SPWID 2020

The Sixth International Conference on Smart Portable, Wearable, Implantable and Disability-oriented Devices and Systems

September 27th – October 1st, 2020

SPWID 2020 Editors

Sandra Sendra, Universitat Politècnica de València, Spain

SPWID 2020

Foreword

The Sixth International Conference on Smart Portable, Wearable, Implantable and Disability-oriented Devices and Systems (SPWID 2020), held between September 27 – October 1st, 2020, is an inaugural event bridging the concepts and the communities dealing with specialized implantable, wearable, near-body or mobile devices, including artificial organs, body-driven technologies, and assistive services

Mobile communications played by the proliferation of smartphones and practical aspects of designing such systems and developing specific applications raise particular challenges for a successful acceptance and deployment.

We take here the opportunity to warmly thank all the members of the SPWID 2020 Technical Program Committee, as well as the numerous reviewers. The creation of such a broad and high quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and efforts to contribute to SPWID 2020. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

Also, this event could not have been a reality without the support of many individuals, organizations, and sponsors. We are grateful to the members of the SPWID 2020 organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that SPWID 2020 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in the areas of smart portable devices and systems.

SPWID 2020 Chairs:

SPWID 2020 Publicity Chair
Mar Parra, Universitat Politecnica de Valencia, Spain

SPWID 2020

COMMITTEE

ACCSE 2020 Publicity Chair

Mar Parra, Universitat Politecnica de Valencia, Spain

ACCSE 2020 Technical Program Committee

Safa'a AbuJarour, University of Potsdam, Germany
Kishwar Ahmed, University of South Carolina Beaufort, USA
Maxim Bakaev, Novosibirsk State Technical University, Russia
Abdul Basit, State Bank of Pakistan (Central Bank of Pakistan), Pakistan
Carlos Becker Westphall, Federal University of Santa Catarina, Brazil
Jagadeesha R Bhat, St. Joseph Engineering College, Mangalore, India
Freimut Bodendorf, Institute of Information Systems - University of Erlangen-Nuremberg, Germany
An Braeken, Vrije Universiteit Brussel, Belgium
Arun Das, Visa Inc., USA
Alessandro Farasin, Istituto Superiore Mario Boella (ISMB), Turin, Italy
Barbara Gili Fivela, University of Salento, Italy
Aviel Glam, Technion - Israel Institue of Technology | RAFAEL - Advanced Defence System Ltd., Israel
Josefa Gómez, University of Alcalá, Spain
Robert C. Green II, Bowling Green State University, USA
Béat Hirsbrunner, University of Fribourg, Switzerland
Fu-Hau Hsu, National Central University, Taiwan
Michael Huebner, BTU Cottbus-Senftenberg, Germany
Sergio Ilarri, University of Zaragoza, Spain
Ilias Iliadis, IBM Research - Zurich Laboratory, Switzerland
Tomayess Issa, Curtin University, Australia
Keiichi Kaneko, Tokyo University of Agriculture and Technology, Japan
Hyunbum Kim, University of North Carolina at Wilmington, USA
Ratan Lal, Kansas State University, USA
André Langer, Chemnitz University of Technology, Germany
Yiu-Wing Leung, Hong Kong Baptist University, Kowloon Tong, Hong Kong
Shigang Li, Hiroshima City University, Japan
Christopher Mansour, Mercyhurst University, Erie, USA
Alfonso Mateos Caballero, Universidad Politécnica de Madrid, Spain
Vinod Muthusamy, IBM T.J. Watson Research Center, USA
Hidemoto Nakada, AIST, Japan
Isabela Neves Ferraz, Universidade de Brasília, Brazil
Isabel Novo Corti, University of A Coruña, Spain
Petra Perner, Institute of Computer Vision and applied Computer Sciences IbaI, Germany
Xose Picatoste, University of A Coruña, Spain
Jim Prentzas, Democritus University of Thrace - School of Education Sciences, Greece
Yenumula B Reddy, Grambling State University, USA
Daniel Ritter, SAP, Germany

Claudio Rossi, Istituto Superiore Mario Boella (ISMB), Turin, Italy
Maya Sappelli, HAN University of Applied Sciences, Netherlands
Mukesh Singhal, University of California, Merced, USA
Dimitrios Skoutas, University of the Aegean, Greece
Young-Joo Suh, POSTECH, Korea
Abdelhamid Tayebi, University of Alcalá, Spain
David Tormey, Institute of Technology Sligo, Ireland
Yuehua Wang, Texas A&M University-Commerce, USA
John Woodward, Queen Mary University of London, UK
Ning Wu, School of Computer Science and Engineering - Beihang University, China
Wen-Chi Yang, NeuHelium Co. Ltd., Shanghai, China
Aleš Zamuda, University of Maribor, Slovenia
Ye Zhu, Cleveland State University, USA
Jason Zurawski, Lawrence Berkeley National Laboratory / Energy Sciences Network, USA

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

Mood Detection for Improving Lifestyle of Older Adults in Ambient Assisted Living Contexts

Andrea Caroppo, Alessandro Leone, Pietro Siciliano

National Research Council of Italy - Institute for Microelectronics and Microsystems
Via Monteroni, c/o Campus Università del Salento, Palazzina A3, Lecce, Italy
email: andrea.caroppo@cnr.it email: alessandro.leone@cnr.it email: pietro.siciliano@le.imm.cnr.it

Abstract— **Ambient Assisted Living (AAL) has the ambitious goal of improving the lifestyle or quality of life of elderly or vulnerable people through the use of technology. In this research, area considerable efforts have been made for the design and development of automatic solutions for the recognition of mood trough patterns of facial expressions, even using low cost and commercial vision sensors. However, there are still some open issues to be faced like the age or any situation of disability of the observed subject, the pose of the face and the environment illumination conditions. A lot of progress has been made in this topic with the emergence of deep learning methods. In the proposed work, the performance of two recent deep convolutional neural networks models are evaluated on the CIFE and FER-2013 datasets that include also facial expressions of older adults performed in uncontrolled conditions. A thorough session of experiments focused on the concept of "Transfer Learning" was carried out. The results obtained demonstrated that both deep architectures reach levels of accuracy higher than 67.8% grouping expressions of older adults into 3 categories: positive, negative and neutral. The latter classification may be sufficient for a mood detection module that could be the input of a future e-coaching platform to be integrated in the AAL context.**

Keywords- Mood Detection; Ambient Assisted Living; Older Adults; Disability; Lifestyle Improvements.

I. INTRODUCTION

In the last few years, the number of older adults is increasing. In addition, there are currently more than 2 billion disabled people in the world, that is 37.5% of the world's population. Consequently, many efforts have been made in the past and are currently being made by researchers with the aim to improve the quality of lifestyle for these categories of subjects with different frailties. Related to this issue, an Ambient Assisted Living (AAL) works to create better living conditions. AAL systems are able to continuously monitor the health status of the frailty subject through data coming from heterogeneous sensors.

In this context many potential applications, such as robotics, communications, security, medical and assistive technology, would benefit from the ability of automatically recognize facial expression [1][2], because different facial expressions can reflect the mood, the emotions and also mental activities. An automatic system capable of recognizing facial expressions could be the input of an e-coaching system, useful for example to change the living environment, implementing devices (music or video players) or changing lighting conditions based on the mood of the observed subject.

