Facial Depth Maps: Detecting Sleep Apnea with Deep Learning

L. Crownie Majestica[1], Akshay Singh[2], Bhavya Gajam[3],B. Rishitha[4], Chandra Shekar[5], Dr. A. Satyanarayana[6]

[1,2,3,4,5]Student , [6]Associate Professor
Department of Artificial Intelligence and Machine Learning[1,2,3,4,5,6]
Siddhartha Institute of Technology and Sciences, Hyderabad ,Telangana, India[1,2,3,4,5,6]

Abstract: Obstructive Sleep Apnea (OSA) presents a significant health concern, characterized by repetitive airway obstruction during sleep, often leading to symptoms such as snoring, disrupted sleep, and daytime fatigue. Despite its prevalence, OSA diagnosis remains challenging and costly, contributing to under diagnosis and untreated cases. This study explores the potential of deep learning techniques applied to facial depth maps for OSA diagnosis, leveraging the relationship between facial morphology and the condition. By utilizing depth maps, which offer enhanced spatial information compared to 2D images, we aim to improve diagnostic accuracy. Our approach incorporates transfer learning to achieve promising results with limited data, achieving an efficient validation accuracy. We focus on predicting OSA severity, categorizing patients into above-moderate and below-moderate groups based on an apnea-hypopnea index threshold of 15. This research signifies a step toward non-invasive and efficient OSA diagnosis, potentially facilitating timely interventions and improved patient outcomes

Index Terms: Obstructive Sleep Apnea, Deep learning, VGG-19

I. INTRODUCTION

OSA disrupts the normal breathing patterns during sleep, leading to recurrent episodes of upper airway obstruction lasting over 10 seconds each. These interruptions deprive the lungs of oxygen, prompting the individual to awaken briefly to restore airflow. Diagnosis of OSA typically occurs when more than 15 such apneas transpire during sleep. The diagnostic process entails a comprehensive assessment encompassing patient history, physical examination, and specialized tests like polysomnography. PSG serves as the gold standard for diagnosis, requiring individuals to undergo monitoring in a hospital unit equipped with sensors that track breathing patterns, oxygen levels, heart rate, and body movements.

Amidst the intricate web of diagnostic procedures and economic burdens, the personal toll of OSA and other sleep disorders looms large. Individuals grappling with poor sleep quality often find themselves navigating a labyrinth of daytime fatigue, impaired concentration, and mood disturbances. Moreover, the ripple effects extend beyond the individual, impacting interpersonal relationships, professional obligations, and overall quality of life. Thus, understanding the multifaceted impact of sleep disorders is imperative, not only for individuals but also for healthcare systems and economies at large

II. LITERATURE REVIEW

Poor sleep imparts a significant personal and societal burden. Therefore, it is important to have accurate estimates of its causes, prevalence and costs to inform health policy. A recent evaluation of the sleep habits of Australians demonstrates that frequent (daily or near daily) sleep difficulties (initiating and maintaining sleep, and experiencing inadequate sleep), daytime fatigue, sleepiness and irritability are highly prevalent (20%-35%). These difficulties are generally more prevalent among females, with the exception of snoring and related difficulties. While about half of these problems are likely to be attributable to specific sleep disorders, the balance appears attributable to poor sleep habits or choices to limit sleep opportunity. Study of the economic impact of sleep disorders demonstrates financial costs to Australia of $5.1 billion per year. This comprises $270 million for health care costs for the conditions themselves, $540 million for care of associated medical conditions attributable to sleep disorders, and about $4.3 billion largely attributable to associated productivity losses and non-medical costs resulting from sleep loss-related accidents. Loss of life quality added a substantial further non-financial cost. While large, these costs were for sleep disorders alone. Sleep apnea (SA) is a common sleep disorder that is not easy to detect. Recent studies have highlighted ECG analysis as an effective method of diagnosing SA. Because the changes caused by SA on the ECG are imperceptible, the need for new methods in diagnosing this disease is required more than ever.Machine Learning (ML) is recognized as one of the most successful methods of computer aided diagnosis. ML uses new methods to diagnose diseases using past clinical results. The purpose of this study is to evaluate studies using ML algorithms based on ECG characteristics to assess people suffering from SA. In this study, systematically-reviewed articles written in English before October 2020 and indexed in PubMed, Scopus, Web of Science, and IEEE databases were searched with no lower time limit. From these articles, 48 were selected for further review. The selected articles adopted different ML methods for classification. All of these studies were binary where SA was detected from the normal state based on a full ECG stripe (per record), or based on one-minute segments (per segment). Our analysis show that the most common features used in the studies were frequency, time series, and statistical features. Support-Vector Machine (SVM) and deep learning-based neural network (i.e. CNN, DNN) performed best in full record data detection. The highest accuracy, sensitivity, and specificity reported among the selected studies were 100%, which was obtained by an SVM. In another study, the classification was conducted based on ECG segments, and accordingly, the highest classification accuracy was observed in the residual neural network algorithm (RNN). The accuracy, sensitivity, and specificity of this algorithm were reported to be 99%. In general, it can be stated that ML techniques based on ECG characteristics have a high capability in diagnosing SA. These techniques can increase the diagnosis of patients with SA or the detection of SA episodes on ECG record, and can potentially prevent complications of the disease at later stages.

