In-Bed Posture Classification Using Deep Autoencoders

Mehrdad Heydarzadeh[1], Mehrdad Nourani[1], Sarah Ostadabbas[2]

***Abstract*— Pressure ulcers are high prevalence complications among bed-bound patients which are not only extremely painful and difficult to treat, but also impose a great burden in our health-care system. We target automatic posture detection which is a key module in all pressure ulcer monitoring platforms. Using data collected from a commercially-available pressure mapping system, we applied deep neural networks to automatically classify in-bed posture using features extracted from the histogram of gradient technique. High accuracy of up to 98% was achieved in classifying five different in-bed postures for more than 60,000 pressure images.**

I. Introduction

A. Pressure Ulcer

A pressure ulcer or bedsore is an injury to skin and/or underlying tissues which mainly occurs over bony prominences like sacrum, coccyx and heels. The ulcer development is due to a prolonged pressure applied to the soft tissues which totally or partially blocks the blood flow to the region.

An ulcer may develop in a few hours for example during a multi-hour surgery or in a few days in post surgery recovery in bed-bound patients. Ulcers are very uncomfortable and expensive to treat and they can lead to amputation or even death [1]. Statistics show that pressure ulcer caused 29,000 deaths in 2013 alone [2]. Based on a report from the national pressure ulcer advisory panel (NPUAP), the total annual pressure ulcer care cost in the United State is over $11 billion and can cost up to $70,000 per individual wound treatment [3].

Instead of lengthy and expensive treatment, redistributing the pressure by regularly turning the person can effectively prohibit the development of ulcers in bed-bound patients. Therefore, monitoring the patients' posture change over time and alerting the caregivers if a repositioning needed can play a key role in the prevention of pressure ulcers in hospitals. However, monitoring patients in a hospital with thousands of patients and already over worked nursing staff requires an intelligent system which is designed to automatically monitor, record, and report the posture change history of each patient to the caregivers and provide them with decision support in scheduling each patient's repositioning tasks. Such a system would have a unique value for the researchers in the pressure ulcer study field to understand the ulcer etiology and its contributing factors as well as for hospitals by proposing a better resource management technique [4].

[1]Department of Electrical Engineering, University of Texas at Dallas, {mehhey,nourani}@utdallas.edu
[2]Department of Electrical and Computer Engineering, Northeastern University, ostadabbas@ece.neu.edu

B. Prior Work

Electronic in-bed pressure mapping systems have already been commercially available in the forms of pressure mats for a few years [5]. Monitoring patients in bed using pressure mats does not put a threat on patient privacy unlike using regular cameras. The pressure mat measures the interface pressure applied to the patient's body parts which can be used to calculate the risk of ulceration by monitoring under pressure limbs. In order to adapt the pressure mat as an ulcer prevention tool, the first step is detecting patient's postures on the bed to keep track of the time of pressure exposure in each body part.

Since the pressure mat provides a two-dimensional array of pressure values, posture detection can be considered as an image processing task. Researchers have applied different image processing techniques to solve this problem. Authors in [6] considered a template-based approach using a modified Gaussian Mixture Model (GMM) and by fitting the pressure images to different templates and finding the best match, detected the postures. A machine learning approach for posture detection using Principle Component Analysis (PCA) and Support Vector Machine (SVM) was presented in [7]. Use of Bayesian inference and a pictorial structure were proposed in [8] and [9], respectively. In [10], kurtosis and skewness were suggested as features for each pressure image.

In spite of the number of suggested algorithms, the current posture classification methods only offer an accuracy in the range of 60% to 90% in discriminating three to nine distinct postures [6], [7], [8], [9], [10]. The difficulty arises from the fact that person-specific factors such as height, weight and body form can cause a high variability in pressure images obtained even from a specific posture. Another important limiting factor is the low resolution of the existing pressure mapping systems. These limitation can be addressed by applying a powerful machine learning and image processing algorithm which is based on a deep neural network architecture [11]. Deep neural networks have been already successfully applied in posture detection from camera images [12].

C. Key Contribution

In this paper, we apply deep autoencoder neural networks to the histogram of gradient (HoG) of the pressure image. The proposed algorithm has a high accuracy and it is robust against different variabilities in the pressures images from different individuals. Deep neural network provides a rich learning capabilities which can capture different inter class variabilities in each posture.

The rest of this paper is organized as follows: Section II reviews background which is used for developing proposed system. In Section III, the system architecture is presented. Experiment results and conclusions are in Sections IV and V, respectively.

