g., iris, face, fingers, and signature), or they are expensive and rarely used (e.g., DNA). The increased security measures called for modern biometrics that is more secure, less expensive, and cannot be falsified. As a result, the goal of this paper is to create a system for distinguishing people based on their tongue prints. It will contribute to solving many forensic issues and increasing electronic security because it has features suitable for identification and biometrically distinguishing between people. In this paper, the tongue is located based on the fixed window size method. After tongue localization, feature extraction using the VGG-16 model, and a classification system that uses both transfer learning and machine learning as VGG-16, XGBoost, KNN, and random forest classifiers, extracted features are then trained for personal identification. The dataset consisted of 1085 tongue images of 138 people with a test ratio of 20%, and the results achieved an accuracy of 92%. The process of distinguishing people through tongue prints has proven to be effective and accurate.

Keywords: pattern recognition; ROI; texture recognition; tongue; VGG-16

I. INTRODUCTION

Identifying persons is one of the most important processes that has recently interested academics due to its numerous uses [1]. It can be used for protection in electronic attendance recording or person verification systems. Researchers invented several methods, but they need assistance with cost, safety, and speed issues [2]. There was a research gap in this subject until biometric-based solutions were established. Biometrics is a method that automatically uses a person's unique behavioral and physiological characteristics to identify and verify that person's identity [3]. Humans are identified based on their bodily characteristics, not the exterior items they must give. Since it is so difficult to recreate a person's unique characteristics, outcomes are usually almost accurate. Numerous well-known approaches have previously been used to identify human (fingerprint, face, retina, iris, etc.) [4]. Personal identity is the most crucial one. Conventional identification systems use badges, keys, cards, PINs, passwords, and biometric, behavioral, and physiologic identifiers. There is a risk that the user's identification password may be lost, forgotten, or shared with others. The user encounters difficulties such as theft and memory dependence.

Biometrics is increasingly vital for verifying a person's identification in the current digital world; in every field, the use of biometrics to verify a person's identification is becoming more prevalent. The technology is mainly used for identification and access control or to identify people and offer surveillance-controlled access. These bodily segments, including fingerprints, the iris, voice, facial geometry, hand geometry, ear geometry, and

DNA, can be utilized as biometrics. This method provides a high degree of accuracy for identifying a person. Although these approaches can guarantee that the person authenticating either possesses the token or knows the password, they cannot guarantee the individual's personality [5]. Because these authentication systems have some flaws, fingerprints may be degraded, altered by labor, surgically altered, and damaged by burns and traumas. Therefore, they are not stable. Retinal scanning is the susceptible users can be affected by bright light and diseases such as cataracts and astigmatism. Users with light-sensitive conditions, such as cataracts or astigmatism, avoid examining the retina in bright light. In intense emotional conditions, the voice may be influenced by a cold or cough, increasing the likelihood of uttering incorrect words. A biometric system should support information security's identity, authentication, and non-repudiation aspects. Because they are susceptible to forgery, conventional biometric systems cannot satisfy these authentication criteria. As a result, tongue prints are gaining prominence as a new biometric technology for biometric authentication [6].

The tongue is an essential internal organ shielded from the outside world by the protective covering of the oral cavity. Individuals and identical twins have noteworthy variances in the tongue's defining characteristics. It is no surprise that the tongue has its network of nerves, muscles, and blood arteries, just like every other organ. It also has taste buds and papillae. Observing the tongue's properties, such as its color and form, is crucial in traditional Chinese medicine for illness diagnosis [7]. Recently, it has seen a rise in use as a biometrics tool. They aimed to create a 3D representation of the tongue that described the texture and form of the tongue images. They constructed a database of 3D tongue images that defined the images' texture and form [8]. Every day, all money transactions and payments are made through the Internet.

Corresponding author: Mohammed Y. Kamil (e-mail: m80y98@uomustansiriyah.edu.iq).

As mentioned above, much other biometrics can be used. To this day, tongue recognition technology is novel in biometrics; authentication is the superior method for providing top-tier security. Some examples of such uses are biometrics has great potential in many areas, including account access, criminal identification, online banking, ATM use, employee access, personal data access, medical identification, and air travel [9].

The rest of the paper is organized as follows: Section II discusses the related work. Section III presents the classification methods that are used in the work. Section IV provides the materials and methods as a dataset. Section V shows the findings and discussion of the models, while Section VI states the most critical conclusions.

