Development of a Smart Hospital Bed Based on Deep Learning to Monitor Patient Conditions

Sarra Ayouni¹, Mohamed Maddeh², Shaha Al-Otaibi¹, Malik Bader Alazzam³*, Nazik Alturki¹ and Fahima Hajjej¹

¹Department of Information Systems, College of Computer and Information Sciences, Princess Nourah Bint Abdul Rahman University, Riyadh 11671, Saudi Arabia
²Department of Information Systems, College of Applied Computer Science, King Saud University, Riyadh 11451, Saudi Arabia
³Information Technology Department, Information Technology College, Ajloun National Private University, Ajloun, Jordan

Correspondence to:
Malik Bader Alazzam*, e-mail: malikbader2@gmail.com, Tel.: 0096779358694

Received: May 21 2023; Revised: June 22 2023; Accepted: June 23 2023; Published Online: July 19 2023

ABSTRACT
An Internet of Things-based automated patient condition monitoring and detection system is discussed and built in this work. The proposed algorithm that underpins the smart-bed system is based on deep learning. The movement and posture of the patient’s body may be determined with the help of wearable sensor-based devices. In this work, an internet protocol camera device is used for monitoring the smart bed, and sensor data from five key points of the smart bed are core components of our approach. The Mask Region Convolutional Neural Network approach is used to extract data from many important areas from the body of the patient by collecting data from sensors. The distance and the time threshold are used to identify motions as being either connected with normal circumstances or uncomfortable ones. The information from these key locations is also utilised to establish the postures in which the patient is lying in while they are being treated on the bed. The patient’s body motion and bodily expression are constantly monitored for any discomfort if present. The results of the experiments demonstrate that the suggested system is valuable since it achieves a true-positive rate of 95% while only yielding a false-positive rate of 4%.

KEYWORDS
smart bed, deep learning, convolutional neural network, regression loss, region of interest, overfitting, threshold component evaluation

INTRODUCTION
When integrated with computer vision, machine learning, and deep learning methods, the Internet of Things (IoT) delivers rapid and precise monitoring of patients and detection systems. This is one of the many ways in which the IoT is helping to advance the area of smart healthcare (Pace et al., 2018). A few of the traditional patient monitoring systems often include approaches based on sensors that the patients wear on their body. The challenging and time-consuming task of providing the best medical care is a tough one. Patients are often checked in when they are experiencing significant health issues. They could possibly enter for conditions that are not life-threatening but do need frequent monitoring. Also, numerous therapy practices may not be administered to a patient until after he or she has been admitted to the hospital in a private room at a medical facility for a certain amount of time. Approximately 40% of these newly admitted patients are members of a population that is mostly old, with an average age of >65 years (Mattison et al., 2013). Additionally, there are those who are disabled who are admitted to hospitals (Carpenter et al., 2007). The activities and relationships between hospitalised patients and the community in which they are housed are becoming prohibitive because of major health concerns, factors such as age, infirmity, and pre-established medical directives. Several recent studies concluded that the healthcare system sector suffers from a serious lack of available personnel, and healthcare experts carry a heavy burden of labour during the whole process (Akter et al., 2019; Qureshi et al., 2019). This contributes to a further decline in the overall quality of medical attention for people who are ill (Farid et al., 2020). Additionally, research demonstrates that the healthcare industry is under a significant amount of strain that is already developed, in addition to those that are still developing (Akter et al., 2019). To reiterate, the situation is the way it is because the widespread COVID-19 epidemic has reached a new level of decline (McGarry et al., 2020; Sterpetti, 2020). However, developing nations such as Finland, Canada, England, Ireland, and New Zealand are examples. Both New
Zealand and Japan have put a lot of effort into developing innovative approaches to patient care with the interest of enhancing the standard of medical treatment (Reinhold et al., 2019). However, despite the fact that scholars have concentrated on building assistance programmes for patients who are hospitalised, bedridden, and individuals who are unable to move about and have limited ability for engagement (for example, monitoring systems for patients based on brainwave activity interface systems, namely interaction systems based on hand gestures, etc.) (Gao et al., 2003; Raheja et al., 2014; Saha et al., 2018), just a little amount of study has been carried out thus far to determine the needs of patients in this category and then come up with a plan to meet their specific needs. To mitigate the effects of this emergency on the healthcare system, the necessities that a patient in a hospital smart bed is expected to have are clearly perceived and well comprehended. A viable solution may be developed as well as created to assist the hospitalised patients alongside the professionals in the healthcare industry.