II. Computational Model

A. Deep Autoencoders

Autoencoders or Diabolo network is an artificial neural network used for learning an efficient dimension reduction [13]. It contains two modules: first, an encoder which maps the input feature vector $\boldsymbol{x}$ to the hidden node vector $\boldsymbol{h}$ through a function $f : \boldsymbol{h} = f(\boldsymbol{x})$, in which $\boldsymbol{h}$ has a lower dimensionality than $\boldsymbol{x}$. Second, a decoder which maps from the hidden node vector to input space through another function $g : \boldsymbol{x} = g(\boldsymbol{h})$. These two functions are trained by a neural network architecture which minimizes the error of reconstruction of input space from hidden nodes, $\|\boldsymbol{x} - g(f(\boldsymbol{x}))\|^2$, where $\|\cdot\|$ is the Euclidean norm.

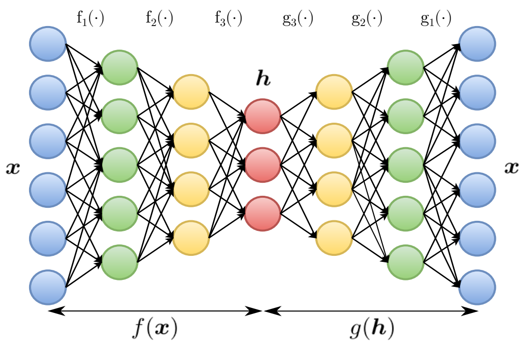


Fig. 1. An autoencoder with five layers (blue nodes are input nodes which are not counted). The architecture narrows in the middle. Encoder and decoder function are shown by $f(\boldsymbol{x})$ and $g(\boldsymbol{h})$ as part of total architecture.

Figure 1 shows such neural architectures. These networks are wide in input and output layers and narrow in the middle. After training the network, it is split in two parts. The first part represents f which represents a dimension reduction function or an encoder and the second part represents the reconstruction function or decoder. Intuitively, if the reconstruction error is small, the autoencoder architecture can represent the input vector with a more compact vector.

An autoencoder can be built by neurons with any type of activation function. In this work, we use sigmoid function defined as $\varphi(z) = \frac{1}{(1+e^{-z})}$ so, the output of the jth node in the ith layer is obtained from the output of previous layer's nodes as:

$$h^i_j = \varphi(net^i_j) = \varphi(\sum_{k=1}^{N_{i-1}} w^i_{j,k} h^{(i-1)}_k + b^i_j) \qquad (1)$$

where $w^i_{j,k}$ is the synaptic weights which connect the jth node in the ith layer to the kth node in previous layer, b^i_j is bias scalar and N_i represents the number of nodes in the ith layer.

A deep autoencoder, or in general a deep neural network architecture, is a neural architecture with several layers, usually more than three hidden layers [14]. Adding more layers increases the complexity of network and its evaluation time as well as its ability to learn more sophisticated patterns. For the application of this paper, computation time is not a bottleneck since the pressure in the monitoring framework is sampled at a very slow rate ($<$ 2 Hz). However, training such network is expected tp be computationally intensive. The next sub-section discusses the training method used in this work.

B. Greedy Training

In the training phase, the matrix of all synaptic weights ($\boldsymbol{W}$) and vector of all bias terms ($\boldsymbol{b}$) are tuned to minimize the mean square error of network over a training set with $\mathcal{N}$ samples:

$$\mathcal{E} = \sum_{i=1}^{\mathcal{N}} \|\boldsymbol{x} - \hat{\boldsymbol{x}}\|^2 = \sum_{i=1}^{\mathcal{N}} \|\boldsymbol{x} - g(f(\boldsymbol{x}))\|^2 \qquad (2)$$

For tuning the network, one can apply a method like gradient descent algorithm or its variants to train the weights. The optimization using gradient descent is usually called error back propagation algorithm in the context of neural networks. However, training a deep neural network using traditional back propagation algorithms practically is not feasible due to a phenomena usually referred as *vanishing gradient* [15].