In 1971, Ekman and Friesen reported that facial expression acts as a rapid signal that varies with contraction of facial features like eyebrows, lips, eyes and cheeks. Moreover they determined that there were six basic classes in Facial Expression Recognition (FER): anger, disgust, fear, happiness, sadness and surprise [3]. A classical automatic facial expression analysis usually involves three steps: face acquisition, facial data extraction and representation (feature extraction), and classification.

FER systems can be divided into two main categories according to the feature representations: static image FER and dynamic sequence FER. In static-based methods [4], the feature representation is encoded with only spatial information from the current single image, whereas dynamic-based methods [5] consider the temporal relation among contiguous frames in the input facial expression sequence.

For the feature extraction step, the majority of the traditional methods have used handcrafted features. Generally, they are divided into the following categories: geometric-based, appearance-based and hybrid-based approaches. In particular, geometric-based features are able to depict the shape and locations of facial components such as mouth, nose, eyes and brows using the geometric relationships between facial points to extract facial features [6]; appearance-based descriptors aim to use the whole-face or specific regions in a face image to reflect the underlying information in a face image [7]; hybrid-based approaches combine the two previous types of features in order to enhance the system's performance and it might be achieved either in features extraction or classification level.

Geometric-based, appearance-based and hybrid-based approaches have been widely used for the classification of facial expressions even if it is important to emphasize how all the aforementioned methodologies require a process of feature definition and extraction very daunting. In addition, this category of methodologies easily ignores the changes in skin texture such as wrinkles and furrows that are usually accentuated by the age of the subject. Last but not least recent studies have pointed out that classical approaches used for the classification of facial expression are not performing well when used in real contexts where face pose and lighting conditions are broadly different from the ideal ones used to capture the face images within the benchmark datasets. As highlighted above, among the factors that makes FER very difficult (and consequently mood

detection), one of the most discriminating is the age [8][9]. In particular, expressions of older individuals appeared harder to decode, owing to age-related structural changes in the face. Consequently, state-of-the-art approaches based on handcrafted features extraction may be inadequate for mood detection of older adults. In recent years, machine learning techniques based on convolutional neural networks (CNNs) have achieved great success in the field of computer vision. It is also a promising approach for the research of FER. Different from traditional techniques, CNNs can perform tasks in an end-to-end way, associating both feature extraction and classification steps together by training. For example, Zhang et al. [10] proposed a deep neural network (DNN) with the scale-invariant feature transform (SIFT) feature, which achieved the accuracy of 78.9% on the multi-view BU-3DFE dataset [11]. Lopes et al. [12] proposed a combination of CNN and special image pre-processing steps (C-CNN) to recognize six expressions under head pose at 0∘, whose accuracy was 90.96% on the same aforementioned dataset but its robustness was unknown under different head poses. To reduce the influence of various head poses, Jung et al. [13] proposed a jointly CNNs with facial landmarks and colour images, which achieved the accuracy of 72.5%, but the network consisted of only three convolutional layers and two hidden layers, making it be difficult to accurately learn facial features. Li and Deng [14] also presented a very interesting survey on CNN based FER techniques.

Lack of training samples is a big problem for FER in the wild using deep CNNs. In the proposed work, this problem is accentuated by having very few datasets available containing facial expressions of elderly subjects. To solve this issue, some methods used pre-trained network for classification or re-trained a network model to re-initialize the weights for new datasets. The techniques are regarded as "transfer learning" [15].

Based on the above discussion, in this paper, two pre-trained CNN, who have been successful in image recognition, are evaluated for mood detection in the wild and tested on two benchmark datasets that contain facial expressions divided by age group.

The rest of this paper is structured as follows. Section II reports the proposed pipeline emphasizing some details for pre-processing steps. The same section describes also the deep architectures and the algorithmic procedures used in this work in order to adapt the aforementioned models to the problem of mood detection in older adults. Section III presents datasets, experimental procedures and results obtained, while Section IV concludes the paper by providing an overview of the themes highlighted and the need to monitor the future.

II. OVERVIEW OF THE PROPOSED MOOD DETECTION SYSTEM

A representation of the proposed mood detection system for older adults is given in the block diagram shown in Figure 1. First, the implemented pipeline performs a pre-processing task on the input images (face detection, cropping & resizing). Once the images are pre-processed

they can be classified into three categories (positive, negative or neutral expressions) using pre-trained deep networks.
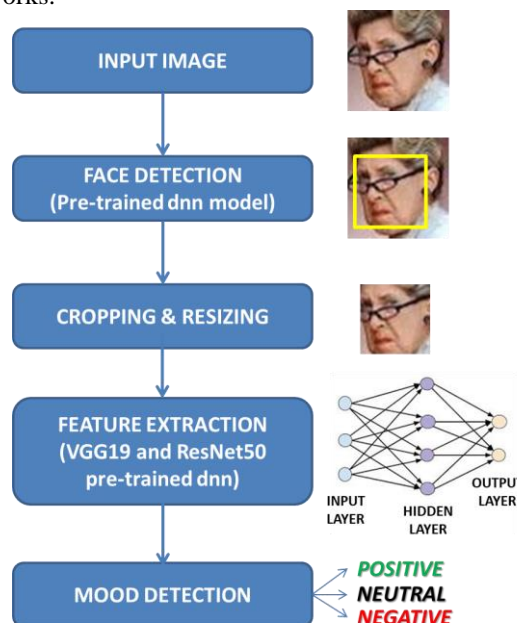

Figure 1. Overview of the proposed mood detection system for older adults in AAL context

A. Pre-Processing

The task of detecting a face in the considered scenario is not an easy problem because many difficulties arise and must be taken into account. In a living environment, for example, the face of observed subject could occupy very little area in most images if the camera is not positioned near the end user. Moreover faces can look very different depending on orientation, pose and age. Face detection went mainstream in the early 2000's when Viola and Jones [16] invented a way to detect faces that was fast enough to run on cheap cameras. Today it is very close to being the de facto standard for solving face detection tasks. Given that the original Viola-Jones face detector has limitations for multi-view face detection, in the implemented pipeline a recently methodology that has met much success is proposed to tackle this problem.

Here the library with functions that mainly aiming real-time computer vision (i.e. last version of OpenCV) is selected. In particular, starting from OpenCV 3.3 a deep neural network (DNN) architecture that performs an accurate face detector is included. The main advantage of this module is the ability to detect faces "in the wild" in real-time even if a PC without GPU is used for the processing. The aforementioned module is based on Single Shot MultiBox Detector (SSD) framework, using a reduced ResNet-10 model [17] and its output is constituted by the coordinates of the bounding box of the facial region accompanied by a confidence index, useful in case it is necessary to set a reliability threshold with respect to the detection of the facial region. Once the face has been detected a simple routine was written in order to crop the

facial image. This is achieved by detecting the coordinates of the top-left corner, the height and width of the face enclosing rectangle, removing in this way all background information and image patches that are not related to the expression. Since the facial region could be of different sizes after cropping, in order to remove the variation in face size and keep the facial parts in the same pixel space, the algorithmic pipeline provides both a down-sampling step and an increasing resolution step that generate face images with the specific size required by tested deep architectures. For the down-sampling a simple linear interpolation was used, whereas a nearest-neighbor interpolation was implemented in order to increase the size of the facial images.