Recent research suggests that systems based on deep learning may improve the effectiveness, precision, and adaptability of healthcare systems (LeCun et al., 2015; Shrestha and Mahmood, 2019). Several systems created for use in the medical industry have been put into use by using technology based on deep learning (for instance, medical image analysis, illness prognostication, and other similar topics) (Shen et al., 2017; Bakator and Radosav, 2018). Recently, researchers have focused on utilising deep learning for smart beds. Consequently, the goals of this work are as follows:

1. To gain an understanding of the smart-bed control mechanism for patients who are bedridden based on the graphical user interface (GUI), Node microcontroller unit (MCU), and 12V 6-channel relay.
2. To design and construct the control system for the smart-bed system in accordance with the five key points targeted, which correspond to the heart rate, blood pressure, body temperature, motion detector, and bed occupancy for patient body monitoring and an internet protocol (IP) camera device for monitoring the smart bed by using sensory data and deep learning technologies.
3. To analyse the system using the proposed deep learning-based model through empirical research and analysis.

According to the research that was carried out, it was found that quite a few different studies have been conducted to help the healthcare industry. In addition, most of the interface systems created in earlier studies concentrated on implementing a single mode of operation without considering the patients in their various states of varying degrees throughout the range of the impairment. It was discovered that just a few patient care systems had been established using various forms of deep learning technology. Because of this, the study focuses on developing a solution for the smart-bed system as well as the difficulties confronted by its challenges.

This paper develops an IoT-based automated non-invasive patient detection and monitoring system using an algorithm based on deep learning by using line-of-sight cameras and cumbersome wearable sensing devices. The suggested system makes use of a trained and pre-trained Mask Region Convolutional Neural Network (Mask-RCNN) model for most of its detection work, particularly for key points that are identified and connected to a particular sensor.

The remainder of the paper is laid out as follows: this research paper has five sections: Introduction, Literature Review, Proposed Technique, Experimental Findings and Discussion, and Conclusion.

**LITERATURE REVIEW**

The underlying foundation for the algorithms that are employed in this system is deep learning, and hence the effectiveness and precision of the functioning of the system in real time have been improved. There have been several different digital systems created to provide medical treatment to those who are unable to care for themselves due to illness or disability and those in advanced years (Lakkis and Elshakankiri, 2017; Khan et al., 2018; Hossain et al., 2019; Hasan et al., 2021; Islam et al., 2021). In Kanase and Gaikwad (2016), the authors suggested using a cloud-based approach that uses a variety of sensory data. To get information from the patient’s room, sensors were used. To exhibit this, both an application and a website were built, and readings from the sensors and functions for regulating the patient’s environment were used by the relevant medical professionals, nurses, and staff members.

In addition, the research demonstrated how the various types of data that are essential may be transported from the perspective of the sick to that of the carer. But in this planned system, the patient was not given any control over what was happening to them.

Some system proposals were made by Khan et al. (2018), Saha et al. (2018), and Kamruzzaman (2020) for the monitoring and care of patients during emergency alarm generation. In Aadeeb et al. (2020), authors have constructed a brain-computer interface (BCI) and Deep management system for hospital patient rooms that is based on assisting individuals who are immobile, ill, or incapacitated in exercising control over their immediate environment.

Tam et al. (2019) created a controller that is based on hand gestures. The operator was responsible for the collection of electromyographic sensors that assisted in the detection of muscular motions. Sensors for electromyography may be found here and were utilised as features for training a convolutional neural network (CNN).

There are some alternative methods for computer vision and face feature analysis for individual patients. In Khan et al. (2017), the authors developed and conducted tests on a control system for people who are unable to use their hands because of their condition.

The developments that have been made in the IoT in recent years have made it possible for us to create healthcare systems that are more intelligent and predictable. It is necessary to have a gateway (also known as a bridge point) to link the internet and the sensor network architecture in most of the technology based on the IoT in the medical field, notably in smart hospitals. The gateway is often located on the periphery of the network and is responsible for performing critical tasks utilised by the internet and networks of sensors (Alam et al., 2017; Bakator and Radosav, 2018). Recently, research based on deep learning (for instance, medical image analysis, illness prognostication, and other similar topics) (Shen et al., 2017).
et al., 2019; Chen et al., 2019). These gateways provide for streamlined management of the sensor network and access to relevant data over the internet, which is the medium via which data are being transferred (Casadei et al., 2019).

The IoT is well positioned to play an important role in trends in healthcare due to technology on a variety of different levels. Unlocking the potential of the IoT to improve healthcare facilities is made possible when a hospital is equipped with a real-time smart healthcare system (Savaglio et al., 2020). It makes the experience better for patients, enhances the experience for carers, has a positive impact on health and clinical outcomes, and reduces costs.

The IoT-based systems, in conjunction with data mining methods (Amato et al., 2018; Piccialli et al., 2019), monitor and synthesise patient information, crosscheck these records against registered patterns, and analyse illness indications. The IoT delivers patient monitoring solutions that are efficient and quick, playing a crucial role in the process of remote patient monitoring in the absence of a medical personnel. Currently, it offers a supporting monitoring system that makes use of non-intrusive equipment (IoT-based sensors or cameras), in addition to computer vision and deep learning algorithms (Piccialli et al., 2020; Qureshi et al., 2020; Piccialli et al., 2021). Monitoring patients becomes more effective when it is based on the ability to detect their symptoms based on visual signals, such as how they stand, how they move, and how they seem. In recent years, scientists have experimented with systems for tracking patients that include specialist equipment, pressure beds, and sensors (Birku and Agrawal, 2018), but this has resulted in incurring extra costs. According to the available research, most of the developed approaches make use of several different camera devices or sensors for monitoring patients.

