2018 International Conference on Machine Learning and Data Engineering (iCMLDE)

Deep Learning of Facial Depth Maps for Obstructive Sleep Apnea Prediction

Syed MS Islam
School of science
Edith cowan university
Joondalup, Australia
syed.islam@ecu.edu.au

Hassan Mahmood
School of science
Edith cowan university
Joondalup, Australia
hmahmood@our.ecu.edu.au

Adel Ali Al-Jumaily
School of Biomedical Engineering
University of Technology Sydney
Sydeny, Australia
Adel.Al-Jumaily@uts.edu.au

Scott Claxton
Specialist Medical Centre
Joondalup Health Campus
Joondalup, Australia
Scott.Claxton@genesiscare.com

***Abstract*—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. Previous research has shown the link between facial morphology and OSA. In this paper, we investigate the application of deep learning techniques to diagnose the disease through depth map of human facial scans. Depth map will provide more information about facial morphology as compared to the plain 2-D color image. Even with very less amount of sample data, we are able to get around 69validation accuracy using transfer learning. We are predicting patients with above moderate $>$ 15 or below moderate $\leq$ 15 OSA.**

***Index Terms*—Obstructive Sleep Apnea, Transfer learning, Deep Learning, Facial Depth Map**

I. 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 [1] 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 [2]. Normally during sleep, our upper airway remain 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 [3]. 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 patients 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 [4]. 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 [5].

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.

II. Related Work

To make diagnosis easier and quicker imaging techniques have been used. Like Jose et. al [6] examined the relationship between anatomic measures of upper airway structures and OSA. They used dental X-ray and magnetic resonance imaging (MRI) scans to do the anatomic measurements. [7] used volumetric analysis on MRI of the upper airway soft tissue structures to calculate physiological differences between OSA- and non-OSA subjects. [8] also assessed the relationship between craniofacial structures through tomography of lateral cephalometry, tongue, soft palate, and upper airway size in control subjects and sleep apnea patients.

These procedures mainly based on advanced imaging techniques, enabling a thorough analysis of facial structures. And also these procedures are expensive and time-consuming. Some studies like [9] shows the link between craniofacial and intraoral photography in OSA subjects. Also, work like [10] investigate the relationship between the facial dimension and upper airway structures. [11] discovered that main face variations were found at the lower part of the face and in the upper part of the neck. These studies encouraged researchers to devise facial image-based diagnosis of OSA.

978-1-7281-0404-1/19/$31.00 ©2019 IEEE
DOI 10.1109/iCMLDE.2018.00036

Lee et al. [12], [13] have used features, from digital photographs, of craniofacial surface structures to predict OSA. They analyzed frontal and profile photographs of 114 subjects. Using photographic measurements they obtained an accuracy of 76.1% with an area under Receiver Operating Characteristic (ROC) curve of 0.82. Also [14], [15] have explored the prediction capability of OSA by analyzing facial profile and frontal images. These results show that features capturing the composite elements of craniofacial structures and regional adiposity can predict OSA better than demographic data (e.g. BMI or neck circumference) collected by clinical observations. However, digital photographs are two dimensional in nature and hence neither non-linear measurements nor measurements of the shape of craniofacial anatomy can be obtained. All of these approaches need manual landmark detection, which is time-consuming and also depend upon the experience of the person who will mark the key point on facial data.

To our knowledge, there is only one work for sleep apnea detection using automated landmark detection. AT Balaei, et al. [16] uses a regression approach to find 21 profile and 14 front face landmarks. An accuracy of 70% was achieved using trained classifier over detected landmarks. They also try to directly predict OSA form colored facial data of 50x50 pixel from frontal and profile images through training of neural network. In that regard, they achieve an accuracy of 62

Three-dimensional surface imaging technologies have recently been developed that are well suited for imaging the human head, face and neck. Such imaging has been previously used to analyze craniofacial changes before and after various treatment modalities for OSA treatment [17], [18], however to date only one study has used this technique to obtain 3D surface images of the face from 40 OSA and 40 non-OSA subjects, analyzing only the association of craniofacial obesity with the OSA severity. No comprehensive study has been undertaken with this technique to compare the discriminatory capacity of facial morphometry between individuals with and without OSA [19].

Our goal is to take leverage of 3d information in facial depth maps and to devise a framework to detect OSA in subjects without human intervention. For that, we trained a deep network for facial depth maps. In section 3, we describe the datasets and methodology section 4 is about results. In the end, we make the conclusion of our paper.

III. Dataset & Methodology

Sleep data and 3D scans were collected from the patients appearing to Genesis SleepCare 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 frigure 1.

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. As shown in figure 2.

