Prediction of Fatigue Using Deep Learning Models

Ph.D. Thesis

ELSAYED ABDELGHANY MOHAMED ABOUSHARARA

March 2021

Prediction of Fatigue Using Deep Learning Models

By

ELSAYED ABDELGHANY MOHAMED ABOUSHARARA

Doctoral thesis approved by the Graduate School of Advanced Technology and Science for the degree of

Doctor of Engineering

in

Systems Innovation Engineering

March 2021

Department of Information Science and Intelligent Systems

Tokushima University, Japan

ACKNOWLEDGEMENTS

In the name of Allah, Thanks Allah, the creator of human and universe. “He taught man that which he knew not". Allah who ordered us never stop reading, thinking and exploring ourselves, and the marvelous world around us. My appreciation and greetings to all of you my dearest guests, to Professor Kenji Terada my supervisor who show huge support to my thesis. to Assoc. Prof. Stephen Karungaru for his very helpful advice, to Dr. Akinori Tsuji for his powerful assist, to my international colleagues for taking an important part and accepting participation in this research's experiments, to whoever respectfully took from his valuable time to believe in this research, considered correcting and altering it and made the effort to improve its content. This is dedicated to my Dad who left this world before I even realize what death means! To my Mother who passed away before reaping what she saw! to my small family, my beloved wife Mona and the apple of my eye Rowan, Rana and Hamza, to every single person in my bigger family, my long time best friends, to everyone who encouraged me to start my research journey in my forties. Here I am, standing in front of you today, dedicating this Ph.D. to all of you.

CONTENTS

ABSTRACT

A headache, pain and drowsiness are the most common symptoms of fatigue caused by a long duration of work using a visual display terminal (VDT). A sign of the headache generally involves placing a hand on the head, eyes, nose, or face. The recognition of these gestures is a challenging problem due to the difficulty in similar skin color of hands and face. Drowsiness detection is one of the interesting research on computer vision regarding driver's fatigue. Previous studies of drowsiness detection have mainly focused on tracking of the eyes and mouth conditions like eyes blinking, yawing, and so on. The detection of hand gesture used in this work as the sign of drowsiness such as the hand cover on mouth and with the closure of the eyes by this hand gesture sign. The applications of this research can expand from only the observation of drowsiness for drivers to observing the status of drowsiness at other environments, for example, call center agents, control room observers, and air traffic controllers. Facial expressions and hand gestures recognized as a part of human emotions, especially in a feeling of fatigue signs. In computer vision research, a positioning of hand over face is one of the challenging problems. Therefore, the challenges summarize as follows: The difficulty of the difference of skin color for

hands and face with the same texture, the limitation of illumination change, and the performance of traditional feature extraction methods.

In this research, modeling the symptom characteristics of fatigue like headaches, upper pain, and drowsy using the hand pose e.g. The hand on the forehead or covering the eye or nose and develop a system for recognizing the fatigue sign and alert the computer users to take rest and do some optimum exercise to staying healthy and safe. The proposed system is easy to install on a general PC because it only uses an RGB camera. There is no need for distinct gadgets and markers. Multiple methods used to improve the classification accuracy and avoid the face detection failure in the case of the hand over face gestures by performing the skin and face detection and the deep learning method for classification. In the proposed method, a deep learning based on a convolutional neural network (CNN) for the classification of the hand poses is applied. In addition, a class activation map (CAM) to visualize the prediction of the classification network for localization of the hand over face poses implemented. The CAM used for three tasks: the first one is to visualize the important area guided the prediction decision, it found the conflict caused by the error of prediction and modified the data and trained the network again until the higher classification accuracy without overfitting. The second task is to avoid under fitting by designing the deep network based on the visual explanation of prediction on every convolutional layer. In the third task, CAM used for localization of the hand over face poses. The proposed method was designed to reduce the type-I error (false-positive result) for the classification. The main Contributions in this research includes a framework for fatigue detection using computer vision and deep learning, computer vision Model for recognizing the sign of fatigue using hand pose, a method for hand localization without annotations dataset and a method to solve the hand over face occlusion problem. From the experimental results, the hands over face gestures achieved classified under illumination change (50–500 lx), head orientation (∼45◦ left or right) and solved the occlusion problem of the hand over face. The drowsiness successfully recognized in the following cases:

(1) The eyes closed and opened the mouth.
(2) The eyes closed, and the hand covered on the mouth. The system achieved the classification ratio of drowsiness average 96.6 % and pain signs average 99.5%.

A headache prediction the hand poses as the signs of frontal and unilateral headaches without the classification overfitting and data biasing errors successfully classified with high accuracy recognition ratio of 98.5% for classification of the hand over face poses as the prediction of headaches.

Finally, as a social impact, this research can help the computer vision and physical therapy researchers to design new algorithms to recognize the visible fatigue signs using images captured by computer cameras, mobile phones, surveillance cameras in the workplace, and care for the elderly, children and disability humans.

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

1.1 Motivation

A headache, pain and drowsiness are the most common symptoms of fatigue caused by a long duration of work using a visual display terminal (VDT) [1-3]. The motivation of this research described in the following subsections.

Health Phenomena: Laborer in various business domains are rapidly increasing the use of visual display terminals (VDT) such as computer screens and liquid crystal display (LCD) monitors. Sitting for a long time in front of VDTs changes the working habits and reduces the physical movements of workers. It has negative effects on the health of the workers [4-8]. In addition, the era of Covid-19 as a side impact of electronic activities because of increasing dependence of online service in education, business and social communication, researchers are talking about new a phenomena called “zoom fatigue“, “e- fatigue “ to describe the fatigue caused by using computers and smartphones for a long time without taking rest or doing exercise[9]. The danger of this phenomenon is that, if fatigue continues, it lead to “burnout” and death due to overwork named in Japan “KAROSHI”. [10]

Impact of Quality Control: Despite the efforts to automate quality control processes in companies, the utilization of a machine besides a human is still the best solution because of the quality and cost [11]. Several factors such as lighting, noise, and working hours directly affect the performance of the work. In addition, Mental and physical condition, may appear in the form of stress exhibited as a sign of fatigue, e.g., covering the eyes by hands or closing the eyes. These signs lead to some risk and error of the quality control process when vision obscures for a second. These actions become a dangerous situation in permanent vigilance such as control room observers, air traffic controllers, etc.

Computer Vision Challenges: Facial expressions and hand gestures recognized as a part of human emotions especially in a feeling of fatigue signs. In computer vision research, a positioning of hand over face is one of the challenging problems. Therefore, the challenges summarized as follows:

- The difficulty of the difference of skin color on hands and face with the same texture.

- The limitation of illumination change.
- The performance of traditional feature extraction methods.

1.2 Problem Statement

Modeling the symptom characteristics of fatigue like headaches, upper pain, and drowsy using the hand pose e.g. the hand on the forehead or covering the eye or nose and develop a system for recognizing the fatigue sign and alert the computer users to take a rest and do some optimum exercise to stay healthy and safe.

1.3 Research Approach

In the proposed method, a deep learning based on a convolutional neural network (CNN) for the classification of the hand poses is applied. A class activation map (CAM) to visualize the prediction of the classification network for localization of the hand over face poses was implemented.

1.4 Key Contributions and Research impact

- Framework for fatigue detection using computer vision and deep learning.
- Computer Vision Model to recognize the sign of fatigue using hand pose.
- A method for hand localization without annotations Dataset.
- A method to solve the hand over the face occlusion problem.
- The system is robust against skin color, illumination change and head orientation.
- For sustainable development goals: Goal No.3 good health and well-being. Goal No.8 decent work and economic growth. This research contributes to improve workers health and environments in order to save lives.
- Help the computer vision and physical therapy researchers to design new algorithms to recognize the visible fatigue signs using images captured by computer cameras, mobile phones, surveillance cameras in the workplace, and care for the elderly, children and People with disabilities.

1.5 Structure of Thesis

The rest of thesis is structured as follows: **Chapter 2** discusses the literature review. **Chapter 3** explains the implementation of pre-processing. **Chapter 4** describes the proposed method framework, model and dataset. **Chapters 5, 6 and 7** presents the experimental results, discussion, and comparative results. **Chapter 8** discusses the Comparison, Evaluation and Conclusion as shown in Fig 1.1.

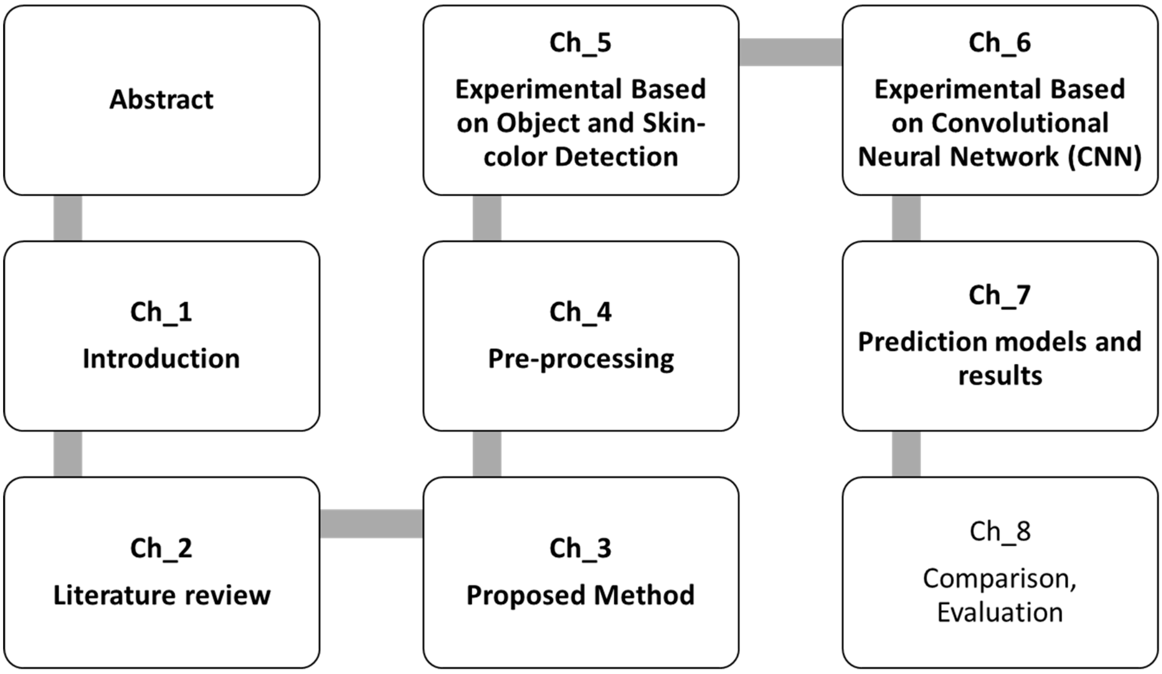


Fig. 1.1. The structure of thesis.

Chapter 2

Literature Review

The related work in the computer vision topics of fatigue detection, the hand over face, and visualization will discuss in this chapter.

2.1 Fatigue Detection

The most popular computer vision topic related to this research is the **driver's** fatigue detection as shown in Fig. 2.1. The driver's fatigue detection methods focus on **drowsiness** detection as a sign of fatigue, but Fatigue described as the lack of energy and motivation, both physical and mental [12][13]. This is different from drowsiness, a term that describes the need to sleep. [14-19] In addition, the research of fatigue detection for VDT workers focused only on the eyestrain, so-called

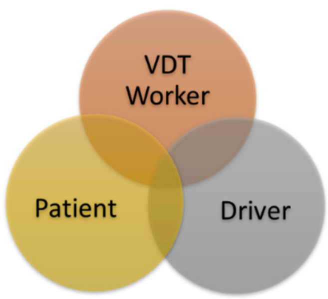


Fig. 2.1. The main fatigue detection applications in computer vision research.

computer vision syndrome. This research combined with research to recognize driver fatigue that the area of interest (ROI) is the eyes. Other studies revealed that VDT workers often suffer from **neck and shoulder pain**. [20]. Lin, et al. (2008) published a study about **visual fatigue** as one of the VDT worker problems [21]. Hua Chen, et al. (2015) proposed an approach for **chronic fatigue syndrome** diagnosis using feature extraction based on image processing [22].The following review in table 1.1 of previous research in fatigue detection for the **non-intrusive methods** for **behavioral analysis** using **vision-based** approaches was included:

1- Fatigue parameters

2- Region of interested (ROI).

Table 1.1. Summary of fatigue detection methods.

Reference	**Parameters**	**ROI**	**Detection techniques**	**Classification methods**	**Accuracy**
Lew, M.; et al. [23]	Blinking and yawn	Face and eyes	Gabor filters	SVM	96.0 %
Saradadevi, et al [24]	Yawn	Mouth	Haar-wavelet	SVM	81.0 %
Xiao, F.; et al [25]	Yawn	Mouth	Gabor wavelets	LDA	91.0 %
Zhang, Z.; et al [26]	Eye closure	Eye	Haar algorithm	SVM	99.0 %
Flores, M.; et al [27]	Drowsiness	Face and eye	Gabor filters	SVM	93.0 %
Yin, B.-C.; et al [28]	Dynamic features	Face	Gabor filters	Ada boost	98.3 %
D'Orazio, T.; et al [29]	Eye closure	Eye	Hough transform	Neural network	95.0 %

3- Detection techniques.

4- Classification methods.

