2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)

On the Development of Smart Home Care: Application of Deep Learning for Pain Detection

Hermawan Nugroho
Center for Digital Future
Faculty of Engineering, Computing and Science
Swinburne University of Technology, Sarawak Campus
Kuching, Malaysia
hnugroho@swinburne.edu.my

Dani Harmanto
University of Derby
Derbyshire, United Kingdom
d.harmanto@derby.ac.uk

Hamada Rasheed Hassan Al-Absi
Center for Digital Future
Faculty of Engineering, Computing and Science
Swnburne University of Technology, Sarawak Campus
Kuching, Malaysia
halabsi@swinburne.edu.my

***Abstract*—In Malaysia, as it is the case in many other countries, the population of the elderly is increasing, and as they age, unfortunately, many suffer of chronic diseases. Smart homes can enable continuous and remote monitoring of elderly health and wellbeing at a low cost. The technologies will allow the elderly to stay in their home instead of expensive and limited healthcare facilities. In smart home settings, video and camera are usually used to monitor. In this paper, we present a progress on smart home system to detect pain through the analysis of facial expressions using camera or video. The system uses deep learning techniques and is developed to work with portable devices. Preliminary result shows an accuracy of 93% in detecting pain on a standard facial dataset.**

Keywords—Deep Learning, Pain Detection, Smart Home Care, Facial Expressions

I. INTRODUCTION

In our society today, the population of the elderly is increasing for the past few decades. In Malaysia, it is reported that the number of Malaysians aged 60 years and above are estimated to be around 1.4 million people. This number is projected to reach an all-time high, which are approximately 3.3 million individuals by the year 2020. The number is double in comparison from the year 1990 [1], [2].

Living independently can be a problem for the elderly, especially those with chronic diseases where they experience pain that in many cases goes underrated or undetected [3]. Undetected pain by relatives or even health centers' personnel could affect the quality of life of the elderly especially those who are not able to communicate the pain such as those with dementia [3], [4].

Detecting pain using a computer system is one way of tracking pain non-invasively and it could be able to provide valuable insights on the frequency of pain felt by them. In fact, while people around the patient may miss a felt pain (regardless of its degree), a smart system that analyses facial expressions could be able to detect pain and alert caregivers [5].

Liu et al [6] reported that the use of cameras and video as one of the main tools used in smart home for monitoring the elderly. Smart home is an automation of home functions with embedded systems. Generally, in a smart home, internet-connected devices are used to monitor and manage appliances such as lighting and air conditioning systems. Smart home for elderly is a smart home that focuses not only to assist the elderly living independently but also to monitor the elderly and to contact the nearest caregivers if something happens. Smart home for elderly can provide the opportunity for old people to live comfortably in their homes without the need for an expensive caring nurses or caring centers and therefore ensures their privacy and empowers them to keep being independent and contribute to their well-being and better quality of life [7] [8].

Currently, we are developing a smart home system tools, which can be installed on a portable device, which could be fixed anywhere (i.e bedroom, living room…etc) with the capabilities of recording and analyzing images instantly. With the recent progress in deep learning, we believe that such system can be built. In this paper, we use deep learning method to detect pain in face images.

II. APPROACH

A. Pain – Is it visually discernible ?

Pain is defined as an unpleasant subjective sensory and emotional experience associated with present or potential injury [9]. Pain is often accompanied by changes in behavior. Like expressions of emotion, researcher believe that facial expressions during pain could give a critical indication in communicating information about the pain experience [10].

The ability to measure pain is important in any clinical settings for determining proper treatment, assessment of treatment efficacy or the appropriate dosage of medicine. The Facial Action Coding System (FACS), proposed in 1978 by Ekman and Friesen is the most popular system on facial expression definition [11]. At first, FACS was designed to provide objective descriptions of six (6) basic emotions, namely joy, disgust, anger, fear, surprise and sadness. In FACS, each facial action is described in terms of one of 44 individual action units (AUs). To determine objectively, both the frequency as well as the intensity of each AU are classified in a 5-point scale [9].

