Transfer Learning Based Prediction Model for Obstructive Sleep Apnea from Facial Depth Maps

Dr. A. Swetha[1*], Syed Junaid[2], M. Ram Reddy[2], M. Sandeep[2], G. Mahesh[2]

[1]Assistant Professor, [2]UG Student, [1,2]Department of CSE(AI&ML)

[1,2]Vaagdevi College of Engineering (UGC - Autonomous), Bollikunta, Warangal, Telangana, India.

*Corresponding Email: Dr. A. Swetha (swetha_a@vaagdevi.edu.in)

ABSTRACT

Stress levels are rising at an alarming rate in today's society as a direct result of the increased level of competition in both the educational and professional spheres. This stress is a contributing factor in the development of a wide variety of ailments, including obstructive sleep apnea. Relaxation of the tongue and the muscles that line the airway may cause obstructive sleep apnea (OSA), which occurs when there is a recurring blockage in the airway during sleep. Snoring, difficulty sleeping because of choking or gasping for breath, and waking up feeling exhausted are typical symptoms of obstructive sleep apnea (OSA). The OSA diagnosis is time-consuming and expensive, both financially and in terms of lost productivity. Because of this, a significant number of patients continue to go untreated and are uninformed of the nature of their illness. Through a depth map of human face scans, the application of deep learning algorithms is employed to identify the condition. In comparison to a standard 2-D colour picture, the depth map offers much more information on the morphology of the face. The traditional machine learning models did not succeed in producing the best possible results in terms of prediction and classification accuracy. Following the extraction of deep face map features using the proposed VGG-19 method and the subsequent training of both the algorithm and a module that was learned on the IMAGENET dataset, transfer learning is used to train the algorithm on OSA facial pictures. The deep learning algorithm known as VGG-19 is trained with the use of the photos from the 3D face scan. In order to predict OSA from fresh test photos, a trained model of VGG-19 is used.

Keywords: Obstructive Sleep Apnea (OSA), Stress-related disorders, VGG-19, Depth map imaging, 3D face scan.

1.INTRODUCTION

Social and personal activities are significantly affected by poor sleep. There are different types of sleep disorders, and it is costing us at different levels. As shows that only in Australia sleep disorder costs the economy around $5.1 billion per year that comprises health care, associated medical conditions, productivity, and non-medical costs. And among all sleep disorder, OSA is the most common cause. Normally during sleep, our upper airway remains open due to relaxed but strong enough muscles, lining the upper throat. But in OSA, someone can have a recurring blockage in upper airway due to different reasons, for more than 10 sec for each blockage, which causes the lungs out of oxygen and person to wake, which will restore the airway. If more than 15 apneas occur, then the diagnosis of OSA is made. History of the patient, physical examination, polysomnography (PSG) test, and imaging are being used to diagnose OSA. The gold standard to diagnoses is PSG test. In which a person needs to sleep in a unit in a hospital with some sensors to monitor breathing patterns, Oxygen level, heart rate, and body movements. Some devices are also helping to conduct these tests at the patient's own home, but there will be a question mark on the reliability of the test and have not been proved to be as accurate as PSG . After the test Apnea-Hypopnea Index (AHI) is computed. This index points out the severity of sleep apnea. Due to cost in term of money and time, invasiveness of the PSG, non-specific nature of

symptoms associated with OSA and the limited access to sleep clinics, many OSA patients remain undiagnosed until significant symptoms appear. Many attempts have been made in the past to predict OSA based on questionnaires. For example, the Berlin questionnaire predicts the level of risk based on snoring, tiredness, blood pressure and body mass index information while the Epworth Sleepiness questionnaire assesses sleepiness in various situations during the day. Although they are self-administered and low-cost, they have shortcomings in accurately identifying affected individuals.

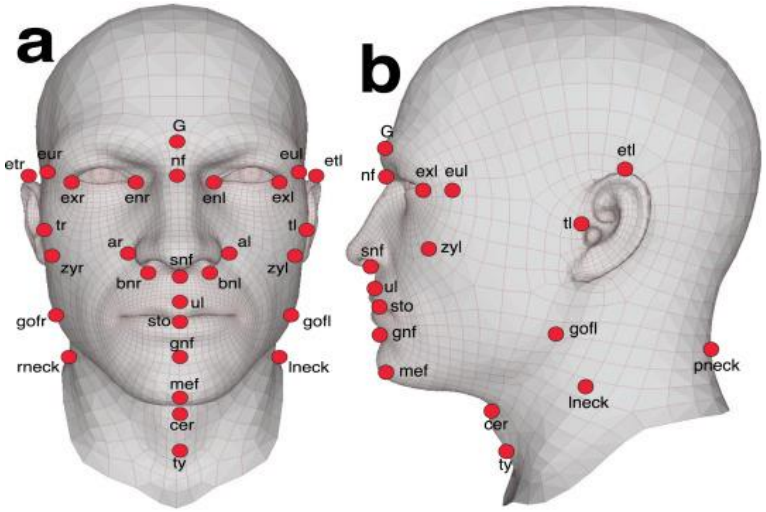


Fig. 1: Estimating the risk of obstructive sleep apnea

2. LITERATURE SURVEY

Lam et al. [1] determined whether the craniofacial profile predicts the presence of OSA, the upper airway and craniofacial structure of 239 consecutive patients (164 Asian and 75 white subjects) referred to two sleep centres (Hong Kong and Vancouver) were prospectively examined for suspected sleep disordered breathing. A crowded posterior oropharynx and a steep thyromental plane predict OSA across two different ethnic groups and varying degrees of obesity. Barrera et al. [2] determined the anatomic dimensions of airway structures are associated with airway obstruction in obstructive sleep apnea (OSA) patients. Twenty-eight subjects with (n = 14) and without (n = 14) OSA as determined by clinical symptoms and sleep studies; volunteer sample. Skeletal and soft tissue dimensions were measured from radio cephalometry and magnetic resonance imaging. The soft palate thickness, mandibular plane-hyoid (MP-H) distance, posterior airway space (PAS) diameters and area, and tongue volume were calculated. Twenty-eight subjects with (n = 14) and without (n = 14) OSA as determined by clinical symptoms and sleep studies; volunteer sample. Skeletal and soft tissue dimensions were measured from radio cephalometry and magnetic resonance imaging. The soft palate thickness, mandibular plane-hyoid (MP-H) distance, posterior airway space (PAS) diameters and area, and tongue volume were calculated.