2.2 The Hand over Face

Many research domains in computer vision related to hands over face gestures as shown in Fig. 2.2. A hand over face gesture has an occlusion problem caused by similar skin colors and texture of the face and hands. One solution is to use a depth camera in order to detect the distance of hands and face. Even though it is easy to recognize the gestures, the depth camera cost is high and difficult to use for office computers in working environment. Marwa Mahmoud et al. used a force field method with edge detection and a Gabor filter [30][31]. Jun Xa and Xiong Zhang used a force field method with Camshift tracking [32]. These force field methods with edge detection and Camshift applied after the face detection process to segment the hands from the face. In this case, if the face detection fails once, then the post process also generally failed. Behanz Nojavanasghari et al. presented an improved

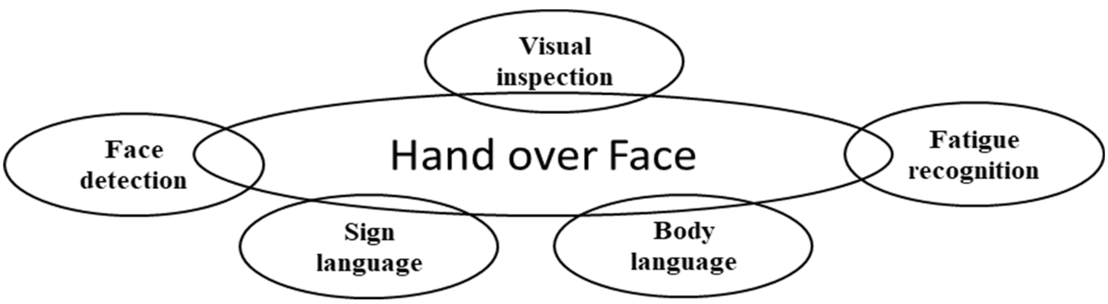


Fig. 2.2. Research domains in computer vision related to hands over face gestures.

method divided a face into eight regions. It applied a convolutional neural network for each small region [33]. This method also has a limitation at detecting the hand over face poses, especially failed in the face borders and nose area. Elsayed A. Sharara et al. used the skin color and facial feature detection to classify the hand poses on the face or shoulders, however it has limitation of illumination change. [34]

2.3 Visualization

Visualizing a deep network is also an interesting topic of classification. Visual explanation and highlighting the specific region improve the efficiency of the classification. Bolei Zhou et al. introduced the Class Activation Mapping (CAM) for the object localization without any bounding box annotations. Based on CNN with a global average pooling [35]. Ramprasaath R. Selvaraju et al. proposed a visual explanation for CNNs called Gradient-weighted Class Activation Mapping (Grad-CAM) [36]. Zhang, Xufan, et al. Designed interactive visualization tool "NeuralVis "amid to analyzing and understanding structure and behaviors of neural network models [37]. Meng, Fanman, et al. Generate CAM by using a few representative classes, to extracting discriminative cues. [38]

2.4 Discussion

Fatigue Detection: In the related research, many algorithms have been proposed to monitor the face features especially eyes and mouth for purpose of fatigue detection using a wide range of feature extraction and classification methods such as color-space transformation, edge detection, Haar-like, template matching, and neural network, etc. the main challenges ; the change of illumination and face occlusion. **The Hand over Face:** The challenges of hand pose gestures summarized as follows: Identifying the difference of the hands and face with similar skin color and texture, and Lack of the performance due to the traditional feature extraction methods. To overcome these problems, recent research mainly classifies three categories of gesture recognition:

1- Uses a distinct camera such as a depth camera.

2- Applies the traditional feature extraction methods, e.g., edge detections, Gabor filters, and CAMShift.

3- Uses a deep learning based on a convolutional neural network (CNN).

Visualization: The output of CNN visualization methods used for Visualizing; the feature in CNN layers, prediction decision area and object localization.

2.5 Contributions

- **Fatigue Detection :** Referring to medical research [1-10] that define the fatigue caused by staying a long time in front of the screen and its signs, **In this research we introduced a new framework to identify fatigue, including headaches, pain, and drowsiness.**
- **The hand over Face:** A combination between deep learning and pre-image processing used to classify the **hand over face gestures for purpose of fatigue detection** for the VDT workers. These consist of the hand on the forehead, eyes, nose, mouth, neck and shoulders.
- **Visualization:** Detection of highlighted area that guided the classifier prediction decision based on the Class Activation Mapping (CAM), used in this work to rearrange the dataset and network design in the **preprocessing stage and localization** of the hand pose.

Chapter 3

Proposed Method

The sign of fatigue can be found by the hand pose on the head, neck, shoulders, heart, and chest, in addition, facial expressions of feeling pain or drowsiness. Also, change in the color of the skin like redness of the cheek or the darkness below the eyes. In this section, we introduced a framework for fatigue detection using hand pose. The framework has five layers: Image capture, Pre-processing, Feature extraction, Classification and Models as shown in Fig. 3.1

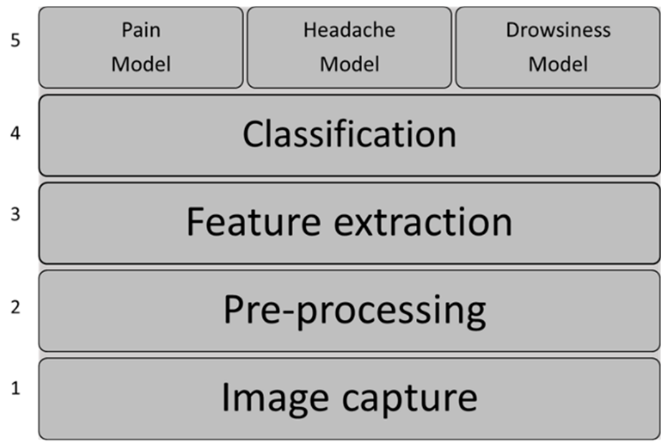
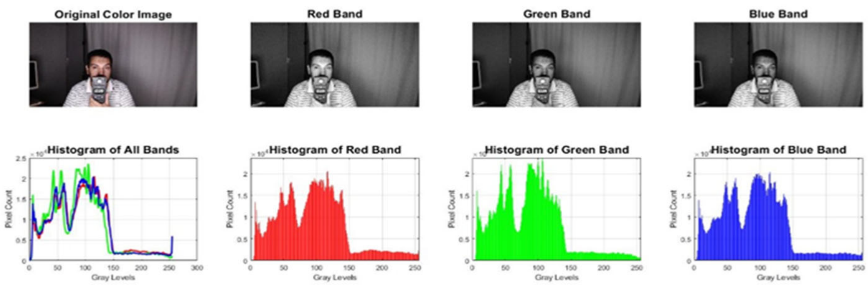


Fig. 3.1 A computer vision framework for fatigue detection

3.1 Image acquisition and representation

The first step in the computer vision pipeline is image capture using image sensors and representation the data after captured using computational imaging to create new image in 2D or 3D i.e. RGB , Depth and HDR images. A hand over face gesture has an occlusion problem caused by similar skin colors and texture of the face and hands. A depth camera used in order to detect the distance of hands and face. Even though it is easy to recognize the gestures. However, **the depth camera cost is high and difficult to use for office computers in working environment**. In this research, front RGB camera was used with indoor light rang as shown in Fig. 3.2 .IR camera with LED introduced to solve the challenges of light changes for driver fatigue detection.

(a)



(b)

Fig. 3.2 Image acquisition (a) Depth camera used for segment hand over face. (b) RGB image used in this research.

3.2 Pre-processing of Image

In general, the image pre-processing have many operations to enhancements the computer vision results like segmentation, color space conversions ...etc. Image pre-processing used in presented framework for enhancement the quality of feature extraction and classification by guided the classifier to search in the closed area including the target poses to reduce the type-I error (false-positive result) for object detection and classification. Image pre-processing stage has five sub steps as following:

1- Skin –color detection.

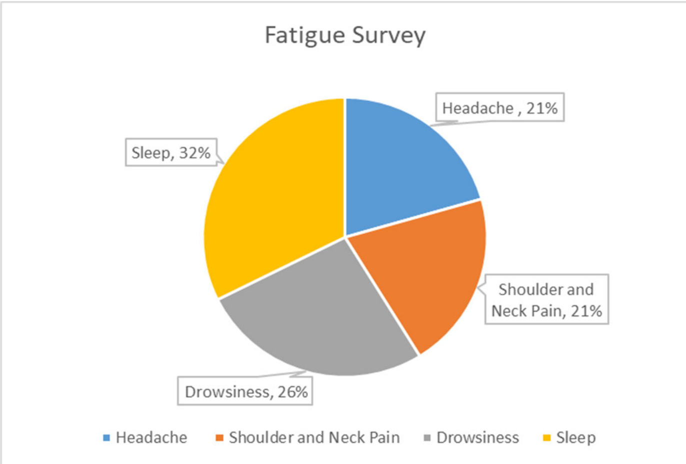
2- Face detection.

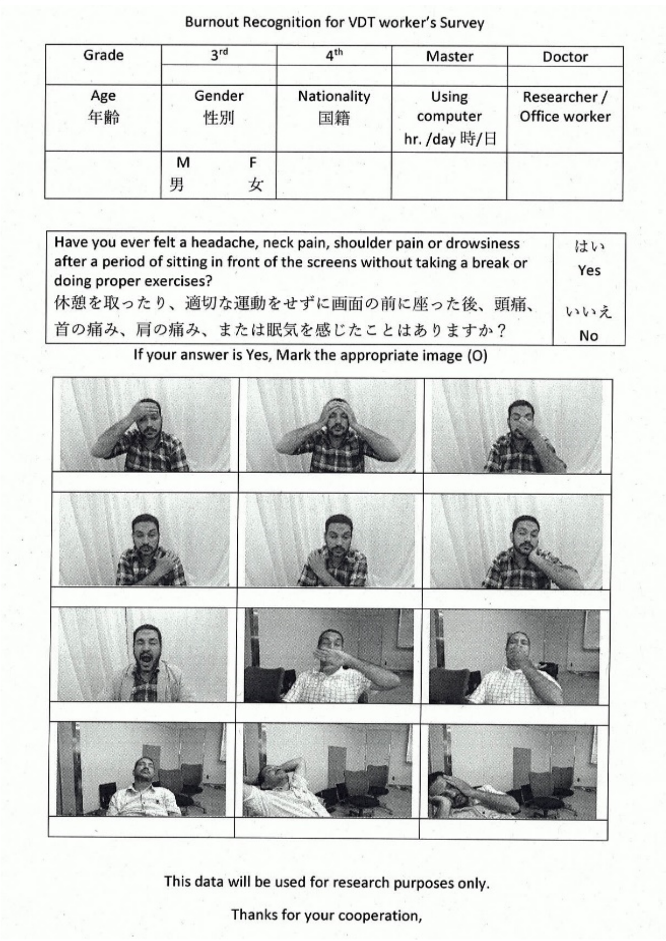
3- Eye detection.

4- Mouth detection.

As a pre-processing step before the image dataset collected, we asked the laboratory students to answer the fatigue questionnaire to confirmed the fatigue sign based on the medical review. The results are shown in Fig. 3.3. Graduated and undergraduate students participated in this survey age from 21 to 43 years old from Japan, Egypt, Mongolia, and Bangladesh. Working in front of screen 2 - 6 hours per day. About

32 % sleep sometimes, 21 % felt headache, 21 % neck and shoulder pain, and 26 % drowsiness.





Burnout Recognition for VDT worker's Survey

Grade	3rd	4th	Master	Doctor
Age 年齢	Gender 性別	Nationality 国籍	Using computer hr. /day 時/日	Researcher / Office worker
	M 男 F 女			

Have you ever felt a headache, neck pain, shoulder pain or drowsiness after a period of sitting in front of the screens without taking a break or doing proper exercises? 休憩を取ったり、適切な運動をせずに画面の前に座った後、頭痛、首の痛み、肩の痛み、または眠気を感じたことはありますか？	はい Yes いいえ No

If your answer is Yes, Mark the appropriate image (O)

This data will be used for research purposes only.

Thanks for your cooperation,

Fig. 3.3. The questionnaire and results of the fatigue survey.

3.3 Feature Extraction

The computer vision research have two approaches, traditional feature descriptors (i.e. SIFT, Surf, and FAST) shown in Table 3.1, and deep learning for feature extraction. Hybrid techniques used in this research, Haar cascaded used for face, eye and mouth detection and deep learning for hand pose detection. Four fundamental families of feature description methods [39]:

1. Local Binary Descriptors (LBP, ORB, FREAK, others)
2. Spectra Descriptors (SIFT, SURF, others)
3. Basis Space Descriptors (FFT, wavelets, others)
4. Polygon Shape Descriptors (blob object area, perimeter, centroid)

Table 3.1. Some characteristics of traditional feature description methods [39]

Method	Feature shape	Feature pattern	Feature density	Search method	Robustness
SIFT [40]	Square, with circular weighting	Square with circular-symmetric weighting	Sparse at local 16x16 DoG interest points	Sliding window over scale space	brightness, contrast, rotation, scale, affine transforms, noise
HAAR [41]	Square, rectangle	Dense	Variable-sized kernels	Grid search typical	illumination
SURF [42]	HAAR rectangles	Dense	Sparse at Hessian interest points	Dense sliding window over scale space	scale, rotation, illumination, noise
HOG [43]	Rectangle or circle	Dense 64x128 typical rectangle	Dense overlapping blocks	Grid over scale space	illumination, viewpoint, scale, noise

3.4 Image Classification

In traditional computer vision, the feature extraction and classification are two stages the deep learning approaches achieved high accuracy in feature extraction and classification tasks. As shown in Fig. 3.4. In this, work a Convolutional Neural Network (CNN) used for feature extraction and image classification.