In the published study, most researchers investigated a variety of sensor-based approaches to fall detection. In addition, several scholars resorted to relying on vision in their investigation of fall detection (Birku and Agrawal, 2018).

Numerous researchers, such as Merrouche and Baha (2016), extract visual information from scenes that only include a single room by using numerous camera systems. Researchers monitored the pace, depth, and consistency of patients’ breathing using sensors and information based on signals. The analysis was based on the inhalation and exhalation duration and ratio, for example Sathyanarayana et al. (2018). Very few of them created vision-based systems that used information about patients’ facial expressions to assess pain, such as Jan et al. (2018), which requires the patients’ faces to be aligned directly with the camera. Monitoring measures for respiration or breathing include keeping an eye on things like the patient’s breathing rate, stability, and depth, as well as their exhalation/inhalation time ratio and any indications of sleep apnoea (Uysal and Filik, 2018). In Dhillon (2017), a method was proposed that is responsible for the development of a monitoring and warning system for the diagnosis of epileptic seizures. Researchers from many institutions looked at a patient’s behaviour and utilised the data they gathered to conduct an analysis of the patient’s medical problems (Sathyanarayana et al., 2018). In Liu and Ostadabbas (2017), the authors suggested an in-bed patient posture monitoring system that does not involve any intrusive procedures. Techniques for posture-based monitoring using many cameras have been researched by Liu and Ostadabbas (2017), but they mostly concentrate on the patient’s upper body region. In Deng et al. (2018), the authors presented the idea of computerised eye tracking for sleep analysis that makes use of motion sensors and infrared cameras. In the published study, most researchers concentrated their attention on a single patient and/or bed, in addition to certain contact-based devices. For these methods to work, it is necessary to connect sensors to patient bodies or beds, which is not only inconvenient but also expensive to do and often something the patient does not want. Several studies included visual methods, although most of the time different types of cameras were used for monitoring breathing and detecting falls. Several researchers have created techniques that are based on patients’ facial expressions; however, these techniques require the patient to have his face oriented squarely towards the camera.

This research work develops an automated system for patient monitoring or detection. This system makes use of specialised hardware, sensors, or line-of-sight cameras in order to monitor patient conditions. The suggested method identified the patient body’s important spots using a model that was based on deep learning. This model was known as Mask-RCNN. The information from the key sites is then utilised for further investigation of the data from the smart bed, which corresponds to the heart rate, blood pressure, body temperature, motion detector, and bed occupancy. As a metric of performance, we use the average distance between two successive critical points that have been identified.

In this study, in contrast to earlier similar efforts, a technique based on deep learning is developed to classify and investigate the IP-based top-view camera for the smart bed and five different key points of the patient’s body connected to sensors.

DEEP LEARNING-BASED SMART-BED MONITORING SYSTEM

Proposed methodology

Deep learning, an artificial intelligence technique, uses deep neural networks to analyse medical data and provide accurate diagnoses and prognoses for patients. Deep learning works in this domain; image analysis: deep neural networks analyse magnetic resonance imaging (MRIs), X-rays, and histopathological samples. The deep model recognises complex characteristics and illness patterns to help diagnose cancer and neurological ailments early. Prediction and prognosis: deep learning processes massive amounts of patient data, including medical records, clinical reports, and diagnostic findings (Arshad et al., 2023). These data help the deep model anticipate illness progression and therapy responses. Predictions may enhance treatment planning and identify outcome-affecting factors. Deep learning may uncover subtle medical data patterns and changes that healthcare practitioners overlook. It may identify health issues and disease development by finding complicated data linkages (Islam et al., 2022).
Deep learning might improve medical diagnostics, especially in terms of accuracy. Medical data may teach advanced models disease patterns and indicators. This improves diagnosis and treatment. Deep learning may anticipate and diagnose illnesses. Advanced models may use prior data to predict patient response to therapy and sickness progression. This dataset might improve and customise treatment regimens. Deep learning reduces medical mistakes, improving patient safety. Deep models may identify problems and provide solutions by studying data and risk. Deep learning algorithms may discover anomalies early, improving patient monitoring and health outcomes.

The utilisation of deep learning enables the provision of tailored treatment to individual patients. The utilisation of deep models has the potential to facilitate individualised treatment plans and therapies through the analysis of medical history, genetic information, and lifestyle factors. The implementation of a tailored approach has the potential to enhance both the physical well-being and overall satisfaction of patients (Bhardwaj et al., 2022; Sujith et al., 2022).