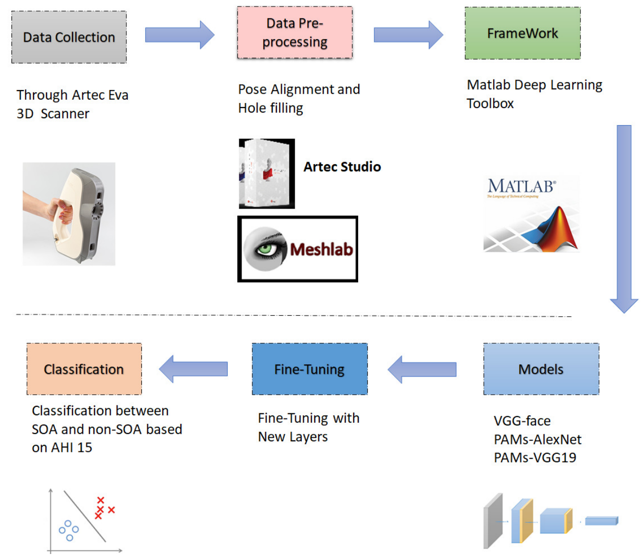

Fig. 1. Block diagram of the proposed deep learning based OSA prediction framework.

Fig. 2. Sample raw images in the collected 3D dataset.

While converting these 3D scans to frontal 2D depth maps, we want to reduce these unwanted variations. We use Artec Studio to make the corrections in all the 3D scans. As shown in figure 3.

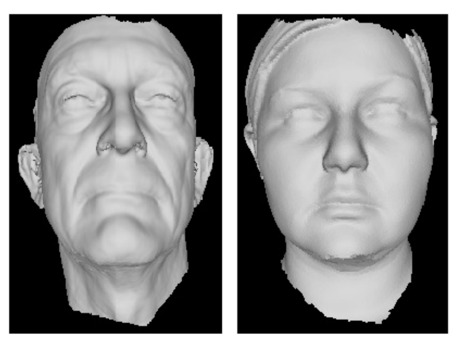

Fig. 3. Sample pre-processed images corresponding to raw images in Fig. 2.

After making corrections in default poses and other factors, we use MeshLab [21] to create 2D facial depth map of the frontal face. We choose the maximum and minimum scale to get higher resolution across depth values as shown in figure 4.

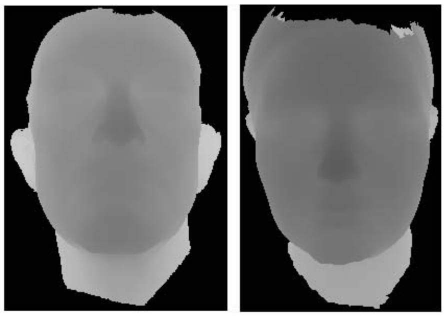

Fig. 4. 2D Facial depth maps

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 [22] Pose-Aware CNN Models (PAMs) for Face Recognition [23] 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-16 CNN architecture. [23] provided two pre-trained networks for face recognition with AlexNet [24] and 19 layer VGGNet [25]. To make these network classify for two classes, last fully connected layers are replaced by new fully connected layer and a softmax layer in each network type. After Adding the last layers below is the block diagram of all three networks.

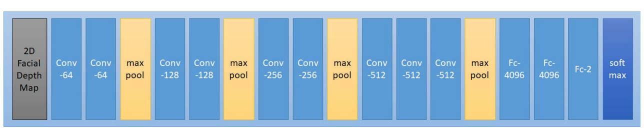

Fig. 5. Pre-Trained VGGFace with edition in last layers

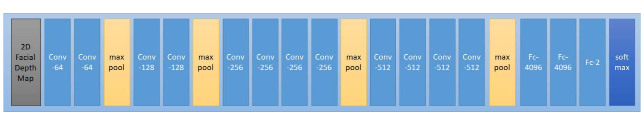

Fig. 6. Pre-Trained PAMs-vgg19 with edition of last layers

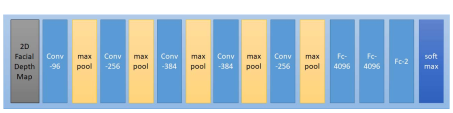

Fig. 7. Pre-Trained PAMs-AlexNet with edition of last layers

Matlab deep learning toolbox is used to fine-tune these models. All pre-trained models are available in caffe and also importable to Matlab. To fine-tune the model usual practice is to keep the learning rate of learned layers much less than the layers we are adding. We tried different learning rates and also tried by freezing different layers in each network. We set the batch size of 10 and fine-tuned the models through stochastic gradient descent (SGD). We started from pre-learned weights on VGGface, PAMs-VGG19, and PAMs-AlexNet and initialize fc8 layer from scratch. The initial learning rate is set to 0.0001. We fine-tune all the layers with this learning rate but the new fc8 layer learning rate is kept 20 times higher than the original one.

we separate 14 samples for Test Data and rest of 30% of data is separated for Validation and 70% is used for Training the models. Training is done on NVIDIA GeForce GTX 1080 using Matlab Deep learning toolbox. We trained the network for end-to-end classification, where we will give the facial depth map of the patient and it will classify between sleep apnea and non-sleep apnea.