One way of tunning such a deep architecture is a greedy layer by layer training. In the first step of the greedy training, the first and the last layers of an autoencoder are trained as a network with two layers. In this step, the vector $\boldsymbol{h}_0 = \boldsymbol{x}$ is fed to network as input and network is tuned to produce the same vector as the output. In the next step, the training dataset is mapped by obtained encoder in the first step to obtain the next level of features, $\boldsymbol{h}_1 = f_1(\boldsymbol{x})$. Then, using the new feature vector, another pair of layers is trained for obtaining the second layer. This process continues until all layers are tuned one after another. The final autoencoder will be obtained by stacking trained layers. Figure 2 symbolically shows the process of greedy training.

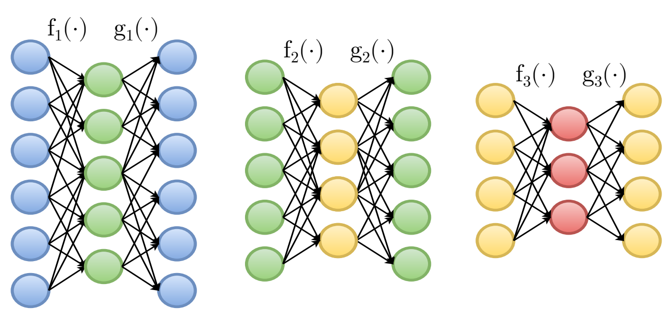


Fig. 2. Greedy training of autoencoders. At each step a pair of layers (f_i, g_i) are trained. This picture shows the training process for three layers. After training, the encoder function is obtained by stacking trained layers as $f(\boldsymbol{x}) = f_3(f_2(f_1(\boldsymbol{x})))$.

C. Softmax Regression

After finding the desired autoencoder, one can use the encoder function for feature extraction. For classifying these features, various types of classifiers can be used. However, the softmax regression is usually used in conjunction with deep neural architecture. In this way, it is possible to integrate the classifier to encoder function and apply a fine tuning.

Softmax (multinomial logistic) regression generalizes logistic regression. Logistic regression is used for binary classification, however softmax is used for classification with K classes ($K > 2$). It estimates the posterior probability of the i^{th} class as:

$$P_i = P(y = i|\boldsymbol{z}) = \frac{e^{\boldsymbol{w}_i^t \boldsymbol{z}}}{\sum_{k=1}^{K} e^{\boldsymbol{w}_k^t \boldsymbol{z}}} \qquad i = 1, \cdots, K \tag{3}$$

where $\boldsymbol{z}$ is the input to the classifier and $\boldsymbol{w}_i$'s are training weights for the ith class. Obviously, $0 < P_i < 1$ for any i and $\sum_{i=1}^{K} P_i = 1$. Then, the classification is done by comparing P_i's as $y = \arg\max_i P(y = i|\boldsymbol{Z})$. The softmax can easily be stacked to an encoder function for fine tunning the overall deep architecture by traditional back-propagation algorithms.

III. Proposed Monitoring System

In this section the architecture of proposed monitoring system is introduced. Figure 3 shows the architecture of proposed monitoring system.

Fig. 3. The proposed system methodology.

A. Pressure Mat

In this study, a Vista Medical [Manitoba, Canda] force-sensing array (FSA) pressure mat is used to collect pressure data. The FSA sensor is a flexible mat with $2048(32 \times 64)$ resistive sensors uniformly distributed. The sensor can measure pressure between 0 and 100 mmHg and is sampled at the rate of 1.7Hz. The sensor mat is flexible and can cover the total contact area between the body and the bed. A software interface is developed for interfacing the pressure mat and sending information to a computer. Figure 4(a) shows a sample pressure image.

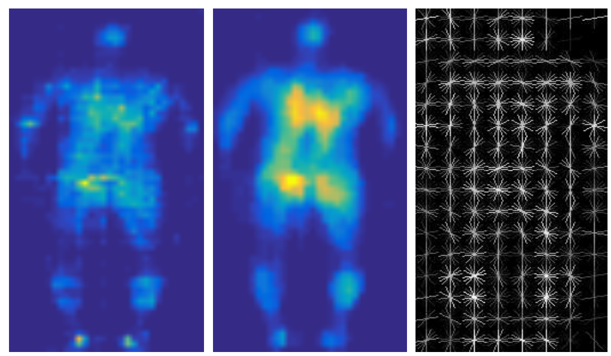
B. Preprocessing

To reduce the noise and increase the contrast of the sensed pressure image, our preprocessing module filters the input image with a $2D$ median filter with the block size of $(4, 2)$ pixels. Median filter is considered as a nonlinear filter which reduces the noise while preserves edges and avoids blurring. After filtering, the input range is compressed with the $log(\cdot)$ function. This range compression increases the quality of visualization by increasing the contrast and at the same time it increases the accuracy of posture classification module by preventing the saturation in the feature extraction module. Figure 4(b) shows the output of this module which is an improved image for the next step in dB scale.