Researchers found that the expression of pain is unique. Prkachin found that four facial units (AUs) in FACS carried most important information about pain. There are brow lowering (AU4), orbital tightening (AU6 and AU7), levator contraction (AU9 and AU10) and eye closure (AU43). Prkachin and Solomon reported that pain can be defined as the sum of intensities of brow lowering, orbital tightening, levator contraction and eye closure [12]. The Prkachin and Solomon pain intensity (PSPI) metric is defined as,

$$Pain = AU4 + (AU6\ or AU7) + (AU9\ or AU10) + AU43 \tag{1}$$

The sum of AU4, AU6 or AU7 (whichever is higher in intensity), AU9 or AU10 (whichever is higher in intensity) and AU43 will generate a 16-point scale, which can be used as pain assessment scale.

The PSPI FACS pain scale is found to be the only metric, which can define pain on a frame-by-frame basis [12]. Therefore, it is a suitable metric to monitor pain from facial expression in smart home.

B. Related Works – Focus on Deep Learning Approaches

In this work, we will look at previous works, which used PSPI metric in their pain assessment with focus on deep learning approaches. Lucey et al developed pain detection algorithm using Active Appearance Model (AAM) based computer vision system, which detect pain based on facial expressions coded of FACS [13]. In their work, face is firstly tracked using an AAM. Next, the AAM will provide shape and appearance features that are used for classifying AUs using Support Vector Machine (SVM). The output of SVMs for AUs are fused together using linear logistical regression (LLR). Lastly, the fusion score is used for pain detection. The system achieved best accuracy of 90% for AU43 detection and the worst (53%) for AU4.

Reza et al, used Convolutional Deep Belief Network (CDBN), which is a multilayer generative model for fully unsupervised feature learning and Support Vector Machine as the classifier to detect pain in frames [11]. Their system runs on Matlab 8.4.0 on a laptop with i7 2.4 GHz CPU and 8 GB memory and achieved 87.2% accuracy.

Jing Zhou et al developed different deep learning approach to detect pain. In their work, they used recurrent convolutional neural network (RCNN) [14]. A convolutional neural network (CNN) is a class of deep, feed-forward artificial neural networks and a recurrent neural network (RNN) is a class of artificial neural network where connections between nodes form a directed graph along a sequence. CNNs are usually used to map image data to an output variable and RNNs are usually designed to work with sequence prediction problems. The idea of RCNN is to add recurrent connections within every convolutional layer of the feed-forward CNN. It is reported that their proposed method achieved 1.54 MSE. The algorithm runs on two 2.30GHz Intel(R) Xeon(R) E5-2650 v3 CPU, 320GB RAM, and an NVIDIA(R) Tesla K80 GPU.

Feng Wang et al also developed different deep learning approach. Their method is based on modified face verification network [15]. The face verification network is a CNN with a new loss function, namely center loss [16]. The new loss function will minimize the intra-class variations while keeping the features of different classes separable. This can be done by 1) updating the loss function on mini-batch (instead of entire training set) and 2) using a scalar α to control the learning rate of the loss function. The algorithm achieved 0.80 MSE and runs on system with Matlab and GPU with CUDA support.

For smart home care, we need pain detection algorithm, which can run off-the-shelf in mobile scenarios. As trade-off, the algorithm may have less accuracy than the state-of-the-art. However, the accuracy still need to be in an acceptable rate.

C. OpenFace Approach

OpenFace is a face recognition algorithm built in mobile scenarios. The algorithm can give high accuracy with low training and prediction times [17]. Fig. 1 shows the project structure of OpenFace.

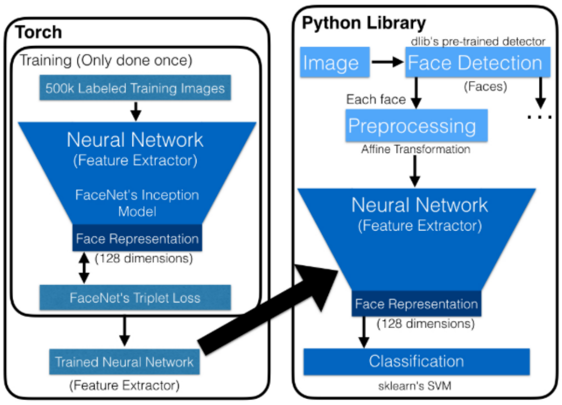


Fig. 1 Project structure of OpenFace (reproduced from [17])

As shown in the structure, OpenFace use neural network architecture taken from Google's FaceNet [18] as the main feature extractor tools.