The potential of deep learning is substantial; however, it is imperative that it is employed in conjunction with medical expertise and clinical discernment to optimise patient care.

**Design specifications**

For the time being, we can determine each of the data sets that are currently available and each of the in-lab trials that can be found in the published research. The data set was compiled with the assistance of a top-view IP-based camera that was positioned at a height of 4 m. The images from the IP camera have been considered to have a resolution of 640 × 480. The system keeps a record of the patient’s smart-bed image using an IP camera and sensory data from five key points of the patient’s body, that is, heart rate, blood pressure, body temperature, motion detector, and bed occupancy. The findings of their investigation revealed that the detection and monitoring of smart beds can be associated with a GUI to assist in making healthcare better. A personal computer camera was attached to the smart bed of the patient. Then, the images are displayed on a computer for further monitoring.

**The proposed Internet of Things-based system**

An IoT-deep learning-based smart-bed system has been proposed in this study. The system uses a deep learning algorithm to monitor patients’ activity. An IP-based camera is used to capture images from the top view of the patient’s smart bed, and a deep learning-based model is employed to recognise information about the patient’s key points connected to body sensors, which correspond to the heart rate, blood pressure, body temperature, motion detector, and bed occupancy. Together, these two components make up the total system. The proposed system’s block diagram is shown in Figure 1. The workflow of the proposed model with hardware specifications is shown in Figure 2.

**Development of the CNN model**

The preparation of the dataset was the first step that needed to be taken before the CNN models could be trained. This required several stages, such as the gathering of picture data, the tagging of image data, the resizing of images, the cropping of images, the scaling of images, and several other procedures. After that, the picture data were split into two groups, one of which will be used to train the CNN models and the other to verify them. In addition, the pictures were shuffled around and divided up into batches before being used for training the CNNs. The Results section provides a more in-depth analysis of the final picture collection that can be credited to the previously indicated actions and procedures. There is one input layer (Wang, 2003), several convolutional layers (Albawi et al., 2017), and many pooling layers (Sun et al., 2017) that together form the classification system that is the CNN. The input layer is where a picture is fed into the CNN before it is processed further. This layer is responsible for providing an input feature map to the succeeding CNN layers so that the picture may be processed (Wu and Lin, 2018). With a three-dimensional feature map as the input, the convolutional layer applies the convolution operation (Ludwig, 2013) defined by Eq. 1. F stands for the feature map, I for the convolution kernel, x and y for the location on the feature map where the kernel’s centre lies, and i and j for the surrounding coordinates.

\[
F_{i}I(x, y) = \sum_{j-x}^{j} \sum_{i-y}^{i} F(i, j)I(x+i, y+j). 
\]  

By using the pooling layer, shrinkage of a feature map’s total width is possible. Each feature map that is supplied to it is processed by a single kernel. Because of its ability to reduce the data size, feature maps may likewise be converted using the pooling layer into one-dimensional feature maps.

The input parameter x for the dense layer is a one-dimensional feature map, and it multiplies that parameter by the weight values w of the layer. In addition to this, the dense layer is having a bias value denoted by b, which is applied to the result generated when weights are multiplied together such that \( y = wx^2 + b \). The dense layer will conduct the multiplication of weight, as well as the addition of bias, with the prescribed feature inputs. An activation function (Sharma et al., 2017) is included in each layer of a CNN. This function is responsible for modifying the output that is produced by the CNN layer. The SoftMax equation determines the parameters for the SoftMax type of activation function, which is utilised in CNNs to analyse feature maps from convolutional layers, and is written as

\[
y = \max(x, 0) 
\]  

\[
\sigma(x_i) = \frac{e^{x_i}}{\sum_{j} e^{x_j}}. 
\]  

An activation function (the rectified linear unit activation function and the SoftMax type of activation function) with the parameters has been implemented in the final defence layer for the purposes of classification. The CNNs that are
The proposed Mask-RCNN model

The images from the IP camera are sent to the monitoring model, which is based on deep learning, so that they may be further processed. Mask-RCNN (Islam et al., 2021) is performed to monitor and identify any signs of patient discomfort, as shown in Figure 1. The Mask-RCNN model receives the images that are being inputted into it. The model can identify the patient as well as critical places on their bodies by using sensors. The identified critical spots are then used for further investigation. Figure 3 presents the block diagram for the proposed Mask-RCNN network for predicting patient conditions.