II. RELATED WORK

In 2007, David Zhang *et al.* [10] presented a verification framework and new biometrics based on the tongue-print tongue image database TB06 contains 134 subjects. The method employed extracts the textual features and shape and the Gabor filter. The actual accept rate (GAR) was 93.3%, while the false accept rate (FAR) was 2.9%. In 2009, K.H. Kim *et al.* [11] suggested using a digital tongue diagnosis system (DTDS). The method used is graph-based segmentation, the database of tongue image Private Contain 56 People with properties 1280F960 and RGB 24 bit-BMP. The findings demonstrate that the segmented region contains helpful information, omitting a non-tongue area, diagnosis of tongue surface spots, and accurate coating, and the difference rate is 5.5%. In 2014, Miao-Jing Shi *et al.* [12] suggested initializing the tongue area into the level set matrix for the under and binary upper parts. The DGF was presented for segmenting the tongue area and detecting the tongue edge in the image. The Geo-GVF stood assessed in the upper part and evaluated the geodesic outpour in the under part. A Shanghai University for TCM has already collected a clinical tongue image database. Using 100 images, compared to previous studies, the DGF demonstrated superior outcomes. Its real-positive volume percentage reached 98.5%, while its faulty-negative volume fraction was 1.42%. The faulty-positive volume fraction was 1.5%. In 2016, R. S. Chora [13] proposed some image tongue recognition methods and used a private tongue image collection consisting of 30 images to test the effectiveness of various methods of tongue recognition. A method presented combines the results of tongue recognition using color moments features (RGB and YCbCr) and Gabor filters. The principal component analysis (PCA) technique preserved the Gabor features that were the most helpful. In 2017, Meo Vincent *et al.* [14] suggested a tongue print biometric identification system capable of employing SIFT and BRIEF keypoint descriptor techniques. To ascertain which of the two algorithms is more efficient in recognizing images. Used a database from 30 stock images for testing the suggested system. The test findings reveal that applying the SIFT algorithm for tongue print identification resulted in an average recognition time of 13.829 seconds, compared to 7.644 seconds for the BRIEF algorithm. The accuracy test results suggest that applying the BRIEF algorithm enhances recall and accuracy compared to the SIFT algorithm. In 2020, Wen Jiao *et al.* [15] proposed PCA and the K-Means algorithm for tongue color classification and diagnosis and clustering into four groups. In total, 595 images of tongues were studied in total. The clustering accuracy (CA), adjusted rand index (ARI), and Jaccard similarity coefficient (JSC) were used to assess the results and successfully classified the group tongue images into four groups. The CA, ARI,

and JSC were 89.04%, 0.721%, and 0.890%, respectively, and in the color space L*a*b*, they kept 89.63% of the original information.

Traditional biometrics represents a challenge and an obstacle as they can be falsified, and duplicates can be made (iris, face, fingers, and signature) or be expensive and rarely used (DNA). The increased security measures called for modern biometrics that is more secure, less expensive, and cannot be falsified. As a result, this paper aims to create a system for distinguishing persons based on their tongue prints. It will contribute to solving many forensic issues and increase electronic security because it has features suitable for identification and distinguishing between people biometrically. Traditional biometrics represents a challenge and an obstacle as they can be falsified, and duplicates can be made (iris, face, fingers, and signature) or be expensive and rarely used (DNA). The increased security measures called for modern biometrics that is more secure, less expensive, and cannot be falsified. As a result, this paper aims to create a system for distinguishing persons based on their tongue prints. It will contribute to solving many forensic issues and increase electronic security because it has features suitable for identification and distinguishing between people biometrically.

III. CLASSIFICATION PROCESS

Most machine learning and deep learning operations involve studying images, extracting their features, and storing them in a database. Then the algorithm is grafted with a new image; its features are extracted and compared with the features stored in the database to identify whether the image belongs to one of the previously stored classes. This process is known as classification. The process used in this study belongs to supervised machine learning. Many classifiers were used in this study, and they will be discussed in detail to gain a deep understanding of this study [16].

A. XGBOOST

XGBoost is defined as a large-scale machine-learning system for tree boosting. You can get the system as an open-source package that can be installed on the python environment. The influence of the system on various machine learning and data mining problems has received widespread recognition. While domain-specific data analysis and feature engineering are important components of our solutions, the fact that XG Boost is the learners' preferred option demonstrates the effect and significance of our system and tree boosting. The most critical component in XG Boost's success is its scalability. The system operates ten times quicker than current popular solutions on a single computer and grows to billions of instances in distributed or memory-constrained environments. The scalability of XG Boost may be attributed to several significant algorithmic and systemic advances. These innovations include a novel tree learning method for managing sparse data and a theoretically supported weighted quantile sketch procedure for managing instance weights in approximation tree learning. Computers that are parallel and distributed accelerate learning, enabling faster model exploration. Using out-of-core processing, XG Boost enables data scientists to analyze millions of instances on a single workstation [17].