B. CNN Architectures For Feature Extraction

CNN is a type of deep learning model designed to automatically and adaptively learn spatial hierarchies of features, from low to high-level patterns. A typical implementation of CNN for FER encloses three learning stages in just one framework. The learning stages are: 1) feature learning, 2) feature selection and 3) classifier construction. Creating a CNN from scratch is not an easy task. So, in order to save ourselves from this over-head, in the present work the concept of "transfer learning" was adopted [15]. Transfer learning is a common and recent strategy to train a network also on a small dataset, (which is one of the main problems in the case of recognition of facial expressions of the elderly) where a network is pre-trained on an extremely large dataset, such as ImageNet [18], which contains 1.4 million images with 1000 classes, then reused and applied to the given task of interest.

In this work two deep CNNs (DCNNs), namely VGG19 and Resnet50, were evaluated as the feature extractors of the proposed method for mood detection in older adults. These DCNNs are pre-trained on a nature image dataset (ImageNet) for distinct generic image descriptors and then applied to extract discriminative features from facial images based on transfer learning theory. Below is a brief description of the two architectures.

VGG19 - The VGG networks with 16 layers (VGG16) and with 19 layers (VGG19) [19] were the basis of the Visual Geometry Group (VGG) submission in the ImageNet Challenge 2014, where the VGG team secured the first and the second places in the localization and classification tracks respectively. The VGG architecture is structured starting with five blocks of convolutional layers followed by three fully-connected layers. Convolutional layers use 3×3 kernels with a stride of 1 and padding of 1 to ensure that each activation map retains the same spatial dimensions as the previous layer. A rectified linear unit (ReLU) activation is performed right after each convolution and a max pooling operation is used at the end of each block to reduce the spatial dimension. Max pooling layers use 2×2 kernels with a stride of 2 and no padding to ensure that each spatial dimension of the activation map from the previous layer is halved. Two fully-connected layers with 4096 ReLU activated units are then used before the final 1000 fully-connected softmax layer. A downside of the VGG16 and

VGG19 models is that they are more expensive to evaluate and use a lot of memory and parameters. VGG16 has approximately 138 million parameters and VGG19 has approximately 143 million parameters. Most of these parameters (about 100 million) are in the first fully-connected layer, and it was since found that these fully-connected layers could be removed with no performance downgrade, significantly reducing the number of necessary parameters.

ResNet50 - Residual Networks (ResNets) [20] are deep convolutional networks where the basic idea is to skip blocks of convolutional layers by using shortcut connections to form blocks named residual blocks. These stacked residual blocks greatly improve training efficiency and largely resolve the degradation problem present in deep networks. In ResNet50 architecture, the basic blocks follow two simple design rules: 1) for the same output feature map size, the layers have the same number of filters; and 2) if the feature map size is halved, the number of filters is doubled.

The down sampling is performed directly by convolutional layers that have a stride of 2 and batch normalization is performed right after each convolution and before ReLU activation. When the input and output are of the same dimensions, the identity shortcut is used. When the dimensions increase, the projection shortcut is used to match dimensions through 1×1 convolutions. In both cases, when the shortcuts go across feature maps of two sizes, they are performed with a stride of 2. The network ends with a 1,000 fully-connected layer with softmax activation. The total number of weighted layers is 50, with 23,534,592 trainable parameters.

C. Classification

The methodology for automated mood detection using transfer learning is shown in Figure 2. The structure illustrated consists of layers from the pre-trained model and few new layers.
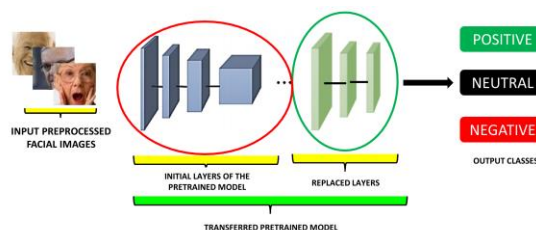

Figure 2. Mechanism of transfer learning using pre-trained models for mood detection of older adults

For the present work only the last three layers of VGG19 and ResNet50 were replaced to accommodate the new image categories. Since an age-related decline in decoding facial expressions has been repeatedly reported in literature [21][22], in the proposed work the traditional facial expression classification has been modified by grouping the expressions into 3 main groups: positive (happiness), negative (fear, disgust, anger and sadness) and neutral. The expression "surprise" was not considered since it can have

any valence; that is, it can be neutral/moderate, pleasant, unpleasant, positive, or negative.

Consequently, the step-by-step procedure for expression classification is outlined below:

1) resize the input images so that they are consistent with the size of the input layer of the pre-trained network model;

2) partition the data into training and test sets; 70% of images per category are taken for training and 30% as a test dataset to test the network;

3) review the network architecture (replace the final layers/modify the layers): for VGG19 and ResNet50 alter the last three layers of the pre-trained networks with a set of layers ("fully connected layer," a "softmax layer," and a "classification output layer") to categorize the images into the respective classes;

4) train the network and test the new model on the target datasets.

III. RESULTS

The proposed mood detection system was tested on two benchmark datasets among the few present in the literature that include facial expressions acquired in uncontrolled conditions and containing subjects of different age groups.

The CIFE dataset [23] is composed of facial images with seven different types of expressions. The expression subsets have the following sizes: 3636, 1905, 975, 2485, 1994, 1381 and 2381 for Happiness, Anger, Disgust, Sadness, Surprise, Fear and Neutral respectively. The images are extracted from the web with gathering techniques permitted to collect in total 14757 images containing candid expression images that are randomly posed. This last detail redeems the process of expression recognition more difficult than the two previous datasets, which are made up of facial frontal expressions acquired in controlled environments.

Since CIFE dataset contains images with only the label of facial expression and without any indication about the age of the subject, it was necessary to perform an age estimation technique consolidated in the literature. The approach used in the present work is inspired from the algorithm described in [24] that permitted to group the images into four different subgroups. Table 1 presents the total number of images in the CIFE dataset, divided according to the estimated age, whereas in Figure 3 some examples of expressions performed by ageing adults are represented.

TABLE I.   CIFE IMAGES BROKEN DOWN BY AGE

	Age (years)				
	<35	35-55	56-68	>68	Total
#images	5587	4828	2263	2079	14757

Happiness *Anger* *Disgust* *Sadness*

Surprise *Fear* *Neutral*

Figure.3. Some examples of expressions performed by aging adults from the CIFE dataset.

FER-2013 [25] is a large-scale FER dataset used in the ICML 2013 workshop's facial expression recognition challenge. The dataset has seven expressions including anger, disgust, fear, happy, sad, surprise, and neutral. It is comprised of 48×48 pixel grey-scale images of human faces. The training set consists of 28,709 examples, while both the test and validation sets are composed of 3,589 examples. The images of FER-2013 were collected from the Internet and the faces greatly vary in age, pose and occlusion, thus resulting in that the accuracy of human recognition is only approximately 65 ± 5%. As a powerful machine learning tools, the CNN can now surpass human beings on the FER-2013 task, and the state-of-the-art accuracy on FER-2013 is 75.42% by combining CNN extracted features and hand-crafted features for training.