The fundamental design is additionally expanded so that human posture may be estimated by leveraging information from critical spots (He et al., 2017). The process consists of two phases. In the first step, a CNN is used to extract a features map from the input image using the region proposal network (Lin et al., 2017). CNN’s extracted candidates
for the bounding box choose a spot on the features map at random. It is possible to extract bounding box candidates of varying sizes; hence, in the second step, a layer known as RoI (Region of Interest) Align is used to reduce the dimension of the extracted features to ensure that their sizes are all comparable to one another, as shown in Figure 3. The collected characteristics are then supplied to the parallel branches of CNNs to make predictions about the segmentation masks and the bounding boxes. The loss function for the Mask-RCNN is arrived at by adding together the results in $L = L_{cls} + L_{loc} + L_{mask}$. Here, the symbol $L_{cls}$ stands for classification loss (He et al., 2017), while the symbol $L_{loc}$ denotes regression loss for the observed bounding box. The formula for determining the $L_{cls}$ of a RoI is as follows: $L_{cls}(u, p) = -\log p$. In the equation $u$ stands for the actual class of the item, and $p = (p_0, ..., p_k)$ displays the projected probability distribution over $k + 1$ categories of objects. The equation for determining the regression loss inside the box’s confines $L_{loc}$ for RoI is as follows: $L_{loc}(t^*, v) = \sum_{x \in [v_{w}, v_{h}, v_{x}, v_{y}]} smooth L_{1}(t^* - v)$ in which $v = (v_{w}, v_{h}, v_{x}, v_{y})$ represents the true RoI regression inside a box and $t^* = (t^*_{w}, t^*_{h}, t^*_{x}, t^*_{y})$ represents the anticipated regression inside a class’s bounding boundary $u$. In the third equation, the formula for smooth $L_1$ is written as in Ren et al. (2015), and it is as follows: $smooth L_1 = \begin{cases} 0.5x^2, & \text{if } |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases}$. The loss function for binary data is the value of $L_{mask}$ and is found by calculating the average loss in mask branch binary cross-entropy. The output of this mask branch is the value $K^u$ for each RoI, and $K$ stands for a binary mask, with a resolution of $m \times n$ and $K$.
denotes the number of classes. The $L_{mask}$ is determined as He et al. (2017) which is written as follows:

$$L_{mask} = \frac{1}{m^2} \sum_y (Q_y \log P_y^u + (1 - Q_y \log(1 - P_y^u))).$$  \hspace{1cm} (3)$$

In this equation, $Q$ represents the actual mask, while $P_u$ represents the anticipated mask for the class $u$ of the RoI. These two masks are represented by the expressions $Q_u \in [0, 1]$ and $P_u \in [0, 1]$, respectively. The framework is enhanced so that facial expression estimation may take place.

Next, the values of the sensory data from identified key point is represented as a one-hot mask in the model. The model forecasts a total of $K$ masks, with one representing each of the $K$ key point types. While this is happening, the object detection algorithm is being taught to determine whether the patient’s key point is normal or needs attention. Both steps, “patient expression detection” and “key point detection,” may be completed in isolation from one another. The pre-trained model that is shown in Figure 3 initially locates and segments patients in the input picture. In addition to this, the model can identify five distinct important data records from the patient’s body.

The information about the patient’s expression is utilised, and other sensory data are used to determine whether the patient is experiencing any discomfort and needs some assistance.

The patient’s motion is characterised by such a lack of control and a high frequency that it is regarded as a marker of pain. These motions and the detection of pain may be different for different illnesses, as other aspects may be linked to them. In this context, it is presumed that patients are experiencing standard discomfort. It is thought that there is discomfort present when there is a greater frequency of unknown situations and frequent moves over a longer period. The continuous motions of a specific body part are used as the basis for conducting an analysis of pain in that body part of the patient.

This technique monitors the movement associated with a motion detection sensor over time by utilising key point

---

**Figure 3:** Block diagram for the proposed Mask-RCNN network in predicting patient conditions. Abbreviation: Mask-RCNN, Mask Region Convolutional Neural Network.
coordinates \((x, y)\), and it determines if the patient’s state is normal or whether he is experiencing pain. Using information about the patient’s distance from the sensor, every movement that may have happened in any organ of the patient’s body may be quantified. Calculating the distance between consecutive images of the video sequences using the Euclidean distance allows for the related key spots of the concerned organ to have their distances determined. The equation that follows has been taken into consideration to determine how to compute the distance between the \((x, y)\) coordinates of all critical locations in two subsequent images such that \(D = \sqrt{(k_{y_j} - k_{y_{j-1}})^2 + (k_{x_j} - k_{x_{j-1}})^2}\). The term \(D\) denotes the key point to pivot Euclidean distance \(k\) in subsequent images \(j\) and \(j - 1\) correspondingly in the equation that was just presented. The notation \((k_{x_j}, k_{y_j})\) and \((K_{x_{j-1}}, K_{y_{j-1}})\) are used to indicate the \((x, y)\) location of pivotal pixel coordinates \(k\) in images \(j\) and \(j - 1\) respectively. An equation is used to determine whether a threshold \(T\) should be used such that if \(T = \left\{ \begin{array}{ll} 1, & D \geq d_{\text{e}} \\ 0, & \text{otherwise} \end{array} \right.\). Regarding the picture’s pixels that make up the picture or frame, the distance between consecutive key points, also known as the movement of key points, is what the threshold value \(T\) measures. The \(d_{\text{e}}\) value for this work has been increased all the way up to 25 pixels. The distance threshold is used to assess whether there has been movement in any of the identified key points of the connected patient body key point associated with motion detection. This allows doctors to determine whether a body part has been moved. For instance, if the \((x, y)\) coordinates of any of the identified key points on a patient’s body changed would cause a change in the value of the \((x, y)\) coordinates. An equation has been used to do an analysis on the Euclidean distances for rapid movement in the patient’s body (Mov_body) to achieve this goal and is represented as follows:

