

Automated Patient Discomfort Detection Using Deep Learning

Imran Ahmed¹, Iqbal Khan¹, Misbah Ahmad¹, Awais Adnan¹ and Hanan Aljuaid^{2,*}

¹Center of Excellence in Information Technology, Institute of Management Sciences, Peshawar, Pakistan

²Computer Sciences Department, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University (PNU), Riyadh, Saudi Arabia

*Corresponding Author: Hanan Aljuaid. Email: haaljuaid@pnu.edu.sa

Received: 28 June 2021; Accepted: 30 August 2021

Abstract: The Internet of Things (IoT) has been transformed almost all fields of life, but its impact on the healthcare sector has been notable. Various IoT-based sensors are used in the healthcare sector and offer quality and safe care to patients. This work presents a deep learning-based automated patient discomfort detection system in which patients' discomfort is non-invasively detected. To do this, the overhead view patients' data set has been recorded. For testing and evaluation purposes, we investigate the power of deep learning by choosing a Convolution Neural Network (CNN) based model. The model uses confidence maps and detects 18 different key points at various locations of the body of the patient. Applying association rules and part affinity fields, the detected key points are later converted into six main body organs. Furthermore, the distance of subsequent key points is measured using coordinates information. Finally, distance and the time-based threshold are used for the classification of movements associated with discomfort or normal conditions. The accuracy of the proposed system is assessed on various test sequences. The experimental outcomes reveal the worth of the proposed system' by obtaining a True Positive Rate of 98% with a 2% False Positive Rate.

Keywords: Artificial intelligence; patient monitoring; discomfort detection; deep learning

1 Introduction

The IoT begins smart healthcare systems in the medical sector, generally comprised of smart sensors, a remote server, and the network. In smart healthcare, it has many applications, including early warning service (emergency, first aid, medical assessment), real-time supervision services (patient monitoring, elderly care), scheduling and optimization service (medical staff allocation, bed allocation, resources allotment). A patient monitoring system has been gaining the consideration of researchers in the field of advanced computer vision and machine learning. It is one of the ongoing research fields because of its broad range of applications, including respiration monitoring, pain detection, depression monitoring, sleep monitoring, patient behavior monitoring, posture monitoring, epilepsy seizure detection, etc. Researchers have developed different systems for patient monitoring systems, e.g., some use specialized hardware, pressure mattresses, and



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

sensors but at the cost of additional expense. Similarly, connecting sensors to the body of the patient is unwilling from the patient's point of view. Few used signal based approaches to observe breathing, depth rate, and steadiness of breath besides monitoring the breath time and the ratio. Even though pain detection techniques exist, they mainly use facial expressions. The major drawback of such systems is that they require the patient to align face directly to the camera. A sleep monitoring system has been developed to detect sleep apnea and sleep disorders; such a system is also based on hardware and sensors installed in patients' beds. Some techniques monitor patient behavior, which helps to analyze their medical condition. However, such developed techniques are based on the installation of multiple cameras.

Multiple camera posture-based monitoring techniques have also been developed, e.g., mainly focusing on the upper body part of the patient. Because of these limitations, a non-invasive discomfort detection system has been proposed in this work, which neither utilizes specialized hardware/sensors nor a line of sight vision devices or any constrained/ specialized environment. The introduced system is principally based on ten layers of the Convolutional neural network (CNN). It is the class of deep learning containing input, output as well as some hidden layers. The layers are fully connected, which helps to detect and recognize features and patterns. A pretrained model is used to test/evaluate the patient's discomfort using our newly recorded data set. The CNN model's output is 18 keypoints detected on different patient body locations using confidence maps. The detected keypoints are further utilized to shape six major body organs. This formation is based on association rules and part affinity fields. The distance of all detected keypoint is estimated from each succeeding keypoint of successive frames. The distance and time-based thresholds are considered to recognize discomfort in a specific organ of the body of the patient. Finally, experimental evaluation is made using manually created ground truths. The work presented in the paper has the following main contributions;

- An automated system is introduced for detection of patient discomfort using a deep learning-based model.
- By utilizing CNN architectures, confidence maps and 18 different keypoints are detected at various locations of the patient's body,
- The detected keypoints are then converted into six main body parts/organs based on association rules and part affinity fields, and the distance of the following key points is measured using coordinates information,
- Finally, distance and the time-based threshold are utilized for the classification of movements as either discomfort or normal conditions.

The proposed system could have many possible applications such as analysis, monitoring, detection of pain, discomfort, automatic patient monitoring in hospitals or homes, and elderly monitoring. The presented work is organized as follows: A review of the related work has been presented in Section 2. Then, the proposed system is introduced in Section 3. While Section 4 explains experimental results. Lastly, Section 5 concludes the presented work and provides future directions.

2 Literature Review

In recent years automated patient monitoring has been gaining the interest of researchers. Different signal, image processing, and computer vision techniques have been developed in the last decade. Some of the techniques have been discussed in this section which has been categorized as follows:

2.1 Respiration Monitoring Approaches

Respiration monitoring aims to observe the depth and steadiness of breath besides monitoring the inhalation and exhalation time and the ratio. Cho et al. [1] used a thermal image-based approach to respiration rate monitoring by specifying a region of interest under the nose. In [2], a radio frequency-based method is proposed, which helps to estimate the rate of respiration using a Multiple Signals Classification (MUSIC) algorithm. Authors in [3], presented a contactless breathing monitoring system using single camera approach. Ostadabbas et al. [4] have proposed a respiration monitoring system for estimating airway resistance non-intrusively using depth data obtained from the Microsoft Kinect sensor. Fang et al. [5] proposed a system for detecting sudden infant death syndrome. Al-Khalidi et al. [6] used facial thermal images of children to monitor their respiration rate. Janssen et al. [7] use the intrinsic respiratory features for finding the region of interest for respiration and motion factorization to extract respiration signals. Braun et al. [8] divide the input images into blocks and then estimates motion for each block. These block motions are then classified to find the respiration activity. Wiede et al. [9] introduce a method for remotely monitoring respiration rate using RGB images. This approach finds the region of interest and applies principal component analysis and frequency finding methods to determine the respiration rate. Frigola et al. [10] produced a video-based non-intrusive technique for respiration monitoring, which detects movement applying optical flow and quantifies the detected movement. Monitoring a patient's respiration can provide insights and help diagnose many diseases like lung problems and abnormal respiration rates.

2.2 Pain Detection and Depression Monitoring Approaches

In the literature, pain detection and depression monitoring has been handled mostly by analyzing facial expressions. Authors in [11] exploited facial appearances for pain detection by using a feature-based method similar to [12–16], i.e., pyramid histogram of oriented gradients and pyramid local binary pattern. They used these features to extract the shape and appearance of patients' faces, respectively. Authors in [17] used Prkachin and Solomon Pain Intensity (PSPI) metric. Other approaches that consider facial emotions to detect pain and/or depression are proposed in [18–22]. Each of these movements is categorized as a different action unit. The authors extract the face's canonical appearance using Active Appearance Models (AAMs), filtered to extract features. These features are then fed to different SVMs, each trained to measure a separate level of pain intensity. In [23], authors suggested a system using AAM to detect patients' pain in videos. In [24,25], authors introduced a system that could discriminate facial emotions of pain from other facial emotions and applied SVM for severity score of pain. The system has been tested on UNBC-McMaster Database [26] using four different classifiers, namely SVM, Random Forest, and two neural networks. For assessment of the system, they applied the HI4D-ADSIP data set [27]. Nanni et al. [28] classify pain states by proposing a descriptor named Elongated Ternary Patterns (ELTP), which combines the features of Elongated Binary Pattern (ELBP) [29] and Local Ternary Patterns (LBP).