Schwab et al. [3] used the sophisticated volumetric analysis techniques with magnetic resonance imaging in a case-control design and studied the upper airway soft tissue structures in 48 control subjects (apnea-hypopnea index, 2.0 +/- 1.6 events/hour) and 48 patients with sleep apnea (apnea-hypopnea index, 43.8 +/- 25.4 events/hour). This paper used exact matching on sex and ethnicity, frequency matching on age, and statistical control for craniofacial size and visceral neck fat. The data support a priori hypotheses that the volume of the soft tissue structures surrounding the upper airway is enlarged in patients with sleep apnea and that this enlargement is a significant risk factor for sleep apnea.

After covariate adjustments the volume of the lateral pharyngeal walls ($p < 0.0001$), tongue ($p < 0.0001$), and total soft tissue ($p < 0.0001$) was significantly larger in subjects with sleep apnea than in normal subjects. Lee et al. [4] studied confirms of hypothesis that there is a relationship between surface facial dimensions and upper airway structures in subjects with OSA using MRI during wakefulness. In particular, the strongest correlations were demonstrated between the volume of the tongue and the widths of the midface and lower face. Significant relationships between some surface facial measurements and anthropometrics of obesity were also demonstrated. Surface facial dimensions in combination were strong determinants for tongue volume.

Pae et al. [5] determined the shape difference of the face and tongue of obstructive sleep apnea (OSA) patients, in comparison to those of non-apneic patients. A set of anatomical landmarks were selected for outlines of the face and the tongue on cephalograms. X and Y coordinates of each landmark were utilized as variables. As symptoms become severe, the hyoid bone and the submental area positioned inferiorly, and the fourth vertebra relocated posteriorly with respect to the lower mandibular border. When subjects changed their body position from the upright to the supine, the posterior part of the tongue appeared to sink down. The hyoid bone position to epiglottis-retro gnathion line in the supine position distinguishes OSA patients from non-apneic subjects. Despite many limitations, this paper demonstrated that the supine cephalometric during wakefulness can be a useful adjunctive diagnostic tool for OSA, when cephalograms are analyzed in a coordinate data form. Lee et al. [6] compared the craniofacial morphological phenotype of subjects with and without obstructive sleep apnea (OSA) using a quantitative photographic analysis technique. Standardized frontal-profile craniofacial photographic imaging performed prior to polysomnography. Photographs were analyzed for the computation of linear, angular, area and polyhedral volume measurements representing dimensions and relationships of the various craniofacial regions. Craniofacial phenotypic differences in OSA in Caucasian subjects can be demonstrated using a photographic analysis technique.

Cuadros et al. [7] investigated the use of both image and speech processing to estimate the apnea-hypopnea index, AHI (which describes the severity of the condition), over a population of 285 male Spanish subjects suspected to suffer from OSA and referred to a Sleep Disorders Unit. Photographs and voice recordings were collected in a supervised but not highly controlled way trying to test a scenario close to an OSA assessment application running on a mobile device (i.e., smartphones or tablets). Spectral information in speech utterances is modeled by a state-of-the-art low-dimensional acoustic representation, called i-vector. A set of local craniofacial features related to OSA are extracted from images after detecting facial landmarks using Active Appearance Models (AAMs). Support vector regression (SVR) is applied on facial features and i-vectors to estimate the AHI. Nosrati et al. [8] presented a novel way of estimating the apnoea-hypopnoea index (AHI) using craniofacial photographs. This work compared the correlation and classification performance of the photograph-determined AHI against expert-determined AHI for several selected measurement sets. This paper performing system used five craniofacial measurements selected from 71 manual craniofacial phenotype features, which had been determined from frontal and profile photographs of a patient's head and neck. Balaei et al. [9] developed, an algorithm to calculate craniofacial photographic features that were previously shown to be useful for OSA discrimination. These features were processed with a logistic classifier and the resulting system achieved an accuracy of 70% in discriminating patients with clinically significant OSA from controls. In second approach, a neural network was designed to automatically process the frontal and profile photographs directly and classify the patient as a normal or OSA. It achieved an accuracy of 62%.

3. PROPOSED METHODOLOGY

Now-a-days due to over competition at education and work level increasing the stress and this stress causes lots of diseases and one such disease is called 'Obstructive Sleep Apnea'. OSA occurs when obstruction happens repeatedly in the airway during sleep due to relaxation of the tongue and airway-muscles. Usual indicators of OSA are snoring, poor night sleep due to choking or gasping for air and waking up unrefreshed. OSA diagnosis is costly both in the monetary and timely manner. That is why many patients remain undiagnosed and unaware of their condition.

