2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Video monitoring over anti-decubitus protocol execution with a deep neural network to prevent pressure ulcer

Ivan Danilovich, Vyacheslav Moshkin, Alexander Reimche, Mikhail Tevelevich, Nikolay Mikhaylovskiy

Abstract — **Video monitoring of the patient position in the intensive care units is complicated by the obstacles covering the patient body. Conventional posture detection algorithms do not work in this case. A reformulation of the posture detection problem for the case as an object detection/image classification problem and the use of recent deep learning techniques allowed us to achieve 94.5% accuracy on a pre-clinical test classifying 4 postures using imagery from an off-the-shelf camera and edge processing, which is a 60% improvement over the result previously known in literature. This in turn allowed us to build a ready for the clinical trials system based on inexpensive off-the-shelf cameras.**

Clinical Relevance — **A cheap and practical system of automatic video monitoring of bedridden patients allows to minimize the risks of pressure ulcer in ICU.**

I. Introduction

In most cases, a pressure ulcer is a consequence of a failure to comply with the anti-decubitus protocol — a systematic change in body positions of an immobilized patient [1].

Nationally, the U.S. spends somewhere from $3.6 billion to about $26.8 billion a year on treatment costs for pressure ulcers, according to different estimates [2][3][4]. Thus, tools that help hospitals and care providers to comply with the anti-decubitus protocol can be very advantageous.

Here we describe a low-cost, computer-vision-based anti-decubitus protocol execution monitoring solution that can be implemented in an intensive care unit with little or no disruption to the existing workflow. The system identifies that the position of the patient has not changed for too long, and alerts the nursing staff to intervene.

A. Previous work

Various methods have been previously proposed for the monitoring of the anti-decubitus protocol implementation. These include systems from checklists to pressure sensors and computer vision systems.

Computer vision-based methods are the most similar to the purpose of the current work. Most works use the depth cameras, for example, Chang et al. [5] developed a patient monitoring system that detects the postures and monitorы the patient activity, Grimm et al. [6] classify three postures with an accuracy of 94%, Li et al. [7] classify 10 postures with 93% accuracy. On the other hand [7] notes that when a quilt was laid on top of the subject, the accuracy was reduced to 89%. Martinez et al. [8] claim that their depth-sensor based Bed-Aligned Map technology delivers 100 % accuracy in sleeping position recognition in a 4-class test, which seems to be a deficiency in the experimental setup. Other similar works include Yoshino and Nishimura [9].

Several works use a camera in addition to other sensors. Achilles et al. [10] used both motion capture cameras and a Kinect depth sensor to estimate the patient position under a blanket.

In addition to previously mentioned systems, Liu et al. [11] used a camera with a polarizer in addition to two types of infrared cameras to monitor the size of a pressure ulcer. Liu and Ostadabbas [12] used thermal imaging and achieved pose estimation performance of 98.0% and 96.0% in a living room and in a hospital room, building upon previous work of Liu et al. [13].

The work that is possibly most similar to ours is that by Liu and Ostadabbas [14], but they have used HOG-based approach for classification.

B. Our contribution

We have developed a posture change detection system that can be used in the clinical practice and is currently ready for clinical trials. Unlike known systems, it

- uses a single, off-the-shelf, general-purpose, not a stereo/depth/RGBD, camera;
- uses inexpensive, off-the-shelf IR lighting for night vision;
- processes the video on an edge device with no need to transfer the sensitive personal video data;
- uses classification and object detection approaches instead of the full-blown posture detection;
- achieves 94.5% accuracy on a pre-clinical test, which is a 60% improvement over the results of [14];
- does not degrade the accuracy when the subject is covered by a quilt, bedspread or medical equipment.

II. Solution approach

A. Dataset

Although there are several public human pose datasets available such as MPII [16], LSP [17], FLIC [15], Buffy [18],

N. Mikhaylovskiy and I. Danilovich are with NTR Labs, Moscow, Russia, 129594, and Tomsk State University, Tomsk, Russia, 634050 (phone: +7-903-790-5815; e-mail: nickm@ntr.ai, idanilovich@ntr.ai).