2.3 Sleep Monitoring Approaches

Sleep monitoring encompasses recording and analyzing chest and abdomen movements, as is the case with respiration monitoring. In [30], Al-Naji et al. developed a system for detecting sleep apnea and monitoring respiration rate in children by using the Microsoft Kinect sensor. Li et al. [31] proposed a non-invasive system for cardiopulmonary signals monitoring in various sleeping positions. The infrared light source and Infrared sensitive camera are used in this approach. Metsis et al. [32] proposed sleep patterns monitoring system. They have investigated

many factors corresponding to sleep disorders. Malakuti et al. [33] address the problem of sleep irregularities based on pressure data. Liao et al. [34] designed and measured the sleep quality using infrared video. They have used the technique of motion history image [35] for analyzing videos to recognize the patterns of patients' movements. Nandakumar et al. [36] introduced a smartphone-based sleep apnea detection system, which analyzes chest and abdominal motion. Saad et al. [37] proposed a device for finding sleep quality using several sensors in the room. The sensors are used for determining heart rate, temperature, and movement of the body. Hoque et al. [38] attach WISPs [39] to the bed's mattress to know about the positions of the body and thereby monitor sleep. Accelerometer data is used for movement detection.

2.4 Behavior Monitoring Approaches

Human behavior understanding also plays a vital role in knowing much about people. Borges et al. [40] tried to recognize individual activities associated with psychiatric patients by utilizing blob detection, and optical flow analysis, and applied decision rules to analyze patients' activities. Authors in [41] proposed a system based on monitoring patients' vital signs to prevent incidents such as falling, injuries, and pain. The system uses the Canny Edge Detector and Hough Transform for detecting beds. Once a bed is detected, the system determines whether or not a patient is present in the bed by detecting the patient's head. Martinez and Stiefelhagen [42] have applied multi-cameras for observing the behavior of patients' in an ICU irrespective of the environmental conditions. By examining a patient's behavior, much information can be collected about his medical conditions [43].

2.5 Posture Monitoring Approaches

Knowing about patient posture proves helpful for purposes like fall detection, pressure ulcer detection, and activity recognition. Chang et al. [44] introduced a system based on depth videos for restricting pressure ulcers in the bed of patients by investigating their movement and posture. In [45], the authors introduced a non-invasive patient posture monitoring method. This approach extracts HOG features for the classification of postures. The system also tracks the postures of the patient and generates a report accordingly. Wang et al. [46] have introduced a monitoring system for recognizing a person's pose while covered with a blanket. In another approach, [47] proposed a system for determining the top body parts of the human under a blanket utilizing an overhead camera [48–50]. Brulin et al. [51] suggested a technique for monitoring the elderly at home. The proposed method is based on posture recognition. This technique detects the individual body and then utilizes posture identification methods on the human silhouette based on Fuzzy Logic.

2.6 Epilepsy Monitoring Approaches

Many attempts have been made towards vision-based detection and prediction of epilepsy seizures. In [52] proposed a method for eyeball detection. The main purpose is to track the movement of eyes for knowing about the presence or absence of epileptic seizures. Lu et al. [53] used color videos and proposed a method for quantification of limb movement occurring in seizures associated with epilepsy. Cuppens et al. [54] apply the optical flow method and detect epilepsy movement. Kalitzin et al. [55] used the optical flow method to find movements associated with epileptic seizures.

All of the above discussed approaches focus either on a single patient and/or a single bed, and specialized hardware is used. Also, the intrusive approaches among these need connecting sensors to the body or bed to record various measurements that are both costly and unwanted from the patient's point of view. Even though pain detection approaches are there, they wholly solely

depend on facial expressions, restraining the patient from retaining his/her face directly towards the camera. On the other hand, the proposed system may work in the existing wards setups monitor more than one patient simultaneously, lacking advanced beds or functional equipment, etc., except a single camera. Being non-invasive, it makes no contact with the patient while recording their movements. Recently, scholars also utilized deep learning based methods [56–59] for patient discomfort monitoring [60]. In this work, we also used a deep learning based method for automated patient discomfort detection.

3 The Proposed Method

In this section, a deep learning based sustainable discomfort detection system is introduced. The flow chart presented in Fig. 1 highlights the main steps of the proposed method. The proposed method is mainly based on Convolution Neural Network (CNN) based architecture [61]. Firstly, the input images of the patient from the IMS-PDD-II data set are transmitted to the pre-trained model, which detects key points at various locations on the body of a patient. Then, the information of detected key points is then applied for the formation of the patient body organs using defined association rules. Finally, a distance threshold has been applied to recognize discomfort or pain in the organs of a patient's body. The detailed explanation of the proposed method exhibited in Fig. 1 has been described with the help following steps:

- The pre-trained model used non-parametric representation, which is called parts affinity fields. The parts affinity fields contain the orientation and position information used to identify human body parts in the input image. The model employs CNN architecture, shown in Fig. 1 [62]. The input images from the data set are given to the pre-trained model. The trained model mainly has two branches—the top branch is used for predicting the confidence maps and detection of human body parts, while the bottom one is for predicting part affinity fields, which are used to link human body parts, as shown in Fig. 2. Each of the two branches is repetitive prediction architecture refining the predictions via the number of successive stages.