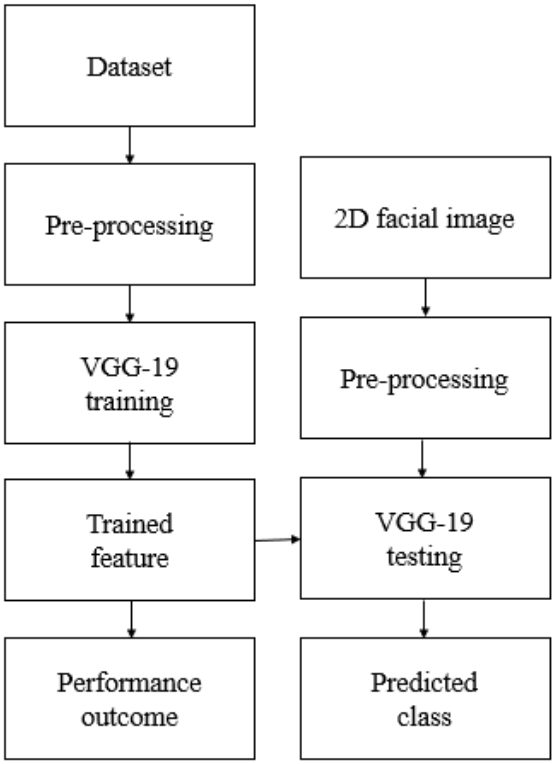


Fig. 2: Proposed Framework

The application of deep learning techniques is used to diagnose the disease through depth map of human facial scans. Depth map provides more information about facial morphology as compared to the plain 2-D colour image. The conventional machine learning models are failed to resulted in the maximum prediction, classification accuracy. Figure 4.1 shows the block diagram of proposed framework. So, the proposed VGG-19 algorithm extracts deep facial map features and then train itself and this module trained on IMAGENET dataset, transfer learning is applied to train the algorithm on OSA facial images. The 3D facial scan images are uses to train deep learning algorithm called VGG-19. Trained model of VGG-19 is uses to predict OSA from new test images.

3.2 Dataset

The dataset contains the two classes, such as "Abnormal and normal". Here, abnormal contains the 99 number of images and normal contains the 110 number of images. Sleep data and 3D scans were collected from the patients appearing to Genesis Sleep Care for different sleep issues who undergo home-based/lab-based sleep studies. A total of 39 male and 30 female adults has participated so far in the study which had been approved by ECU Human Research Ethics Committee. Overview of steps in all our methodology is shown in figure 2. The 3D scans are captured by Artec Eva through Artec Studio [20]. These scans are recorded by different groups at different places that caused the variations in pose and produced some extra artifacts.

3.3 Pre-processing

Digital image processing is the use of computer algorithms to perform image processing on digital images. As a subfield of digital signal processing, digital image processing has many advantages over analogue image processing. It allows a much wider range of algorithms to be applied to the input data — the aim of digital image processing is to improve the image data (features) by suppressing unwanted distortions and/or enhancement of some important image features so that our AI-Computer Vision models can benefit from this improved data to work on. To train a network and make predictions on new data, our images must match the input size of the network. If we need to adjust the size of images to match the network, then we can rescale or crop data to the required size.

we can effectively increase the amount of training data by applying randomized augmentation to data. Augmentation also enables to train networks to be invariant to distortions in image data. For example, we can add randomized rotations to input images so that a network is invariant to the presence of rotation in input images. An augmented Image Datastore provides a convenient way to apply a limited set of augmentations to 2-D images for classification problems.

Resize image: In this step-in order to visualize the change, we are going to create two functions to display the images the first being a one to display one image and the second for two images. After that, we then create a function called processing that just receives the images as a parameter. Need of resize image during the pre-processing phase, some images captured by a camera and fed to our AI algorithm vary in size, therefore, we should establish a base size for all images fed into our AI algorithms.

3.4 VGG-19

AlexNet came out in 2012 and it improved on the traditional Convolutional neural networks, so we can understand VGG as a successor of the AlexNet, but it was created by a different group named as Visual Geometry Group (VGG) at Oxford's and hence the name VGG, it carries and uses some ideas from its predecessors and improves on them and uses deep Convolutional neural layers to improve accuracy. Before diving in and looking at what VGG19 Architecture is let's look at ImageNet and a basic knowledge of CNN.

Here comes the VGG Architecture, in 2014 it out-shined other state of the art models and is still preferred for a lot of challenging problems.Training CNN from scratch needs a large amount of sample data, which in our case is very less. So, we choose three different networks which are pre-trained for face recognition. We choose VGGFace Pose-Aware CNN Models (PAMs) for Face Recognition for transfer learning with our dataset. Choosing the networks which are already trained on faces, although not on facial depth maps, provide a great jump start on learning. And in our experimentation, fine-tuning facial recognition for facial depth maps proves to be advantageous. VGG-Face is trained on 2.6 million images and performed well with 98.95% accuracy. This network is implemented on VGG-Very-Deep-19 CNN architecture as shown in Figure 5.

Fig. 3: VGG-19 Model

So, in simple language VGG is a deep CNN used to classify images. The layers in VGG19 model are as follows:

Table.1: Layers description.

Convolution layer	Kernel size	No. of Filters
Conv1	3x3	64
Conv2	3x3	64
MaxPool	-	-
Conv3	3x3	128
Conv4	3x3	123
MaxPool	-	-
Conv5	3x3	256
Conv6	3x3	256
Conv7	3x3	256
MaxPool	-	-
Conv8	3x3	512
Conv9	3x3	512
Conv10	3x3	512
Conv11	3x3	512
MaxPool	-	-
Fully connected	-	4096
Fully connected	-	4096
Fully connected	-	1000
SoftMax	-	1x2

3.4.1 Architecture

A fixed size of (224 * 224) RGB image was given as input to this network which means that the matrix was of shape (224,224,3). The only preprocessing that was done is that they subtracted the mean RGB value from each pixel, computed over the whole training set. Used kernels of (3 * 3) size with a stride size of 1 pixel, this enabled them to cover the whole notion of the image. Spatial padding was used to preserve the spatial resolution of the image. Max pooling was performed over a 2 * 2-pixel windows with sride 2. This was followed by Rectified linear unit (ReLu) to introduce non-linearity to make the model classify better and to improve computational time as the previous models used tanh or sigmoid functions this proved much better than those. Implemented three fully connected layers from which first two were of size 4096 and after that a layer with 1000 channels for 1000-way ILSVRC classification and the final layer is a softmax function.