B. K-NEAREST NEIGHBOR ALGORITHM

The k-nearest neighbor (k-NN) algorithm was first presented by [18] as a nonparametric technique and allocates query data to the

class to which most of its k-nearest neighbors belong. For instance, the k-NN method classifies data independently without requiring an explicit model. How well k-NN works depends on how many neighbors are closest to it. In general, there is no way to determine the best k. However, a trial-and-error technique is typically employed to determine its best value. The k-NN algorithm's key benefit is its ability to explain categorization results. On the other hand, the main disadvantage is determining the best k and creating a suitable metric for measuring the distance between the query instance and the training examples. The standard k-NN algorithm is (1) Determine the Euclidean distances between an unknown object (o) and each object in the training set. (2) Based on the calculated distances, choose the k objects from the training set that are most similar to object (o). (3) Assign object (o) to the group to which the majority of K objects belong.

C. RANDOM FOREST

A random forest (RF) resampling strategy employs the bootstrap method, which involves randomly picking k samples from the original training sample set N to produce new training sample sets. Subsequently, to produce RFs, k-classification trees are constructed according to the bootstrap sample set. The new data classification outcomes depend on the score determined by the classification tree's vote. The algorithm is described as [19]: (1) The initial training set is N, and the bootstrap method is used to randomly select K new self-help sample sets and build K classification trees. K data outside the bag are samples not drawn at each time point. (2) Mall variables are set in total, and more variables are chosen at random at each node of each tree. The variable with the greatest classification ability is chosen. Each classification point is examined to determine the variable classification threshold. (3) Pruning allows each tree to reach its full potential. (4) RFs are formed by the generated multiple-classification trees. New data are collected and classified using the RF classifier. The classification results are determined by the number of votes provided by the tree classifiers. RFs enhance the algorithm for decision trees by merging many decision trees. Each tree is created using a sample that was individually extracted. The distribution of trees in the forest is uniform. Classification error depends on each tree's classification capacity and correlation. Utilizing a stochastic strategy, feature selection is used to partition each node. The mistakes created under different conditions are then compared. The number of selected characteristics is determined by detecting internal estimating errors or the classification and correlation capabilities. A single tree's capacity for categorization may be poor. After generating a large number of decision trees at random and doing statistical analysis based on the classification outcome of each tree, the classification most likely to apply to a test sample is picked.

III. MATERIALS AND METHODS

A. TONGUE IMAGE ACQUISITION

The first step in a computerized tongue identification system is obtaining digital tongue images. Researchers have investigated the use of digital cameras in tongue examination for years as digital image technology has advanced. Different types of color cameras obtained images of the tongue under white illumination. Researchers have paid increased attention to this development since these imaging devices are inexpensive and quick to install. Nearly 10
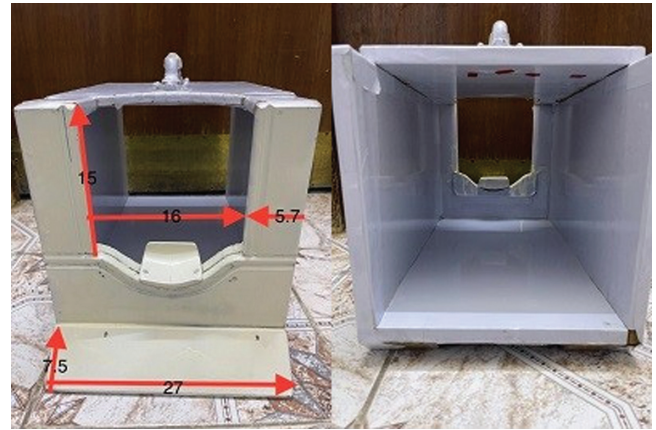

Fig. 1. Proposed acquisition image tool.

imaging systems with different imaging characteristics have been developed. These advanced acquisition devices differ mainly in the light source and camera choice. Because these devices were not all the same, the quality of the tongue images they took was different.

Before starting the system, the dataset containing the image of the own tongue must be loaded. All images are captured for 138 individuals, with eight images per individual, using an iPhone 12 pro max camera with a resolution of (9 megapixels). We set up an appropriate environment for image capture. This environment can be created with a high-density fibreboard box measuring $45.5 \times 17 \times 27.5$ cm^3. The light source was positioned on one side of the camera (corresponding to the person). The fixed distance between both the camera and the candidate is 15 cm. In addition to uniform lighting, this system is designed to ensure that every image has the same environment and conditions. Totaling 138 individuals, 87 females and 51 males had their tongues collected. Each individual was photographed eight times, and a total of 1104 photographs were obtained. Due to the presence of deformation caused by the subject's motion, nineteen images were excluded. The remainder was 1085 images of 138 items (person). Figure 1 illustrates the proposed acquisition image tool.