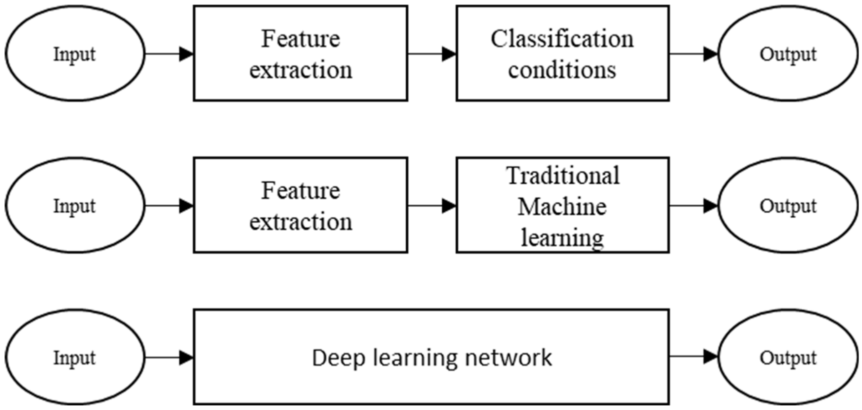


Fig. 3.4. Block diagram for image classification Approaches

AlexNet used for ImageNet classification [45], it is considered a new birth for the use of The Convolutional Neural Networks (CNN) for image classification as powerful artificial intelligence (AI), and since then CNN has proven their ability to solve many challenges in the field of computer vision, especially the feature

Table 3.2. Comprehensive review of CNN models used for image classification using transfer learning

CNN models	Conv. Layers	Top-5 error rate	Total weights
AlexNet [44]	5	15.3 %	61M
VGG -16 [46]	16	7.3 %	138M
GoogleNet [47]	21	6.7 %	7M

extraction and classification. Much CNNs architectures presented to enhancement the feature extraction and classification accuracy, Table 3.2 shows a comprehensive review of CNN models used for image classification using transfer learning

3.5 Prediction Models

The fatigue presented in this framework by three main symptoms as shown in Fig. 3.5: Pain, a headache and drowsiness Models using hand pose as a sign of fatigue. Every model defined by the symptoms characteristics related to the hand pose like the location L, the duration t, and the frequency f. and class c. as following:

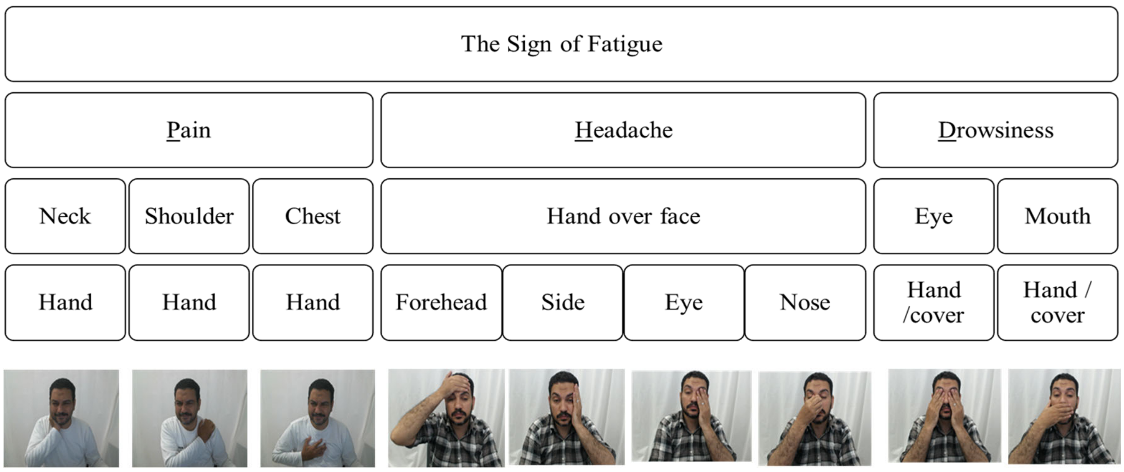


Fig. 3-5. Target hand pose presents the main symptoms of fatigue pain, headache and drowsiness

1. **A pain Model (p)**: The hand Location on neck, shoulder, Face expressions and duration.
2. **A headache Model (H)**: The hand Location over face, duration and frequency.
3. **A drowsiness Model (D)**: Closed Eye (duration) and hand covered Mouth.

Chapter 4
Implementation of Pre-processing

The preprocessing is guiding the classifier to search in the closed area including the target poses. The preprocessing steps are face detection, skin-color detection, and visualizing deep networks.

4.1 Face detection

The work proposed by Viola and Jones was the motivation for all the subsequent Adaboost applications [48]. The algorithm procedure classifies images based on the value of simple features and has mainly four stages: Haar feature selection, creating integral images, Adaboost training algorithm, and cascaded classifiers. The integral image at a location (x, y) contains the sum of the pixels above and to the left of x and y inclusive of

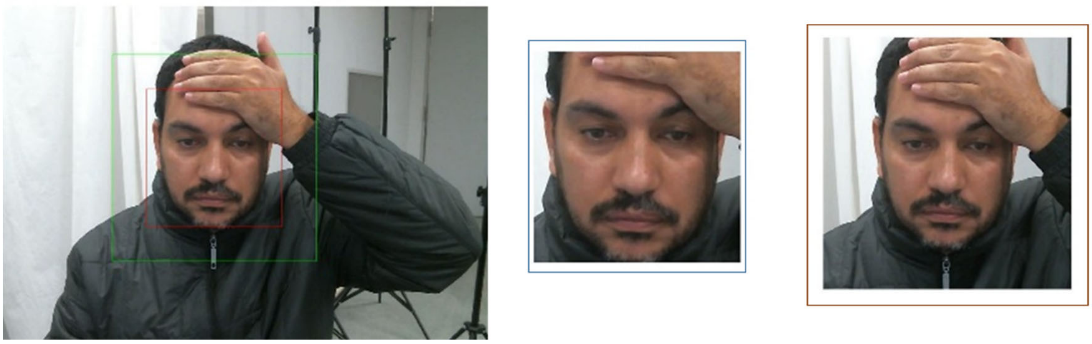
$$ii\,(x, y) = \sum_{x` \leq x, y` \leq y} i(x`, y`). \tag{4.1}$$

The followings are the pair of recurrences:

$$s(x, y) = s(x, y - 1) + i(x, y), \tag{4.2}$$

$$ii(x, y) = ii(x - 1, y) + s(x, y) \tag{4.3}$$

where s(x, y) is the cumulative row sum, s(x, −1) = 0, and ii (−1, y) = 0 is the integral image, which can be computed in a single pass over the original image. This system can detect a face using two features. The first feature is that the region of the eyes is frequently darker than the regions of the nose and cheeks. The second feature is that the eyes are darker than the bridge of the nose, for the face detection, the object detector presented by Paul Viola and Michael Jones was used in our algorithm. Because it is accurate, low computational cost, and performs extremely well in real-time processing compared to others such as SIFT, SURF, and ellipse detection [49]. Face detection is applied to the hand over face gesture image. From the results of face detection, the detected face area is not enough to see the hand over face poses, therefore, the area is increased by 25% around as shown in Fig. 4.1.

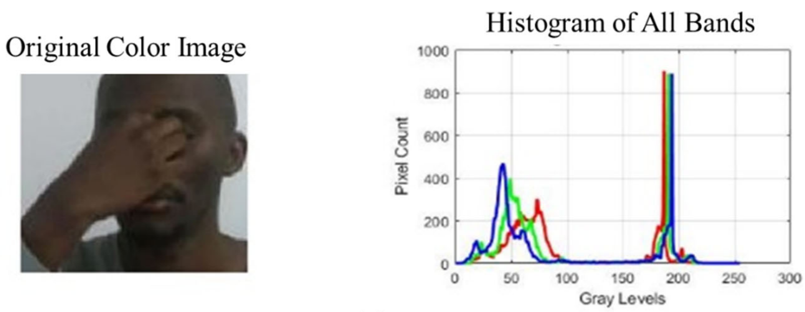


(a) (b) (c)

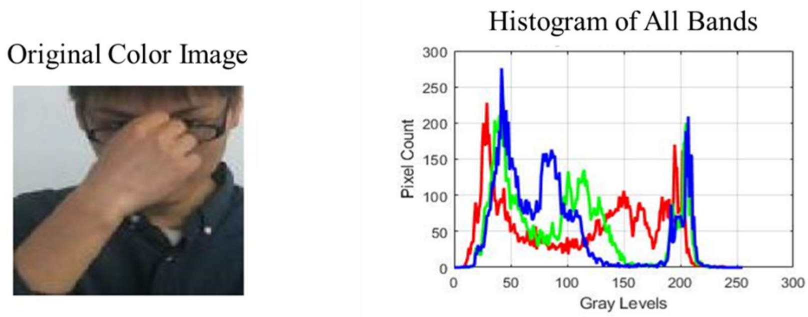
Fig. 4.1. Results of the face detection (a) the original image, (b) detection of the face area, (c) detection of the face area increased by 25% around.

4.2 Skin color detection

The human skin color is a useful feature for detecting a face region. The advantage of using color is the simplicity of skin detection rules because a face indicates different color distribution. The color information is used as the property in order to identify the face from the image. In the literature, there are many color spaces to classify the pixels as the masks. The common color spaces are RGB, YCbCr, and HSV. A skin detection method detects different types of skin under various lighting conditions, such as white, warm-white, inside, or outside with backgrounds. In this research, the dataset is based on the Fitzpatrick skin type (FST) scale [50]. The



(a)



(b)

Fig. 4.2. Histogram of the RGB bands of the candidates form Rwanda and Japan showed the difference of skin color.

participants are from several countries, Mongolia (Type I), Japan (Type II), Egypt (Type III), Bangladesh (Type IV and V), and Rwanda (Type VI) to cover a wide range of the skin tones [51],[52]. Fig. 4.2 shows the histogram of the RGB bands of two different skin type candidates. In our method, the YCbCr color space model is used to transform from the RGB image to the YCbCr color space in Eq.4.4 :

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112 \\ 112 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{4.4}$$

The YCbCr color space is used for skin color detection and localization of the hand. For the skin color detection, the intensity image Y is discarded, then only Cb and Cr are used for the skin color features as the following range:

$$104 \le Cb \le 135 \text{ and } 122 \le Cr \le 132. \tag{4.5}$$

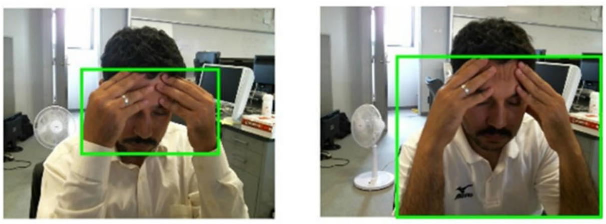
For the localization of hand, the range of Y, Cb, and Cr after applying the CNN-CAM method is

$$18 \le Y \le 91,\ 96 \le Cb \le 116 \text{ and } 130 \le Cr \le 209. \tag{4.6}$$

These parameters are determined by our experiments. Fig. 4.3 shows the results of the skin color detection of the different hand poses. The skin color detection tends to be sensitive in the clothes type, color, and illumination condition. The general system fails to identify the face detection because of the covering of the facial features such as eyes or nose, therefore, it cannot further classify the position of hands on face. In contrast, our proposed method detects the skin-color with the face recognition by performing before the classification process.

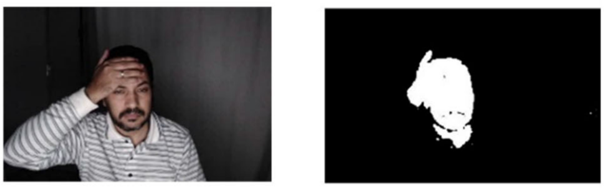
4.3 Visualizing deep networks

The main purpose of improving headache detection should be more robust. Therefore, Class Activation Map (CAM) are implemented to highlight the target region of the image. The specific target region is the most important for the prediction of poses, and the decision by the deep learning network. Thus, for the



(a)

(b)



(c)

Fig. 4.3. Results of the skin color detection sensitive in (a) Clothes type, shirt and T-shirt (b) Clothes color, and (c) Illumination condition (50 lux)

results based on the CAM, the training and design are rearranged in the preprocessing. The implementation is a pipeline of the generated CAM described as the following steps.

1. modified a CNN architecture,
2. Retrain the network,
3. Extract the features,
4. Predict the class labels, and finally generating the class activation map.

The CAM is a specific class by the activation of unit k in the last convolutional layer at location (x, y):

$$F_k = \sum_{x,y} f_k(x, y). \tag{4.7}$$

The input of softmax for class c is

$$S_c = \sum_k w_k^c \; F_k. \tag{4.8}$$

The output of softmax for class c is

$$P_c = \frac{\exp(S_c)}{\sum_c \exp(S_c)}. \tag{4.9}$$

The class activation map for class c is

$$M_c(x,y) = \sum_k w_k^c \; f_k(x,y). \tag{4.10}$$

The CAM used for three tasks: the first one is to visualize the important area guided the prediction decision, it found the conflict caused by the error of prediction and modified the data and trained the network again until the high classification accuracy without overfitting. The second task is to avoid under fitting by designing the deep network based on the visual explanation of prediction on every convolutional layer. In the third task, CAM used for localization of the hand over face poses. The proposed method was designed to reduce the type-I error (false-positive result) for the classification. Fig. 4.4 shows Visualize the important area guided the prediction decision

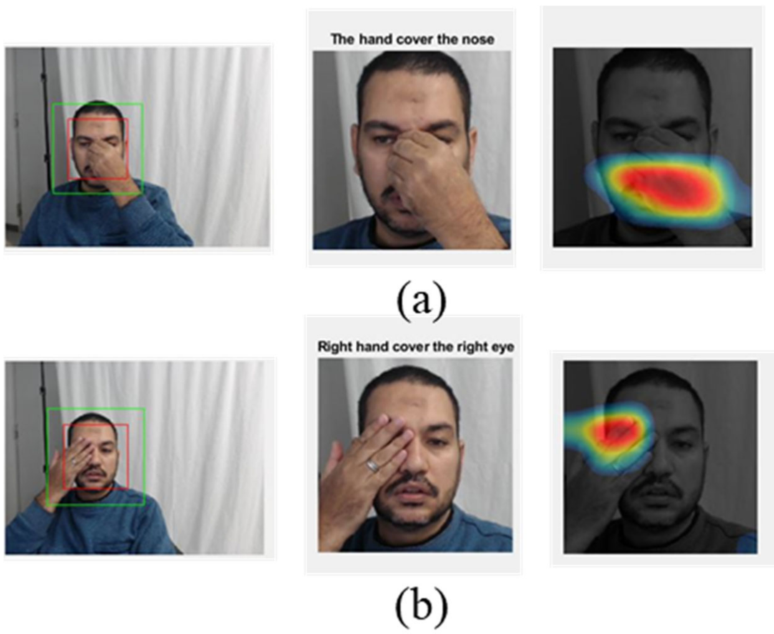


Fig. 4.4. Visualize the important area guided the prediction decision (a) The hand cover the nose, and (b) Right hand cover the right eye.

4.4 Localization of the hand over face.

The color maps commonly used for network visualization. In this work, a jet color map used to highlight the area guided to the classifier prediction decision. By applying the YCbCr color space, parameters were determined by our experiments.

In this work, the prediction model used the classification result label to define the hand pose location and used the localization method presented to confirm the hand detection result. This result can be extended for object detection in the scene without an annotated labeled training dataset as shown in Fig. 4.5.

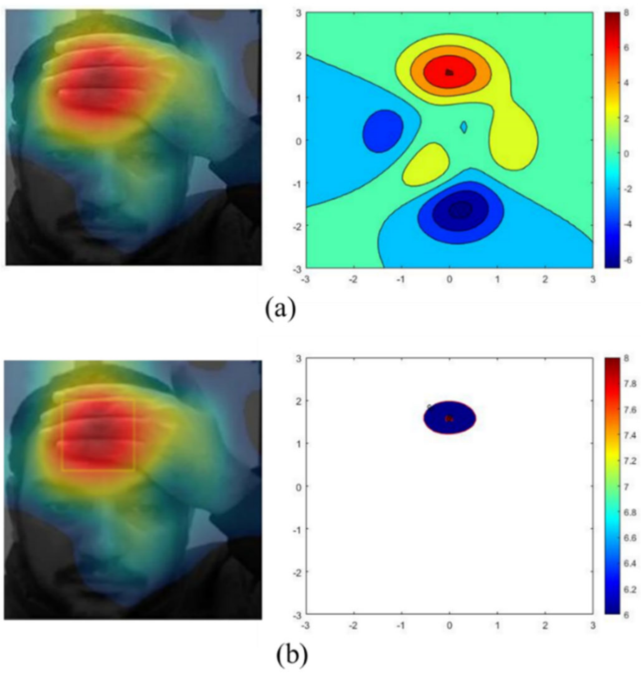


Fig. 4.5. Experimental results of the localization of the hand pose using the CNN-CAM with YCbCr color space. (a) Jetcolor map distribution, (b) Segment the highest peak of Jetcolor map.