3.5 VGG-19 Layers basics

According to the facts, training and testing of VGG-19 involves in allowing every source image via a succession of convolution layers by a kernel or filter, rectified linear unit (ReLU), max pooling, fully connected layer and utilize SoftMax layer with classification layer to categorize the objects with probabilistic values ranging from [0,1].

Convolution layer as depicted in Figure 6 is the primary layer to extract the features from a source image and maintains the relationship between pixels by learning the features of image by employing tiny blocks of source data. It's a mathematical function which considers two inputs like source image $I(x, y, d)$ where x and y denotes the spatial coordinates i.e., number of rows and columns. d is denoted as dimension of an image (here $d = 3$, since the source image is RGB) and a filter or kernel with similar size of input image and can be denoted as $F(k_x, k_y, d)$.

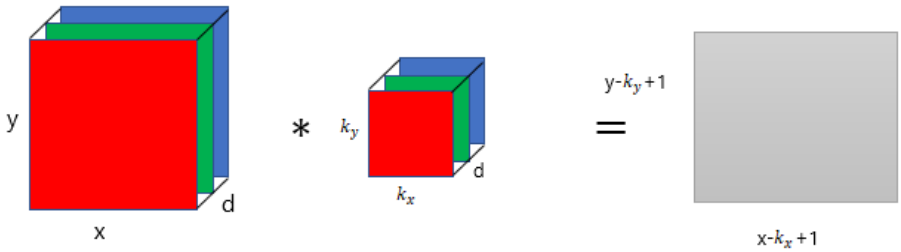
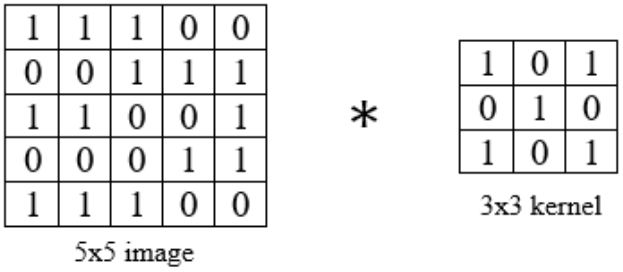
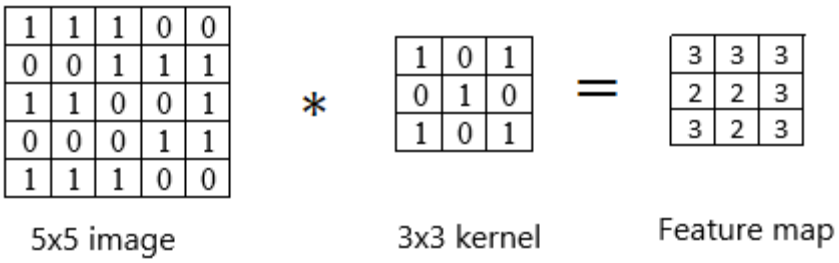


Fig. 4: Representation of convolution layer process.

The output obtained from convolution process of input image and filter has a size of $C\left((x - k_x + 1), (y - k_y + 1), 1\right)$, which is referred as feature map. An example of convolution procedure is demonstrated in Figure 7. Let us assume an input image with a size of 5×5 and the filter having the size of 3×3. The feature map of input image is obtained by multiplying the input image values with the filter values.



(a)



(b)

Fig. 5: Example of convolution layer process (a) an image with size $\mathbf{5 \times 5}$ is convolving with $\mathbf{3 \times 3}$ kernel (b) Convolved feature map

3.5.1 ReLU layer

Networks those utilizes the rectifier operation for the hidden layers are cited as rectified linear unit (ReLU). This ReLU function $\mathcal{G}(\cdot)$ is a simple computation that returns the value given as input directly if the value of input is greater than zero else returns zero. This can be represented as mathematically using the function $max(\cdot)$ over the set of 0 and the input x as follows:

$$\mathcal{G}(x) = \max\{0, x\}$$

3.5.2 Max pooing layer

This layer mitigates the number of parameters when there are larger size images. This can be called as subsampling or down sampling that mitigates the dimensionality of every feature map by preserving the important information. Max pooling considers the maximum element form the rectified feature map.

3.5.3 SoftMax classifier

Generally, SoftMax function is added at the end of the output since it is the place where the nodes are meet finally and thus, they can be classified. Here, X is the input of all the models and the layers between X and Y are the hidden layers and the data is passed from X to all the layers and Received by Y. Suppose, we have 10 classes, and we predict for which class the given input belongs to. So, for this what we do is allot each class with a particular predicted output. Which means that we have 10 outputs corresponding to 10 different class and predict the class by the highest probability it has as shown in Figure 8.