B. PREPROCESSING

Following the tongue image acquisition process and before the execution of the feature extraction procedure, the native tongue images are put through a series of image preprocessing operations. These operations include converting colors and the crop of the same tongue region from each image. Lips, teeth, and other facial features are frequently affected. Region of interest (ROI) extraction is part of the preprocessing stage. That describes a region in the shape of a rectangle whose primary component is the body of the tongue [20]. Before feature extraction, the initial pictures of the tongue are exposed to a series of image-processing processes, such as extraction ROI from original tongue images. Traditional segmentation methods are force complex to segment an object from the image, revealing the face's central part (mouth region). Because the mouth area contains not just the tongue but also teeth, lips, the mouth itself, and other objects, how to get the tongue, among other things, in that mouth area is the problem of this image segmentation [21]. Many methods are used to extract ROI by window (fixed size window): In the preprocessing of a tongue print identification system, the main task is the localization and segmentation of

ROI. Extraction ROI seeks to determine which portion of the picture fits the center of the tongue body, maintaining the helpful information in the ROI, deleting the unnecessary information in the background, and taking fixed-size windows from the center of the image. The essential rule in extracting ROI is that ROI should automatically be extracted for all tongue images in the dataset. Perfect ROI extractions of tongue pictures will considerably minimize the computing complexity of any future processing and increase the tongue print identification system's performance [22]. Cropping is an essential technique for improving collected images' visual quality. Its purpose is to enhance an image by improving its composition, adjusting its aspect ratio, and removing any extraneous information from the image. Automatic image cropping has gained significant attention from academia and business over the last decades because cropping is a frequent necessity in photography but a tiresome task when multiple photographs have to be cropped. The study on image cropping concentrated mainly on cutting the primary subject or most crucial section of an image for usage in miniature displays or to make image thumbnails; this reflects the no uniqueness of image cropping. Start with an image as your source. You may get several different excellent crops (denoted with "p") by experimenting with various aspect ratios (e.g., 1:1, 4:3, 16:9). Even while maintaining the same aspect ratio, there are a few different cropping choices that are acceptable. If we take the crop in the middle of the three with a 16:9 aspect ratio as the ground truth, the bottom one, which is a poor crop and is designated by the letter "x," would have a more giant IOU (intersection-over-union) than the crop at the top, although having a lower esthetic quality. Demonstrates that IOU is not a trustworthy criterion for assessing crop quality. These approaches were primarily concerned with attention scores or saliency values. Attention-based methods may provide visually unappealing results with little regard for the overall image composition [23].

Furthermore, a user study was used as the primary criterion to subjectively evaluate cropping performance, making it impossible to compare different approaches objectively. Recently, several benchmark databases have been released for image cropping. On these datasets, knowledgeable human participants labeled one or more bounding boxes as "ground-truth" cropping for each image. Two objective measures, boundary displacement error (BDE) and intersection-over Union (IOU), were developed to evaluate the effectiveness of various image-cropping algorithms on these datasets. Numerous academics can now build and evaluate cropping algorithms because of these publicly accessible standards, greatly aiding research on automated picture cropping. Despite several efforts, specific intractable issues remain due to the particular features of picture cropping. Baseline C crops the central part, whose width and height are 0.9 times the source image [23]. The textural characteristics of the tongue are predominantly present on its center surface. To extract this information, we created an area of interest from a subimage of the segmented tongue image (ROI) [10]. Figure 2 depicts the database composed of images of the tongue, and Fig. 3 demonstrates the determination of ROI.

C. FEATURES EXTRACTION

Feature extraction is a critical function in many image applications. A feature is a property of an image that may capture a certain optical quality globally for the whole picture or specifically for areas or objects. Deferent features such as texture, shape, and color can be disengaged from an image. The surface is the variation of

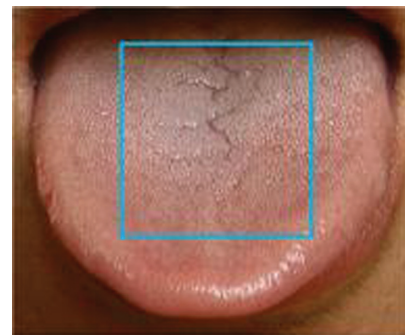**Fig. 2.** Samples of the obtained images.