Also the FER-2013 dataset does not contain the subdivision of the images into age groups, therefore the technique proposed in [24] was used again. Table 2 presents the total number of images in the FER-2013 dataset, divided according to the estimated age, whereas in Figure 4 some examples of expressions performed by ageing adults are represented.

TABLE II.   FER-2013 IMAGES BROKEN DOWN BY AGE

	Age (years)				
	<35	35-55	56-68	>68	Total
#images	13560	7432	6128	5178	32298

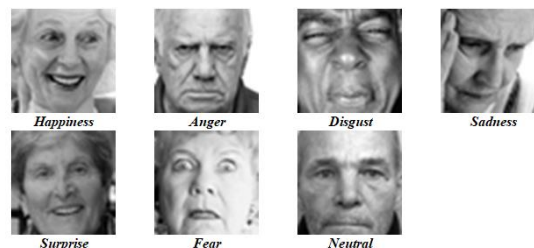*Happiness* *Anger* *Disgust* *Sadness*

Surprise *Fear* *Neutral*

Figure 4. Some examples of expressions performed by aging adults from the FER-2013 dataset.

Various experiments were conducted to assess the FER performance of the pre-trained VGG19 and ResNet50 deep networks with transfer learning. The metric used in this work for evaluating the methodologies is the accuracy, whose value is calculated using the average of n-class classifier accuracy for each group of expressions (positive, negative and neutral):

$$Acc = \frac{\sum_1^n Acc_{\exp r}}{n} \qquad (1)$$

$$Acc_{\exp r} = \frac{Hit_{\exp r}}{Total_{\exp r}} \qquad (2)$$

where $Hit_{\exp r}$ is the number of hits in the group of expression $expr$, $Total_{expr}$ represents the total number of samples of each group of expressions and n is the number of

expressions to be considered (in our case *n*=3). Fig. 5 reports the average accuracy for FER on images of CIFE and FER-2013 datasets. For each dataset the accuracy obtained with VGG19 and ResNet50 is reported, depending on the classifiers used. Here, for the final classifier layer, Random Forest (RF), Support Vector Machine (SVM) and Logistic Regression (LR) were compared.
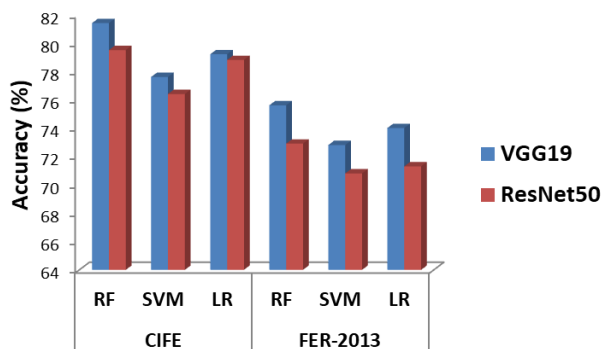

Figure 5. FER average accuracy on CIFE and FER-2013 combining VGG19 and ResNet50 with RF, SVM and LR classifier

From the results reported in the previous figure, it is evident that the recognition performances of three categories of expressions vary significantly as the dataset changes (FER-2013 is more challenging since the images are grayscale and have a resolution of 48x48 pixels) but by using for pre-training VGG 19 greater accuracy is obtained on both datasets, with the RF classifier which tends to provide an improvement in the results (about 3.8% with respect SVM and 2.2% with respect LR for CIFE and 2.8% with respect SVM and 1.6% with respect LR for FER-2013).

It is important to underline how the previous results were obtained by considering all age groups in which facial expressions were divided. Since the main objective of this work is mainly oriented towards the recognition of the facial expression category of the elderly, experiments have been carried out with the aim of measuring the performance of the methodologies by grouping the images of the datasets into different age groups. Table 3 and Table 4 report the final accuracy obtained by the proposed deep architectures with RF as classifier for each age group detected in CIFE and FER-2013 dataset respectively:

TABLE III. FER ACCURACY ON CIFE DATASET

Age Group	VGG19+RF (%)	ResNet50+RF (%)
<35	86.4	85.3
35-55	84.7	81.6
56-68	79.6	77.6
>68	74.9	73.5
Average value	**81.4**	**79.5**

TABLE IV. FER ACCURACY ON FER-2013 DATASET

Age Group	VGG19+RF (%)	ResNet50+RF (%)
<35	82.5	78.8
35-55	77.1	74.7
56-68	73.2	70.3
>68	69.6	67.8
Average value	**75.6**	**72.9**

Previous results demonstrate how the age of the observed subject influences the classification of facial expressions, probably because, as demonstrated by numerous psychological studies, the elderly are less expressive and consequently even with only three classes of expressions there is a gap in the accuracy between young and older subjects.

In a multi-class recognition problem, as the FER one, the use of an average recognition rate (i.e., accuracy) among all the classes could be not exhaustive since there is no possibility to inspect what is the separation level, in terms of correct classifications, among classes (in our case positive, negative and neutral facial expressions). To overcome this limitation, for each dataset the confusion matrices obtained with VGG19 + RF model are reported in Fig. 6 and Fig. 7 (here only the facial images of older adults with more than 68 years were considered). The notation POS in tables referred to positive expression (happiness); the notation NEU referred to neutral expression whereas NEG referred to negative expressions (fear, disgust, anger and sadness).

	POS	NEU	NEG
POS	**88.2**	**6.3**	**5.5**
NEU	**6.1**	**72.1**	**21.8**
NEG	**3.7**	**31.9**	**64.4**

Figure 6. Confusion Matrix on CIFE dataset (performed by older adults with more than 68 years) using the proposed VGG19 + RF architecture.

	POS	NEU	NEG
POS	**81.5**	**11.2**	**7.3**
NEU	**5.3**	**68.4**	**26.3**
NEG	**4.5**	**37.6**	**58.9**

Figure 7. Confusion Matrix on FER-2013 dataset (performed by older adults with more than 68 years) using the proposed VGG19 + RF architecture.

The confusion matrices provide, in relation to elderly subjects, a very important information to be taken into consideration in the case of implementation of e-coaching platforms: negative expressions are confused considerably with neutral expression. This observation poses a very important problem, as negative expressions are symptomatic of the onset or aggravation of diseases.

IV. CONCLUSION

The purpose of this paper was to explore and evaluate two deep transfer learning approach for mood detection in older adults considering that the majority of the works in the literature that address this topic are based on benchmark datasets that contain facial images with a small span of lifetime (generally young and middle-aged subjects). Among the two different deep architectures tested, the pre-trained VGG19 architecture in combination with an RF classifier yielded the best performance for each considered dataset and for each age group in which the dataset has been divided considering only three main classes of facial expressions: positive, negative and neutral. It was chosen to classify the expressions in these categories because they are sufficient for the development of an integrated system capable of implementing e-coaching systems based on the mood detected.