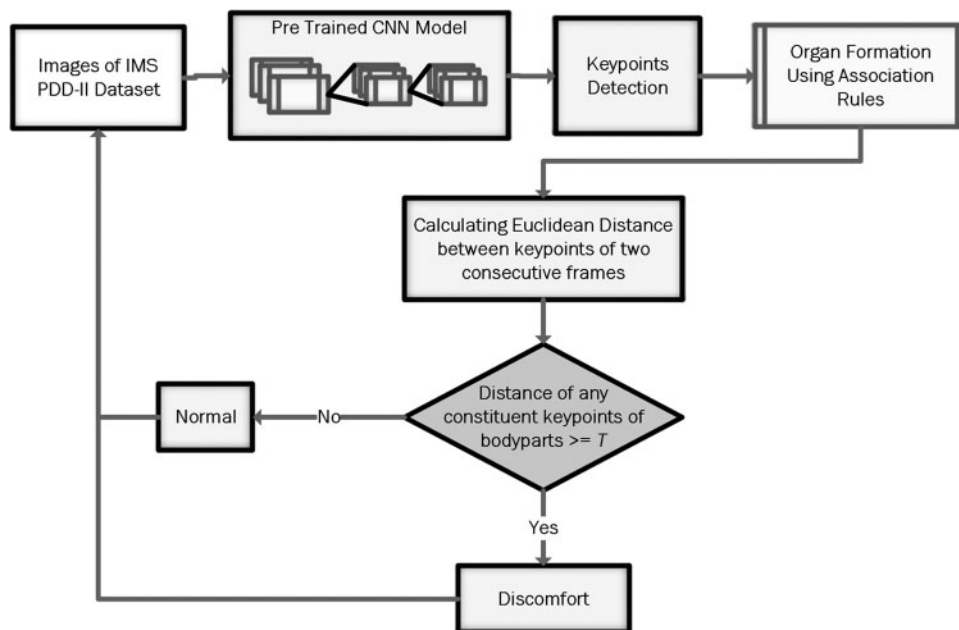


Figure 1: Flowchart of the CNN-based discomfort detection method

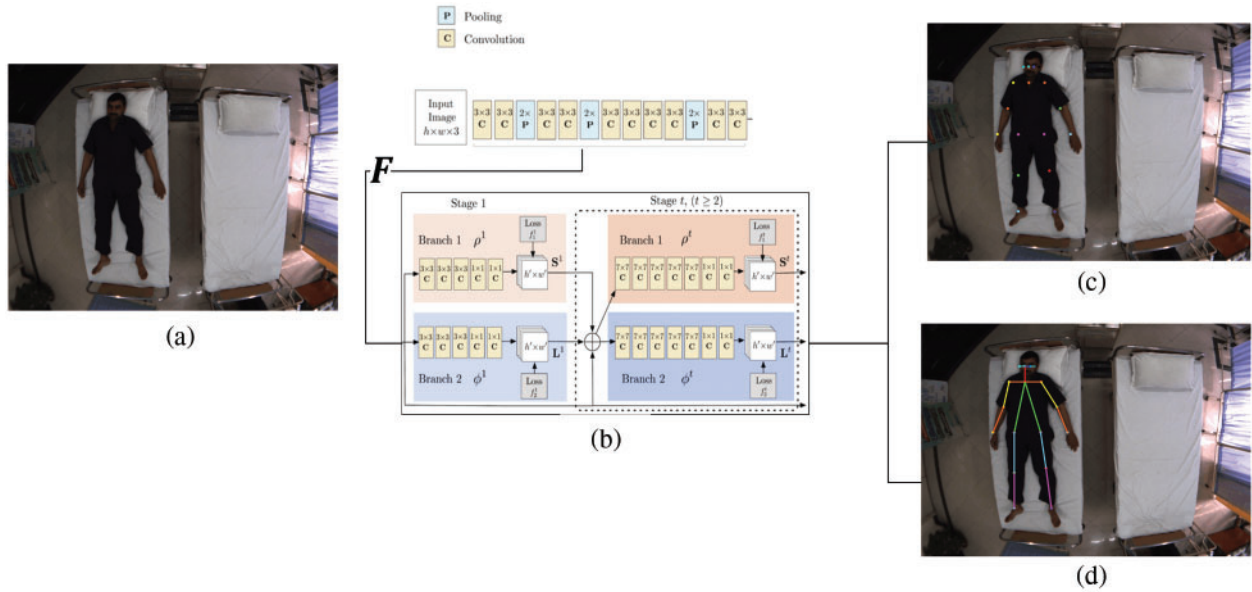


Figure 2: Proposed Model Architecture. (a) shows input image, (b) CNN model and (c) shows detected key points and (d) shows detected patients' body organs

- A set of feature maps represented by F are extracted for each input image using CNN. The F is used as input features to the initial stages of both branches, as shown in Fig. 2. At these initial stages, the network generated a set of detection confidence maps. The detected confidence maps for the initial stage is given as;

$$S^1 = p^1(F) \quad (1)$$

While for t^{th} stage the confidence, maps have been calculated as;

$$S^t = p^t(F, S^t, L^{t-1}), \forall t \geq 2 \quad (2)$$

In Eq. (2), t is the CNN for inference at the initial stage to t^{th} stage of branch 1 as shown in Fig. 2.

- The part affinity fields have also been generated along with confidence maps S^1 . The part affinity field for the initial stage is calculated using the below Equation:

$$L^1 = \phi^1(F) \quad (3)$$

Moreover, for t^{th} stage, the part affinity fields are shown in Eq. (2).

$$L^t = \phi^1(F, S^t, L^{t-1}), \forall t \geq 2 \quad (4)$$

Here ϕ^1 represent the CNN for inference at the initial stage to t^{th} stage of branch 2. After every succeeding stage, the model concatenates both branches' predictions and generates image features. These features are used for refined predictions calculated in Eqs. (2) and (4), as shown in Fig. 2.

- For iterative prediction of confidence map of the human body part at the first branch and part affinity fields at the second branch of each stage, loss function has been calculated. As there are two branches, so two-loss functions are calculated and applied on each stage. These loss functions are given by Eqs. (5) and (6) [62]. The first loss function for the first branch and calculated as f_t^S ;

$$f_t^S = \sum_{j=1}^J \sum_p W(p) \cdot \|S_j^t(p) - S_j^*(p)\|^2 \quad (5)$$

In Eq. (5) S_j^* is ground truth confidence map of human body. The second loss function f_t^L for ground truth of part affinity fields is given as:

$$f_t^L = \sum_{c=1}^C \sum_p W(p) \cdot \|L_c^t(p) - L_c^*(p)\|^2 \quad (6)$$

where L_c^* is ground truth of part affinity vector. In Eqs. (5) and (6), p is the location at input image, W is a binary mask equal to 0, in case annotation is missing there at location p . The calculated loss function at each stage is to minimize the distance between predicted and real confidence maps for each affinity part.

- The main objectives of the calculated loss function L for full architecture shown in Fig. 2 are obtained by adding Eqs. (5) and (6) is given by.

$$L = \sum_{t=1}^T (f_t^S + f_t^L) \quad (7)$$

- The pretrained model shown in Fig. 3 gives 18 detected key points on the body, as determined in Fig. 4a. The key points information is moreover utilized to form body organs, as shown in Fig. 4b. Finally, using association rules, six organs of body are formed and have been manually highlighted in Fig. 4d.
- When a patient feels any type of discomfort, frequently movement occurs in any such part of the patient's body. For example, the patient may touch/hold his/her head with hands or moving legs or arms. Furthermore, in few cases, patients may move his/her legs, arms, or any other part in a disruptive way. For instance, he/she sometimes sits or lies or switch sides frequently. All such random and frequent changes are considered as discomfort signs. If the frequency of these frequent and random movements lasts for a long duration, it is considered as a discomfort condition. The discomfort investigation in a body is based on constant movements of the specific part of body. The presented system determines a change in the body organ utilizing key points information across time and categorizes the condition as discomfort or normal. The coordinates information of detected key points is used to identify pain. The movement in any body part or organ is measured using distance information that is determined by applying Euclidean distance across consecutive video frames.