Fig. 3. The ROI.

data at different scales. Various techniques have been created for texture extraction, such as VGG 16. They can be extracted from wavelet transform coefficients and co-occurrence matrices [24]. Extracting features from an image and storing them as feature vectors in a feature database is the primary goal of feature extraction. The image's data are discovered by analyzing the value of this feature (or group of values), known as the feature vectors. Images in the database are compared to the query image using these feature vectors. Features are a method used in pattern recognition to differentiate between different types of objects—extraction of features such as shape, texture, and color. When designing an image retrieval system, it is possible for there to be many representations of each feature [25]. Accordingly, the features retrieved at each level are the results of the final layer, resulting in a feature pyramid [26]. The VGG16 Model for feature extraction comprises a thousand-node output layer, three dense layers for the FC layer, and sixteen convolutional layers with five Max Pooling (for 1000 classes). In the context of this investigation, the top layers (FC and output layers) have been removed [27]. As feature extractors from images, transfer learning (TL) algorithms are applied from DL. Several studies have used the pretrained DL model to classify

breast cancer based on images from the BreaKHis histopathology. For example, the authors of [28] employ VGG16 as pretrained DL models to categorize the BreaKHis dataset. The fundamental objective of this research is to investigate whether or not pretrained deep learning VGG16 can be used effectively as a feature extractor for binary and multiclass breast cancer histopathology image classification tasks once VGG16 has been modified specifically for these tasks. These breast cancer histology images are from the BreakHis 40× magnification dataset, which is open to the public [28]. BreakHis categorization was performed using VGG16 at 40× magnification. On multiclass classification, they scored 89.6%. One explanation for these poor results is because both employed CNN as a classifier model, which may need precise parameter adjustment. Consequently, supplying the characteristics to the classifiers may provide excellent results [29]. Several considerations are presented: Rather than integrating feature extraction and classification inside a single model, it is preferable to use a CNN that has already been trained for feature extraction. Second, the problem of the multiclass sort in BreaKHis is complex and requires the input of several experts in their respective fields. Third, a data augmentation approach needs to be implemented to lessen the imbalance in the dataset [27].

D. CLASSIFICATION FROM PRETRAINING DEEP LEARNING

Due to the success of CNNs in many areas, there is now a large variety of CNN designs with distinct deep characteristics and learning requirements. Here, we investigated three CNN architectures that are traditional and representational of contemporary feature extractors with potential intermediate representations for capturing complicated visual representations. Figure 4 shows the pretraining of our proposed system.

The CNN architectures used here were chosen for their proven success in natural image recognition tasks. It is possible, however, that they are not well suited for representing and discriminating patterns. Therefore, we used TL to acquire a representation. TL is a well-known method that applies learned weights from meaningful broad picture representations by adapting many layers to a particular domain, such as tongue images. The Mk model ($k = $ VGG16) formally specified the learned picture representation as $Mk = [F, P(F)]$, where F is the feature space and $P(F)$ is the marginal probability distribution. In this scenario, the generic picture domain was represented hierarchically with respect to the job at hand by the notation $F = [Fi (Fi\text{-}1)$. Consequently, the problem's initial definition of classes $Dt = [d1, \ldots, dn]$ was included in the scope of the task Tt (ImageNet). As such, TL's goal was to transform the generic codified learning task Tt into a different task Ts, as $Tt = [Dt, Mt] \rightarrow Ts [Ds, Ms]$. TL (Tt→Ts) is an adaptive, iterative approach that utilizes a relatively modest learning rate and batches from a new domain, in this instance, trained CT slices. Finally, a
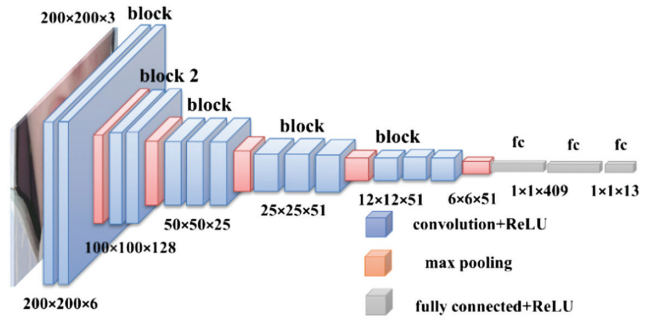

Fig. 4. Pretraining of our proposed system.