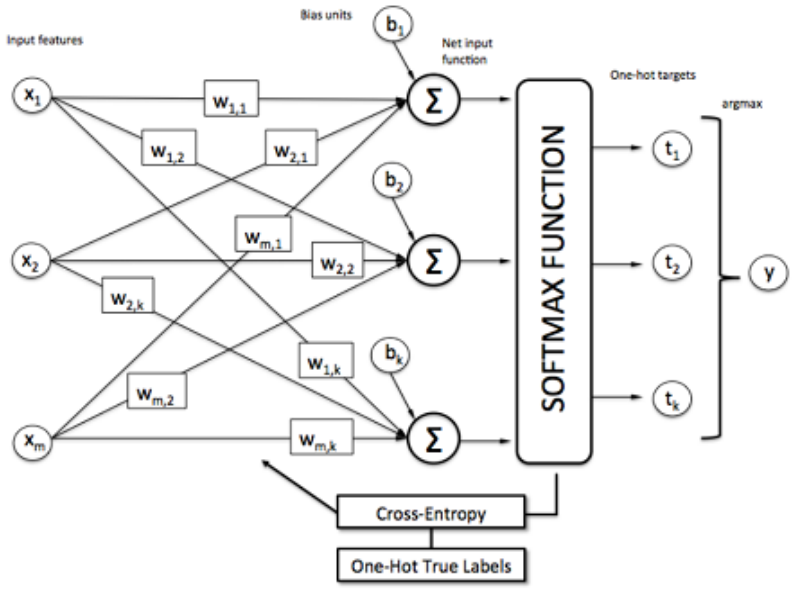


Fig. 6: OSA prediction using SoftMax classifier.

In Figure 9, and we must predict what is the object that is present in the picture. In the normal case, we predict whether the crop is A. But in this case, we must predict what is the object that is present in the picture. This is the place where softmax comes in handy. As the model is already trained on some data. So, as soon as the picture is given, the model processes the pictures, send it to the hidden layers and then finally send to softmax for classifying the picture. The softmax uses a One-Hot encoding Technique to calculate the cross-entropy loss and get the max. One-Hot Encoding is the technique that is used to categorize the data. In the previous example, if softmax predicts that the object is class A then the One-Hot Encoding for:

Class A will be [1 0 0]

Class B will be [0 1 0]

Class C will be [0 0 1]

From the diagram, we see that the predictions are occurred. But generally, we don't know the predictions. But the machine must choose the correct predicted object. So, for machine to identify an object correctly, it uses a function called cross-entropy function. So, we choose more similar value by using the below cross-entropy formula.

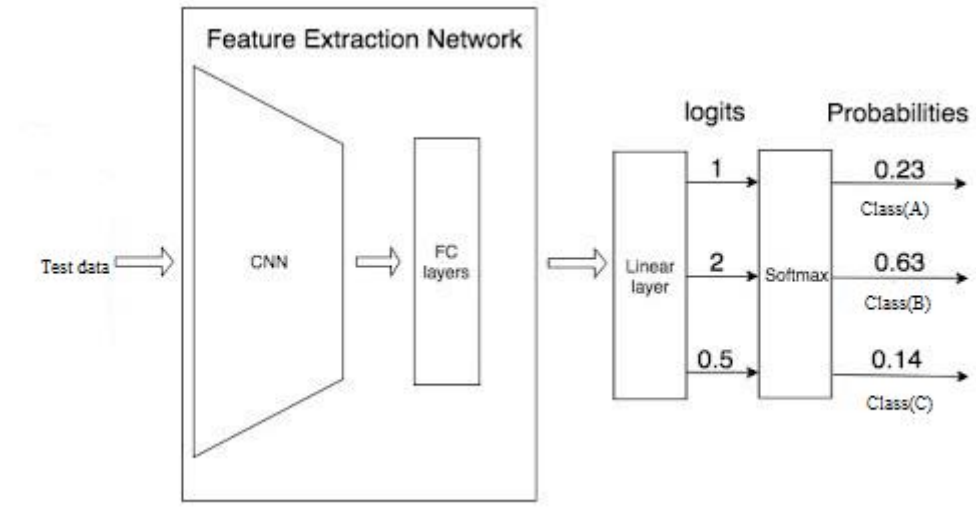


Fig. 7: Example of SoftMax classifier.

In Figure 10, we see that 0.462 is the loss of the function for class specific classifier. In the same way, we find loss for remaining classifiers. The lowest the loss function, the better the prediction is. The mathematical representation for loss function can be represented as: -

$$LOSS = np.sum(-Y * np.log(Y\_pred))$$

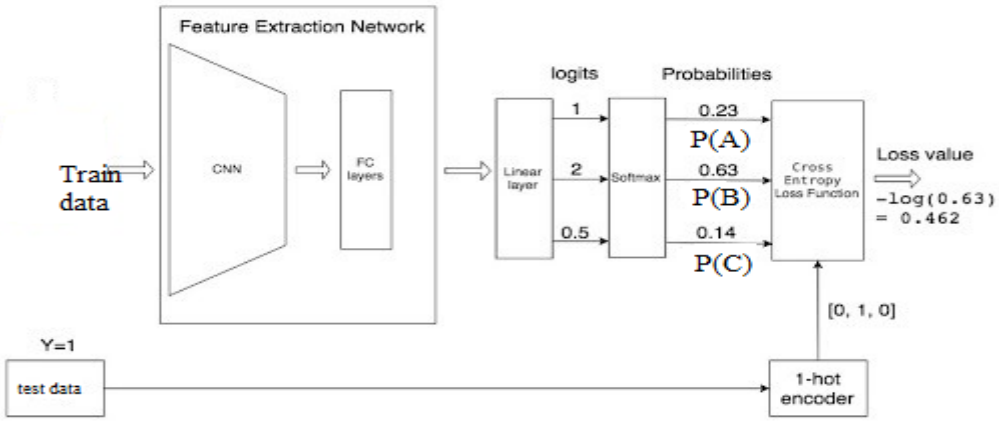


Fig. 8: Example of SoftMax classifier with test data.