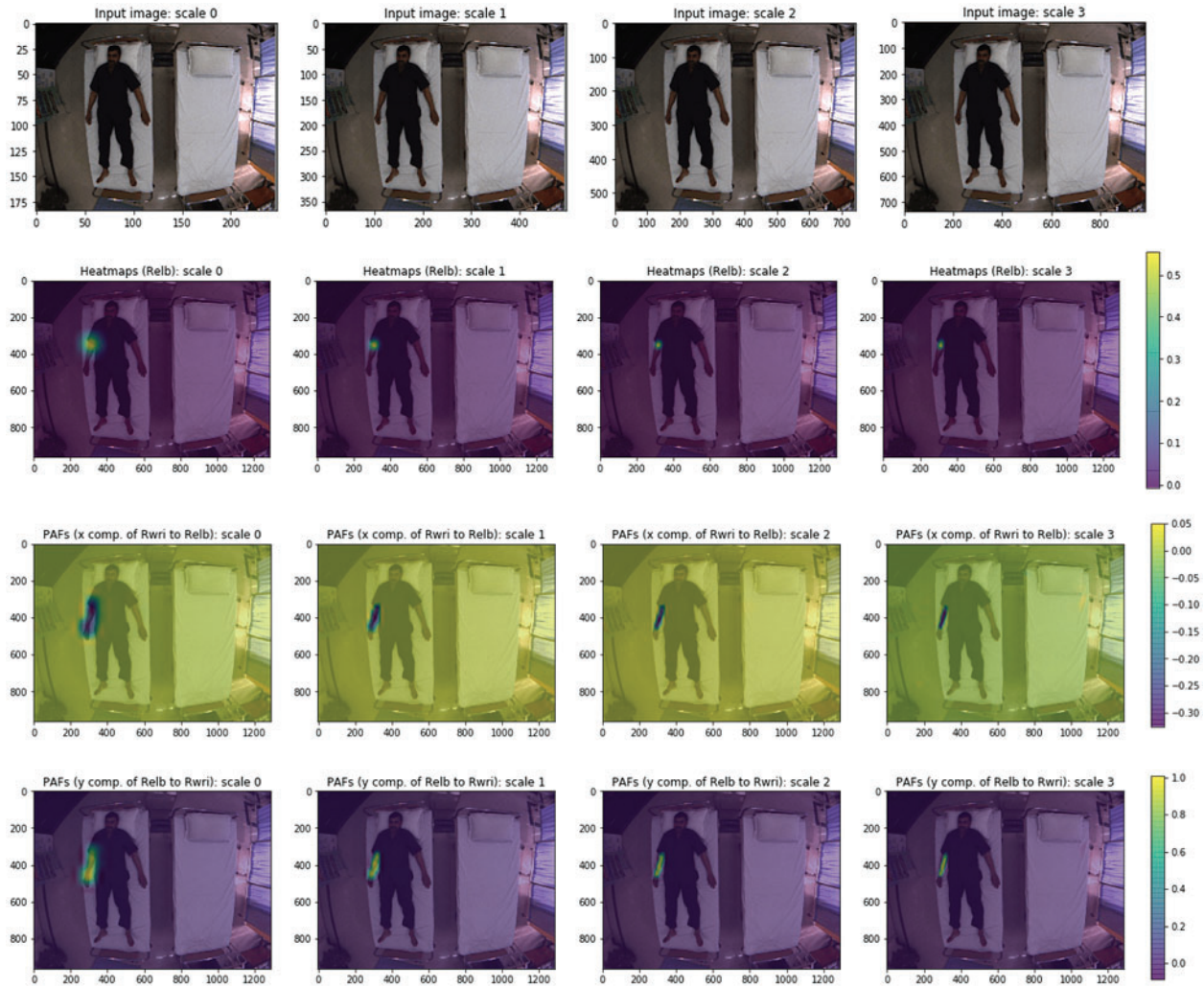


Figure 3: The sample images show Heatmap and PAF's for the right elbow. The body part is encoded in the 3rd channel, so in this case, the right knee is at index 9

- The threshold measuring distance T of consecutive key points is used in terms of the number of pixels. T have been set as 25 pixels. The threshold decides movements in the patient body organ or part b . For instance, a variation in the coordinates (x,y) of detected key points e.g., 5, 6 and 7 on a body of the patient will cause a movement in the left arm and change in the (x,y) coordinates of joints 8, 9 and 10 would mean a movement in the right leg. For this reason, the Euclidean distances for all detected key points of that body organs have been examined using Eq. (8).

$$M_{body_part_b} = \begin{cases} 1, & \text{if } \forall_{i=1}^n = 1d \geq T \\ 0, & \text{Otherwise} \end{cases} \quad (8)$$

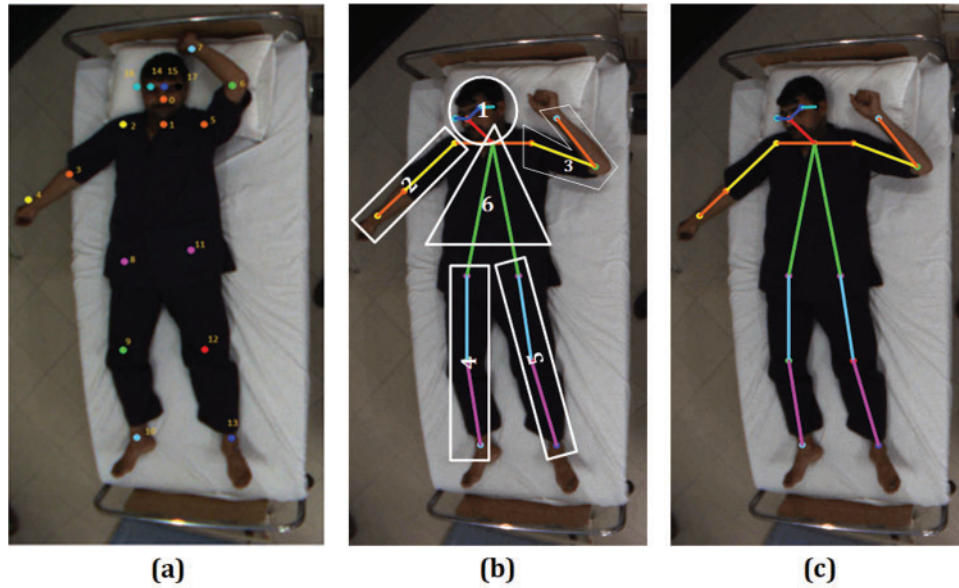


Figure 4: Organ formation (a) shows detected key points on the patient's body, (b) shows six different body organs formed using part affinity method and association rules, (c) shows linking of six body organs

- Lastly, to investigate either a patient is feeling normal or having some discomfort problem, video frames are examined for frequent movements of occurrences using a time-based threshold T_t as shown in Eq. (9). (This threshold can be changed depending upon on size and variety of data set. In this work, ten frames per second have been practiced due to limited data set).

$$C_{patient} = \begin{cases} Discomfort, & M_{body_part_b} \sim T_t \\ Normal, & Otherwise \end{cases} \quad (9)$$

where $C_{patient}$ represent the condition of the patient P . T_t is the time threshold representing the span of time that discriminates between the normal and discomfort movements.

4 Experimental Results and Discussion

The proposed method has been evaluated on a recorded IMS-PDD-II data set. A brief description of the video clips considered in this work is given in Tab. 1. Experiments have been performed on an HP core i3 Laptop with 8 GB RAM. The frames of the video clips are given as input to the pre-trained model to identify key points and organs of the patient's body. A few output images of detected organs can be observed from Fig. 5. After detecting the key points, the movement frequency of the patient organs has been analyzed using key points coordinates information. Based on movement frequency, the discomfort in the patient's body has been decided. In this section, the result of different video clips has been discussed; each video clip contains movement in the different organs of the patient body. The results of the different video clips show movements in different organs are briefly discussed in this section.

Table 1: Description of video clips used

Video clip no.	No. of patients	Movement in patient organs	
		Bed 1	Bed 2
1	One	Left Arm	–
2	One	Right Arm	–
3	One	Right Leg	–
4	One	Both Arms	–
5	Two	Any Part	Any Part

**Figure 5:** Sample of output images shows organs of the patients body

In video 1, the patient moved his left arm over many times, as noted in [Fig. 6](#). To be exact, left arm involves movement in frames 21–40, 43–54, 57–82, 84–135, 137–157, 160–174, 176–188, 190–207, 209–229, 239–258. All these changes occurs continuously and are greater than the defined threshold. Also, these sequences of movements in the left arm are separated by no movement in one or two frames, indicating that there is continuous movement in the left arm. It determines that there is severe pain (discomfort) in the left arm. The movements in the left arm are also

accompanied by movement or changes in the right arm in some frames because the patient retains touching his left arm with his right hand, as explained in Fig. 7. In video 2, excessive movements have occurred in the patient’s right arm, i.e., in frames 14–22, 24–68, 73–84, 93–102, 148–180, 185–202, 207–218, and 227–236 consecutively. The frequency of the movement in the right arm is large compared to other organs. In addition, the patient has moved his head, left arm, and both legs in some of the frames, as is seen from Fig. 7. However, as the pattern or frequency of the right arm is greater than the threshold, this indicates discomfort in the right arm. The reason is that most of the time, discomfort in one part of the body also causes movements in other parts besides the concerned body part.