FaceNet is a unified system for face verification (is this the same person), recognition (who is this person) and clustering (find common people among these faces). It trains the outputs to be a compact 128-D embedding using a triplet-based loss function. Fig. 2 illustrates the 128-D embedding of an image.

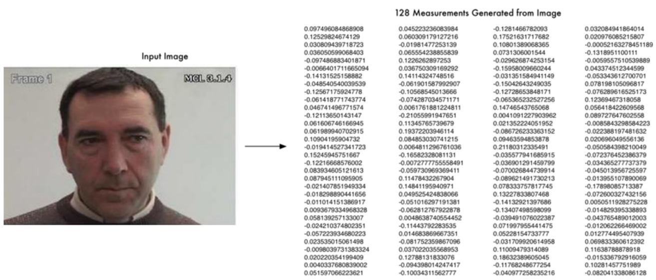


Fig. 2 128-D of an image

The triplets consist of two images, which comes from similar source and a non-matching image, which comes from different source. The triplet loss function objective is to separate the positive pair from the negative pair by a distance margin.

OpenFace use a modified FaceNET's NN4 network, which is based on GoogLeNet Inception models [19]. Table I shows the deep neural network architecture used in FaceNET

Table I Inception Model of FaceNet

Type	Output size	Depth
conv1	112 x 112 x 64	1
max pool +norm	56 x 56 x 64	0
inception (2)	56 x 56 x 192	2
norm + max pool	28 x 28 x 192	0
inception (3a)	28 x 28 x 256	2
inception (3b)	28 x 28 x 320	2
inception (3c)	14 x 14 x 640	2
inception (4a)	14 x 14 x 640	2
inception (4b)	14 x 14 x 640	2
inception (4c)	14 x 14 x 640	2
inception (4d)	14 x 14 x 640	2
inception (4e)	7 x 7 x 1024	2
inception (5a)	7 x 7 x 1024	2
inception (5b)	7 x 7 x 1024	2
avg pool	1 x 1 x 1024	2
fc	1 x 1 x 128	1
L2 normalization	1 x 1 x 128	1

The modified OpenFace network will have inception layers of 4b, 4c, 4d removed and smaller inception layers of 5a and 5b.

D. Triplet Loss For Pain Detection.

In Triplet Loss, we minimize the distance between anchor and a positive, both of which come from same source and e maximizes the distance between the anchor and a negative of a different source.

The deep learning network will produce embedding from an image. It embeds an image x into a d-dimensional Euclidean distance. The method wants to ensure that an image, x_i^a (anchor) is closer to all other images, x_i^p (positive), which comes from similar person, than to any images, x_i^n (negative) of any other person.

It can be written as,

$$\left\|x_i^a - x_i^p\right\|^2 + \propto \ll \|x_i^a - x_i^n\|^2 \tag{1}$$

$\propto$ is a margin, which can be used to set whether the loss function is strict or loose.

The objective function is to find L, such that

$$L = \max(\left\|x_i^a - x_i^p\right\|^2 - \|x_i^a - x_i^n\|^2 + \propto ,0) \tag{2}$$

which indicates that ideally,

$$\left\|x_i^a - x_i^p\right\|^2 - \|x_i^a - x_i^n\|^2 + \propto < 0 \tag{3}$$

For pain detection, we hypothesized that by set the margin, $\propto$, small enough, the OpenFace algorithm can be modified to detect pain in images.

III. Results and Discussion

To test our hypothesis, we used UNBC-MacMaster Shoulder Pain Expression Archive in our experiment [12]. The dataset contains a recorded video of the face of patients suffering from should pain. There are 25 subjects comprising of 48,398 total frames. They also provided the PSPI (Prkachin and Solomon Pain Intensity) scores for most of the subjects.

A. Experiment Settings

From the dataset, we consider only subjects that had PSPI scores. From these subjects, we looked at their PSPI scores and selected subjects that had minimal PSPI scores of 10. As an initial study, we would like to ensure that the data used in the experiment is representative. There were 14 subjects, which follows our requirements.