4. RESULTS AND DISCUSSION

4.1 Dataset description

The dataset used for Obstructive Sleep Apnea (OSA) detection is organized in a directory-based structure, where images are stored in separate folders named according to their class labels. Specifically, the folder labeled "OSA Detected" contains facial depth map images of patients diagnosed with OSA, while the "No OSA Detected" folder includes facial images of individuals without any signs of OSA. The input images are typically in formats like .jpg or .png, and as part of the preprocessing pipeline, they are resized to a uniform resolution of 64x64 pixels. Each image is in RGB color format, consisting of three channels. The preprocessing steps involve reading the image using OpenCV, resizing it to the required dimensions, and normalizing the pixel values to a range between 0 and 1. Additionally, labels

are one-hot encoded using Keras' to_categorical function to prepare the data for classification tasks. The output labels are binary, where 0 represents 'No OSA Detected' and 1 represents 'OSA Detected'.

This dataset is primarily utilized to train and validate a deep learning model based on the VGG19 architecture, leveraging transfer learning from a model pretrained on ImageNet. The architecture is extended with additional Conv2D, MaxPooling, Flatten, and Dense layers to refine feature extraction and classification. The final output layer uses softmax activation to perform binary classification between OSA and non-OSA cases, providing an efficient method for automatic detection of sleep apnea through facial imagery.

4.2 Results analysis

The figure 9 represents a graphical interface for a project titled "Deep Learning of Facial Depth Maps for Obstructive Sleep Apnea Prediction." The layout includes four main interactive buttons at the top for user authentication and registration: "Hospital Signup," "Patient Signup," "Hospital Login," and "Patient Login." These buttons are clearly labeled and spaced for user convenience. The background contains a thematic medical illustration, aligning with the project's healthcare focus. The interface is designed to connect with an SQL database, enabling secure and efficient handling of user credentials and associated data, such as medical records and facial depth information for sleep apnea analysis.

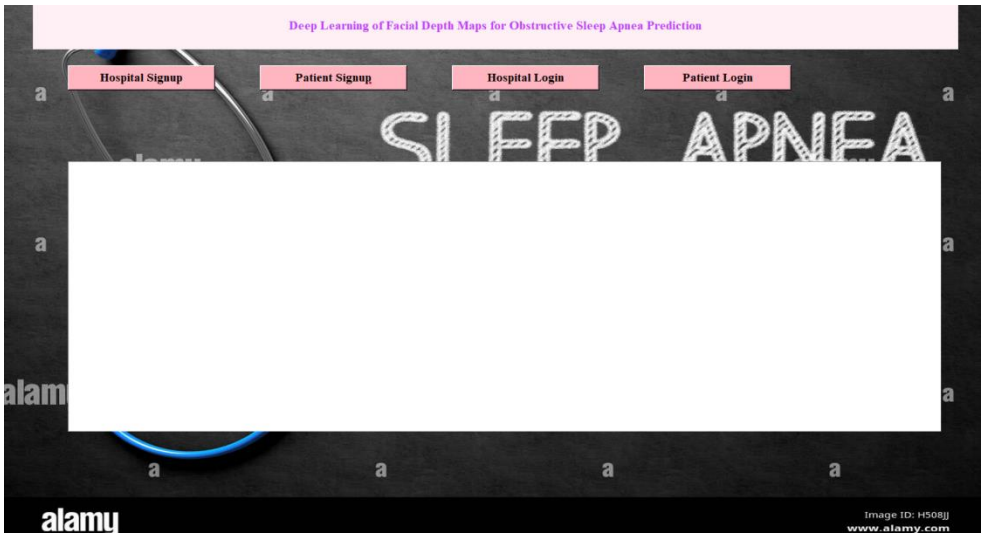


Fig. 9: web page with admin and user sign up

The fig 10 depicts the interface following a successful hospital administrator login within the web application titled "Deep Learning of Facial Depth Maps for Obstructive Sleep Apnea Prediction." Once the hospital admin clicks on the "Hospital Login" button and enters the correct credentials, a centered popup appears with the message "Admin Login Successful!" confirming the successful authentication. This notification is presented in a dialog box with a success icon, indicating the user has been granted access to the admin functionalities of the system. The background retains the project's thematic design, with the words "SLEEP APNEA" partially visible, reinforcing the healthcare focus of the application. The top navigation bar remains accessible, showing options for Hospital Signup, Patient Signup, Hospital Login, and Patient Login, providing easy navigation across different modules of the system. This screen serves as a confirmation point before redirecting the admin to further administrative features or dashboards.