III. METHODOLOGY

- **Data Collection and Data Pre-Processing**

Use depth-sensing cameras to capture facial depth maps during sleep.Place the camera in a position where it can capture the subject’s face clearly while they sleep. Ensure consistent lighting and minimal movement to avoid data noise. Effective in removing salt-and-pepper noise. Smoothens the image while preserving edges. Use algorithms to isolate the face from the background.

- **Model Selection**

Convolutional Neural Networks (CNNs) used for spatial feature extraction. Typical CNN layers include convolutional layers, pooling layers, and fully connected layers. Individual frames or short sequences of depth maps. LSTM layers can be stacked and combined with CNN layers for hybrid models. Particularly effective in handling long-term dependencies in time series data. VGG-19 is a deep convolutional neural network (CNN) architecture that was developed by the Visual Geometry Group (VGG).

- **Parameter**

VGG-19 has a large number of parameters, primarily due to its deep architecture and the large number of filters in the convolutional layers. Training such a deep network requires a substantial amount of computational resources and data.

- **Model Training and Evaluation**

Manual annotation or using signals from polysomnography (PSG) for ground truth.Binary cross-entropy for binary classification (apnea vs. normal). Adam or RMSprop optimizers for training the neural network. Use k-fold cross-validation to ensure robustness and prevent overfitting. Analyze false positives and false negatives. Evaluate using accuracy, precision, recall.

- **Feature Extraction**

Monitor vertical and horizontal movements of facial landmarks over time to infer breathing. Identify key points on the face using algorithms like Dlib's facial landmark detector. Points such as the corners of the eyes, tip of the nose, corners of the mouth. Track the rise and fall pattern which corresponds to inhalation and exhalation cycles. Combining facial depth maps with physiological data from wearables can enhance accuracy and provide a comprehensive monitoring system.

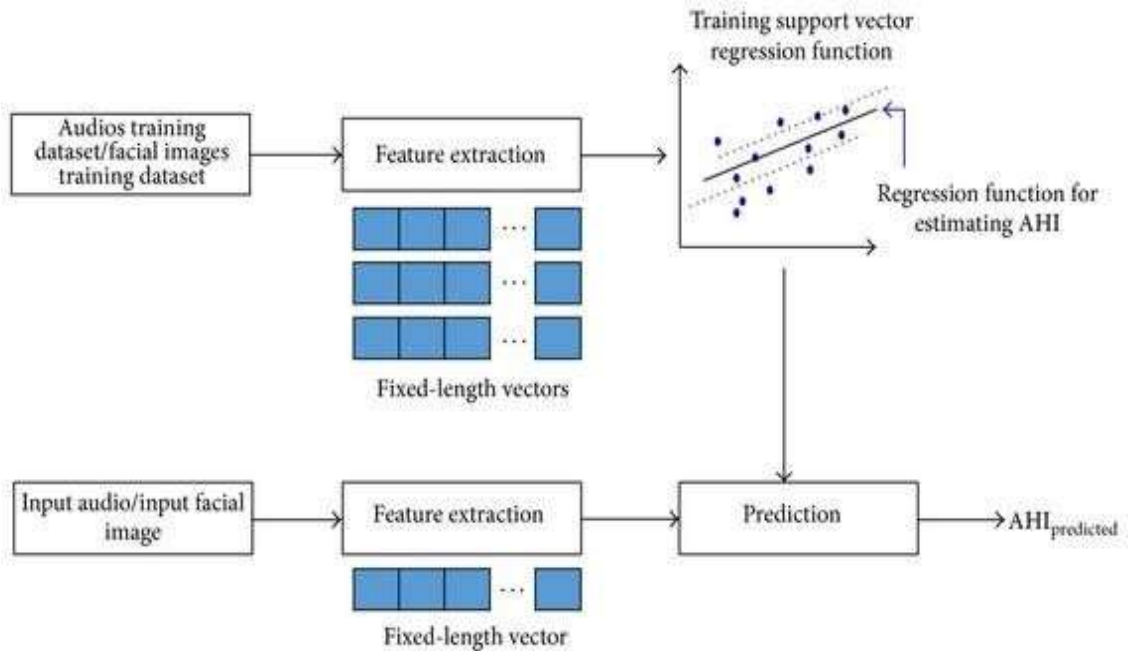
- **Prediction**

The trained model predicted OSA severity, categorizing patients into above-moderate (AHI ≥ 15) and below-moderate (AHI < 15) groups. The hybrid CNN-LSTM model leveraged temporal sequences of depth maps to predict apnea events more accurately over the entire sleep period. The model's prediction capabilities were tested for potential real- time monitoring applications, providing immediate feedback and intervention options.