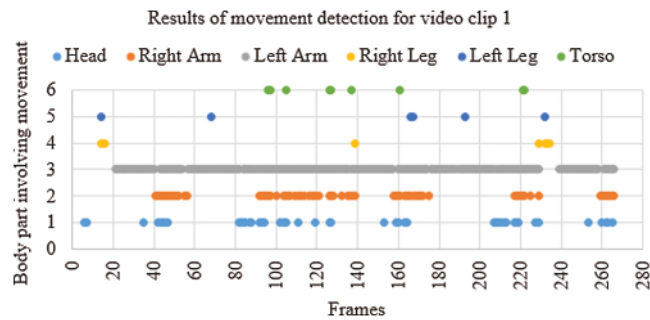


Figure 6: Movement detection in video 1

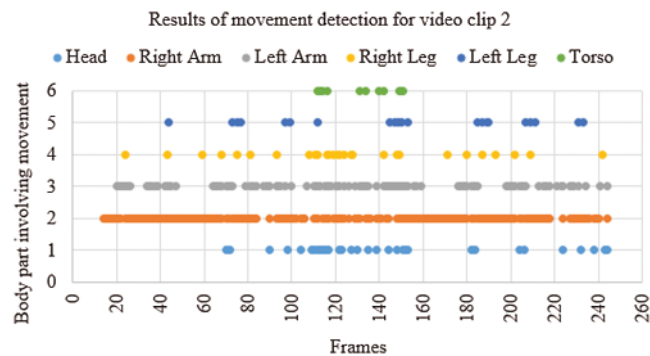


Figure 7: Movement detection in video 2

Video 3 contains movement in the patient’s right leg almost continuously throughout the video with the exception of a few frames gaps. The movement in the right leg is accompanied by movement in the right arm in most of the frames. The patient has also moved his head and left arm, but the movement in the head is a bit more frequent, as depicted in Fig. 8. Ten or more consecutive frames involving movement in the right leg are 3–30, 62–82, 84–102, 134–149, and 158–178, 180–209, and 241–255. This situation can be classified as a discomfort in the right leg. In video 4 the patient moved both his arms frequently throughout the video, but movements in the right arm are more substantial and last for a longer duration, as is clear from Fig. 9. Here, consecutive frames with movements in the right arm include 2–32, 62–78, 81–105, 120–139, 155–191, and 197–213, 216–240 and 255–274. Movements in the left arm also occur almost parallel to those in the right arm in most of the frames. The patient has also moved his head and right leg

in some frames. The frequent movements in the right arm help reach the conclusion that in this video, movements in both arms caused discomfort in the right arm of the patient.

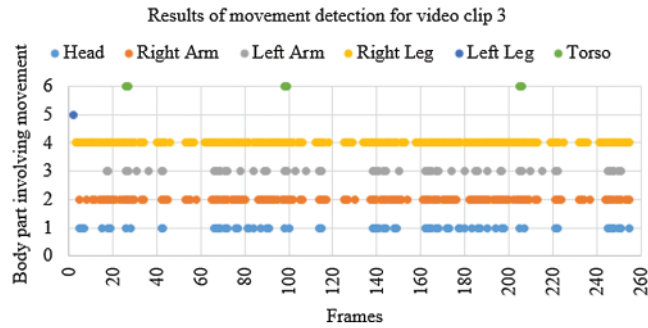


Figure 8: Movement detection in video 3

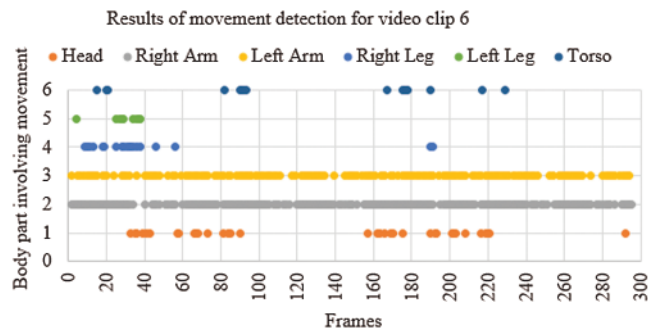


Figure 9: Movement detection in video 4

Video 5, on the other hand, comprises of two patients—the first patient is lying on bed 1 (left side), while the second patient is lying on bed 2 (right side). The results of movement in various organs of both patients are presented in Figs. 10 and 11, respectively. Patient in bed 1 has largely moved his head and both arms, particularly in frames 38–77 and 103–244. All these movements satisfies the time-based threshold hence intimating that the patient feels some pain in his body. On the other hand, the patient lying in bed 2 also moved various parts of his body. Fig. 12 also shows that for patient 2 most of the frames containing a change in various body parts, although the frequency of movement is less than the defined threshold, which shows that the movement is normal.

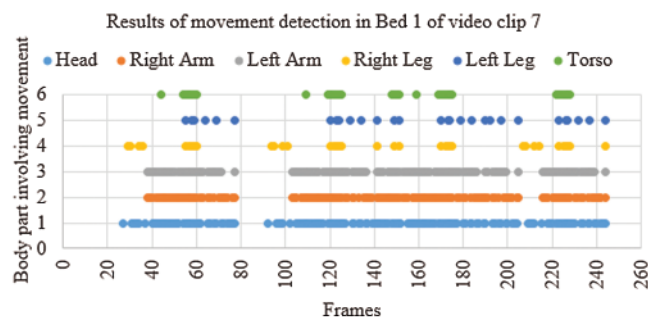


Figure 10: Movement detection in video 5 Bed 1

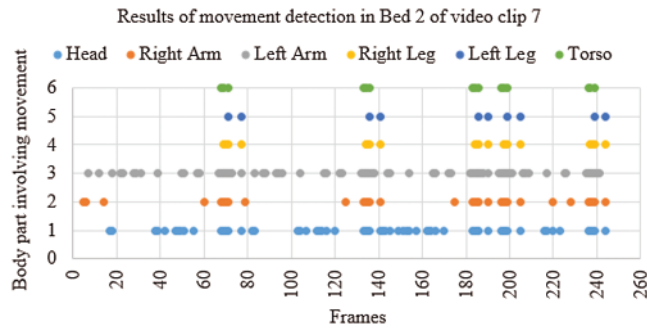


Figure 11: Movement detection in video 5 Bed 2

The evaluation of the proposed system is made for which ground truth is labeled manually for each of the video clips, whereby each frame of the video was inspected for the (x, y) coordinates of detected key points. To measure movement in a particular key point, Euclidean distance has been calculated between the coordinates of the same point in successive frames. Finally, for knowing about which body part was moved, the quantified movements in all the key points associated with the organ of the body of patients were examined against the threshold. The results produced by the system for each video clip is compared to those in the ground truth. The confusion matrices and the derived performance measures have been measured as follows:

- TP: Movement occurs in a particular organ, and the method also detects it.
- TN: Movement does not occur in a particular organ, and the method also does not detect it.
- FP: Movement does not occur in a particular organ, but the method detects it.
- FN: Movement occurs in a particular patient, but the method does not detect it.