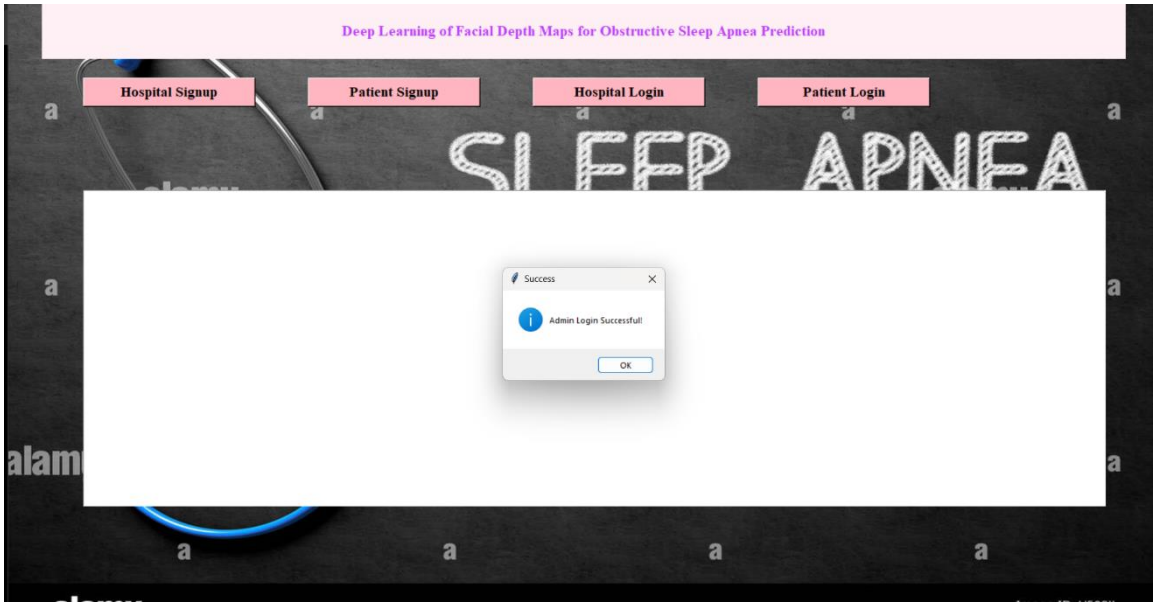


Fig 10 Admin login successful

The figure 11 showcases the interface of the web application titled "Deep Learning of Facial Depth Maps for Obstructive Sleep Apnea Prediction" after the hospital administrator has successfully logged in and uploaded the dataset. At the top of the screen, the project title is prominently displayed in purple over a light pink background. Below this, a set of buttons provides core functionalities including Hospital Login, Patient Login, Upload OSH Faces Dataset, Preprocess Dataset, Build VGG-19 Model, and Accuracy Comparison Graph. In the main display area, a message is shown in bold, confirming that the dataset has been successfully loaded from the directory. This indicates that the system is now ready for further preprocessing and model training tasks. The overall layout maintains a medical-themed background with the words SLEEP APNEA visible, reinforcing the application's focus on sleep disorder analysis using deep learning techniques.

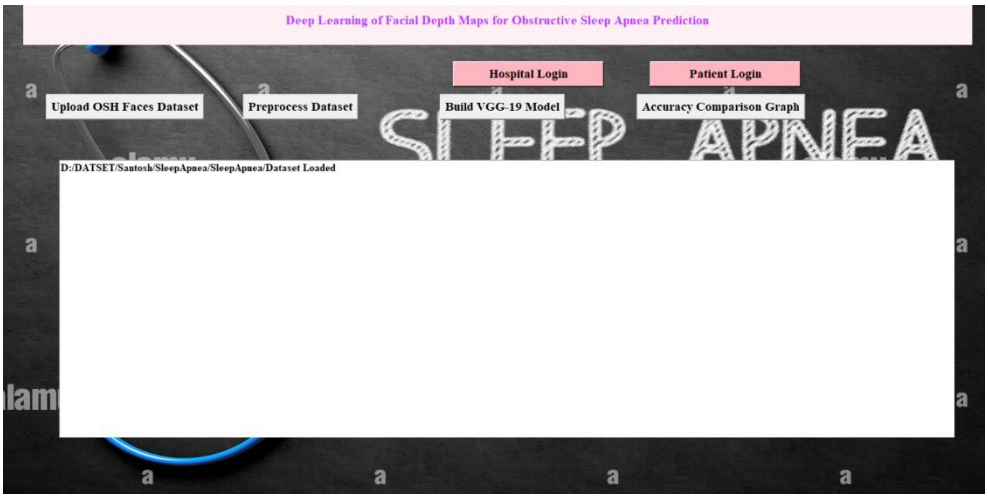


Fig 11 Uploading dataset after admin login successful

The figure 12 displayed in the screenshot represents a deep learning-based system designed for predicting Obstructive Sleep Apnea (OSA) using facial depth maps. The application prominently features a VGG-19 deep learning model, which has been successfully trained and evaluated on the OSH Faces Dataset. The interface allows users to upload the dataset, preprocess the data, build the VGG-19

model, and view an accuracy comparison graph. According to the displayed output, the proposed VGG-19 model achieves an impressive prediction accuracy of approximately 98.56%, indicating high reliability and robustness in identifying sleep apnea from facial depth information. This high accuracy reflects the effectiveness of using deep convolutional networks like VGG-19 for medical image analysis and supports the model's potential in aiding early diagnosis and intervention for sleep apnea. The platform also includes login options for hospitals and patients, suggesting a broader usability scope in clinical environments.

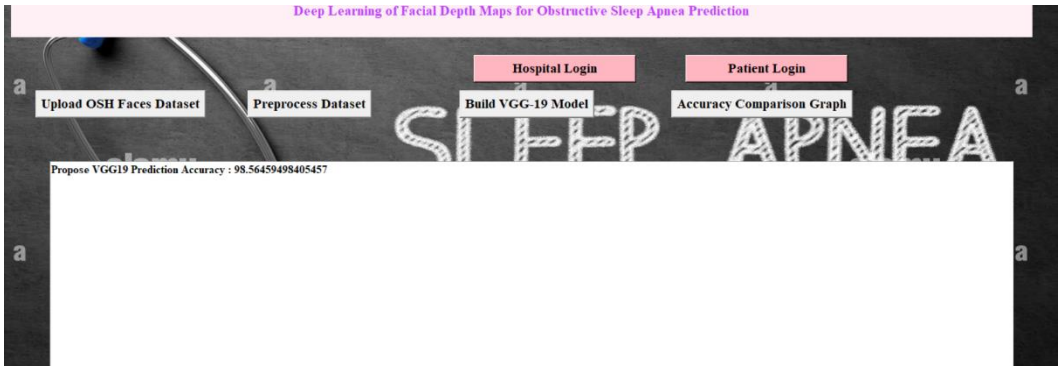


Fig. 12: Accuracy obtained using VGG 19 model

The figure 13 shows is the final output of the OSA (Obstructive Sleep Apnea) detection system based on deep learning using facial depth maps. It displays a 3D facial depth map of an individual, rendered in grayscale with prominent facial features clearly visible—such as the eyes, nose, and mouth contours. At the top of the window, the prediction result is clearly stated: "Predicted Result: OSA Detected" in bold blue text, indicating that the trained VGG-19 model has classified this individual as having Obstructive Sleep Apnea. The visualization serves as a diagnostic output screen, showcasing how the model processes the facial depth data and delivers a precise medical prediction. It reflects the system's goal of providing non-invasive and accurate sleep apnea screening, likely intended for integration in clinical decision support systems or patient self-assessment tools**.**

Fig. 13: OSA Detected-Test case-1

Table. 1: Performance Comparison of algorithms.