From each of these 14 subjects, we identified two sets of images, namely pain image set and normal image. These sets will be also be divided into training set and testing set respectively as shown in Fig. 3. The images were taken randomly from the video frames.

Fig. 3 Normal image set of A divided into training set and testing set

The experiment will first run the training sets of the subjects' normal and pain sets to produce the embedding sets of normal and pain sets. Next, we run the testing sets of subjects' normal and pain sets to evaluate the algorithm.

B. Result – Margin Analyis

In section 2, we hypothesized that the OpenFace algorithm can be modified to detect pain by setting the margin, $\propto$.

To analyse the margin effect, we run pain image sets from the testing sets of each subject with different value of the margin, $\propto$. The initial margin $\propto$ was set at 0.6. To prove our hypothesis, we did another experiment with $\propto$ was set to 0.2. Fig. 4 shows examples that we recorded from three (3) subjects.

Fig. 4 a) Pain image of subject no 42 with margin = 0.6; b) Pain image of subject no 42 with margin = 0.2; c) Pain image of subject no 47 with margin = 0.6; d) Pain image of subject no 47 with margin = 0.2; e) Pain image of subject no 59 with margin = 0.6; f) Pain image of subject no 59 with margin = 0.2

As shown from Fig. 4, if we set the margin $\propto$ small enough (in our case is 0.2). The algorithm was able detect pain in from the testing sets. However, if we set it to its initial value, the algorithm was not able to classify the pain images. Besides subject no 42, 47 and 59, we did the similar experiments to all 14 subjects included in our experiment. Results showed that the algorithm was able to detect pain from the pain sets for all subjects if the margin was set to 0.2. This indicates that our hypothesis is correct.

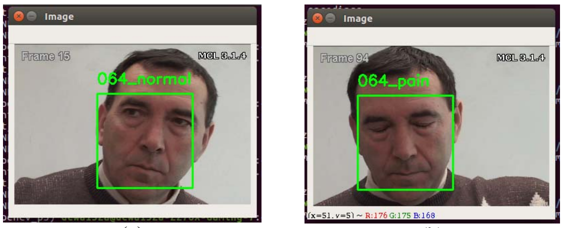
C. Result - Accuracy Analysis

To prove the accuracy of the algorithm, we tested the system the testing sets from 14 subjects. Table II shows the recapitulation of the experiment.

Table II Accuracy Analysis

Subject No	Pain images (Correctly detected / Total images)	Normal image (Correctly detected / Total images)
42	3/3	30/30
43	1/1	30/30
47	4/4	31/31
49	3/3	30/30
52	2/3	38/38
59	4/4	35/35
64	4/4	34/34
66	3/3	32/32
80	6/6	30/30
92	3/3	30/30
95	0/1	30/30
96	2/3	32/32
97	2/2	32/32
107	4/4	33/33
Total	41/44	447/447

Results shows that the propose algorithm achieved 93% accuracy for pain detection (41/44) and 100% accuracy for normal image (447/447). Fig. 5 shows example that the algorithm correctly identified normal and pain from subject no 64.



(a) (b)

Fig. 5 a) Correct detected – normal b) Correctly detected – pain

The result however was lower if we consider individual samples. From 14 subjects, the proposed algorithm can detect pain in 13 out of 14 subjects (92%). The only subject that the algorithm failed was subject no 95 in which we only had 1 test image.

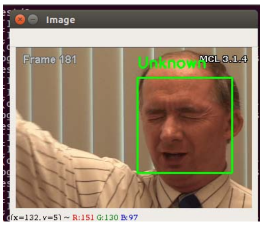


Fig. 6 Failed detection

Fig. 6 shows the failed pain detection sample. We believe this was due to the triplet selection problem. In FaceNet paper, it was stated that correct triplet selection was crucial. So if we had limited number of images for training, it would be hard for the algorithm to perform selection [18].

IV. Conclusion

Smart home is an automation of home function with embedded systems. With the rising number of elderly population worldwide, researchers agree that smart home will enable the elderly to live independently in their home instead of expensive healthcare facilities.

Smart homes for elderly healthcare are usually installed with monitoring and safety devices such as lighting and motion sensors, environmental controls, video cameras, automated timers, emergency assistance systems, and alerts. In this paper, we presented a pain detection system for smart home care.