Chapter 5

Experiments Based on Object and Skin-color Detection

The proposed method based on the objects detection and skin color detection algorithms. The Viola-Jones object detector used to extract the facial features such as the regions of face. The color space conversion applied to the area of extracted regions for skin color detection.

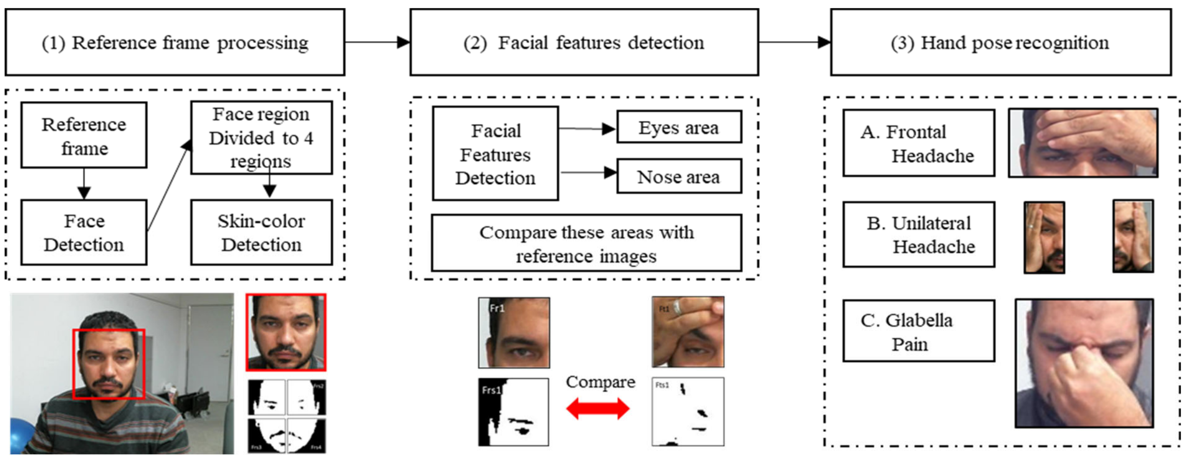


Fig. 5.1. Block diagram of visual fatigue recognition system using facial features and hand poses for VDT workers.

5.1 System block diagram

The system has three main stages as shown in Fig. 5.1. (1) Reference frame processing, (2) Facial features detection, and (3) Headache pose recognition, in order to analyze the VDT worker's behavior and recognize the pose of fatigue. We indicate these three main stages in detail as the following section.

5.2 Reference frame processing

First, we obtain a reference frame from the USB camera that is used for capturing a standard image when a worker exhibits no headache. The reference frame is to detect the difference in the behavioral change of a worker. In this research, we focus on the detection of facial features and hand poses of a worker. This first stage includes three sub-steps: reference frame acquisition, face detection, and skin-color detection. To obtain the reference frame, the user works in front of the display attached to the USB camera. We define the frame under a no headache condition as the reference frame (Fr). To detect the face area, we apply Adaboost cascade of the simple feature method. Then the face area is divided into four regions: upper left (Fr1), upper right (Fr2), lower left (Fr3), and lower right (Fr4), as shown in Fig. 5.2 (a).Furthermore, for skin-color detection, the YCbCr color space conversion is applied to the four regions. Finally, the reference image is divided into four binary regions: Frs1, Frs2, Frs3, and Frs4, as shown in Fig. 5.2 (b). In each region, the skin area can be detected to calculate the summation of the white pixels.

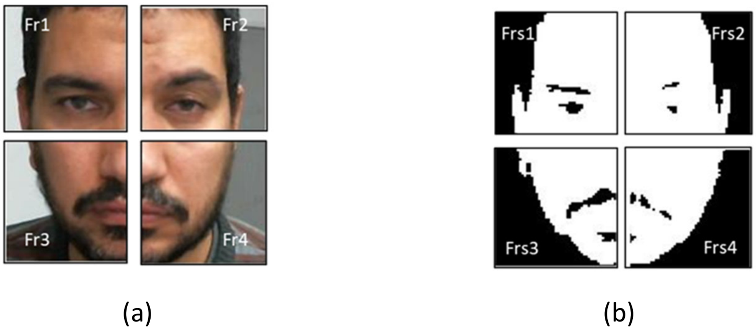


Fig. 5.2 The reference frame (a) The face detection and (b) The Skin-Color detection.

5.3 Facial features detection

The facial feature detection stage has the following sub-steps:

1. Get a new frame and detect the face.
2. Divided the face area into four regions.
3. Apply the skin-color detection as same as in the reference frame processing described in Sec 5.2. The skin color area is calculated new values (Fts1, Fts2, Fts3, and Fts4).
4. Extract the facial features around the eyes and nose, which compared with the skin color area of the reference image.

5.4 Hand pose recognition

In this stage, the system evaluates several types of visual fatigue signs based on a hand pose. The recognition of signs is frontal, unilateral, and vision pain locations as the following.

A. Frontal Headache Pose The hand pose of frontal headache is right or left hand or both hands put on the forehead. To recognize these poses, the system compares summation of white pixels in the skin area detected on upper left and right face region (Fts1, Fts2). It is processed at the facial features detection presented on a section tow and also compared with the corresponding of the reference frame (Frs1, Frs2). Fig. 5-3 (c) and (d) show the results of detecting the frontal headache pose.

B. Unilateral Headache Pose One of the unilateral headache pose is the hand placed on beside the eye of top right or left cheek. On the left side, the system is compared the values of upper and lower left parts (Fts1, Fts3) with the corresponding

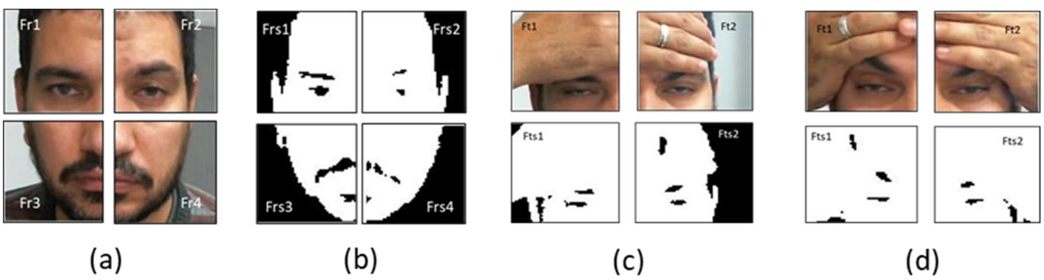


Fig. 5.3 The reference frame and skin-color detection. (a) Face regions after performing face detection. (b) Four binary regions after performing skin-color detection. The results of detecting the frontal headache pose, (c) one hand pose, and (d) both hands pose.

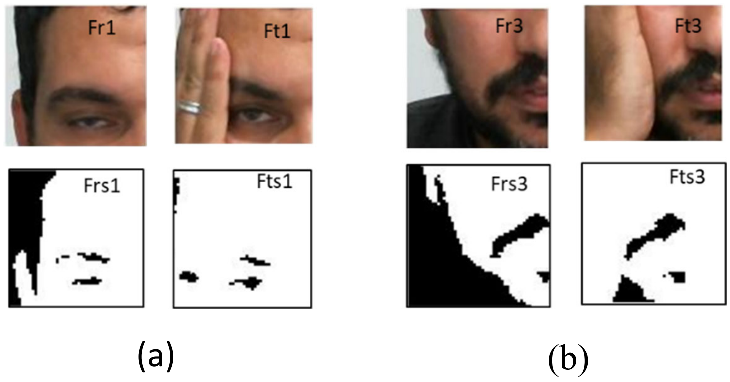


Fig. 5.4. The results of detecting the unilateral headache pose. (a) Left upper side and (b) Left lower side.

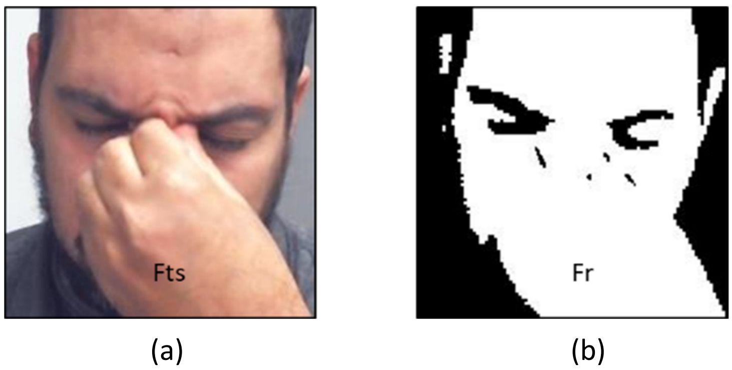


Fig. 5.5 The result of detecting the vision pain hand pose. (a) Color image and (b) Binary image after applying the Skin-color detection.

of the reference frame (Frs1, Frs3). Fig. 5-4 (a) and (b) show the results of detecting the unilateral headache pose.

C. Vision Pain Pose To recognize the vision pain, the hand put on the vision. The system evaluates two conditions to detect the vision pain through the hand pose sign. The first condition is that the detection of nose field. The second condition is that the Skin-Color area detected on the lower left and right face parts in this frame, it is greater than the corresponding ones on the reference frame. Fig. 5-5 shows the result of detecting the vision pain.

5.5 Experimental Results

We have evaluated the visual fatigue recognition system in laboratory environment. We used four dataset simulated visual fatigue conditions related to frontal headache, unilateral headache, and vision pain pose. Table5.1 shows the experimental results of each condition.

In the table, Total indicates total number of frame including each pose in the dataset. The system has achieved 89.0 % for recognize the frontal headache one hand pose, 87.5 % for both hands, 93.0 % for the unilateral headache pose, and 95.0 % for vision pain pose.

Table 5.1. Experimental results of fatigue recognition system.

Dataset	Total	Success	Failed	Results
Frontal headache pose (one hand)	137	122	15	89.0 %
Frontal headache Pose (both hands)	144	126	18	87.5 %
Unilateral headache pose	159	148	11	93.0 %
vision pain pose	162	154	8	95.0 %

The limitations of this method: Change of illumination level affected in skin color detection. In addition, Head orientation affected in feature extraction. The Experimental presented in chapter 6 solved these challenges by using deep learning.

Chapter 6

Experiments Based on Convolutional Neural Network (CNN) with Class Activation Mapping (CAM)

In this research, a hand over face gestures using a deep learning based on a Convolutional Neural Network with Class Activation Mapping (CNN-CAM) are classified. Six typical hand gestures, which signs consist of the hands-on the forehead, hands cover the nose, hands cover the eye, and hands-on the side of the head defined. The proposed method described in detail in the following sections. Fig. 6-1 shows the flowchart of our proposed method.

6.1 Classification of the Hand Poses

A deep learning system aimed to classify the typical six hand over face gestures as the signs of a headache of the VDT worker presented. These consist of the hand on the forehead, eyes, and nose previously used for the classification of these poses.

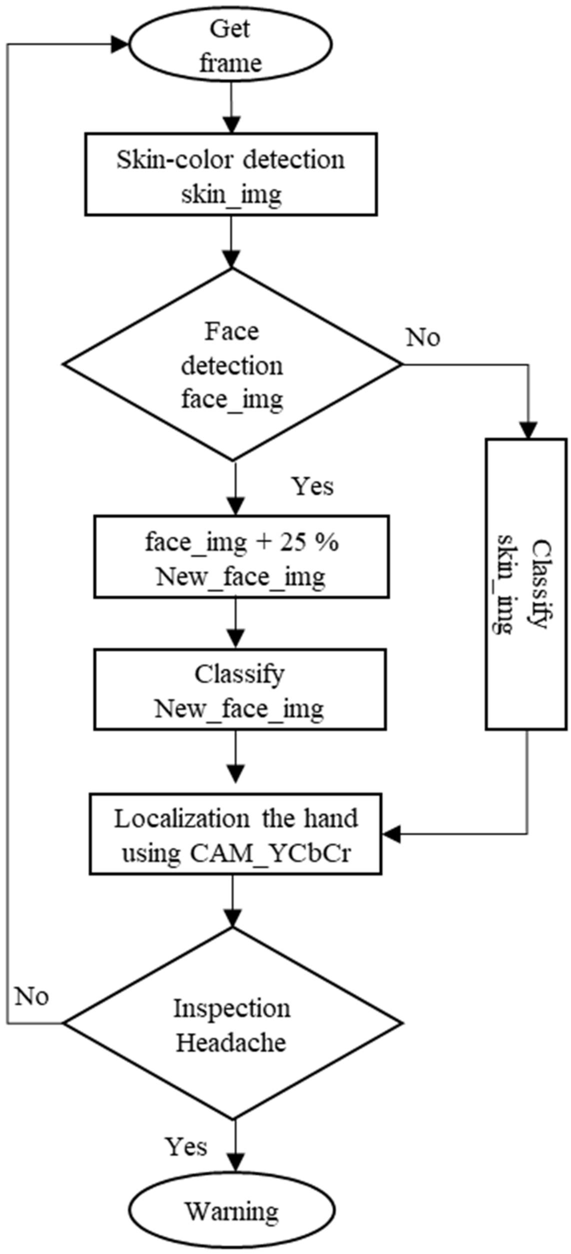


Fig. 6.1. Flowchart of the proposed method.

6.2 The hand over face dataset

In this section, the data set used in our experiments is described. The images selected from the video recorded a laboratory room that simulated the VDT worker's environment. The Logitech HD C720 and C615 camera were used for recording the VDT working. The camera stand is placed on the top of the computer screen in front of the participants. The image size is 320×240, 640×360, and 1280×720, right, and left faces to identify the signs. To cover the wide range of skin color tones, the participants are from different countries, one is from Egypt, Mongolia, and Rwanda, two is from Bangladesh, and three is from Japan. The illumination level in the room is changed from 50 lx to 550 lx to evaluate the different light conditions of the office environment. The background of scene is simply and complex. The participants wear casual clothes (shirt and T-shirt) and formal clothes; this condition affects skin color detection in general. Each recorded video contains six different hands over face poses as the following.