C. Feature Extraction

In order to extract an informative image descriptor from the pressure image, we use *Histogram of Gradient* (HoG), which was originally developed for detecting human pose in regular images [16]. The goal of feature extraction module is extracting highly abstract data from raw sensor data, which is more robust against rotation, translation and changes caused by different patient's body forms.

Technically, HoG normalizes an image and calculates its gradient. The gradient image is then divided into some sub-blocks and the orientation of the gradient is obtained by a voting technique on overlapping sub-blocks. Briefly, HoG gives a vector of feature descriptors for a given image which is robust against different deformations. The details of HoG is beyond of the scope of this paper. Figure 4(c) shows HoG descriptors for a sample pressure image after preprocessing. Such a representation is offered by [16], where each star denotes a block of original image. The intensity of each line in the star shows the votes for direction of gradient in the corresponding direction.



(a) Raw (b) Preprocessed (dB) (c) HoG

Fig. 4. Pressure image at different processing stages of Figure 3.

D. Posture Classifier

To classify in-bed postures, a deep neural network is used. The encoder network is a three layer network which is fed by HoG descriptors of the pressure image. The network has five softmax output neurons corresponding to five posture classes: right foetus, right yearner, supine, left yearner, and left foetus, which are denoted by $\{F_R, Y_R, S, Y_L, F_L\}$, respectively. These postures are shown in Figure 5.

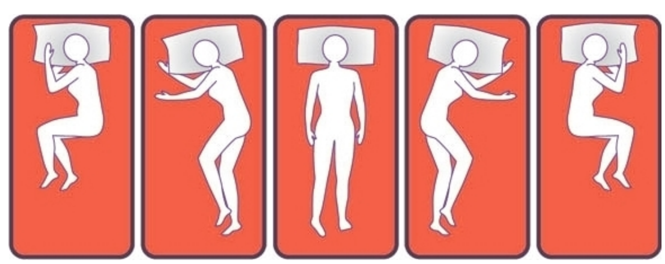


Fig. 5. Main postures studied in this paper (from left to right: F_R, Y_R, S, Y_L and F_L).

For training the system, 10 participants were asked to lie over a pressure mat on a hospital mattress and the pressure data was collected while being in different sleeping postures. The postures of each participant during the experiment were annotated and labelled manually. To increase the size of the dataset and introduce more geometrical variations, random rotations of $\pm 5°$ to $\pm 15°$ and also translation of ± 5 pixels in both x and y directions for each posture image were added to the dataset. 70% of all data was randomly chosen as training set and the rest as test set. The autoencoder was trained by the *stochastic gradient descent* (SGD) algorithm [17]. The Microsoft computation network tool was used for training the network [18].

TABLE I
CONFUSION MATRIX OF CLASSIFIER

Target	Classified (%)				
	F_R	Y_R	S	Y_L	F_L
F_R	96.2	3.8	0	0	0
Y_R	0.8	99.2	0	0	0
S	0	0	100	0	0
Y_L	0	0	0	97	3
F_L	0	0	0	2.1	97.9

IV. EXPERIMENTATION

The proposed system collects data from a pressure mat and enhances the raw image by a noise reduction and range compression preprocessing scheme. Then, the HoG descriptor of enhanced image is extracted and fed to the deep neural network. This network is encoder part of a deep autoencoder. The accuracy of the classifier was validated using the aforementioned test data. The overall accuracy of 98.1% was achieved. The confusion matrix of the test dataset is tabulated in Table I. The proposed method takes $10\,ms$ on average (wall-clock) to classify a frame of pressure image on an i7 PC running Windows 8 at clock of 3.4 GHz which is fast enough to implement as a real-time module.

To investigate the robustness of the proposed algorithm, we randomly rotated each pressure image by $\pm 5°$ to $\pm 15°$ and classified the resulted images, which the accuracy of 94.1% was obtained in this round.

Table II summarizes the comparison of proposed method with results reported by others. Majority of works are limited to a fewer number of postures. Only in [7], their algorithm detects the same number of postures. However, our obtained accuracy is still higher. The results of deep neural network in capturing variabilities in this classification task are better or competitive. This is achieved by paying a higher computational cost in this method.