The system uses deep learning to analyse face images and detect pain from camera or video. The developed algorithm is based on OpenFace and FaceNet topologies. Results show that the preliminary experiment archives 93% accuracy in detecting pain and 100% in detecting normal (no-pain) images. This indicates that the algorithm has a great potential to be implemented in a smart home setting.

For future works, we will perform an accuracy analysis on frame-on-frame basis and develop an unsupervised method

Acknowledgment

The authors would like to thank Dr. Patrick Then for his support. We would like also to thank NVIDIA for the GPU Grant Program. This work was a collaborative work with University of Derby and supported in part by the Center for Digital Futures, Swinburne Sarawak.

References

[1] W. A. Wan-Ibrahim and I. Zainab, "Some demographic aspects of population aging in Malaysia," *World Appl. Sci. J.*, vol. 30, no. 7, pp. 891–894, 2014.

[2] M. Mafauzy, "The problems and challenges of the aging population of malaysia.," *Malays. J. Med. Sci.*, vol. 7, no. 1, pp. 1–3, 2000.

[3] T. Hadjistavropoulos *et al.*, "Pain assessment in elderly adults with dementia," *The Lancet Neurology*. 2014.

[4] A. Karki, "Pain Assessment in Elderly Residents with Severe Dementia – Nurses' Point of View," Helsinki Metropolia University of Applied Sciences, 2018.

[5] L. Y. Mano *et al.*, "Exploiting IoT technologies for enhancing Health Smart Homes through patient identification and emotion recognition," *Computer Communications*. 2016.

[6] L. Liu, E. Stroulia, I. Nikolaidis, A. Miguel-Cruz, and A. Rios Rincon, "Smart homes and home health monitoring technologies for older adults: A systematic review," *Int. J. Med. Inform.*, vol. 91, pp. 44–59, 2016.

[7] S. Majumder *et al.*, "Smart homes for elderly healthcare—Recent advances and research challenges," *Sensors (Switzerland)*, vol. 17, no. 11, 2017.

[8] Q. Lê, H. B. Nguyen, and T. Barnett, "Smart Homes for Older People: Positive Aging in a Digital World," *Futur. Internet*, vol. 4, no. 4, pp. 607–617, 2012.

[9] R. Rojo, J. C. Prados-Frutos, and A. López-Valverde, "Pain assessment using the Facial Action Coding System. A systematic review," *Med. Clínica (English Ed.*, vol. 145, no. 8, pp. 350–355, 2015.

[10] K. M. Prkachin and P. E. Solomon, "The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain," *Pain*, vol. 139, no. 2, pp. 267–274, 2008.

[11] R. Kharghanian, A. Peiravi, and F. Moradi, "Pain detection from facial images using unsupervised feature learning approach," *2016 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 419–422, 2016.

[12] P. Lucey, J. F. Cohn, K. M. Prkachin, P. E. Solomon, and I. Matthews, "Painful data: The UNBC-McMaster shoulder pain expression archive database," *2011 IEEE Int. Conf. Autom. Face Gesture Recognit. Work. FG 2011*, pp. 57–64, 2011.

[13] P. Lucey *et al.*, "Automatically detecting pain in video through facial action units," *IEEE Trans. Syst. Man, Cybern. Part B Cybern.*, vol. 41, no. 3, pp. 664–674, 2011.

[14] J. Zhou, X. Hong, F. Su, and G. Zhao, "Recurrent Convolutional Neural Network Regression for Continuous Pain Intensity Estimation in Video," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, pp. 1535–1543, 2016.

[15] F. Wang *et al.*, "Regularizing face verification nets for pain intensity regression," *Proc. - Int. Conf. Image Process. ICIP*, vol. 2017–Septe, pp. 1087–1091, 2018.

[16] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A Discriminative Feature Learning Approach for Deep Face Recognition," in *ECCV*, 2016, vol. 1, pp. 499–515.

[17] B. Amos, B. Ludwiczuk, and M. Satyanarayanan, "OpenFace: A general-purpose face recognition library with mobile applications,"

Tech. Rep. C. C. Sch. Comput. Sci., vol. 16, no. 118, 2016.
[18] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2015.
[19] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2015.