1. The right hand on the forehead,
2. The left hand on the forehead,
3. The hand(s) on side of the head,
4. The right hand cover the right eye,

5. The left hand cover the left eye,

6. The hand covers the nose.

The dataset includes these six poses for the hand over face gesture. The images were selected after applying the skin color detection and resized every skin color image to 112 × 112. For each hand pose image, the dataset includes 170 images divided into training data 119 images and validation data 51 images. The dataset includes 1360 images (six hand poses, a normal face, and background). The samples of images shown in Fig. 6.2.

Fig. 6.2. Samples of training and validation dataset

6.3 Training network

The modified GoogLeNet shown in Fig. 6.3. was used for the classification of the hand over face poses. The proposed method was implemented using the MATLAB deep learning toolbox. Table 6-1 shows the comparison of the original pertained network GoogLeNet layers with our modified network.

The features of our method are to remove the fully connected, softmax, and classification output layers and add the same three layers with new parameters to reduce the total parameter from 1 025 000 to 8200. Before the softmax layer, added the convolutional layer with Filter size 3 × 3, Num Filters 32, Stride 1 and Padding 1. New convolutional layer Connected with Relu layer and global average pooling layer for CAM operation.

Table 6-1. Comparison of the original and modified network layers

New layers	Original GoogLeNet layers	**Modified layers in our method**
Convolutional layer Filter size: 3×3 Num Filters : 32 Stride 1 Padding 1	Fully connected: 1×1×1000 Weights: 1000×1024 Bias: 1000×1 Total learnable: 1025000	**Fully connected: 1×1×8** **Weights: 8×1024** **Bias: 8×1** **Total learnable: 8200**
Relu layer	Softmax 1×1×1000	**Softmax 1×1×8**
Global Average Pooling layer	Classification output 1000 classes	**Classification output 8 classes**

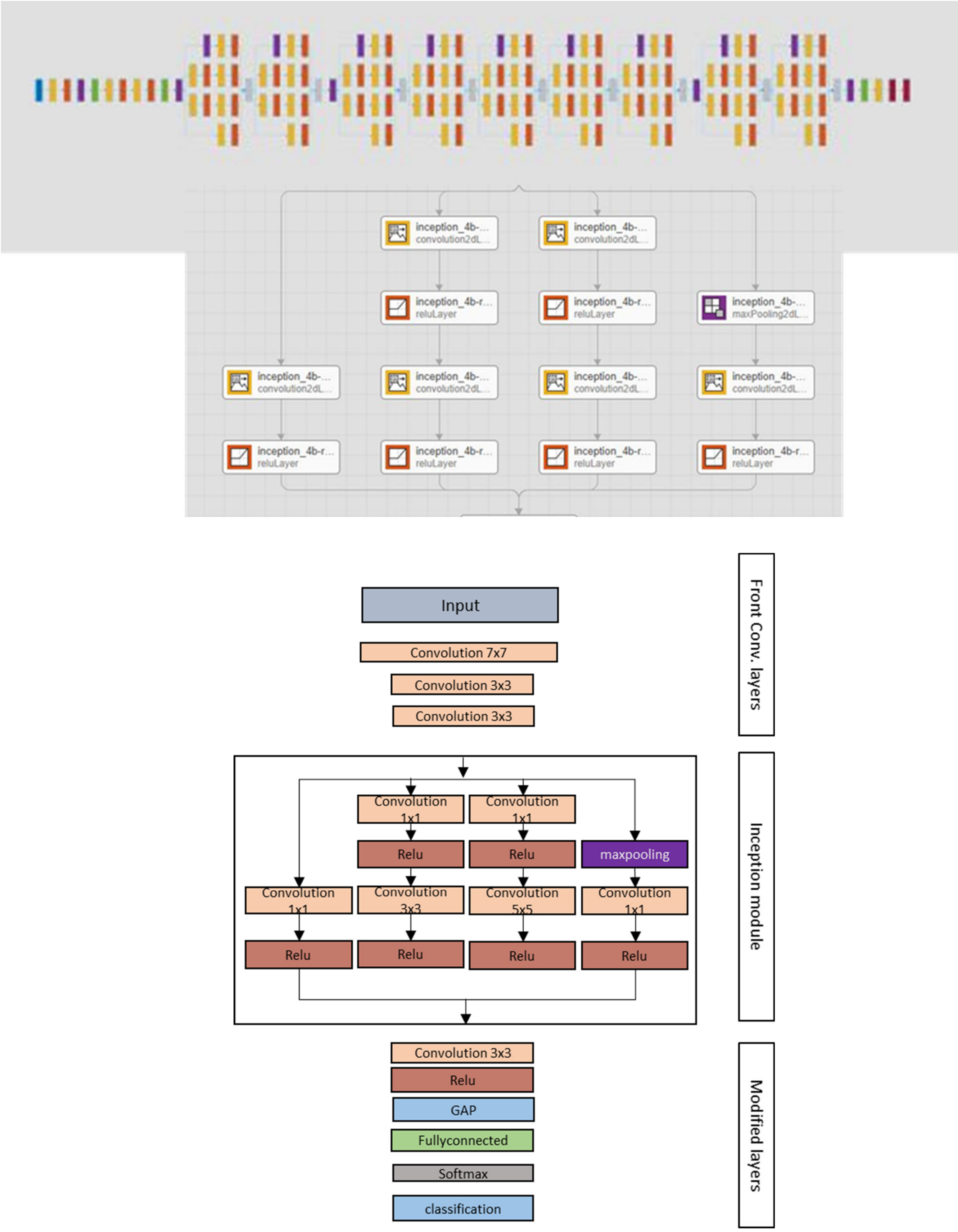


Fig. 6.3. The GoogLeNet (a) Original network and (b) Modified network

6.4 Training progress

In the training process, the parameters of the deep learning network are the following conditions:

1. Input image size 112 × 112, 144 × 1 layers,
2. 170 × 2 connections,
3. Epoch 6,
4. Learning rate 0.0001.
5. The output layer is customized to classify eight classes; which include six hand poses, normal face, and background.

Fig. 6.4 shows the training progress of the dataset using single GPU with our modified GoogLeNet

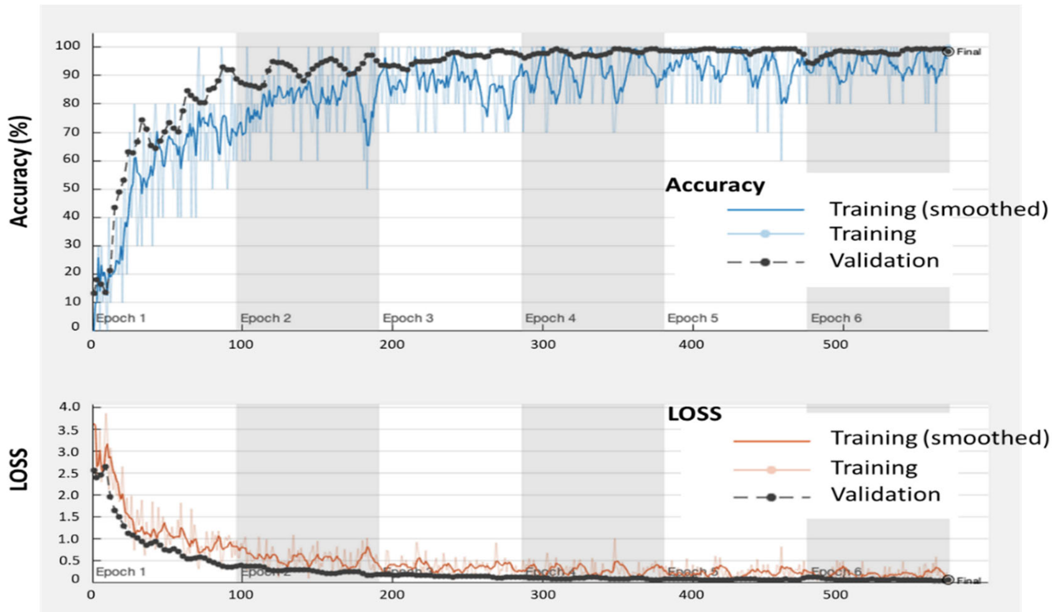


Fig. 6.4. Training progress of the dataset using single GPU with our modified GoogLeNet

The confusion matrix for classification of the evaluation dataset per class as shown in Fig. 6.5.

1. Background: 100 %
2. Face: 100 %
3. Left hand cover the left eye : 96.6%
4. Left hand on the forehead : 92.7%
5. Right hand cover the right eye : 100 %
6. Right hand on the forehead : 100%
7. Right hand cover the nose : 100 %
8. Hand on the side of the head : 100 %

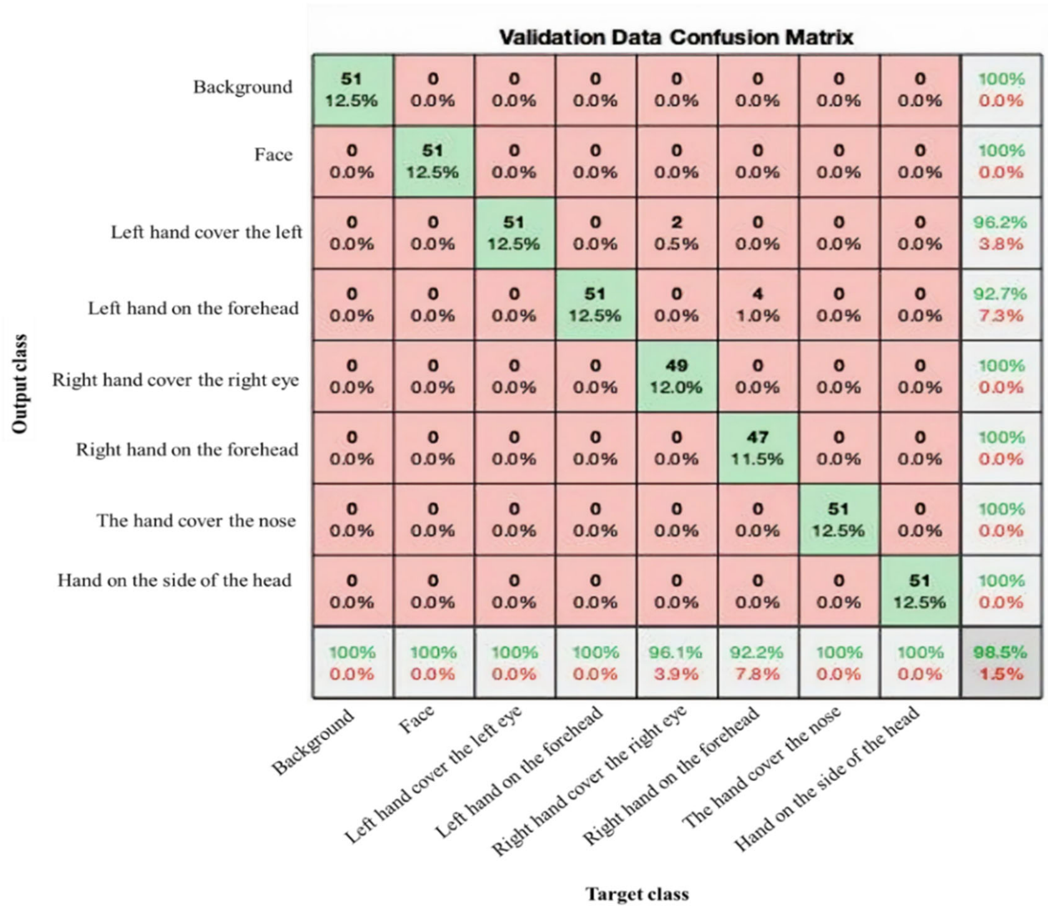


Fig. 6.5 The confusion matrix of the headache model for evaluation dataset.

Chapter 7

Prediction models and results

In this chapter, I will introduce the prediction models for headache, pain and drowsiness, and discussing the results in detail.

7.1 Prediction of the Headaches

There are specific poses when feeling headaches according to TTH on the bilateral location of the head as shown in Fig. 7.1. Migraine appears on the unilateral location of the head. The international classification of headaches diagnostic criteria section for every type of headache determines the duration of the feeling of headache from the international classification, for example, less than 5 min for some headache types

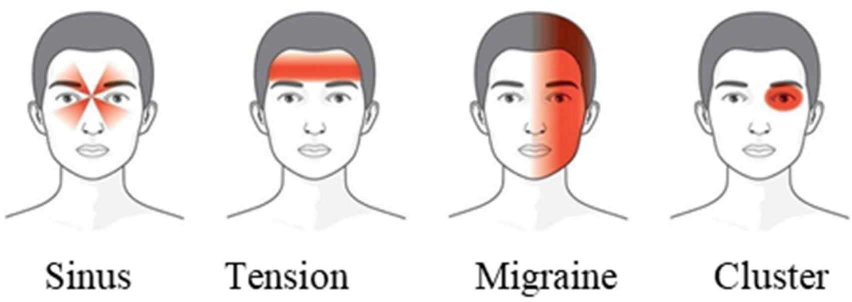


Fig. 7.1. Main four types of headache.

to several hours, once a day to several times a day repeated for days or months. In this section, the characteristics of the headache based on the hand poses as the signs of headache defined our headache prediction model and presented a method for localization of the hand over face pose evaluated.

7.1.1 A Headache prediction model

The headache prediction model defined from three main characteristics of the headaches as the following. The location L, the duration of headache pain t, and the frequency of attack f.

$$H = \{\mathrm{L(x,y)}, \mathrm{t(s)}, \mathrm{f(m)}\} \tag{7-1}$$

Where H is the headache detection at the location L(x,y), x,y is the location coordinates, f (m) is frequency, m is the number of times and t(s) is the time duration, s is second. The location of the hand over face L defined by segment the largest region in the class activation map M c (x, y). t is the difference between the time of the first frame, it has a hand pose of the headaches and the last one. The experiments performed involved three main stages to classify the hand over face gestures as the signs of a headache. The first stage is the preprocessing step, it included the face and skin color detection. In the second stage, the hand over face poses are classified using the pertained convolutional neural network. In the final stage, the signs of headache were predicted. In our proposed method, a combination of the face and skin color detection becomes more robust for illumination change from 50 lx to 500 lx and the head orientation (45◦ left or right) as shown in Fig. 7.2.