Future work will deal with the following main aspects. A first development might be to perform the pre-training of deep architectures on datasets different from ImageNet and more specific for the topic considered. Moreover, it will be necessary to extend the number of compared deep learning approaches since a limitation of the present work is the evaluation of only two pre-trained deep architectures which have already been overcome in terms of image classification from more deeper architectures like Inception-v4 [20] and Inception-ResNet-V2 [20].

ACKNOWLEDGMENT

REFERENCES

[1] M. Pantic and L. J. M. Rothkrantz, "Automatic analysis of facial expressions: The state of the art", IEEE Transactions on pattern analysis and machine intelligence, vol. 22, no. 12, pp. 1424-1445, 2000.

[2] B. Fasel and J. Luettin, "Automatic facial expression analysis: a survey", Pattern recognition, vol. 36, no. 1, pp. 259-275, 2003.

[3] P. Ekman and W.V. Friesen, "Constants across cultures in the face and emotion", Journal of personality and social psychology, vol. 17, no. 2, p. 124, 1971.

[4] A. Mollahosseini, D. Chan, and M.H. Mahoor, "Going deeper in facial expression recognition using deep neural networks", in 2016 IEEE Winter conference on applications of computer vision (WACV), pp. 1-10, IEEE, 2016, March.

[5] X. Zhao et al., "Peak-piloted deep network for facial expression recognition", In European conference on computer vision (pp. 425-442). Springer, Cham, 2016, October.

[6] R. Shbib and S. Zhou, "Facial expression analysis using active shape model", International Journal of Signal Processing, Image Processing and Pattern Recognition, 2015, vol. 8, no. 1, pp. 9-22.

[7] J. Chen, Z. Chen, Z. Chi, and H. Fu, "Facial expression recognition based on facial components detection and hog features", In International Workshops on Electrical and Computer Engineering Subfields, 2014, (pp. 884-888).

[8] G. Guo, R. Guo, and X. Li, "Facial expression recognition influenced by human aging", IEEE Transactions on Affective Computing, 2013, vol. 4, no. 3, pp. 291-298.

[9] S. Wang, S. Wu, Z. Gao, and Q. Ji, "Facial expression recognition through modeling age-related spatial patterns", Multimedia Tools and Applications, 2016, vol. 75, no. 7, pp. 3937-3954.

[10] T. Zhang, W. Zheng, Z. Cui, Y. Zong, J. Yan, and K. Yan, "A deep neural network-driven feature learning method for multi-view facial expression recognition", IEEE Transactions on Multimedia, vol. 18, no. 12, pp. 2528-2536, 2016.

[11] L. Yin, X. Wei, Y. Sun, J. Wang, and M. J. Rosato, "A 3D facial expression database for facial behavior research", In 7th international conference on automatic face and gesture recognition (FGR06) (pp. 211-216). IEEE, 2006, April.

[12] A. T. Lopes, E. de Aguiar, A. F. De Souza, and T. Oliveira-Santos, "Facial expression recognition with convolutional neural networks: coping with few data and the training sample order", Pattern Recognition, vol. 61, pp. 610-628, 2017.

[13] H. Jung, S. Lee, J. Yim, S. Park, and J. Kim, J, "Joint fine-tuning in deep neural networks for facial expression recognition", In Proceedings of the IEEE international conference on computer vision (pp. 2983-2991), 2015.

[14] S Li and W. Deng, "Deep facial expression recognition: A survey" IEEE Transactions on Affective Computing, 2020.

[15] S. J. Pan and Q. Yang, "A survey on transfer learning", IEEE Transactions on knowledge and data engineering, vol. 22, no. 10, pp. 1345-1359, 2010.

[16] P. Viola and M. J. Jones, "Robust real-time face detection", International journal of computer vision, vol. 57, no. 2, pp. 137-154 , 2004.

[17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition", In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778), 2016.

[18] O. Russakovsky, J. Deng, H. Su et al., "Imagenet large scale visual recognition challenge", International journal of computer vision, vol. 115, no. 3, pp. 211-252, 2015.

[19] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition", arXiv preprint arXiv:1409.1556, 2014.

[20] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning", In Thirty-first AAAI conference on artificial intelligence, 2017, February.

[21] A.J. Calder et al., "Facial expression recognition across the adult life span", Neuropsychologia, vol. 41, no. 2, pp. 195-202, 2003.

[22] N. C. Ebner and M. K. Johnson, "Young and older emotional faces: are there age group differences in expression identification and memory?", Emotion, vol. 9, no. 3, p. 329, 2009.

[23] W. Li, M. Li, Z. Su, and Z. Zhu, "A deep-learning approach to facial expression recognition with candid images", In 2015 14th IAPR International Conference on Machine Vision Applications (MVA) , 2015, (pp. 279-282). IEEE, 2015

[24] T. Wu, P. Turaga, and R. Chellappa, "Age estimation and face verification across aging using landmarks", IEEE Transactions on Information Forensics and Security, 2012, vol. 7, no. 6, pp. 1780-1788.

[25] I. J. Goodfellow, D. Erhan, P. L. Carrier et al, "Challenges in representation learning: A report on three machine learning contests", In International Conference on Neural Information Processing, 2013, (pp. 117-124). Springer, Berlin, Heidelberg

WatChair: AI-Powered Real-time Monitoring of Sitting Posture and Corrective Suggestions using Wearable Motion Sensor System

Shayan Fazeli
Computer Science Department
University of California, Los Angeles
shayan@cs.ucla.edu

Mohammad Kachuee
Computer Science Department
University of California, Los Angeles
mkachuee@cs.ucla.edu

Majid Sarrafzadeh
Computer Science Department
University of California, Los Angeles
majid@cs.ucla.edu

Afshin Aminian
Orthopaedic Institute
Children's Hospital of Orange County
afshinaminian@hotmail.com

Abstract—The majority of the population around the globe spend a considerable portion of their days seated. This fact can be associated with several factors, such as the circumstances of most of the current jobs and the prevalence of the use of computer systems. One could argue that this knowledge indicates that the impacts of maintaining proper posture while sitting can be observed more than before. Therefore, it is critical to be able to observe, correct, and control our sitting posture throughout the day. Monitoring and correcting our short-term and long-term sitting habits over time can lead to significant improvement in our physical well-being. In this work, we propose WatChair, an AI-powered remote subject monitoring system that assists in short-term sitting posture recognition, activity-level tracking, long-term monitoring, and providing corrective suggestions. Our platform consists of a small wearable component, an application, and a cloud-based back-end. Our framework has been evaluated in practice, and the results of empirical validation and the user-friendliness questionnaire points to a simple, accurate, and user-friendly system for remote sitting posture monitoring. This framework also presents an adaptable solution for general dynamic posture recognition and tracking using wearable systems based on motion sensors.