- **Integration with Clinical Workflow**

Monitor vertical and horizontal movements of facial landmarks over time to infer breathing. Identify key points on the face using algorithms like Dlib's facial landmark detector. Points such as the corners of the eyes, tip of the nose, corners of the mouth. Track the rise and fall pattern which corresponds to inhalation and exhalation cycles. Combining facial depth maps with physiological data from wearables can enhance accuracy and provide a comprehensive monitoring system.

IV. SYSTEM DESIGN



V. IMPLEMENTATION

VGG-19 is a deep convolutional neural network (CNN) architecture that was developed by the Visual Geometry Group (VGG) at the University of Oxford. It is a variant of the VGG network, which was originally proposed for image classification tasks. VGG-19 is characterized by its deep architecture consisting of 19 layers, including convolutional layers, pooling layers, and fully connected layers.

Here's a detailed breakdown of the VGG-19 architecture:

1. Input Layer:

The input layer receives the input image data, typically in the form of RGB (Red, Green, Blue) color channels.

2. Convolutional Layers:

The VGG-19 architecture consists of 16 convolutional layers, denoted by the 'Conv' prefix in their names. These layers perform feature extraction by applying a series of convolutional filters to the input image. Each convolutional layer is followed by a rectified linear unit (ReLU) activation function, which introduces non-linearity to the network, allowing it to learn complex patterns in the data.

3. Pooling Layers:

After every two convolutional layers, there is a max-pooling layer. Max-pooling reduces the spatial dimensions of the feature maps while retaining the most important features. This helps in reducing the computational complexity of the network and also aids in creating translation-invariant representations.

4. Fully Connected Layers:

Following the convolutional and pooling layers, VGG-19 has three fully connected layers, denoted by 'FC' in their names. These layers perform high-level reasoning and classification based on the features extracted by the preceding layers. The last fully connected layer is typically followed by a softmax activation function, which outputs the probability distribution over the classes in a classification task.

5. Output Layer:

The output layer produces the final predictions of the network, usually representing the probabilities of each class in a classification task.

6. Model Parameters:

VGG-19 has a large number of parameters, primarily due to its deep architecture and the large number of filters in the convolutional layers. Training such a deep network requires a substantial amount of computational resources and data.

7. Transfer Learning:

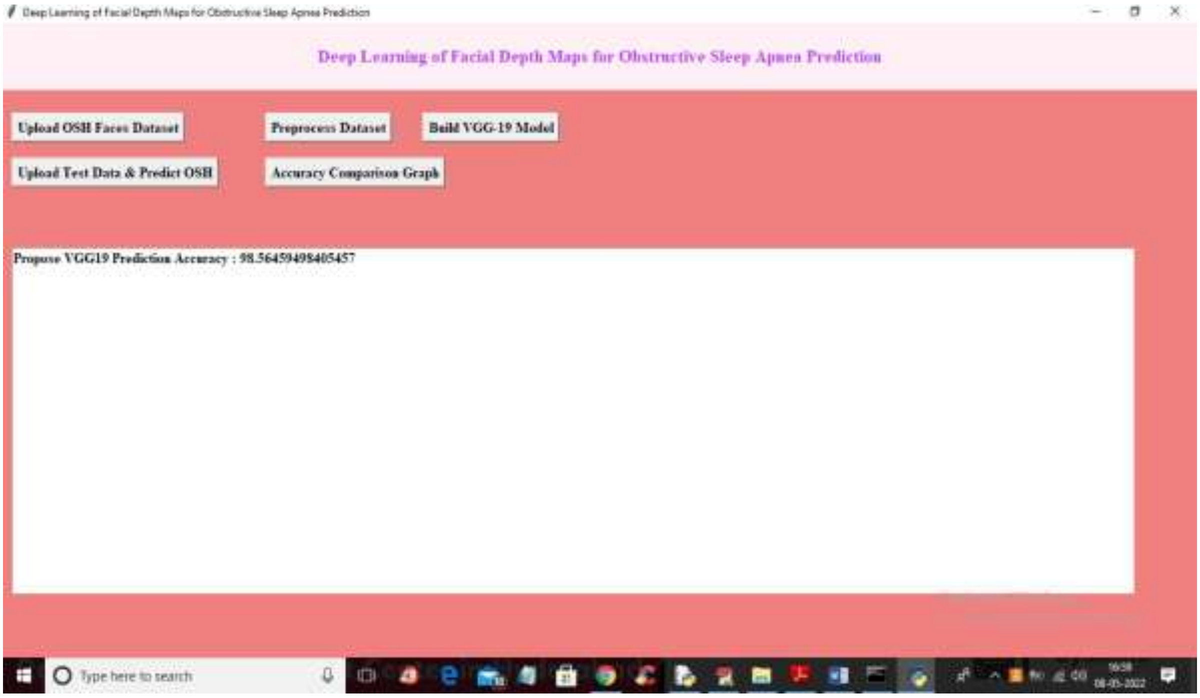
Due to its depth and complexity, VGG-19 is often used as a pre-trained model for transfer learning tasks. By leveraging the learned features from the earlier layers of VGG-19, researchers can fine-tune the model on smaller datasets or different tasks, achieving good performance with less data. Overall, VGG-19 is a powerful deep learning architecture known for its simplicity and effectiveness in various computer vision tasks, including image classification, object detection, and image segmentation