7.2 A Headache prediction results

The headache poses were recognized as the following cases:

a. The hand(s) on the forehead.

b. The hand covers the nose.

c. The hand(s) cover the eye.

(a)

(b)

Fig.7.2. Experimental results of the hand over face pose (a) with illumination change to 50 lx and (b) head orientation to 45°.

d. The hand(s) on the side of the head

The proposed method warned to the workers if the position of hands has repeatedly categorized as a sign of headache during a certain period. For example, the frontal headache characteristics is the hand on the forehead (L(x, y),) 20 s (t(s)) frequented five times (f (m)) in the headache prediction model. Fig. 7.3 shows the hand on forehead, nose, right eye, and the hand on the left side of the head as the examples of a sign of the Tension (forehead), Sinus headaches (the area between eyebrows) Cluster, (behind the eye) and Migraine headaches (unilateral side), respectively. To

(a)

(b)

(c)

(d)

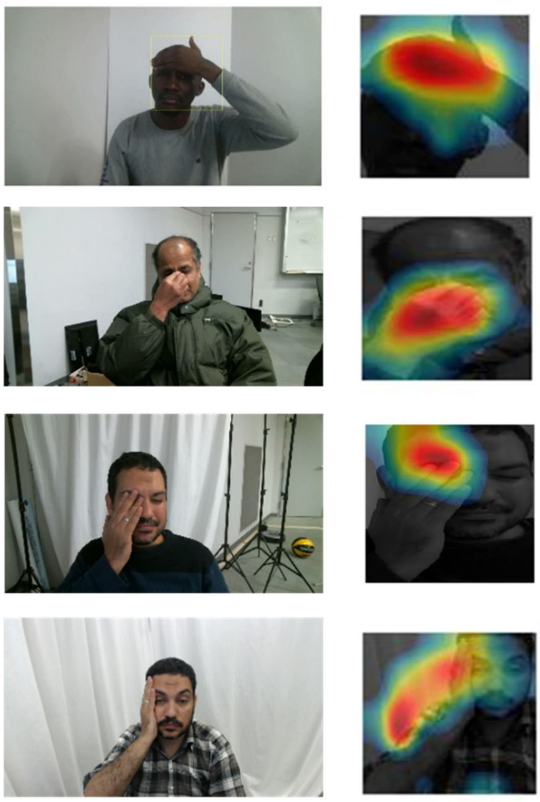


Fig. 7.3. Experimental results of the headache prediction with the hand poses applied our proposed method. (a) on forehead, (b) cover nose, (c) cover right eye, and (d) on the left side of face.

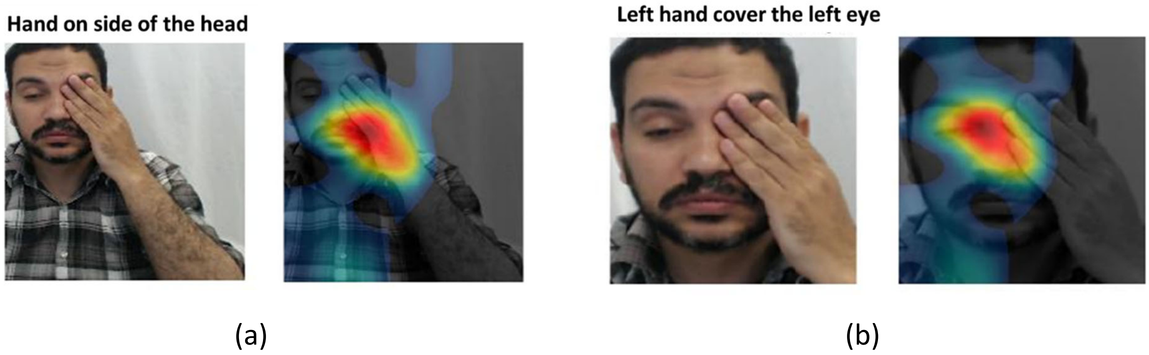


Fig. 7.4. Improvement of the prediction using CAM. (a) The hand on side of head (False-positive result), (b) The correct result of the hand cover left eye (True-positive result).

avoid the classification overfitting and data biasing errors, the proposed system used the CAM to visualize the prediction of classification network, and then found the confusing area to redefine the classifier input area and rearrange the data into optimum classes. The proposed method was designed to reduce the type-I error (false-positive result) for the classification. The preprocessing stage guided the classifier to search the area of target pose as shown in Fig. 7.4 (a). The classification result is the hand on side of the head (false-positive result for hand pose). Fig. 7.4 (b) shows the corrected result of the hand cover left eye (true-positive result).

7.3 Prediction of Pain and Drowsiness

In this section, modeling the characteristics of drowsiness and pain as fatigue symptoms using the hand pose, eye and mouth status based on poses in Fig. 7.5.

- The hand on the shoulder or neck, as a sign of neck pain and upper pain.

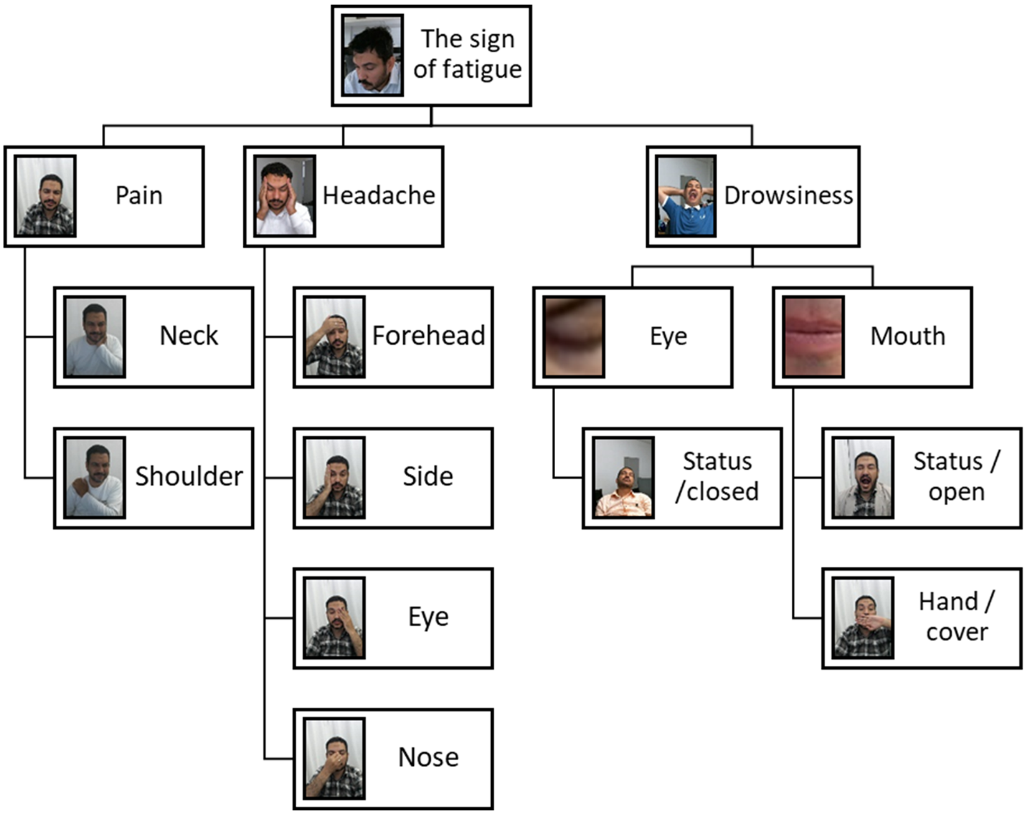


Fig. 7.5. Headache, pain and drowsiness as fatigue symptoms using the hand pose, eye and mouth status

- The hand cover the mouth as a sign of drowsiness,

Prediction of the sign of pain and drowsiness using the hand pose proposed in this work can expand the applications of this research to observing the status of fatigue in different environments, for example, call center agents, control room observers, and air traffic controllers.

7.4 Pain and drowsiness prediction models

Drowsiness detection is one of the interesting research for computer vision regarding driver's fatigue. Previous studies of drowsiness detection have mainly focused on tracking of the eyes and mouth conditions like eyes blinking, yawing, and so on. In following sections our prediction model and proposed method discussed in details.

7.4.1 Prediction models

- The Pain prediction model has two parameters; location of the hand on the neck or shoulder (L) and the time duration t(s) as shown in eq. 7.2

$$P = \{\mathrm{L(x, y), t(s)}\} \tag{7.2}$$

- The drowsiness prediction model recognize the d in two cases; the hand coved the mouth $\{\mathrm{L(x, y), t(s)}\}$ or in case of the eyes closed E_c and the mouth open $M_o$at same time as shown in eq. 7.3

$$D = \{\mathrm{L(x, y), t(s)}\} + \{E_c \times M_o\} \tag{7.3}$$

- The proposed system integrates traditional machine learning and deep learning algorithms to increase the performance of accurate detection using four stages as shown in Fig. 7.6:

    1. Detection of eyes and mouth in the scene.

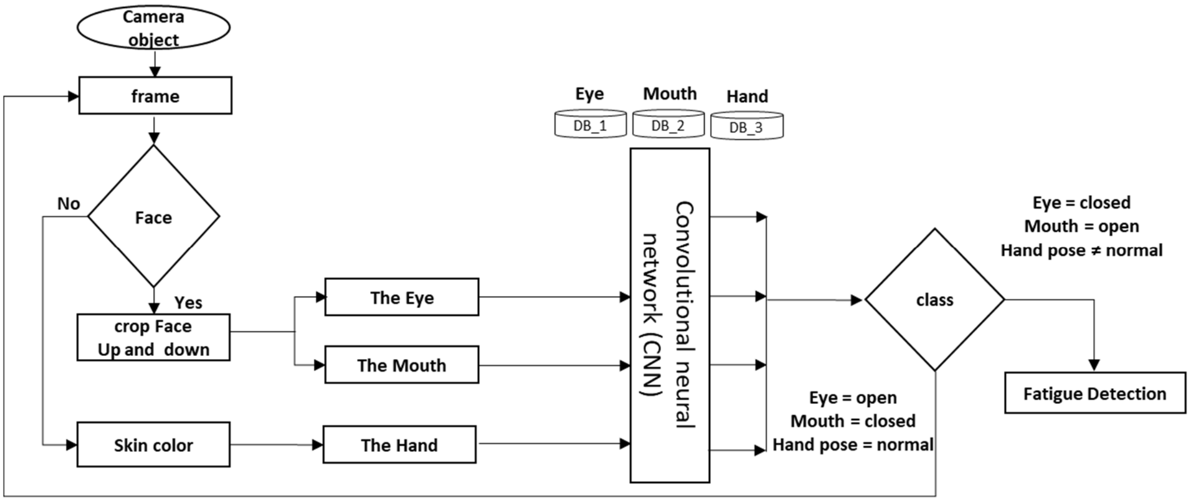


Fig. 7.6. Flowchart of the proposed method for drowsiness and pain detection.

Table 7.1. The datasets used for training the model

Dataset	Training	Validation	Total	Accuracy
Opened eye	210	90	300	97.6 %
Closed eye	210	90	300	
Opened mouth	105	45	150	96.7 %
Closed mouth	105	45	150	
Hand pose	750	525	1275	99.5 %

2. Skin color detection to extract the face and hands area.
3. Classification using deep learning to recognize the status.
4. Detect the signs of drowsiness and pain automatically.

7.4.2 The datasets

The images used in drowsiness and pain detection model selected from recorded videos; for eyes and mouth opened, closed, and hand pose in mouth, neck and shoulder as shown in Fig. 7.7. 70% of dataset images used for training and 30% used for validation. The environment limitations as following:

1. The illumination level 50 lx to 550 lx,
2. The background of scene is simply and complex,
3. The head orientation (~45◦ left or right).

- Eye dataset (DB_1): two classes, open eye and closed eye. Training progress shown in Fig. 7.8
- Mouth dataset (DB_2): two classes, open mouth and closed mouth Training progress shown in Fig. 7.9
- Hand pose dataset (DB_3): four classes, the hand cached neck left and right side, the hand on the shoulder left and right side. Training progress shown in Fig. 7.10.

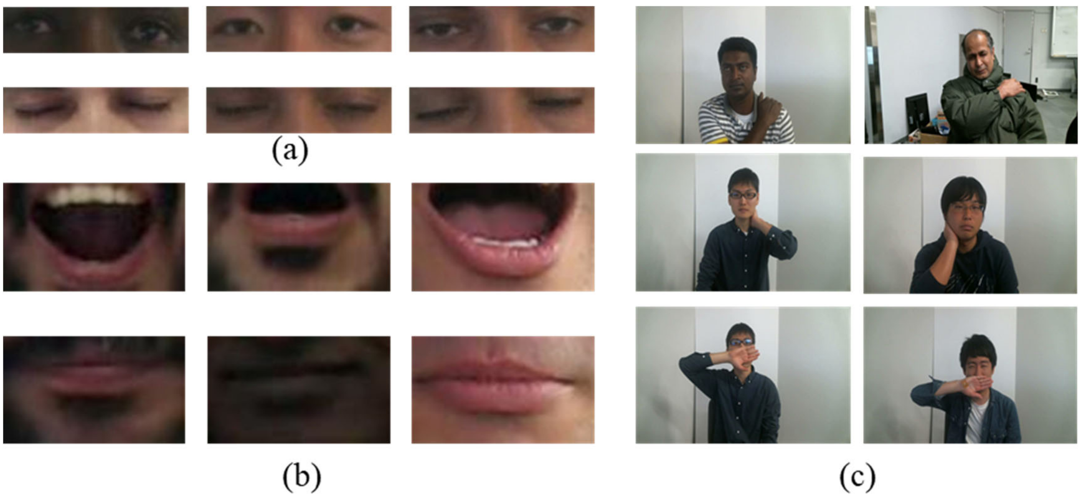


Fig. 7.7. Samples of datasets; (a) open and closed eye (b) open and close mouth and (c) hand pose with shoulder, neck and mouth.