Index Terms—mobile health, remote health monitoring, machine learning, ehealth

I. INTRODUCTION

The sedentary lifestyle has caused people to spend a considerable portion of their lives seated. The act of sitting has the potential to cause severe short-term and long-term health problems if not done correctly. An example of such problems is the feeling of pain and discomfort in the neck and back area [1]–[3]. Most studies have indicated a strong connection between the sitting posture in performing different activities with outcomes including health status and eating, claiming that controlling it and maintaining proper sitting posture can help the subjects [4], [5]. As simple as sitting as an activity might seem, it has also been known for its impacts on certain types of decision-making, such as online grocery shopping [6].
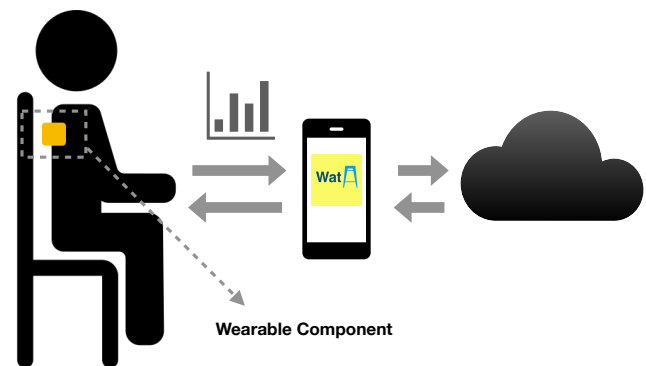

Fig. 1. The system architecture of WatChair platform

For the reasons mentioned above, remote monitoring of sitting posture and attempting to correct it over time is an exciting area of research. This stems from the fact that a convenient and thorough solution for this problem is invaluable as it has the potential to improve the quality of life for many people.

In this work, we introduce WatChair, a platform for continuous monitoring of sitting posture that aids subjects with corrective suggestions. This work attempts to bring a smooth remote monitoring experience to the subjects while attempting to keep the cost and intrusion associated with the data acquisition system as low as possible. This is done by focusing on providing more efficient solutions to the software-related aspects of the problem. Our system will monitor the sitting posture, recognize different sitting posture patterns, and help the subject with a thorough report of the statistics of their sitting posture throughout the day. It also provides them with their history of sitting information, assisting them with easy progress tracking.

This article is structured as follows: In section II, a review of the previous works that are related to sitting habit monitoring is presented. Our system is introduced in section III, and it is followed by the discussion of experiments and results in section IV.

II. RELATED WORKS

Remote monitoring of the human body, postures, and activities is an important area of mobile health research. Many researchers and companies have been working on developing utilities and analytical solutions to monitor, understand, track, and leverage the information related to body posture. These solutions assist subjects in improving physical activities, ranging from general posture to particular therapeutic exercises [7], [8].

There have been numerous research works focusing on monitoring sitting as an important daily activity. These works were mostly focused on analyzing the sitting posture to better fathom its transitions and variations throughout the process of sitting [9], [10]. The system in [11] is designed for sitting behavior tracking and analysis. This work has utilized many stretchable sensors and pressure sensors to prepare a thorough sensor-driven system. Their main contribution is their sitting behavior recognition using neural networks and dynamic time warping.

The system proposed in [12] with the objective of helping mitigate the impacts of poor sitting posture such as pain and discomfort in the back and neck. Their system is based on multiple motion sensors, and it is close to our approach in terms of low cost and affordability. Their system differentiates between sitting, lying, walking, and standing postures. Their framework's main components are cardboard, test pole, and dynamic measure units (DMU) that include accelerometer, gyroscope, and magnetometer.

The framework proposed in [13] follows a different objective of monitoring floor sitting postures. This work proposes the use of a number of pressure sensors in a system to achieve this objective.

A personalized transportable folding device is utilized in [14] to assist with maintaining a better sitting posture while eating. Their results led to the empirical validation of the hypothesis that improving the sitting posture while eating can help with the mitigation of the adverse effects of dysphagia.

The use of sensor-augmented and specialized chairs is also investigated in the literature [15]. In [16], the authors have proposed a system based on a chair embedded with pressure sensors to differentiate between the three main sitting postures of leaning forward, reclining backward or neutral sitting. In terms of the sitting postures that this system attempts to recognize, this work bears considerable similarity to ours. Nevertheless, there is no need for any other circuitry besides a single motion sensor in our approach. A similar grouping of sitting postures can be found in [17] as well. Another work in which a chair is augmented with a large number of electrical sensors is [18]. In a different approach, [19] uses a pressure mat for sitting posture recognition. A sensor-augmented cushion is also used in [20] and [21] for sitting posture detection, which serves as an approach similar to the previously mentioned works.

The work in [22] uses feedback mechanisms for posture correction. The system in [22] is composed of a Kinect device for extracting body landmarks and helping the user by providing feedback, urging them to maintain proper posture. Several other Kinect-based approaches focused on finding unhealthy sitting postures [23]–[26]. In [27], authors combine the information obtained from Kinect with smartwatches to improve the resulting detections. As the sitting posture recognition is especially important for remote health monitoring of elderly patients, a similar work focused on Kinect-based posture recognition for the elderly is presented in [28].

Our contributions in this study and the main differences compared to the mentioned works are as follows:

- The recognition scheme in this work focuses on linking the sensory data to discrete classes corresponding to essential proper and improper sitting postures. The mobile application utilizes the machine learning algorithm and interacts with a single wearable device to associate every time window with the corresponding status and stores the relevant information in the cloud database. It provides the statistics on the short-term evaluation of the subject's posture and activity while seated and shows corrective suggestions. In addition, it reads and visualizes the historical data on the subject's sitting habits, which can help in tracking long-term progress.
- The proposed algorithm is compatible with small sampling frequencies, which leads to a significant reduction in battery consumption on the wearable device and improves the usability of the framework.
- The mobile application provides an easy to use interface to interact with for users and developers. It allows custom labeling for developers, enabling gathering data on custom postures, and storing the data in the cloud database for further analyses.
- The analytical approaches are verified in a Python-based research framework. Our current inference engine for sitting posture recognition is composed of pre-processing steps and linear support vector machine instances, which are implemented in the mobile application as well.
- The experimental setup and the empirical results on our cohort indicate the effectiveness of this system for posture tracking and correction while the subject is seated. As shown in the literature, this can lead to the prevention of complications in short-term and long-term health conditions.

III. PROPOSED SYSTEM

The overall architecture of WatChair is depicted in Figure 1. The first part of the system is a simple wearable component, which includes the motion sensors that we are interested in monitoring. This component would be located on the subjects' back and between the arms and used for sampling data points
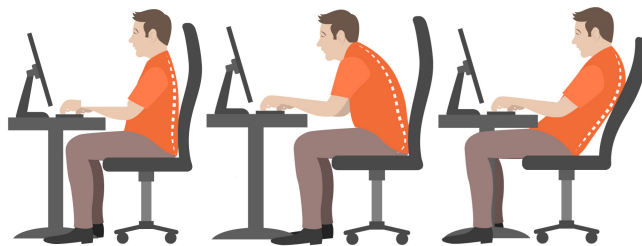
Fig. 2. Three main postures for sitting habit monitoring [29]

Fig. 3. The main interface of our application is shown in this figure.

and transmission. The system obtains its electrical power from a coin battery inside it, and given the frequency of reading, here, there is no need for frequent battery replacement.