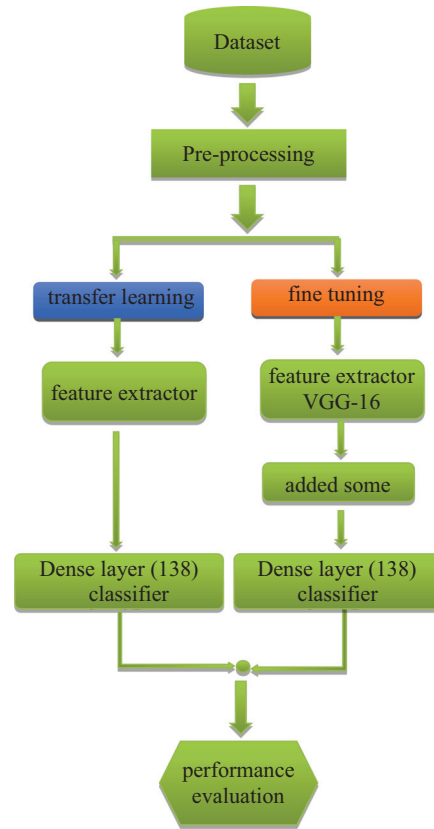**Fig. 5.** The architecture of the VGG16 model.

Fig. 6. Block diagram of the proposed systems.

detailed representation of each CNN architecture was obtained. Figure 5 depicts the architecture of the VGG16 network, and Fig. 6 shows the block diagram of the proposed methods.

IV. RESULTS AND DISCUSSION

The performance of the designed intelligent classification system must be verified by evaluating the classification accuracy. This evaluation accuracy is performed by implementing the designed classification system on the testing dataset and determining the accuracy between the aggregated class obtained by the designed intelligent system and the predefined class of each element in the testing dataset and the second measurement of the relationship among the outputs of the neural network and the targets. Perfect
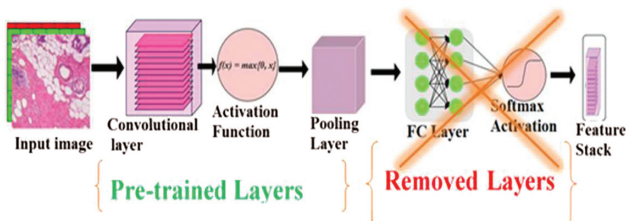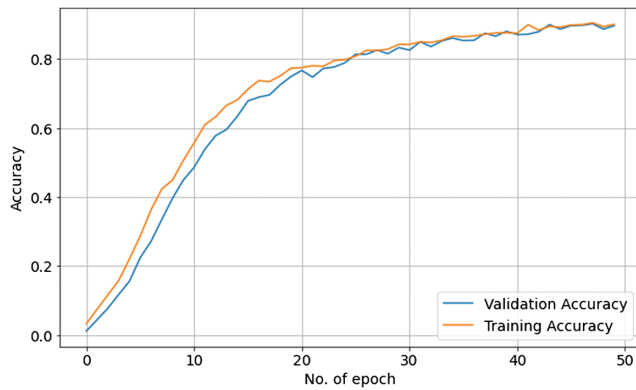
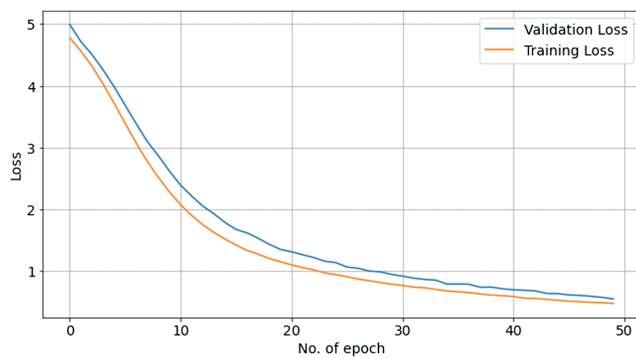**Fig. 7.** Accuracy of transfer learning method.

Fig. 8. Loss of transfer learning method.

training makes the network outputs and the targets would be the same. The VGG16 model was used without a fully connected classifier/layers, and the classifier was added to it by adding three blocks (batch normalization 512, global average pooling 2D 512, and Dense 138). In the implementation of the model, it was determined that the number of epochs was 50, the batch size was 32, the loss function was categorical cross-entropy, and the select optimizer was Adam. To work with pretrained weights, the model's parameters must be untrainable, which is important. Figure 7 represents the accuracy of the learning transfer method. It is noticed from Fig. 8 that the highest value reached overall accuracy in this model was 90.23%. Also, it is noticed that there is high stability between training and validation accuracy, and there is no overfitting, which indicates that the model is perfect. Figure 8 shows the loss of the learning transfer method. It is noticed that the lowest value we reached in this model is 0.478, which indicates the stability of the model used in VGG16.

In this section, the VGG16 model was used as a feature extractor, as shown in Table I, the features were extracted from the images of the tongue. Then, three machine learning models were added to the model as classifier functions: KNN, XGBoost, and RF models. The data were divided into 80% training and 20% testing. Table I shows the accuracy values of the embedding methods for three classifiers, it appears the highest result with the top accuracy and the average of the top 5 accuracies. It is noticed from this method that when deep learning and machine learning are combined, the run speed of the model is higher, but the accuracy is lower compared to the TL and fine-tuning models.