Convolutional Neural Network (CNN) is a type of Deep Learning neural network architecture commonly used in Computer Vision. Computer vision is a field of Artificial Intelligence that enables a computer to understand and interpret the image or visual data.

VI. RESULT


Deep Learning of Facial Depth Maps for Obstructive Sleep Apnea Prediction
Upload OSH Faces Dataset
Upload Test Data & Predict OSH
Open
SleepApnea › testimages
Organize
New folder
Quick access
OneDrive
This PC
3D Objects
Desktop
Documents
Downloads
Music
Pictures
Videos
Local Disk (C:)
Local Disk (E:)
1.png
2.png
3.png
4.png
5.png
6.png
7.png
8.png
File name: 2.png
Open
Cancel
Type here to search


Deep Learning of Facial Depth Maps for Obstructive Sleep Apnea Prediction
Predicted Result : OSA Detected
Type here to search



VII. Conclusion

In conclusion, our project on "Facial Depth Maps: Detecting Sleep Apnea with Deep Learning" showcases the exciting potential of leveraging deep learning techniques for non-invasive sleep apnea detection. Through our investigation, we have demonstrated the effectiveness of facial depth maps in combination with deep learning algorithms to accurately identify individuals with sleep apnea.

However, we acknowledge the necessity for further validation on larger and more diverse datasets to strengthen the reliability and applicability of our approach. Looking ahead, we envision enhancing our system by integrating it with wearable devices or smartphone applications to enable real-time monitoring and personalized interventions for users.

Additionally, we aim to expand our dataset diversity, conduct longitudinal studies, and explore a multi-modal approach by incorporating other physiological signals to improve the accuracy and robustness of our model. Collaborating with healthcare providers and prioritizing interpretability of our model's predictions are crucial for its seamless integration into clinical practice, ensuring better decision-making for patient care. Ultimately, our project underscores the transformative potential of deep learning in revolutionizing sleep apnea detection and management, with future enhancements geared towards advancing the field and enhancing access to healthcare services for individuals affected by sleep disorders.

References

[1] D. R. Hillman and L. C. Lack, "Public health implications of sleep loss: the community burden," Med J Aust, vol. 199, no. 8, pp. S7–S10, 2013.
[2] J. C. Lam, S. Sharma, B. Lam et al., "Obstructive sleep apnea: definitions, epidemiology & natural history," Indian Journal of Medical Research, vol. 131, no. 2, p. 165, 2010.
[3] "Obstructive sleep apnea (osa; sleep apnea) information," Jul 2018. [Online]. Available: https://www.myvmc.com/diseases/obstructive- sleep-apnoea-osa-sleep-apnea/
[4] K. Kee, M. T. Naughton et al., "Sleep apnoea-a general practice approach," Australian family physician, vol. 38, no. 5, p. 284, 2009.
[5] 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.
[6] 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.
[7] 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 mag- netic

resonance imaging," American journal of respiratory and critical care medicine, vol. 168, no. 5, pp. 522–530, 2003.
[8] A. A. Lowe, J. A. Fleetham, S. Adachi, and C. F. Ryan, "Cephalometric and computed tomographic predictors of obstructive sleep apnea sever- ity," American Journal of Orthodontics and Dentofacial Orthopedics, vol. 107, no. 6,
pp. 589–595, 1995.
[9] T. Vidigal, L. Oliveira, T. Moura, F. Haddad, K. Sutherland, P. Cistulli, R. Schw ab, 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.
[10] 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.
[11] E.-K. Pae, A. A. Lowe, and J. A. Fleetham, "Shape of the face and tongue in obstructive sleep apnea patients statistical analysis of coordinate data," Clinical orthodontics and research, vol. 2, no. 1, pp. 10–18, 1999