This wearable component interacts with an android device using the WatChair app. Upon running the application, if the device is in range, the connection is maintained, and the process of data acquisition, recognition, and transmission to our cloud back-end begins.

The application provides a user-friendly interface between the board, the cloud back-end, and the subject. The user can observe long-term and short-term information about their pattern of sitting and score their seated behavior accordingly. This framework renders it possible to fine-tune the models as more data is gathered in the cloud database. In this work, we attempted to evaluate the performance of a general model with no need for per person calibration. Nevertheless, this framework can be used with ease to provide each user with separate user-calibrated models as well.

In what follows, the details of our approach and the components of this framework are elaborated upon.

A. Data

The main and only sensory part of this framework is composed of affordable, portable, and widely available Meta-Tracker boards [30]. These boards are widely used when motion sensors are to be utilized for evaluating physical readiness, an example being training pilots [31].

The single wearable component of this platform is placed on the subject's back and between the arms, which can be easily done using a strap or cord. The main component for the effective tracking of sitting habits is posture recognition. The main three postures that our platform focuses on recognizing are depicted in Figure 2 [32], and empirical results indicate that the data acquired from this location is mostly sufficient for making such determination. In addition, the sensor alignment can be automatically determined, and the system is flexible in terms of small displacements in using the wearable component.

The wearable component uses a Lithium battery, and in our framework, there is no need for frequent battery replacement. This is due to the fact that our framework is focused on seated posture and is compatible with low-frequency sensor readings. This property and being equipped with 2.4 GHz Bluetooth Low Energy chips enables us to maintain a smooth connection route between the wearable component and the app.

The mobile application then handles the data-related routines, including cloud storage, analytics, and transmission.

B. Application

Another essential component in our system is WatChair's mobile application. Upon launching the application, it writes the configuration necessary for the motion sensors, such as the frequency of sampling, and issues the start command for the wearable component. The information then can be monitored continuously using the app's interface, as depicted in Figure 3.

Afterward, the information is continuously retrieved, filtered, and pushed to our cloud database. For the back-end database, we are using Google Cloud Firestore as a NoSQL database composed of documents and collections. The application builds a communication channel to the Firebase instance in order to efficiently perform the transactions. These transactions include pushing the gathered data to the cloud database, an instance of Cloud Firestore, and reading the historical information for updating the visualizations in the app.

Please note that the related intervention information regarding how to correct the posture is also presented in the application, providing the subject with a more in-depth understanding of how to proceed regarding the given personalized posture correction information.

The quality of the application was surveyed through a usability questionnaire developed according to the Usefulness, Satisfaction, and Ease of use questionnaires in [33], [34]. This questionnaire was given to the subjects in our study,

with the idea of receiving feedback on the application as the main component in this framework. It has scored 9.9/ 10.0, indicating the easiness of using and getting familiar with it.

Machine Learning: The machine learning component of WatChair is composed of an inference pipeline that directly links sensor readings to different sitting postures and characteristics. The trained model weights are used in the application to enable efficient use and effortless alterations in the future. The relevant sensor readings that come from motion sensors (mainly, the three-axis sensors of accelerometer and gyroscope) are buffered and transmitted to the application. In our experiments, we considered the time window of 60 seconds; therefore, the seated behavior of the subject for every minute is represented and used by the model for evaluation. The findings are then stored with the timestamps to help with the progress review.

The recognition pipeline is designed using Support Vector Classification which enables differentiating between the labels using one-vs-all classifiers. The training objective uses hinge loss and is as follows:

$$\min_{\theta,b} \frac{1}{2}\theta^T\theta + C \sum_{i=1}^{N} \max(0, y_i(\theta^T\phi(x_i) + b))$$

In the above formula, θ and b are the single model parameters which are trained, and ϕ is the linear kernel. The Support Vector classification with a linear kernel is chosen so the model complexity could fit the problem well, and the trained model could be easily implemented and used in the mobile application as well.

IV. EXPERIMENTS

To obtain a dataset for our analyses, we have asked 6 subjects in the age range of 24-27 to participate in this study. For gathering the training data regarding each type of sitting behavior, a picture was shown to the subjects depicting such behavior. They were then asked to perform as they usually do in such a posture. While performing their normal activities (e.g., working with their laptops, writing), a sufficient amount of data could be gathered along with the supervision signals.

Our analytical platform is implemented in Python3.6, and using Google Cloud SDK, and Sci-Kit learn machine learning library enables us to retrieve the information from our cloud database and perform the corresponding analytical investigations. For evaluation purposes, several experiments were done, and in each experiment, the data from one subject was used as the test set while the model was trained on the rest.

The empirical results indicate that our model is able to recognize and distinguish between our three main pre-defined sitting postures accurately. The micro and macro average F1-Score for our current model over the three sitting postures are 72.12% and 75.09%, respectively.

The normalized confusion matrix for the predictions is shown in Table I. These results suggest that the system exhibits accurate performance, even though the flexibility of the system in terms of the positioning of the wearable component

TABLE I
MICRO-AVERAGED CONFUSION MATRIX FOR EVALUATING OUR MODEL USING LEAVE-ONE-SUBJECT-OUT SCHEME - EACH ROW CORRESPONDS THE PREDICTIONS FOR A GROUNDTRUTH LABEL

	Upright	Leaning Back	Slouching
Upright	0.65	0.22	0.14
Leaning Back	0.05	0.64	0.31
Slouching	0.02	0.0	0.98

can potentially add to the error level. As another example, "Slouching" data, which can be argued as the main posture of interest, has been efficiently captured.

This work aimed to propose an accurate and efficient pipeline to perform monitoring without the need for per-person calibration. However, it is understandable that in specific use-cases (e.g., certain disabilities, the cohort of elderly), the addition of per-person calibration to improve the performance further might be necessary. In such cases, labeled data can be easily obtained and analyzed using WatChair's interface. This will enable the data analysis and preparations for such cases to take place with ease.

To use the model in the mobile application as well, this inference engine was implemented in Java and is included in the application to perform continuous recognition and update the cloud database accordingly.

V. CONCLUSION

We presented WatChair, an AI-powered remote monitoring framework for short-term and long-term tracking of sitting habits and proposing corrective suggestions. The effectiveness of this system in performing this task is then empirically validated. Given the critical health impacts of improper sitting habits, this low-cost and affordable system, combined with accurate machine learning inference, can improve user behavior while seated and therefore help prevent medical complications that are associated with improper sitting.

REFERENCES

[1] C. F. Nooijen, L. V. Kallings, V. Blom, Ö. Ekblom, Y. Forsell, and M. M. Ekblom, "Common perceived barriers and facilitators for reducing sedentary behaviour among office workers," *International journal of environmental research and public health*, vol. 15, no. 4, p. 792, 2018.

[2] L. Smith, A. Sawyer, B. Gardner, K. Seppala, M. Ucci, A. Marmot, P. Lally, and A. Fisher, "Occupational physical activity habits of uk office workers: cross-sectional data from the active buildings study," *International journal of environmental research and public health*, vol. 15, no. 6, p. 1214, 2018.