IV. Results & Discussion

After training the network with a different set of parameters, Fc7 layer features of three models are shown in figure 8

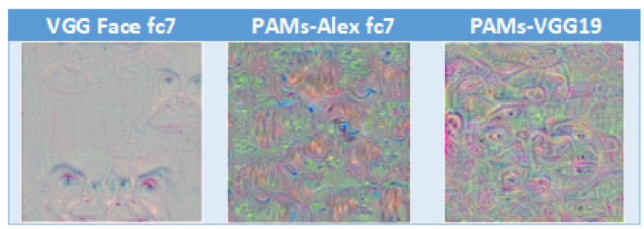

Fig. 8. Netwrok layer features

These feature maps are rightly showing the behavior of the respective models. In both Pose-Aware Models(PAMs), there are different eyes and nose activation at different areas because these model have to deal with pose variation and rotations. And if we look at the output of the VGG Face layer, we generally observe frontal face on it with no rotation. In our experimentation, we observe that VGG Face generally performed well with transfer learning for depth maps.

Table 1

Accuracy of Three Models			
	VGG Face	PAMs-VGG19	PAMs-Alex
Validation Accuracy	68.75	62.5	60.15
Test Accuracy	67.42	57.14	59.37

As we can see in the table 1 that the accuracy of VGG Face is better than PAMs-VGG19 and PAMs-AlexNet. PAMs-VGG19 is deeper than VGGFace but it does not perform well. The advantage of VGGface is that it is pre-trained over a large dataset. Also, our method achieved higher accuracy of 68.75% and 67.42% for validation and test accuracy respectively as compared to [16], which achieved 62% accuracy from directly feeding the input images to the neural network.

V. Conclusion

In this paper, we propose the first facial depth map based sleep apnea detection. Patients dataset is small, to overcome this limitation we took advantage of transfer learning. We analyze three pre-trained models and among them, VGGface performs the best. Our method shows comparable performances to the state-of-the-art results in terms of getting prediction straight from depth facial data using end-to-end deep learning. In future, pose correction problem will be solved through a 3D morphable model. Hole filled depth maps will be created through an automatic procedure. This work gets good results with a very small dataset and with more 3D scans of OSA and non-OSA patients, we will enhance the performance for diagnoses.

Acknowledgment

Authors would like to acknowledge that this work is done with the support of ECU-ECR 2017 and SIF 2017 grants and Higher Education Commission of Pakistan (HEC)-ECU joint scholarship.

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 apnoea: definitions, epidemiology & natural history," *Indian Journal of Medical Research*, vol. 131, no. 2, p. 165, 2010.

[3] "Obstructive sleep apnoea (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 magnetic 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 severity," *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. 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.

[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 patientsstatistical analysis of coordinate data," *Clinical orthodontics and research*, vol. 2, no. 1, pp. 10–18, 1999.

[12] 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.

[13] 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.

[14] F. Espinoza-Cuadros, R. Fernández-Pozo, D. T. Toledano, J. D. Alcázar-Ramírez, E. López-Gonzalo, and L. A. Hernández-Gómez, "Speech signal and facial image processing for obstructive sleep apnea assessment," *Computational and mathematical methods in medicine*, vol. 2015, 2015.

[15] 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.

[16] A. 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.

[17] S. M. Islam, M. S. Goonewardene, and M. Farella, "A review on three dimensional facial averaging for the assessment of orthodontic disorders," in *Innovations and Advances in Computing, Informatics, Systems Sciences, Networking and Engineering*. Springer, 2015, pp. 391–397.

[18] J. Moss, "The use of three-dimensional imaging in orthodontics," *The European Journal of Orthodontics*, vol. 28, no. 5, pp. 416–425, 2006.

[19] S. M. Banabilh, A. Suzina, S. Dinsuhaimi, A. Samsudin, and G. Singh, "Craniofacial obesity in patients with obstructive sleep apnea," *Sleep and Breathing*, vol. 13, no. 1, pp. 19–24, 2009.

[20] "Artec eva." [Online]. Available: https://www.artec3d.com/portable-3d-scanners/artec-eva

[21] "Meshlab." [Online]. Available: http://www.meshlab.net/

[22] O. M. Parkhi, A. Vedaldi, A. Zisserman *et al.*, "Deep face recognition." in *BMVC*, vol. 1, no. 3, 2015, p. 6.

[23] I. Masi, S. Rawls, G. Medioni, and P. Natarajan, "Pose-Aware Face Recognition in the Wild," in *CVPR*, 2016.

[24] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[25] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman, "Return of the devil in the details: Delving deep into convolutional nets," *arXiv preprint arXiv:1405.3531*, 2014.