Various performance measures like accuracy, True Positive Rate (TPR), False Positive Rate (FPR), True Negative Rate (TNR), and Miss-Classification Rate (MCR) are measured from a confusion matrix. The TPR and FPR for each video clip and of each body organ is presented in [Fig. 12](#).

It can be observed that the system reveals good results by identifying discomfort in different body organ. The TPR ranges from 98% to 99 %, while the FPR of the proposed system is between 4% to 1%. The organ-wise average performance measures are shown in [Tab. 2](#). Results show average measures in various videos for different organs of the patient's body, revealing that the proposed system achieves 98% overall average accuracy. The TPR of the proposed system is 99% with 2% of FPR.

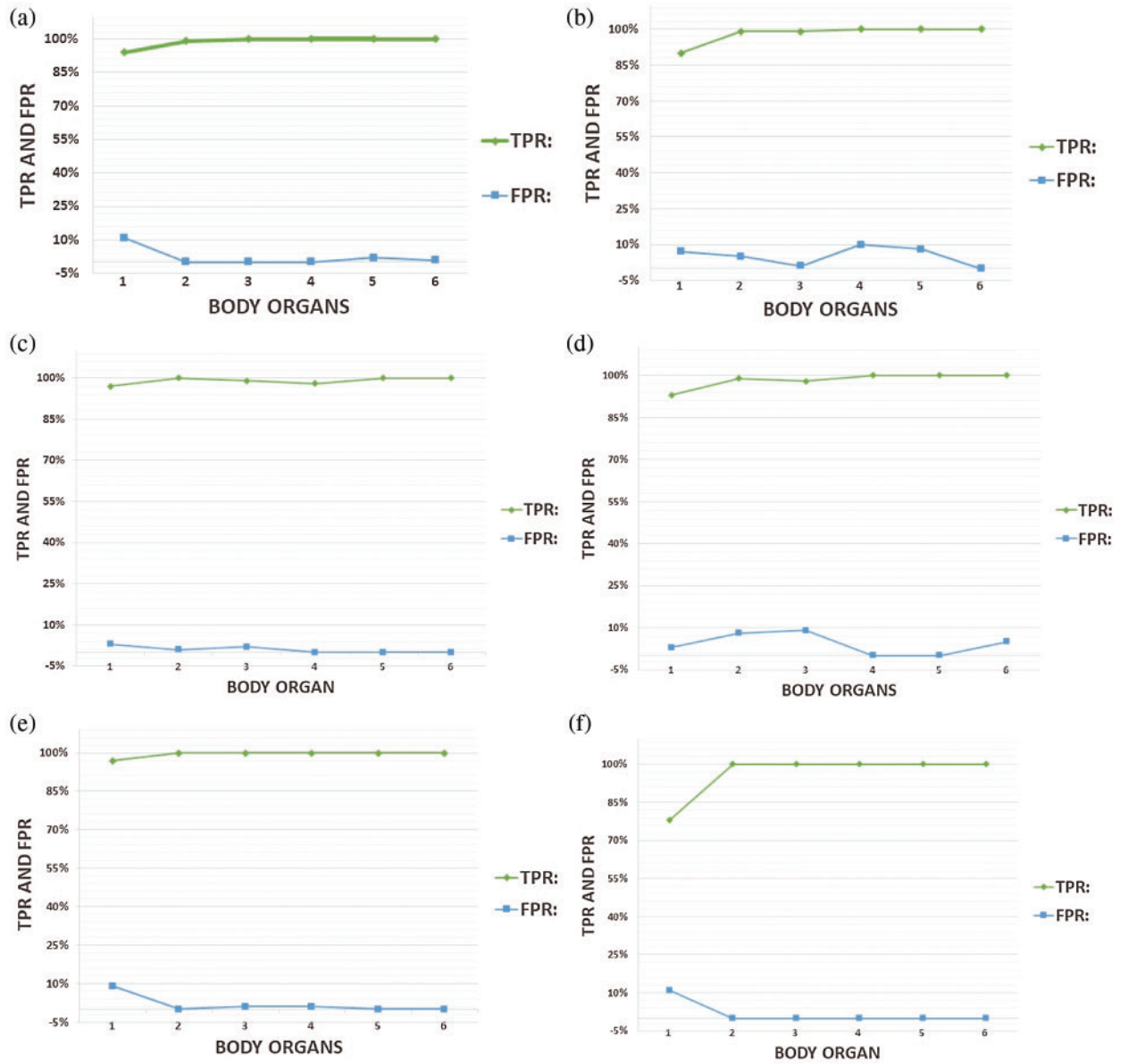


Figure 12: TPR and FPR for different videos against each body organ (1 to 6) representing Head, Right Arm, Left Arm, Right Leg, Left Leg, and Torso, respectively. The TPR and FPR are showing the performance of the proposed method. (a) TPR and FPR of Video 1 (b) TPR and FPR of Video 2 (c) TPR and FPR of Video 3 (d) TPR and FPR of Video 4 (e) TPR and FPR of Video 5 bed 1 (f) TPR and FPR of Video 5 bed 2

Table 2: Average performance measures for all parts in all videos

Performance measures	Head (%)	Right arm (%)	Left arm (%)	Right leg (%)	Left leg (%)	Torso (%)	Average (%)
Accuracy	93	99	99	98	98	99	98
MCR	7	1	1	2	2	1	2
TPR	92	100	100	100	100	100	99
FPR	6	1	2	2	3	1	2
TNR	94	99	98	98	97	99	98

5 Conclusion and Future Directions

In this work, a non-invasive system is developed for automated discomfort detection in the patient body using CNN. The proposed system contains ten layers of the CNN model, which detects key points at different body locations of patient using confidence maps. The key points information is used to form main body organs by applying association rules and part affinity fields. Next, the discomfort in the body's organs of the patient is investigated by estimating the distance between succeeding key points information of consecutive video frames. Finally, the distance and time-based thresholds are used for the classification of movement as discomfort and normal. To investigate the performance, the system is tested on a newly recorded data set. Experiments are evaluated using several performance measures, including TPR, FPR, TNR, MCR, and average accuracy. The TPR and FPR of each body organ are measured for all sequences, revealing the proposed system's robustness. The overall average TPR of the system is 98%, with average FPR of 2%.

This paper provides several future directions. First, new high-quality, pro-long overhead view data sets with multiple patients covering different types of the discomfort of different diseases in consultation with medical experts can be recorded. Second, the proposed work might be continued by recording high resolution data sets, which may capture the facial expressions of patients. This might add a second layer of discomfort detection as facial expressions will be a good way of concluding feelings and emotions. Furthermore, an interactive real-time automated detection system might be introduced for patients' discomfort in which an overhead camera will be accompanied by LEDs installed in the nursing staff room and the medical superintendent's room. The system might generate an alarm in case of the detection of discomfort. This might help the patients immediately receive the attention of the staff on duty.

Funding Statement: This research was funded by the Deanship of Scientific Research at Princess Nourah bint Abdulrahman University through the Fast-track Research Funding Program.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] Y. Cho, S. J. Julier, N. Marquardt and N. Bianchi-Berthouze, "Robust tracking of respiratory rate in high-dynamic range scenes using mobile thermal imaging," *Biomedical Optics Express*, vol. 8, no. 10, pp. 4480–4503, 2017.
- [2] C. Uysal and T. Filik, "Music algorithm for respiratory rate estimation using RF signals," *Electrica*, vol. 18, no. 2, pp. 300–309, 2018.