[3] Y. Kwon, J.-W. Kim, J.-H. Heo, H.-M. Jeon, E.-B. Choi, and G.-M. Eom, "The effect of sitting posture on the loads at cervico-thoracic and lumbosacral joints," *Technology and Health Care*, vol. 26, no. S1, pp. 409–418, 2018.

[4] I.-Y. Moon, C.-H. Yi, I.-W. Park, and J.-H. Yong, "Effects of sitting posture and bolus volume on activation of swallowing-related muscles," *Journal of Oral Rehabilitation*, 2020.

[5] S. Tinitali, K.-A. Bowles, J. L. Keating, and T. Haines, "Sitting posture during occupational driving causes low back pain; evidence-based position or dogma? a systematic review," *Human factors*, p. 0018720819871730, 2019.

[6] D. Krpan and B. Fasolo, "Revisiting embodied approach and avoidance effects on behavior: The influence of sitting posture on purchases of rewarding foods," *Journal of Experimental Social Psychology*, vol. 85, p. 103889, 2019.

[7] 2020. [Online]. Available: https://www.uprightpose.com/

[8] M. Gwak, S. Fazeli, G. Ershadi, M. Sarrafzadeh, M. Ghodsi, A. Aminian, and J. A. Schlechter, "Extra: Exercise tracking and analysis platform for remote-monitoring of knee rehabilitation," in *2019 IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. IEEE, 2019, pp. 1–4.

[9] S. Ma, S. Hong, H.-m. Shim, J.-W. Kwon, and S. Lee, "A study on sitting posture recognition using machine learning," *The Transactions of the Korean Institute of Electrical Engineers*, vol. 65, no. 9, pp. 1557–1563, 2016.

[10] J. E. Estrada and L. A. Vea, "Real-time human sitting posture detection using mobile devices," in *2016 IEEE Region 10 Symposium (TENSYMP)*. IEEE, 2016, pp. 140–144.

[11] Y. Zhang, Y. Huang, B. Lu, Y. Ma, J. Qiu, Y. Zhao, X. Guo, C. Liu, P. Liu, and Y. Zhang, "Real-time sitting behavior tracking and analysis for rectification of sitting habits by strain sensor-based flexible data bands," *Measurement Science and Technology*, 2019.

[12] C.-C. Wu, C.-C. Chiu, and C.-Y. Yeh, "Development of wearable posture monitoring system for dynamic assessment of sitting posture," *Physical and Engineering Sciences in Medicine*, pp. 1–17, 2019.

[13] I. A. Soenandi, M. Ginting, and B. Harsono, "Real time floor sitting posture monitoring using k-means clustering," in *Proceedings of the 2nd International Conference on Software Engineering and Information Management*, 2019, pp. 194–198.

[14] V. Woisard, M. Costes, H. Colineaux, and B. Lepage, "How a personalised transportable folding device for seating impacts dysphagia," *European Archives of Oto-Rhino-Laryngology*, vol. 277, no. 1, pp. 179–188, 2020.

[15] A. Aminian, "Dynamic orthopaedic chair," Jul. 31 2012, uS Patent 8,231,175.

[16] M. Sandybekov, C. Grabow, M. Gaiduk, and R. Seepold, "Posture tracking using a machine learning algorithm for a home aal," in *Intelligent Decision Technologies 2019: Proceedings of the 11th KES International Conference on Intelligent Decision Technologies (KES-IDT 2019)*, vol. 2. Springer, 2019, p. 337.

[17] C. Lim, S. Basah, M. Ali, and C. Fook, "Wearable posture identification system for good sitting position," *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, vol. 10, no. 1-16, pp. 135–140, 2018.

[18] E. Fragkiadakis, K. V. Dalakleidi, and K. S. Nikita, "Design and development of a sitting posture recognition system," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2019, pp. 3364–3367.

[19] S. Suzuki, M. Kudo, and A. Nakamura, "Sitting posture diagnosis using a pressure sensor mat," in *2016 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA)*. IEEE, 2016, pp. 1–6.

[20] K. Ishac and K. Suzuki, "Lifechair: A conductive fabric sensor-based smart cushion for actively shaping sitting posture," *Sensors*, vol. 18, no. 7, p. 2261, 2018.

[21] A. Hwang, "Arduino pressure sensor cushion for tracking and improving sitting posture," *International Journal of Biomedical and Biological Engineering*, vol. 13, no. 11, pp. 454–460, 2019.

[22] R. Baptista, M. Antunes, A. E. R. Shabayek, D. Aouada, and B. Ottersten, "Flexible feedback system for posture monitoring and correction," in *2017 Fourth International Conference on Image Information Processing (ICIIP)*. IEEE, 2017, pp. 1–6.

[23] S. R. D. Maryam and S. Payandeh, "A novel human posture estimation using single depth image from kinect v2 sensor," in *2018 Annual IEEE International Systems Conference (SysCon)*. IEEE, 2018, pp. 1–7.

[24] Y. Xu, J. Chen, Q. Yang, and Q. Guo, "Human posture recognition and fall detection using kinect v2 camera," in *2019 Chinese Control Conference (CCC)*. IEEE, 2019, pp. 8488–8493.

[25] S. He, "A sitting posture surveillance system based on kinect," in *Journal of Physics: Conference Series*, vol. 1026, no. 1. IOP Publishing, 2018, p. 012022.

[26] L. Yao, W. Min, and H. Cui, "A new kinect approach to judge unhealthy sitting posture based on neck angle and torso angle," in *International Conference on Image and Graphics*. Springer, 2017, pp. 340–350.

[27] M. Tariq, H. Majeed, M. O. Beg, F. A. Khan, and A. Derhab, "Accurate detection of sitting posture activities in a secure iot based assisted living environment," *Future Generation Computer Systems*, vol. 92, pp. 745–757, 2019.

[28] N. W.-H. Tsang, K.-Y. Lam, U. M. Qureshi, J. K.-Y. Ng, I. Papavasileiou, and S. Han, "Indoor activity tracking for elderly using intelligent sensors," in *Intelligent Data Sensing and Processing for Health and Well-Being Applications*. Elsevier, 2018, pp. 197–222.

[29] [Online]. Available: https://www.freepik.com/free-vector/illustration-man-with-correct-incorrect-posture_1158025.htm

[30] "Metatracker," Feb 2019. [Online]. Available: https://mbientlab.com/metatracker/

[31] P. Kutilek, P. Volf, J. Hejda, P. Smrcka, J. Adolf, V. Krivanek, L. Lhotska, K. Hana, R. Doskocil, J. Kacer *et al.*, "Non-contact measurement systems for physiological data monitoring of military pilots during training on simulators: Review and application," in *2019 International Conference on Military Technologies (ICMT)*. IEEE, 2019, pp. 1–6.

[32] [Online]. Available: http://ctms0168.epizy.com/2017/12/20/an-all-rounded-helper-to-help-you-improve-your-sitting-postures-2/?i=2

[33] J. R. Lewis, "Ibm computer usability satisfaction questionnaires: psychometric evaluation and instructions for use," *International Journal of Human-Computer Interaction*, vol. 7, no. 1, pp. 57–78, 1995.

[34] A. M. Lund, "Measuring usability with the use questionnaire12," *Usability interface*, vol. 8, no. 2, pp. 3–6, 2001.