Table I. Accuracy of the embedding methods

Classifier	Top accuracy (%)	5 Top accuracy (%)
KNN	75.3	73.2
XGBoost	86.8	85.4
Random Forest	78.1	75.6

Table II. Comparison with other studies

Ref.	No. of classes	Method	Accuracy (%)
[10]	134	Gabor filter	93.30
[18]	174	k-NN	92.00
		Gabor filter	83.00
[30]	34	SVM	97.05
[31]	30	Bhattacharyya distance	93.00
Our work	138	VGG16-transfer learning	90.23
		VGG16-embedding (XGBoost)	85.40

In conclusion, Table II summarizes the findings of many investigations conducted on the tongue image categorization system. The accuracy of the comparison was scored and used as the basis for the evaluation. It is essential to highlight that due to the disparity between the data sets, it is impossible to make direct comparisons (for example, the number of images). However, compared to previous efforts, ours did far better, demonstrating the dependability and robustness of the model provided.

V. CONCLUSIONS

In this work, distinct observations and conclusions were obtained, which could be summarized as follows: The process of distinguishing people through tongue prints has proven effective and accurate, with an accuracy of 92%. Adding and modifying the model gave better results than using the same model without any development or modification because no ready-made model works with all cases. This study can be improved in the future by expanding the database to include the most significant number of people and using deep learning models other than VGG16 and implementing and testing work on other age groups more significant than the ages taken in this work.

Acknowledgments

References

[1] Y. M. Saib and S. Pudaruth, "Is face recognition with masks possible?," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 7, pp. 43–50, 2021.

[2] R. Tamilkodi, "Automation system software assisting educational institutes for attendance, fee dues, report generation through email and mobile phone using face recognition," *Wirel. Pers. Commun.*, vol. 119, no. 2, pp. 1093–1110, 2021. DOI: 10.1007/s11277-021-08252-2.

[3] V. Wati, K. Kusrini, H. Al Fatta, and N. Kapoor, "Security of facial biometric authentication for attendance system," *Multimed. Tools Appl.*, vol. 80, no. 15, pp. 23625–23646, 2021. DOI: 10.1007/s11042-020-10246-4.

[4] V. Reddy, R. Kiranmai, and V. Kumar, "Tongue scanning as a biometric tool for improvising the security in Identification," in *Int. Conf. Recent Trends Eng. Sci. Manag.*, Newton's Institute of Science & Technology, Guntur Dist, Andhra Pradesh, India, 2017, pp. 114–118.

[5] M. V. C. Caya, M. E. C. Caringal, and K. A. C. Manuel, "Tongue biometrics extraction based on YOLO algorithm and CNN inception," in *2021 IEEE 13th Int. Conf. Humanoid, Nanotechnol., Inf. Technol., Commun. Control, Environ., Manag. (HNICEM)*: IEEE, Manila, Philippines, 2021, pp. 1–6.

[6] M. Diwakar and M. Maharshi, "An extraction and recognition of tongue-print images for biometrics authentication system," *Int. J. Comput. Appl.*, vol. 61, no. 3, pp. 36–42, 2013.

[7] M. Y. Kamil Rania and R. Kadhim, "Evaluation of machine learning models for breast cancer diagnosis via histogram of oriented gradients method and histopathology images," *Int. J. Recent Innovat. Trends Comput. Commun.*, vol. 10, no. 4, pp. 36–42, 2022. DOI: 10.17762/ijritcc.v10i4.5532.

[8] S. Sadasivan, T. T. Sivakumar, A. P. Joseph, G. C. Zacharias, and M. S. Nair, "Tongue print identification using deep CNN for forensic analysis," *J. Intell. Fuzzy Syst.*, vol. 38, no. 5, pp. 6415–6422, 2020.

[9] S. Kannan, K. D. Vinod, G. Murali, and D. Baskar, "Tongue print image recognition and authentication using convolutional neural networks," in *2022 7th International Conference on Communication and Electronics Systems (ICCES)*: IEEE, Coimbatore, India, 2022, pp. 1301–1305.

[10] D. Zhang, Z. Liu, J.-Q. Yan, and P.-F. Shi, "Tongue-print: a novel biometrics pattern," in *Int. Conf. Biometr.*: Springer, Seoul, Korea, 2007, pp. 1174–1183.