- [3] C. Massaroni, D. S. Lopes, D. Lo Presti, E. Schena and S. Silvestri, "Contactless monitoring of breathing patterns and respiratory rate at the pit of the neck: A single camera approach," *Journal of Sensors*, vol. 1, no. 11, pp. 1–13, 2018.
- [4] S. Ostadabbas, N. Sebkhi, M. Zhang, S. Rahim, L. J. Anderson *et al.*, "A vision-based respiration monitoring system for passive airway resistance estimation," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 9, pp. 1904–1913, 2016.
- [5] C. Y. Fang, H. H. Hsieh and S. W. Chen, "A vision-based infant respiratory frequency detection system," in *Digital Image Computing: Techniques and Applications (DICTA), 2015 Int. Conf. on Digital Image Computing: Techniques and Applications (DICTA)*, Piscataway, IEEE, pp. 1–8, 2015.
- [6] F. Al-Khalidi, R. Saatchi, H. Elphick and D. Burke, "An evaluation of thermal imaging based respiration rate monitoring in children," *American Journal of Engineering and Applied Sciences*, vol. 4, no. 4, pp. 586–597, 2011.
- [7] R. Janssen, W. Wang, A. Moço and G. de Haan, "Video-based respiration monitoring with automatic region of interest detection," *Physiological Measurement*, vol. 37, no. 1, pp. 100–114, 2015.
- [8] F. Braun, A. Lemkaddem, V. Moser, S. Dasen, O. Grossenbacher *et al.*, "Contactless respiration monitoring in real-time via a video camera," in *Embec & Nbc 2017*, Singapore, Springer, pp. 567–570, 2017.
- [9] C. Wiede, J. Richter, M. Manuel and G. Hirtz, "Remote respiration rate determination in video data-vital parameter extraction based on optical flow and principal component analysis," in *VISIGRAPP (4: VISAPP)*, SCITEPRESS, pp. 326–333, 2017.
- [10] M. Frigola, J. Amat and J. Pagès, "Vision based respiratory monitoring system," in *Proc. of the 10th Mediterranean Conf. on Control and Automation (MED 2002)*, Lisbon, Portugal, pp. 9–13, 2002.
- [11] R. A. Khan, A. Meyer, H. Konik and S. Bouakaz, "Pain detection through shape and appearance features," in *2013 IEEE Int. Conf. on Multimedia and Expo (ICME)*, Piscataway, IEEE, pp. 1–6, 2013.
- [12] I. Ahmed and A. Adnan, "A robust algorithm for detecting people in overhead views," *Cluster Computing*, vol. 21, pp. 1–22, 2017.
- [13] I. Ahmed, A. Ahmad, F. Piccialli, A. K. Sangaiah and G. Jeon, "A robust features-based person tracker for overhead views in industrial environment," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1598–1605, 2018.
- [14] I. Ahmed, M. Ahmad, A. Adnan, A. Ahmad and M. Khan, "Person detector for different overhead views using machine learning," *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 10, pp. 2657–2668, 2019.
- [15] I. Ahmed, M. Ahmad, M. Nawaz, K. Haseeb, S. Khan *et al.*, "Efficient topview person detector using point based transformation and lookup table," *Computer Communications*, vol. 147, no. 1, pp. 188–197, 2019.
- [16] M. Ahmad, I. Ahmed, F. A. Khan, F. Qayum and H. Aljuaid, "Convolutional neural network-based person tracking using overhead views," *International Journal of Distributed Sensor Networks*, vol. 16, no. 6, pp. 1550147720934738, 2020.
- [17] Z. Hammal and J. F. Cohn, "Automatic detection of pain intensity," in *Proc. of the 14th ACM Int. Conf. on Multimodal Interaction*, New York, ACM, pp. 47–52, 2012.
- [18] Z. Hammal and M. Kunz, "Pain monitoring: A dynamic and context-sensitive system," *Pattern Recognition*, vol. 45, no. 4, pp. 1265–1280, 2012.
- [19] J. F. Cohn, T. S. Kruez, I. Matthews, Y. Yang, M. H. Nguyenre *et al.*, "Detecting depression from facial actions and vocal prosody," in *3rd Int. Conf. on Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009*, Piscataway, IEEE, pp. 1–7, 2009.
- [20] M. N. Mansor, S. Yaacob, R. Nagarajan, L. S. Che, M. Hariharan *et al.*, "Detection of facial changes for ICU patients using KNN classifier," in *2010 Int. Conf. on Intelligent and Advanced Systems (ICIAS)*, Piscataway, IEEE, pp. 1–5, 2010.
- [21] G. C. Littlewort, M. S. Bartlett and K. Lee, "Faces of pain: Automated measurement of spontaneous all facial expressions of genuine and posed pain," in *Proc. of the 9th Int. Conf. on Multimodal Interfaces*, New York, ACM, pp. 15–21, 2007.

- [22] A. B. Ashraf, S. Lucey, J. F. Cohn, T. Chen, Z. Ambadar *et al.*, “The painful face-pain expression recognition using active appearance models,” *Image and Vision Computing*, vol. 27, no. 12, pp. 1788–1796, 2009.
- [23] P. Lucey, J. Cohn, S. Lucey, I. Matthews, S. Sridharan *et al.*, “Automatically detecting pain using facial actions,” in *3rd Int. Conf. on Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009*, Piscataway, IEEE, pp. 1–8, 2009.
- [24] L. Nanni, S. Brahnma and A. Lumini, “A local approach based on a local binary patterns variant texture descriptor for classifying pain states,” *Expert Systems with Applications*, vol. 37, no. 12, pp. 7888–7894, 2010.
- [25] P. Werner, A. Al-Hamadi and R. Niese, “Pain recognition and intensity rating based on comparative learning,” in *19th IEEE Int. Conf. in Image Processing (ICIP)*, Piscataway, IEEE, pp. 2313–2316, 2012.
- [26] P. Lucey, J. F. Cohn, K. M. Prkachin, P. E. Solomon and I. Matthews, “Painful data: The unbc-mcmaster shoulder pain expression archive database,” in *2011 IEEE Int. Conf. on Automatic Face & Gesture Recognition (FG)*, Piscataway, IEEE, pp. 57–64, 2011.
- [27] B. J. Matuszewski, W. Quan, L. K. Shark, A. S. Mcloughlin, C. E. Lightbody *et al.*, “Hi4d-adsip 3-d dynamic facial articulation database,” *Image and Vision Computing*, vol. 30, no. 10, pp. 713–727, 2012.
- [28] N. Shafi, F. Bukhari, W. Iqbal, K. M. Almustafa, M. Asif *et al.*, “Cleft prediction before birth using deep neural network,” *Health Informatics Journal*, vol. 26, no. 4, pp. 2568–2585, 2020.
- [29] S. Liao and A. C. Chung, “Face recognition by using elongated local binary patterns with average maximum distance gradient magnitude,” in *Asian Conf. on Computer Vision*, Berlin, Heidelberg, Springer, pp. 672–679, 2007.
- [30] A. Al-Naji, K. Gibson, S. -H. Lee and J. Chahl, “Real time apnoea monitoring of children using the microsoft kinect sensor: A pilot study,” *Sensors*, vol. 17, no. 2, pp. 286, 2017.
- [31] M. H. Li, A. Yadollahi and B. Taati, “Noncontact vision-based cardiopulmonary monitoring in different sleeping positions,” *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 5, pp. 1367–1375, 2017.
- [32] V. Metsis, D. Kosmopoulos, V. Athitsos and F. Makedon, “Non-invasive analysis of sleep patterns via multimodal sensor input,” *Personal and Ubiquitous Computing*, vol. 18, no. 1, pp. 19–26, 2014.
- [33] K. Malakuti and A. B. Albu, “Towards an intelligent bed sensor: Non-intrusive monitoring of sleep irregularities with computer vision techniques,” in *20th Int. Conf. in Pattern Recognition (ICPR)*, Piscataway, IEEE, pp. 4004–4007, 2010.
- [34] W. H. Liao and C. M. Yang, “Video-based activity and movement pattern analysis in overnight sleep studies,” in *19th Int. Conf. on Pattern Recognition, 2008. ICPR 2008*, Piscataway, IEEE, pp. 1–4, 2008.
- [35] A. F. Bobick and J. W. Davis, “The recognition of human movement using temporal templates,” *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 23, no. 3, pp. 257–267, 2001.
- [36] R. Nandakumar, S. Gollakota and N. Watson, “Contactless sleep apnea detection on smartphones,” in *Proc. of the 13th Annual Int. Conf. on Mobile Systems, Applications, and Services*, New York, ACM, pp. 45–57, 2015.
- [37] W. Saad, C. Khoo, S. Ab Rahman, M. Ibrahim and N. Saad, “Development of sleep monitoring system for observing the effect of the room ambient toward the quality of sleep,” *IOP Conf. Series: Materials Science and Engineering*, vol. 210, pp. 012050, 2017.
- [38] E. Hoque, R. F. Dickerson and J. A. Stankovic, “Monitoring body positions and movements during sleep using wisps,” in *Wireless Health*. New York: ACM, pp. 44–53, 2010.
- [39] A. P. Sample, D. J. Yeager, P. S. Powledge, A. V. Mamishev and J. R. Smith, “Design of an RFID-based battery free programmable sensing platform,” *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 11, pp. 2608–2615, 2008.
- [40] P. V. K. Borges and N. Nourani-Vatani, “Vision-based detection of unusual patient activity,” in *HIC*, IOS Press, pp. 16–23, 2011.
- [41] P. Kittipanya-Ngam, O. S. Guat and E. H. Lung, “Computer vision applications for patients monitoring system,” in *2012 15th Int. Conf. on Information Fusion (FUSION)*, Piscataway, IEEE, pp. 2201–2208, 2012.