\[
\text{Mov}_{\text{body}} = \left\{ \begin{array}{ll} 1, & \text{if } \mathbf{V}^{n}_{i} = 1 \\ 0, & \text{otherwise} \end{array} \right.\]

(4)

The variable \(i\) may take on any value between \(1\) and \(n\), and it represents the \(n\) component important points of the patient’s body \(B\). Finally, a time-based threshold known as \(T\) is used to assess each image for the occurrence of movement to determine whether a patient is experiencing normal sensations or some degree of discomfort. The time threshold might change depending on the size of the data collection as well as the kind of data set. The value of Cond_patient reflects the current state of the patient \(P\) which is written as follows:

\[
\text{Cond}_{\text{patient}} = \left\{ \begin{array}{ll} \text{Discomfort}, & \text{Mov}_{\text{body}} \rightarrow T_i \\ \text{Normal}, & \text{otherwise} \end{array} \right.\]

(5)

The time threshold is denoted by the symbol \(T\), and it refers to the amount of time that serves as the dividing line between motions that are considered normal and movements that are produced by any uncomfortable situation. In this piece, the movement for 10 or more consecutive images is referred to as \(T_i\).

**RESULTS AND DISCUSSION**

Examining the Mask-CNNs statistically by calculating their accuracy, precision, F1-Score, Cohen's kappa, recall, receiver operating characteristics area under the curve (ROC AUC), true-positive rate (TPR), false-negative rate (FNR), true-negative rate (TNR), and false-positive rate (FPR). This statistical analysis is done for the overall six key points, such as T1: image from smart bed, T2: heart rate, T3: blood pressure, T4: body temperature, T5: motion detector, and T6: bed occupancy, which were classified based on the prediction values as well as the label values. A few of the statistical measures are shown in Table 1 w.r.t. classification for T1, classification for T2, classification for T3, classification for T4, classification for T5, and classification for T6. In addition, the photos that were used for training were likewise pre-processed using the same procedure and statistically examined using the same metrics for the sake of comparison.

The metrics that have been presented are all considered positive metrics; hence, the performance of the model that is being assessed will be improved in direct proportion to the values of these metrics. The images used in this work went through a series of preparation steps, which included grayscaling, enlarging the image resolution, and rescaling the pixel values. After that, the photos were used by the Mask-R-CNN model, and the values of the predictions were saved in a file. The values that were predicted comprised the category of the items that were found, and the patient is classified as either “normal” or “abnormal.”

This can improve the performance of a CNN model through data augmentation; data augmentation may expand

---

**Table 1: Statistical analysis of classification using the proposed method.**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Classification for T1</th>
<th>Classification for T2</th>
<th>Classification for T3</th>
<th>Classification for T4</th>
<th>Classification for T5</th>
<th>Classification for T6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>0.971/0.963</td>
<td>0.969/0.953</td>
<td>0.971/0.953</td>
<td>0.964/0.944</td>
<td>0.986/0.956</td>
<td>0.976/0.943</td>
</tr>
<tr>
<td>Precision</td>
<td>0.966/0.953</td>
<td>0.953/0.946</td>
<td>0.963/0.941</td>
<td>0.953/0.956</td>
<td>0.963/0.944</td>
<td>0.968/0.922</td>
</tr>
<tr>
<td>Recall</td>
<td>0.987/0.973</td>
<td>0.973/0.964</td>
<td>0.952/0.931</td>
<td>0.969/0.954</td>
<td>0.958/0.923</td>
<td>0.956/0.921</td>
</tr>
<tr>
<td>F1-Score</td>
<td>0.985/0.965</td>
<td>0.963/0.911</td>
<td>0.952/0.921</td>
<td>0.974/0.955</td>
<td>0.966/0.948</td>
<td>0.955/0.934</td>
</tr>
<tr>
<td>Cohen’s Kappa</td>
<td>0.936/0.904</td>
<td>0.921/0.911</td>
<td>0.911/0.901</td>
<td>0.933/0.932</td>
<td>0.943/0.922</td>
<td>0.962/0.932</td>
</tr>
<tr>
<td>ROC_AUC</td>
<td>0.974/0.954</td>
<td>0.974/0.932</td>
<td>0.964/0.941</td>
<td>0.983/0.945</td>
<td>0.974/0.952</td>
<td>0.956/0.939</td>
</tr>
</tbody>
</table>

Abbreviation: ROC_AUC, receiver operating characteristics area under the curve.
and diversify the training dataset. The model may become more resilient and generalise to unknown data by adding random changes like rotation, scaling, and flipping to the input pictures.