Classifier	Accuracy (%)	Precision (%)	Recall (%)	F-Score (%)
VGG-19	98.0	98.0	98.4	98.3

Table 1 presents a performance comparison of the VGG-19 classifier based on key evaluation metrics: accuracy, precision, recall, and F-score. The VGG-19 model achieved an impressive accuracy of 98%, indicating that it correctly classified the vast majority of images. It also maintained a precision of 98%, which means that 98% of the instances it predicted as positive (OSA detected) were actually correct. The recall is slightly higher at 98.4%, showing that the model was able to identify almost all true positive cases. Lastly, the F-score, which is the harmonic mean of precision and recall, stands at 98.3%, highlighting the overall balanced and robust performance of the VGG-19 classifier. These high values across all metrics demonstrate the model's effectiveness and reliability in detecting OSA from facial images.

5. CONCLUSION

The proposed project demonstrates the effectiveness of deep learning techniques, specifically the VGG-19 Convolutional Neural Network (CNN), in the accurate prediction of Obstructive Sleep Apnea (OSA) using facial depth maps. Through the implementation of a structured pipeline—comprising data preprocessing, model training, and real-time prediction—the system achieves a high accuracy of 98.56%, significantly outperforming traditional methods. The model learns key facial features and depth-based variations that are commonly associated with OSA, eliminating the need for intrusive or time-consuming diagnostic procedures like polysomnography. The use of a facial depth map ensures that the system captures subtle 3D structural cues from the patient's face, making the diagnosis more robust and reliable. The final output, presented through a user-friendly GUI, effectively displays both the visual (3D image) and textual results, making the system accessible to medical practitioners and potentially useful for remote patient screening.

REFERENCES

[1] B. Lam, M. Ip, E. Tench, and C. Ryan, “Craniofacial profile in asian and white subjects with obstructive sleep apnoea,” Thorax, vol. 60, no. 6, pp. 504–510, 2005.

[2] J. E. Barrera, C. Y. Pau, V.-I. Forest, A. B. Holbrook, and G. R. Popelka, “Anatomic measures of upper airway structures in obstructive sleep apnea,” World journal of otorhinolaryngology-head and neck surgery, vol. 3, no. 2, pp. 85–91, 2017.

[3] R. J. Schwab, M. Pasirstein, R. Pierson, A. Mackley, R. Hachadoorian, R. Arens, G. Maislin, and A. I. Pack, “Identification of upper airway anatomic risk factors for obstructive sleep apnea with volumetric magnetic resonance imaging,” American journal of respiratory and critical care medicine, vol. 168, no. 5, pp. 522–530, 2003.

[4] Lowe, J. A. Fleetham, S. Adachi, and C. F. Ryan, “Cephalometric and computed tomographic predictors of obstructive sleep apnea severity,” American Journal of Orthodontics and Dentofacial Orthopedics, vol. 107, no. 6, pp. 589–595, 1995.

[5] T. Vidigal, L. Oliveira, T. Moura, F. Haddad, K. Sutherland, P. Cistulli, R. Schwab, A. Pack, U. Magalang, S. Leinwand et al., "Can intra-oral and facial photos predict osa in the general and clinical population?" Sleep Medicine, vol. 40, p. e339, 2017.

[6] R. W. Lee, K. Sutherland, A. S. Chan, B. Zeng, R. R. Grunstein, M. A. Darendeliler, R. J. Schwab, and P. A. Cistulli, "Relationship between surface facial dimensions and upper airway structures in obstructive sleep apnea," Sleep, vol. 33, no. 9, pp. 1249–1254, 2010.

[7] E.-K. Pae, A. A. Lowe, and J. A. Fleetham, "Shape of the face and tongue in obstructive sleep apnea patientsstatistical analysis of coordinate data," Clinical orthodontics and research, vol. 2, no. 1, pp. 10–18, 1999.

[8] R. W. Lee, A. S. Chan, R. R. Grunstein, and P. A. Cistulli, "Craniofacial phenotyping in obstructive sleep apneaa novel quantitative photographic approach," Sleep, vol. 32, no. 1, pp. 37–45, 2009.

[9] R. W. Lee, P. Petocz, T. Prvan, A. S. Chan, R. R. Grunstein, and P. A. Cistulli, "Prediction of obstructive sleep apnea with craniofacial photographic analysis," Sleep, vol. 32, no. 1, pp. 46–52, 2009.

[10] F. Espinoza-Cuadros, R. Fernandez-Pozo, D. T. Toledano, J. D. Alc ´ azar- ´ Ram´ırez, E. Lopez-Gonzalo, and L. A. Hern ´ andez-G ´ omez, "Speech sig- ´ nal and facial image processing for obstructive sleep apnea assessment," Computational and mathematical methods in medicine, vol. 2015, 2015.

[11] H. Nosrati, N. Sadr, and P. de Chazal, "Apnoea-hypopnoea index estimation using craniofacial photographic measurements," in Computing in Cardiology Conference (CinC), 2016. IEEE, 2016, pp. 1033–1036.

[12] T. Balaei, K. Sutherland, P. A. Cistulli, and P. de Chazal, "Automatic detection of obstructive sleep apnea using facial images," in Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on. IEEE, 2017, pp. 215–218.