[11] K. Kim, J.-H. Do, Y. Jeon, and J.-Y. Kim, "Feature extraction methods for tongue diagnostic system," in *13th Int. Conf. Biomed. Eng.*: Springer, Singapore, 2009, pp. 527–531.

[12] M.-J. Shi, G.-Z. Li, F.-F. Li, and C. Xu, "Computerized tongue image segmentation via the double geo-vector flow," *Chinese Med.*, vol. 9, no. 1, pp. 1–10, 2014.

[13] R. S. Choraś, "Automatic tongue recognition based on color and textural features," in *Int. Conf. Image Process. Commun.*: Springer, Poland, 2016, pp. 16–26.

[14] M. V. C. Caya, J. P. H. Durias, N. B. Linsangan, and W.-Y. Chung, "Recognition of tongue print biometric using binary robust independent elementary features," in *2017 IEEE 9th Int. Conf. Humanoid, Nanotechnol., Inf. Technol., Commun. Control, Environ. Manag. (HNICEM)*: IEEE, Manila, Philippines, 2017, pp. 1–4.

[15] W. Jiao et al., "Tongue color clustering and visual application based on 2D information," *Int. J. Comput. Assist. Radiol.*, vol. 15, no. 2, pp. 203–212, 2020.

[16] A. H. Farhan and M. Y. Kamil, "Texture analysis of breast cancer via LBP, HOG, and GLCM techniques," *IOP Conf. Series: Mater. Sci.*

[17] Eng., vol. 928, no. 7, p. 072098, 2020. DOI: 10.1088/1757-899X/928/7/072098.

[17] T. Chen and C. Guestrin, "Xgboost: a scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, Association for Computing Machinery, New York NY United States, 2016, pp. 785–794.

[18] S. Lahmiri, "Recognition of tongueprint textures for personal authentication: a wavelet approach," *J. Adv. Inf. Technol.*, vol. 3, no. 3, pp. 168–175, 2012.

[19] R. Guo et al., "Analysis and recognition of traditional Chinese medicine pulse based on the hilbert-huang transform and random forest in patients with coronary heart disease," *Evid.-Based Complement. Alternat. Med.*, vol. 2015, p. 8, 2015.

[20] N. Jeddy, T. Radhika, and S. Nithya, "Tongue prints in biometric authentication: a pilot study," *J. Oral Maxillofacial Pathol.: JOMFP*, vol. 21, no. 1, p. 176, 2017.

[21] M. Fachrurrozi, "Tongue segmentation using active contour model," in *IOP Conf. Series: Mater. Sci. Eng.*, vol. 190, no. 1: IOP Publishing, p. 012041, 2017.

[22] L. Yang, G. Yang, Y. Yin, and R. Xiao, "Sliding window-based region of interest extraction for finger vein images," *Sensors*, vol. 13, no. 3, pp. 3799–3815, 2013.

[23] H. Zeng, L. Li, Z. Cao, and L. Zhang, "Reliable and efficient image cropping: a grid anchor based approach," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recogn.*, (CVPR), Long Beach, CA, USA, 2019, pp. 5949–5957.

[24] M. Roumi, "Implementing texture feature extraction algorithms on FPGA," M.Sc. Thesis, Delft University of Technology, 2009.

[25] N. Bagri and P. K. Johari, "A comparative study on feature extraction using texture and shape for content based image retrieval," *Int. J. Adv. Sci. Technol.*, vol. 80, no. 4, pp. 41–52, 2015.

[26] S. Wang, G. Sun, B. Zheng, and Y. Du, "A crop image segmentation and extraction algorithm based on Mask RCNN," *Entropy*, vol. 23, no. 9, p. 1160, 2021.

[27] D. Albashish, R. Al-Sayyed, A. Abdullah, M. H. Ryalat, and N. A. Almansour, "Deep CNN model based on VGG16 for breast cancer classification," in *2021 Int. Conf. Inf. Technol. (ICIT)*: IEEE, Amman, Jordan, 2021, pp. 805–810.

[28] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:.01195*, 2014.

[29] V. Kate and P. Shukla, "Breast cancer image multi-classification using random patch aggregation and depth-wise convolution based deep-net model," , vol. 17, no. 1, pp. 83–100, Jan. 2021.

[30] T. T. Sivakumar, S. S. Nair, G. C. Zacharias, M. S. Nair, and A. P. Joseph, "Identification of tongue print images for forensic science and biometric authentication," *J. Intell. Fuzzy Syst.*, vol. 34, no. 3, pp. 1421–1426, 2018.

[31] R. J. Hindi, "Identification system by Tongue based on Bhattacharyya distance," *J. Phys.: Conf. Series*, vol. 1530, no. 1: IOP Publishing, p. 012093, 2020.