**Hyperparameter tuning**

You can try multiple hyperparameter settings to obtain the best CNN model. This covers learning rate, batch size, layer number, filter size, and activation function. Grid or random search can tune hyperparameters.

**Transfer learning**

You can use pre-trained models to start CNN model training. Transfer learning uses ImageNet-trained information. Fine-tuning the pre-trained model for your task and dataset may improve performance with less training data.

**Regularisation**

Regularisation prevents overfitting and improves generalisation. Dropout and L2 regularisation may minimise model complexity and feature dependence. Optimising the CNN model can be done using multiple optimisation strategies. Adam, RMSprop, and momentum may expedite convergence and enhance training efficiency.

**Ensemble methods**

Combine predictions from many CNN models. Averaging or voting may increase model robustness and accuracy by integrating varied viewpoints from different models.

**Architecture changes**

Change the CNN model architecture. This may include adding or deleting layers, modifying the number of filters or neurons, network depth, or convolutional or pooling layer types. A model architecture that fits the job and dataset can be used. Early halting prevents overtraining. The training can be stopped when the model’s performance degrades on a validation set. This helps locate the ideal point when the model has learned meaningful patterns without overfitting the training data.

**Regular model evaluation**

Assess the model’s performance using a distinct validation or test dataset. This monitors model progress, identifies faults, and adjusts. If possible, use multiple data sources. Integrating appropriate sensor data or patient information into the model may improve its performance.

The statistical analysis was carried out to measure the effectiveness of the proposed system in terms of the standard deviation for training and testing data. The standard deviation is measured as the value deviating from the “normal” class. As a result, the usefulness of the proposed model is based on the key points considered.

The findings of the statistical analysis are reported in Table 2 for both the train dataset and the test dataset. Every step of the Mask-RCNNs has been shown to perform better on the training dataset than on the testing dataset, according to the statistical investigation. Since the metrics that were acquired for the testing datasets are not vastly different from the metrics obtained for the training datasets, we can conclude that the models did not suffer from overfitting. The suggested model had the highest value of precision on the training dataset but the lowest value of precision on the testing dataset. The suggested model got the best overall

### Table 2: Measure of the effectiveness of the proposed system in terms of standard deviation for training and testing data.

<table>
<thead>
<tr>
<th>Task</th>
<th>T1: training, testing</th>
<th>T2: training, testing</th>
<th>T3: training, testing</th>
<th>T4: training, testing</th>
<th>T51: training, testing</th>
<th>T6: training, testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>1.63, 1.14</td>
<td>2.54, 1.34</td>
<td>1.54, 1.13</td>
<td>1.58, 1.16</td>
<td>1.79, 1.36</td>
<td>1.83, 1.26</td>
</tr>
<tr>
<td>T2</td>
<td>1.72, 1.68</td>
<td>2.83, 2.41</td>
<td>1.65, 1.21</td>
<td>1.58, 1.45</td>
<td>1.63, 1.31</td>
<td>1.54, 1.21</td>
</tr>
<tr>
<td>T3</td>
<td>1.72, 1.41</td>
<td>2.54, 2.13</td>
<td>1.83, 1.01</td>
<td>1.84, 1.32</td>
<td>1.78, 1.42</td>
<td>1.56, 1.32</td>
</tr>
<tr>
<td>T4</td>
<td>1.52, 1.23</td>
<td>2.47, 2.24</td>
<td>1.92, 1.43</td>
<td>1.77, 1.29</td>
<td>1.65, 1.23</td>
<td>1.83, 1.38</td>
</tr>
<tr>
<td>T5</td>
<td>1.63, 1.21</td>
<td>2.65, 2.74</td>
<td>1.83, 1.21</td>
<td>1.56, 1.32</td>
<td>1.47, 1.27</td>
<td>1.45, 1.03</td>
</tr>
<tr>
<td>T6</td>
<td>1.45, 1.01</td>
<td>2.78, 2.81</td>
<td>1.45, 1.04</td>
<td>1.55, 1.17</td>
<td>1.66, 1.21</td>
<td>1.59, 1.15</td>
</tr>
</tbody>
</table>