TABLE II
COMPARISON OF POSTURE CLASSIFICATION ALGORITHMS

Ref	# of Postures	Accuracy (%)	Algorithm
[6]	3	91.6	GMM
[7]	5	97.0	PCA + SVM
[8]	3	81.4	Bayesian Inference
[9]	3	89.8	Pictorial Structure
[10]	3	95.0	Kurtosis + Skewness
Ours	**5**	**98.1**	**HoG + DNN**

V. CONCLUSIONS

Automatic posture classification helps to monitor the critical body parts by calculating the overall time that a high risk area is under certain amount of pressure. In this paper, an accurate and robust to variation algorithm is proposed for in-bed posture classification using deep neural network. Accuracy of such system can be even improved further if individual limbs are identified separately. Our future work intends to apply advanced machine vision techniques to this problem.

REFERENCES

[1] C Allison Russo, Claudia Steiner, and William Spector. Hospitalizations related to pressure ulcers among adults 18 years and older, 2006. 2008.

[2] Mohsen Naghavi, Haidong Wang, Lozano, et al. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death. *Lancet*, 385(9963):117–171, 2015.

[3] Emily Haesler. National pressure ulcer advisory panel, european pressure ulcer advisory panel and pan pacific pressure injury alliance. *Prevention and treatment of pressure ulcers: quick reference guide.[Internet]*, 2014.

[4] Sarah Ostadabbas, Rasoul Yousefi, Mehrdad Nourani, Miad Faezipour, Lakshman Tamil, and Matthew Q Pompeo. A resource-efficient planning for pressure ulcer prevention. *Information Technology in Biomedicine, IEEE Transactions on*, 16(6):1265–1273, 2012.

[5] Vista Medical. Pressure mapping, 2015. [Online; accessed 1-Dec-2015].

[6] Sarah Ostadabbas, Maziyar Baran Pouyan, Mehrdad Nourani, and Nasser Kehtarnavaz. In-bed posture classification and limb identification. In *Biomedical Circuits and Systems Conference (BioCAS), 2014 IEEE*, pages 133–136. IEEE, 2014.

[7] Rasoul Yousefi, Sarah Ostadabbas, Miad Faezipour, Mehrdad Nourani, et al. A smart bed platform for monitoring & ulcer prevention. In *Biomedical Engineering and Informatics (BMEI), 2011 4th International Conference on*, volume 3, pages 1362–1366. IEEE, 2011.

[8] Chi-Chun Hsia, Yu-Wei Hung, Yu-Hsien Chiu, and Chia-Hao Kang. Bayesian classification for bed posture detection based on kurtosis and skewness estimation. In *e-health Networking, App. and Services. HealthCom 2008. 10th Int. Conf. on*, pages 165–168. IEEE, 2008.

[9] Jason J Liu, Ming-Chun Huang, Wenyao Xu, and Majid Sarrafzadeh. Bodypart localization for pressure ulcer prevention. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pages 766–769. IEEE, 2014.

[10] CC Hsia, KJ Liou, APW Aung, V Foo, W Huang, and J Biswas. Analysis and comparison of sleeping posture classification methods using pressure sensitive bed system. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pages 6131–6134. IEEE, 2009.

[11] Yoshua Bengio, Aaron Courville, and Pierre Vincent. Representation learning: A review and new perspectives. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(8):1798–1828, 2013.

[12] Chaoqun Hong, Jun Yu, Jian Wan, Dacheng Tao, and Meng Wang. Multimodal deep autoencoder for human pose recovery. *Image Processing, IEEE Transactions on*, 24(12):5659–5670, 2015.

[13] Yoshua Bengio. Learning deep architectures for ai. *Foundations and trends® in Machine Learning*, 2(1):1–127, 2009.

[14] Yoshua Bengio. Deep learning of representations: Looking forward. In *Stat. Lang. and Speech Proc.*, pages 1–37. Springer, 2013.

[15] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In *International conference on artificial intelligence and statistics*, pages 249–256, 2010.

[16] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893. IEEE, 2005.

[17] Léon Bottou. Large-scale machine learning with stochastic gradient descent. In *Proceedings of COMPSTAT'2010*, pages 177–186. Springer, 2010.

[18] Dong Yu, Adam Eversole, Mike Seltzer, et al. An introduction to computational networks and the computational network toolkit. Technical report, Tech. Rep. MSR, Microsoft Research, 2016, http://codebox/cntk, 2014.