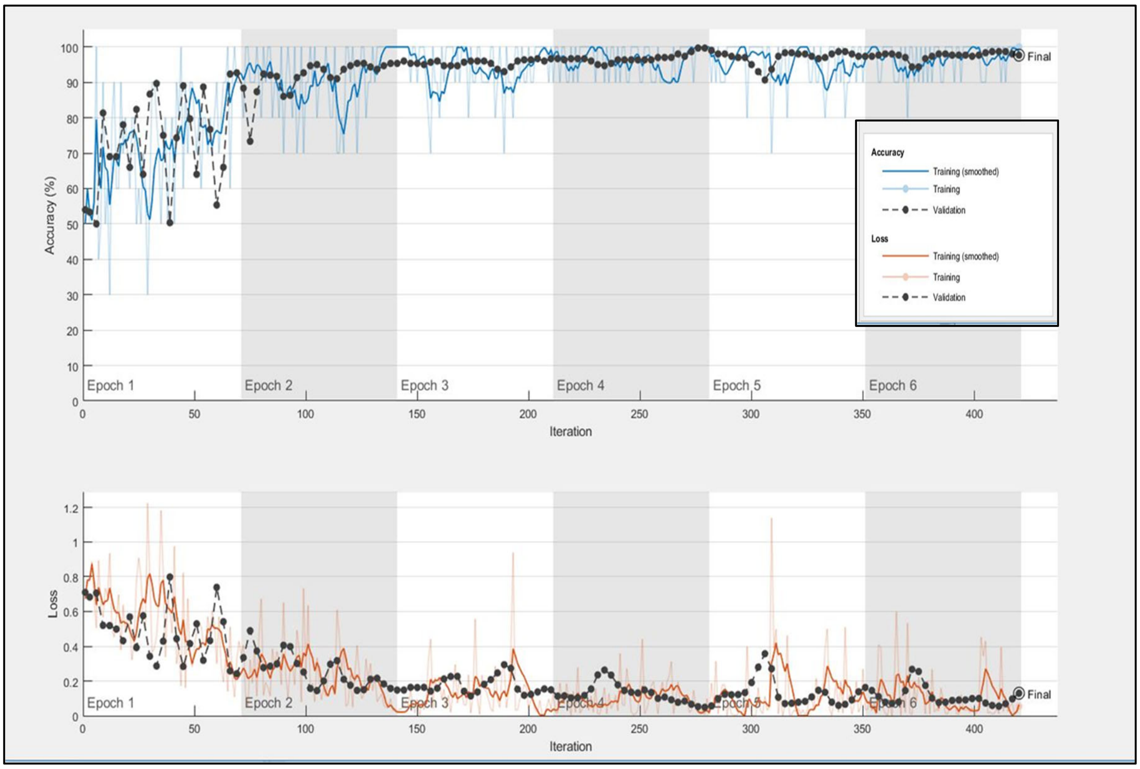


Fig.7.8 Training progress of the eye dataset achieved 97.6 %.

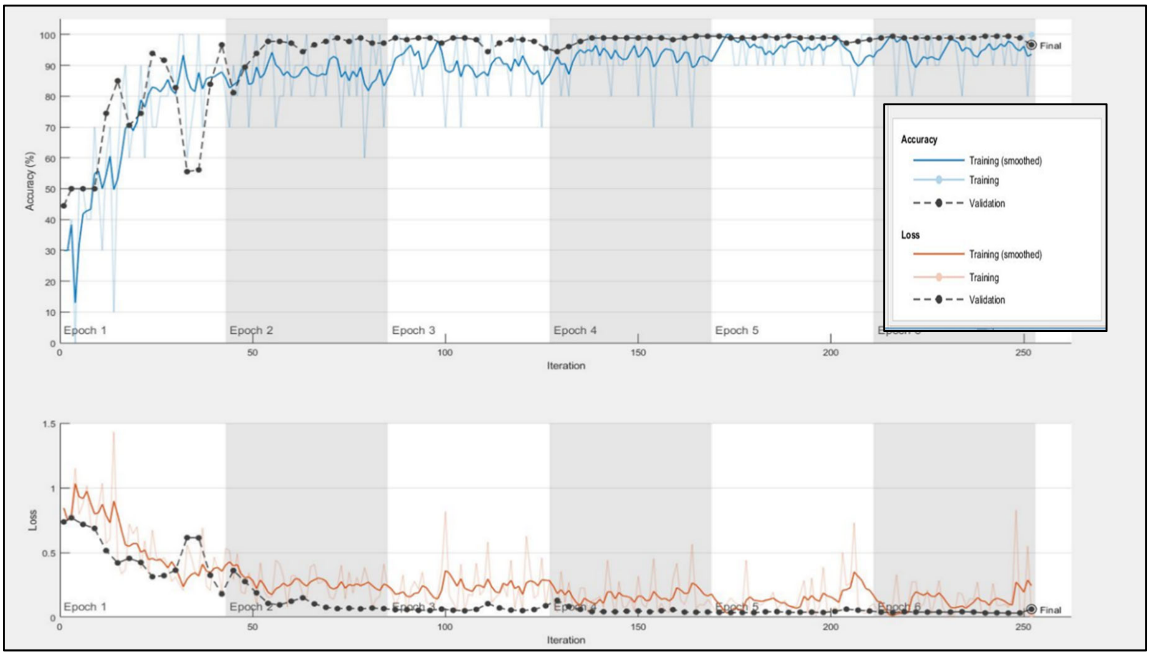


Fig.7.9 Training progress of the mouth dataset achieved 96.7 %

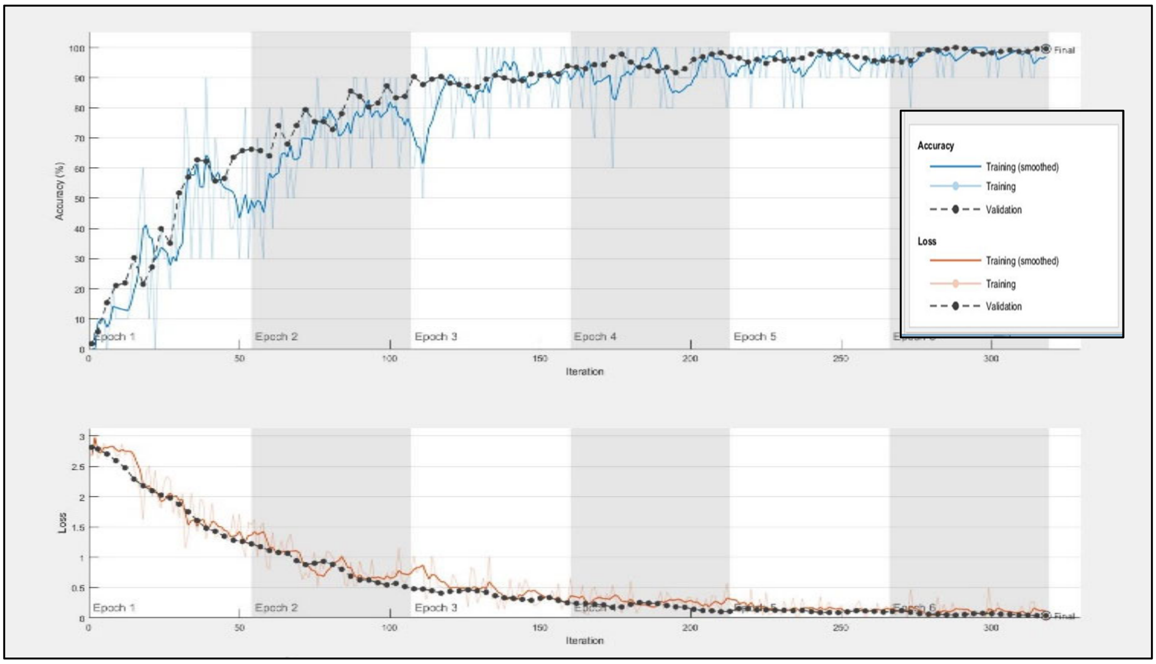


Fig.7.10 Training progress of the hand pose dataset achieved 99.5%.

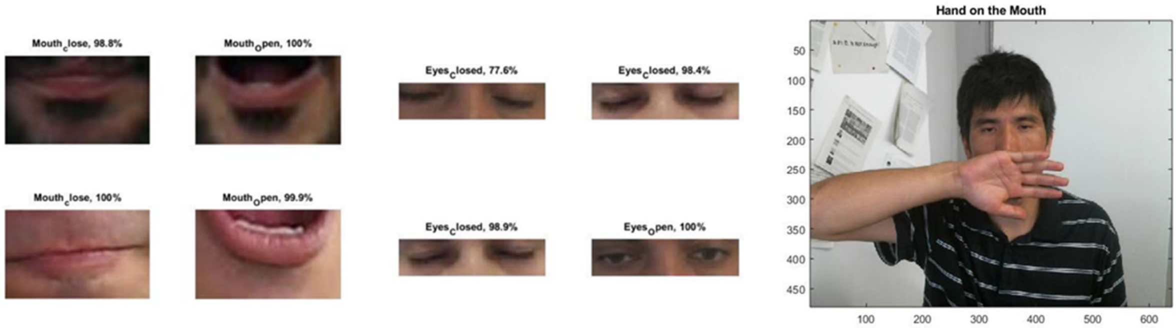
7.5 Drowsiness and pain prediction results

The experiments performed involved three main stages drowsiness and pain. The first stage is the preprocessing step; it included the face, eyes, and mouth and skin color detection. In the second stage, the hand poses classified using the pertained convolutional neural network. In the final stage, drowsiness and pain predicted based on the prediction models in Sec. 7.2 and Sec. 7.3.

7.5.1 Drowsiness prediction results

The system achieved the detection of the drowsiness in the following cases:

1- The eye classified closed and mouth open at same time.



(a) (b) (c)

Fig 7.11. The classification results; (a) open and closed mouth (b) open and closed eyes and (c) the hand covered the mouth.

2- The hands cover the mouth.

The classification results for the mouth, eyes and the hand covered the mouth shown in Fig. 7.11 with average accuracy 96.5 %

7.5.2 Pain prediction results

The system achieved the detection of the neck and upper pain in the following cases:

1- The hand on the right shoulder.
2- The hand on the left shoulder.
3- The hand on the right side of the neck.
4- The hand on the left side of the neck.

The difference between detection of the hand on the shoulder and the hand on the neck detected in the skin color step at the pre-processing stage. In Fig. 7.12 (a) the skin color detection for the full frame of the hand on the shoulder, the biggest connected 8×8 pixel area for face region and the second biggest connected 8×8 pixel area for hand region. In the Fig. 7.12 (b). The hand connected to the face for the hand on the side of the neck and detected as the biggest connected 8×8 pixel area. Example for the hand pose classification results with average accuracy 99.5 % shown in Fig. 7.13. Finally, in this chapter modeling the symptoms characteristics of fatigue pain, headaches, and drowsiness using the hand pose and facial features were presented. The results of four cases of pain and six hands pose over the face as a sign of a headache and finally, two cases for drowsiness predicted.

(a) (b)

Fig 7.12. Skin color detection for (a) the hand on the shoulder and (b) the hand on the neck.

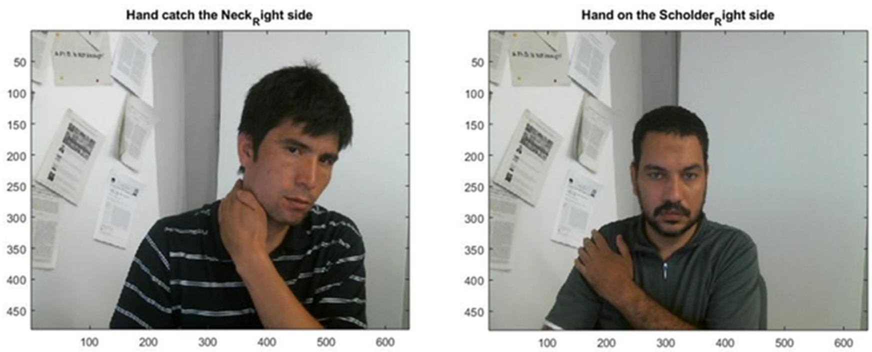


Fig.7.13. The classification results (a) the hand on the neck and (b) the hand on the shoulder.

Chapter 8

Conclusion

8.1 Comparison

In the proposed method, the hand over face gestures related to headache detection is closed domain compared with a general gesture recognition such as research of a sign language, human computer interaction, driver assists, and so on. Therefore, there is no medical standard database including a human action as a sign of headache. In this work, we compared the proposed method with three methods using a face detection, HOG, and SVM [30][31], Force field method and Camshift [32], and CNN with divided the face into 8 regions [33] as shown in Table 8.1. In the previous methods, face detection and head orientation directly affect to the post-processing. The detection of the border of the face and nose is difficult even though the deep learning applied in Hand2face method.

Table 8.1 Comparison of the proposed method with other methods of the hand over face pose with the occlusion problem

Method	Dataset	Feature	Accuracy
HOG and SVM [30]	Custom dataset segments frames form videos and selected another from Cam 3D dataset [53]	Hand cover eye and head rotation are failed in tracking	83.0 %
Force field method and Camshift [32]	Custom dataset taken by CCD camera	Face detention must be success	87.6 %
Convolutional neural network divided the face into 8 regions [11]	Custom dataset selected from LFPW dataset and COFW dataset [54]	Border of the face and nose is difficult for detection.	85.4 %
Proposed Method based on the deep learning with the CNN-CAM in real time processing	**Custom dataset, the participants from different countries and illumination change**	**Robust for illumination change (50～500 lx) and head orientation (~45°)**	**98.5 %**

8.2 Evaluation

Statistical evaluation for the proposed method and comparison with the Hand2face method. The following formulas are used.

Accuracy = (TP + TN)/(TP + FP + FN + TN)

Precision = TP/(TP + FP)

Recall = TP/(TP + FN)

F1 Score = 2 × (Recall × Precision)/(Recall + Precision)

Where TP is True Positives, TN is True Negatives, FP is False Positives, FN is False Positives, respectively. In the evaluation, we have achieved TP = 402, TN = 0, FP = 6 and FN = 0. The model's ability to find all relevant cases in our dataset (Recall) is 100%, Precision = 98.5%, and F1 score = 99.2% as shown in table 8.2. The proposed method has achieved high accuracy of 98.5% for classifying six hands over face gestures under illumination change (50–500 lx) and head orientation (∼45◦ left or right). The proposed method independent of face detection solved the feature extraction and skin color detection by using the deep learning. It enables us to cover a wide range of the skin tones under real office environment. Moreover, the hand over face gestures were classified based on the CNN-CAM, so it has detected the headache signs efficiently because robust to the skin color change, head orientations, and solved the occlusion problem of the hand over face.

8.3 Conclusion

For health monitoring of employees, in this research, modeling the symptoms characteristics of fatigue like headaches, pain, and drowsiness using the hand pose and developed a system for recognizing the fatigue sign and alert the computer users

Table 8.2 A statistical Evaluation comparison between proposed method and Hand2face method.