- [42] M. Martinez and R. Stiefelhagen, "Automated multi-camera system for long term behavioral monitoring in intensive care units," in *MVA*, pp. 97–100, 2013.
- [43] S. Sathyanarayana, R. K. Satzoda, S. Sathyanarayana and S. Thambipillai, "Vision-based patient monitoring: A comprehensive review of algorithms and technologies," *Journal of Ambient Intelligence and Humanized Computing*, vol. 9, no. 2, pp. 225–251, 2018.
- [44] M. C. Chang, T. Yi, K. Duan, J. Luo, P. Tu *et al.*, "In-bed patient motion and pose analysis using depth videos for pressure ulcer prevention," in *IEEE Int. Conf. on Image Processing (ICIP)*, Piscataway, IEEE, pp. 4118–4122, 2017.
- [45] S. Liu and S. Ostadabbas, "A vision-based system for in-bed posture tracking," in *2017 IEEE Int. Conf. on Computer Vision Workshops (ICCVW)*, Piscataway, IEEE, pp. 1373–1382, 2017.
- [46] C.-W. Wang, A. Hunter, N. Gravill and S. Matusiewicz, "Real time pose recognition of covered human for diagnosis of sleep apnoea," *Computerized Medical Imaging and Graphics*, vol. 34, no. 6, pp. 523–533, 2010.
- [47] C. -W. Wang, A. Ahmed and A. Hunter, "Locating the upper body of covered humans in application to diagnosis of obstructive sleep apnea," in *World Congress on Engineering*, pp. 662–667, 2007.
- [48] M. Ahmad, I. Ahmed, K. Ullah, I. Khan, A. Khattak *et al.*, "Person detection from overhead view: A survey," *International Journal of Advanced Computer Science & Applications*, vol. 10, no. 4, pp. 567–577, 2019.
- [49] M. Ahmad, I. Ahmed Misbah and A. Adnan, "Overhead view person detection using YOLO," in *2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, Piscataway, IEEE, pp. 0627–0633, 2019.
- [50] M. Ahmad, I. Ahmed and G. Jeon, "An IoT-enabled real-time overhead view person detection system based on Cascade-RCNN and transfer learning," *Journal of Real-Time Image Processing*, vol. 1, no. 1, pp. 1–11, 2021.
- [51] D. Brulin, Y. Benezeth and E. Courtial, "Posture recognition based on fuzzy logic for home monitoring of the elderly," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 5, pp. 974–982, 2012.
- [52] S. Sathyanarayana, R. K. Satzoda, S. Sathyanarayana and S. Thambipillai, "Identifying epileptic seizures based on a template-based eyeball detection technique," in *2015 IEEE Int. Conf. on Image Processing (ICIP)*, Piscataway, IEEE, pp. 4689–4693, 2015.
- [53] H. Lu, Y. Pan, B. Mandal, H. L. Eng, C. Guan *et al.*, "Quantifying limb movements in epileptic seizures through color-based video analysis," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 2, pp. 461–469, 2013.
- [54] K. Cuppens, L. Lagae and B. Vanrumste, "Towards automatic detection of movement during sleep in pediatric patients with epilepsy by means of video recordings and the optical flow algorithm," in *4th European Conf. of the Int. Federation for Medical and Biological Engineering*, Berlin, Heidelberg, Springer, pp. 784–789, 2009.
- [55] S. Kalitzin, G. Petkov, D. Velis, B. Vledder and F. L. da Silva, "Automatic segmentation of episodes containing epileptic clonic seizures in video sequences," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 12, pp. 3379–3385, 2012.
- [56] I. Ahmed, M. Ahmad, J. J. P. C. Rodrigues and G. Jeon, "Edge computing-based person detection system for top view surveillance: Using CenterNet with transfer learning," *Applied Soft Computing*, vol. 107, no. 1, pp. 107489, 2021.
- [57] M. Ahmad, I. Ahmed, K. Ullah and A. Ahmad, "A deep neural network approach for top view people detection and counting," in *2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conf. (UEMCON)*, pp. 1082–1088, 2019.
- [58] I. Ahmed, S. Din, G. Jeon, F. Piccialli and G. Fortino, "Towards collaborative robotics in top view surveillance: A framework for multiple object tracking by detection using deep learning," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 7, pp. 1253–1270, 2020.

- [59] I. Ahmed and G. Jeon, "A real-time person tracking system based on SiamMask network for intelligent video surveillance," *Journal of Real-Time Image Processing*, vol. 1, no. 1, pp. 1–12, 2021.
- [60] I. Ahmed, G. Jeon and F. Piccialli, "A deep learning-based smart healthcare system for patient's discomfort detection at the edge of Internet of Things," *IEEE Internet of Things Journal*, vol. 8, no. 13, pp. 10318–10326, 2021.
- [61] Z. Cao, T. Simon, S. -E. Wei and Y. Sheikh, "Real-time multi-person 2d pose estimation using part affinity fields," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 7291–7299, 2017.
- [62] S.-E. Wei, V. Ramakrishna, T. Kanade and Y. Sheikh, "Convolutional pose machines," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 4724–4732, 2016.