### Table 3: Evaluation of task completion time using the proposed method.

<table>
<thead>
<tr>
<th>Task</th>
<th>T1: training, testing</th>
<th>T2: training, testing</th>
<th>T3: training, testing</th>
<th>T4: training, testing</th>
<th>T51: training, testing</th>
<th>T6: training, testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>10.16, 9.16</td>
<td>8.46, 9.66</td>
<td>8.14, 7.15</td>
<td>8.99, 8.34</td>
<td>8.93, 8.26</td>
<td>8.34, 7.29</td>
</tr>
<tr>
<td>T3</td>
<td>9.76, 8.45</td>
<td>8.96, 7.57</td>
<td>8.71, 9.41</td>
<td>8.78, 7.46</td>
<td>8.59, 7.45</td>
<td>7.92, 6.74</td>
</tr>
<tr>
<td>T4</td>
<td>7.85, 7.28</td>
<td>8.84, 7.69</td>
<td>8.86, 8.26</td>
<td>9.34, 8.24</td>
<td>8.92, 7.49</td>
<td>8.28, 7.83</td>
</tr>
<tr>
<td>T5</td>
<td>7.68, 8.54</td>
<td>8.61, 9.52</td>
<td>8.63, 7.59</td>
<td>8.59, 7.34</td>
<td>7.65, 6.45</td>
<td>7.92, 6.59</td>
</tr>
<tr>
<td>T6</td>
<td>8.94, 7.45</td>
<td>7.34, 6.54</td>
<td>7.45, 6.34</td>
<td>7.48, 6.22</td>
<td>8.45, 7.93</td>
<td>8.91, 7.48</td>
</tr>
</tbody>
</table>
performance and the highest value for the measure of recall on the training dataset.

Both qualitative and quantitative data were gathered using the strategy for assessing the system described in the methodology section. The information generated by the system has been used to assess its usefulness, efficiency, and user friendliness.

The time it took to complete the network run was used to determine how efficient the suggested method was. Table 3 summarises the results of the network runtime analysis for both the train dataset and the test dataset.

Table 4 illustrates the average performance of various performance metrics like accuracy, recall, precision, FNR, TPR, FPR, and TNR using the proposed method. After the conclusion of the tests, the results were gathered for further consideration. And a decision is to be made about whether the patient needs assistance or whether the condition is normal.

## CONCLUSION

Most patients who are confined to hospital beds need special care from hospital staff. Even though there are many different types of systems designed to help these patients, most of them concentrate on certain duties, such as calling for emergencies or monitoring the patient’s health and activities. This work uses a hospital smart-bed control system with the help of computer vision, smart sensors, and deep learning technologies. The results of the assessment demonstrated, in the end, that the system is successful and efficient. Deep learning has been implemented into the system with the purpose of increasing accessibility. Therefore, the smart-bed system that has been designed, combined with the results of its assessment and the needs that have been specified, presents a viable answer for the crisis that is now occurring in the healthcare sector.

The implementation of smart hospital technology has the potential to enhance patient care through various means, including the improvement of communication and cooperation among healthcare teams. Instant messaging and smartphone apps provide a rapid means of communication for medical professionals such as doctors, nurses, and chemists. Facilitating care coordination and exchanging crucial information are simplified.

The implementation of smart hospital technology enables precise patient monitoring through intelligent monitoring systems. Intelligent monitoring devices and sensors are capable of quantifying vital physiological parameters such as blood pressure, heart rate, and oxygen saturation levels. The utilisation of deep learning algorithms facilitates the analysis of data to detect anomalies or emergencies in a timely manner, enabling the healthcare team to promptly intervene and administer suitable medical interventions.

Smart hospital technology employs data derived from electronic medical records, monitoring devices, and other medical sources to conduct data analysis and generate predictive insights. The utilisation of deep learning and artificial intelligence enables the identification of patterns, trends, and forecasts within the given dataset. The insights have the potential to enhance care planning, facilitate evidence-based decision-making, and optimise treatment outcomes.

The findings of this study shed even more light on the importance of using a GUI for smart beds. The proposed framework demonstrated how computer vision works without the need for any additional hardware sensors for efficient patient care. This framework shows in greater detail how approaches like deep learning may greatly improve the performance of a proposed system by making use of computer vision and several sensors as the components.

## FUTURE TRENDS FOR SMART HOSPITAL BEDS

Smart hospital beds will improve patient care and efficiency. Smart hospital bed development may include:

1. Mobile apps and smart gadgets may help patients, physicians, and nurses communicate. Virtual and augmented reality may also improve patient communication by providing visual information and instructions.
2. Intelligent data analysis: smart hospital beds will need intelligent data analysis. Medical gadgets, computerised medical records, and sophisticated monitoring systems generate massive volumes of data. This analysis improves patient care and allows evidence-based decision-making.
3. Robotics and AI: smart hospital beds may use robotics and AI to simplify operations and help patients. Robots can transport medications and provide basic care, while AI systems can help physicians make data-driven judgements.
The authors extend their appreciation to the King Salman Center for Disability Research for funding this work through Research Group no KSRG-2022-042.

ACKNOWLEDGEMENTS

The authors extend their appreciation to the King Salman Center for Disability Research for funding this work through Research Group no KSRG-2022-042.

REFERENCES