Method	Accuracy	Precision	Recall	F1
Proposed Method	**98.5 %**	**98.5 %**	**100 %**	**99.2 %**
Hand2Face	85.36 %	84.63 %	89.11 %	86.81 %

to take rest and do some optimum exercise to staying healthy and safe. Proposed system is easy to install on a general PC because it only uses RGB camera. There is no need for distinct gadgets and markers. Multiple methods used to improve the classification accuracy and avoid the face detection failure in case of the hand over face gestures by performing the skin and face detection and the deep learning method for classification. The method solved the hand over face occlusion problem. In addition, the system is robust against skin color, illumination change and head orientation. In the proposed method, a deep learning based on a convolutional neural network (CNN) for the classification of the hand poses is applied. In addition, a class activation map (CAM) to visualize the prediction of the classification network for localization of the hand over face poses implemented. From the experimental results, the proposed method was designed to reduce the type-I error (false-positive result) for the classification by applied the pre-processing steps for face and skin color detection. The drowsiness successfully recognized in the following cases: The eye classified closed and mouth open at same time and in case, the hand cover the mouth. The system achieved the detection of the neck and upper pain in the following cases: the hand on the right shoulder, the hand on the left shoulder, the hand on the right side of the neck and the hand on the left side of the neck. The system achieved the classification ratio of drowsiness average 96.6 % and pain signs average 99.5%. A headache predicted for the hand poses as the signs of frontal and unilateral headaches without the classification overfitting and data biasing errors with high accuracy recognition ratio of 98.5% for classification.

8.4 The future work

Simulate of some necessary exercises according to the fatigue sign, such as neck or shoulder pain treatments due to long periods of work such as the call center operators. By robots, a tool in the near future will be present in and spread in the work environment we can simulate the exercise. In addition, the exercises simulate by animation or video for professionals through the computer or mobile devices and tablets. Evaluate of the performance of the exercises by the processing of digital images taken for the person during the exercise, through the depth cameras its available in many gaming devices and work on development of algorithms can evaluated the human action with ordinary cameras.

REFERENCES

1. Dol, Kim Sang. "Fatigue and pain related to Internet usage among university students." *Journal of physical therapy science* 28.4 (2016): 1233-1237.
2. Lin, Yen-Hui, et al. "Visual fatigue during VDT work: Effects of time-based and environment-based conditions." *Displays* 29.5 (2008): 487-492.
3. Iwakiri, Kazuyuki, et al. "VDT worker's posture and workload in free-address office system." *Sangyo Eiseigaku Zasshi= Journal of Occupational Health* 48.1 (2006): 7-14.
4. Fukuta, Kentarou, Teppei Koyama, and Takashi Uozumi. "Representation of visual fatigue during VDT work using Bayesian network." *Soft Computing as Transdisciplinary Science and Technology*. Springer, Berlin, Heidelberg, 2005. 581-590.
5. del Mar Seguí, María, Elena Ronda, and Peter Wimpenny. "Inconsistencies in guidelines for visual health surveillance of VDT workers." Journal of occupational health (2011): 1112090220-1112090220.
6. Segu´ı M d M, Ronda E, Wimpenny P. Inconsistencies in guidelines for visual health surveillance of VDT workers. Journal of Occupational Health 2012; 54:16–24.
7. World Health Organization. Visual Display Terminals and Workers. World Health Organization, Geneva: WHO; 1987.
8. https://apps.who .int/iris/handle/10665/38218?show=full.
9. Reddy SC, Low CK, Lim YP, Low LL, Mardina F, Nursaleha MP. Computer vision syndrome: A study of knowledge and practices in university students. Nepalese Journal of Ophthalmology 2013; 5(10):161–168.
10. Wiederhold, Brenda K. "Connecting through technology during the coronavirus disease 2019 pandemic: Avoiding “Zoom Fatigue”." (2020): 437-438.
11. Asgari, Behrooz, Peter Pickar, and Victoria Garay. "Karoshi and Karou-jisatsu in Japan: causes, statistics and prevention mechanisms." Asia Pac Bus Econ Perspect 4 (2016): 49-72.
12. Kujawinska A, Vogt K. Human factors in visual quality control. ´ Management and Production Engineering Review 2015; 6:25–31.

13. Zapf, Dieter, et al. "Emotion work and job stressors and their effects on burnout." Psychology & Health 16.5 (2001): 527-545
14. Michel Rod, Nicholas J. Ashill, (2013) "The impact of call center stressors on inbound and outbound call center agent burnout", Managing Service Quality: An International Journal, Vol. 23 Issue: 3, pp.245-264.
15. Kusuma Kumari B.W, “REAL Time Detection Derivers Drowsiness Using computer Vision”, IJRET, vol. 3, pp. 147-151, May 2014.
16. Saeid. Fazli, and Parisa. Esfehani, “Tracking Eye state for Fatigue Detection”, ICACEE, pp. 17-20, Nov. 2012.
17. Mihir Jain,Suman K. Mitra and Naresh D.Jotwani, “Eye Detection using Line Edge Map Template”, VISAPP ’08. pp. 152-157, May 2008.
18. Mohd Shamian Bin Zainal,Ijaz Khan and Hadi Abdullah, “Efficient Drowsiness Detection by Facial Features Monitoring”, Research Journal of Applied Sciences, Engineering and Technology. pp. 2376-2380, March 2014.
19. Mitesh Patel , Sara Lal, Diarmuid Kavanagh, and Peter Rossiter, “Fatigue Detection Using Computer Vision”, Jet, 461vol.56,No 4, pp. 457-, July . 2010.
20. Vijayalaxmi and D.Elizabeth Rani , “Driver Fatigue Estimation Using Image Processing Technique”, I.J. Information Technology and Computer science, pp.66-72 , June 2016
21. Michel Rod, Peter Thirkell, Janet Carruthers, (2009) "Job resourcefulness, symptoms of burnout and service recovery performance: an examination of call centre frontline employees", Journal of Services Marketing, Vol. 23 Issue: 5, pp.338-350.
22. Yen-Hui lin, Chih- Yong chen, shih-Yi LU and Yu-chao Lin, “Visual fatigue during VDT work: Effects of time based and environment based conditions”, ELSEVIER, pp. 487-492, 2008.
23. Yun Hua Chen, Weijian Liu, Ling zhang,Mingyu Yan,Yanjun Zeng, “Hybrid facial image feature extraction and recognition for non invasive chronic fatigue syndrome diagnosis”, ELSEVIER, pp. 30-39, 2015.
24. Lew, M.; Sebe, N.; Huang, T.; Bakker, E.; Vural, E.; Cetin, M.; Ercil, A.; Littlewort, G.; Bartlett, M.; Movellan, J. Drowsy driver detection through facial movement analysis. In Human-Computer Interaction; Springer: Berlin, Germany, 2007; Volume 4796, pp. 6–18.

25. Saradadevi, Mandalapu, and Preeti Bajaj. "Driver fatigue detection using mouth and yawning analysis." International journal of Computer science and network security 8.6 (2008): 183-188.
26. Xiao, F.; Bao, C.Y.; Yan, F.S. Yawning detection based on gabor wavelets and LDA. J. Beijing Univ. Technol. 2009, 35, 409–413
27. Zhang, Z.; Zhang, J. A new real-time eye tracking based on nonlinear unscented Kalman filter for monitoring driver fatigue. J. Contr. Theor. Appl. 2010, 8, 181–188.
28. Flores, M.; Armingol, J.; de la Escalera, A. Driver drowsiness warning system using visual information for both diurnal and nocturnal illumination conditions. EURASIP J. Adv. Signal Process. 2010, 2010, 438205.
29. Yin, B.-C.; Fan, X.; Sun, Y.-F. Multiscale dynamic features based driver fatigue detection. Int. J. Pattern Recogn. Artif. Intell. 2009, 23, 575–589.
30. D'Orazio, T.; Leo, M.; Guaragnella, C.; Distante, A. A visual approach for driver inattention detection. Pattern Recog. 2007, 40, 2341–2355.
31. Mahmoud M, El-Kaliouby R, Goneid A. Towards communicative face occlusions: Machine detection of hand over face gestures. ICIAR 2009, LNCS 5627 , 2009; 481–490.
32. Mahmoud, M, Baltrusaitis T, Robinson P. Automatic analysis of ˇ naturalistic hand-over-face gestures. ACM Transactions on Interactive Intelligent Systems (TiiS), 2016; 1–18.
33. Xa J, Zhang X. A real time hand detection system during hand over face occlusion. International Journal of Multimedia and Ubiquitous Engineering 2015; 10:287–302.
34. Nojavanasghari B, Hughes CE, Baltrusaitis T, Morency L-P. Hand2face: Automatic synthesis and recognition of hand over face occlusions. International Conference on Affective Computing and Intelligent Interaction (ACII), 2017.
35. El Sayed, A. Sharara, A. Tsuji, and K. Terada. "Burnout Recognition for Call Center Agents by Using Skin Color Detection with Hand Poses." International Journal of Computer and Information Engineering 11.9 (2017): 1066-1069.
36. Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A Learning deep features for discriminative localization. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.
37. Selvaraju, RR, Das A, Vedantam R, Cogswell M, Parikh D, Batra D Grad-CAM: Why did you say that?. arXiv preprint arXiv:1611.07450, 2016.

38. Zhang, Xufan, et al. "NeuralVis: visualizing and interpreting deep learning models." 2019 34th IEEE/ACM International Conference on Automated Software Engineering (ASE). IEEE, 2019.
39. Meng, Fanman, et al. "Class Activation Map Generation by Representative Class Selection and Multi-Layer Feature Fusion." arXiv preprint arXiv:1901.07683 (2019).
40. Krig, Scott. Computer vision metrics. Berlin, Germany:: Springer, 2016.
41. Lowe, David G. "Distinctive image features from scale-invariant keypoints." International journal of computer vision 60.2 (2004): 91-110.
42. Shih, Frank Y. Image processing and pattern recognition: fundamentals and techniques. John Wiley & Sons, 2010.
43. Bay, Herbert, et al. "Speeded-up robust features (SURF)." Computer vision and image understanding 110.3 (2008): 346-359.
44. Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). Vol. 1. IEEE, 2005.
45. Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Communications of the ACM 60.6 (2017): 84-90.
46. Deng, Jia, et al. "Imagenet: A large-scale hierarchical image database." 2009 IEEE conference on computer vision and pattern recognition. Ieee, 2009.
47. Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).
48. Szegedy, Christian, et al. "Going deeper with convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.
49. Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001. Vol. 1. IEEE, 2001.
50. Kakumanu P, Makrogiannis S, Bourbakis N. Survey of skincolor modeling and detection methods. Pattern Recognition 2007; 40:1106–1122
51. Sachdeva S. Fitzpatrick skin typing: Applications in dermatology. Indian Journal of Dermatology, Venereology and Leprology 2009; 75:93–96.

52. Afifi L, Saeed L, Pasch LA, Huddleston HG, Cedars MI, Zane LT, Shinkai K. Association of ethnicity, Fitzpatrick skin type, and hirsutism: A retrospective cross-sectional study of women with polycystic ovarian syndrome. International Journal of Women's Dermatology 2017; 3:37–43.
53. Kaneko F, Nakamura K, Furukawa H, Oyama N, Nakamura T, Zheng X. Biological characteristics of the sensitive Japanese skin. International Journal of Cosmetic Science 2015; 27:66–67.
54. Mahmoud M, Baltrusaitis T, Robinson P, Riek LD. 3D corpus of ˇ spontaneous complex mental states. In International Conference on Affective Computing and Intelligent Interaction. Springer: Berlin, Heidelberg; 2011; 205–214.
55. Belhumeur PN, Jacobs DW, Kriegman DJ, Kumar N. Localizing parts of faces using a consensus of exemplars. IEEE Transactions on Pattern Analysis and Machine Intelligence 2013; 35(12):2930–2940

Referred Publications

MAIN

PEER-REVIEWED JOURNAL ARTICLE

2020 Abousharara A. Elsayed, Akinori Tuji ,Steohen karungaru and Kenji Terada "Prediction of the VDT Worker's Headache Using Convolutional Neural Network with Class Activation Mapping"IEEJ Transactions on Electrical and Electronic Engineering, Early View 9,2020. **https://doi.org/10.1002/tee.23239**

PROCEEDINGS OF CONFERENCE

5/2019 Abousharara Abdelghany Mohamed Elsayed, Akinori Tsuji and Kenji Terada : Classifictaion of Hand Over Face Gestures using Deep Learning, The Fifteenth International Conference on Quality Control by Artificial Vision 2019(QCAV2019), No.11172, Mulhouse, France, May 2019. doi: 10.1117/12.2521269

6/2018 Abousharara A.Elsayed, Akinori Tuji and Kenji Terada, "Visual Fatigue Recognition Using Facial Features and Hand Poses for VDT workers" , Proceedings of the International Conference on Electrical Engineering (ICEE2018), No.S3-2483, Seoul, Jun. 2018

7/2019 Abousharara Abdelghany Mohamed Elsayed, Akinori Tsuji and Kenji Terada : Drowsiness Detection by Facial Features and Hand Gestures Using Deep learning, The International Council on Electrical Engineering Conference 2019(ICEE2019), No.3, Hong Kong, Jul. 2019.

4/2019 Abousharara A.Elsayed, Akinori Tuji and Kenji Terada,"Advanced prevention for "KAROSHI" using artificial intelligence,The Future of Japanese science", nature's150th Anniversary Symposium.

12/2018 Abousharara A. Elsayed, Akinori Tuji and Kenji Terada ,"Recognition of Fatigue Sings using Artificial Intelligence", Proceedings of the 1st Scientific Conference Advanced and New Approaches in Physical Therapy, Cairo, Dec. 2018

10/2018 Abousharara A.Elsayed, Akinori Tuji and Kenji Terada,"Recognition of Fatigue Signs for VDT Workers Using MATLAB", MATLAB EXPO 2018 Japan, Oct. 2018.

OTHER

Tuniyazi Abudoureheman, Abousharara Abdelghany Mohamed Elsayed, Akinori Tuji and Kenji Terada : Multiple People Tracking Based on Kalman Filter in Complex Background, 平成 30 年度電気関係学会四国支部連合大会, No.13-16, Sep. 2018.

Abousharara Abdelghany Mohamed Elsayed, Akinori Tuji and Kenji Terada : 人工知能の活用による「過労死」の先進的な予防システム, SCI- Tech Festival 2019