V. Moshkin and A. Reimche are with NTR Labs, Moscow, Russia, 129594 (e-mail: vmoshkin@ntr.ai, aremkhe@ntr.ai).

M. Tevelevich is with Ab ovo med, Tomsk, Russia, 634062 (e-mail: mt@abovomed.com).

they are all mainly from scenes such as sports, TV shows, and other daily activities. None of them provides any specific in-bed poses, with the exception of SLP [19], which was not publicly available when this work was started. Thus, we decided to build our own dataset.

The dataset was prepared in collaboration with Yudin Hospital 67, Moscow.

Confidentiality agreements were signed with the hospital and specific patients or their relatives. Two cameras with IR lighting were installed in June 2017 above the beds in the intensive therapy unit, to take pictures every 5 minutes. The photos were collected to a server installed locally in the hospital. All the data were collected and depersonalized in accordance with legislation related to personal and medical data. Photographs not compliant with the above legislation were deleted and the compliant ones were manually transported on a hard drive to the NTR Labs office. Newer cameras were installed instead of the older ones in April 2018, with the same protocol. The cameras were moved several times to different beds to improve the diversity of the dataset.

Data collection was finished in February 2020. The total number of the images collected over 33 months of data collection was 299868. 129886 of them were manually labelled in terms of patient positions by qualified labelers:

- lying on the back 63578,
- on the left side 2554,
- on the right side 2377, and
- half-sitting 3390.

The rest of the images belonged to the two auxiliary classes:

- undefined (say, when the camera was accidentally moved by personnel to miss the bed) 39706
- empty bed 16281

B. *Algorithmic approach*

Dataset images depict patients typically covered with bedspread and/or medical equipment. Pose estimation approaches do not work well in these cases. Due to the concerns above, we have chosen to use classification and object detection approaches to monitor potential pressure ulcer protocol violations instead of full-blown pose detection.

The dataset is highly imbalanced, with “lying on the back” class dominating. The problem of the data imbalance was solved by excluding a large part of the dominant class from the training process.

Cameras took a shot every 5 minutes, and the patients are immobilized, so the consecutive frames are often very similar. Most of the duplicates were excluded from the training process as well.

Position classification pipeline is based on the EfficientNet [20] family of neural network architectures, known for good recognition quality. In addition, the smallest versions of EfficientNet are lightweight enough to be ported to the edge devices.

The recognition pipeline consists of two major parts. First, we detect the location of a patient (bounding box) on the image using EfficientDet [21]. It allows us to eliminate the dependence on the environment at the classification stage. Detector and classifier are combined into a single model.

Often the position of a patient does not change significantly between the shots. It means that consecutive frames are codependent. This fact was used to improve the reliability of classification. Frame sequences were split into series of five frames each. Then the model processes five frames in a single pass. It allows the model to use the frame codependency information and removes the classification jitter.

C. *Training pipeline*

The detector part of the model is a TensorFlow [22] implementation of EfficientDet.

We have used models pretrained on MS COCO dataset [23] and trained on a server with two NVIDIA K80 GPUs for at most 50 epochs. The optimization algorithm used was Adam with custom learning rate, including warmup [24] stage. The training chart is depicted on the Figure 1 below.

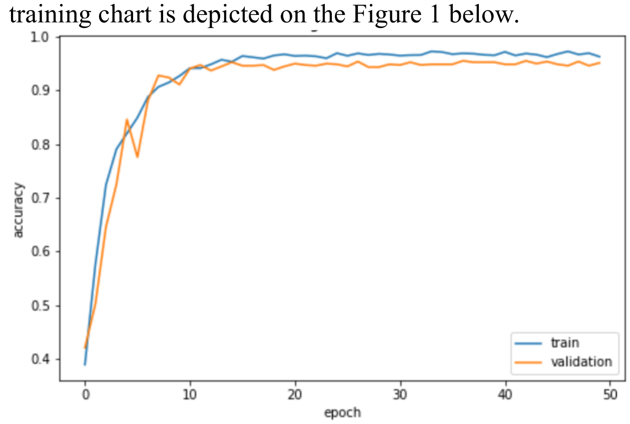

Figure 1. The training chart.

We have used a large assortment of data augmentation techniques provided by the Albumentations library [25]:

- Random horizontal/vertical flip
- Random rotate
- Color conversion
- Blur
- Noise
- Affine transformation

D. *System architecture*

The system architecture consists of:

- a “smart” Odroid-based edge device running Android operating system, with a camera, processor, permanent storage (flash memory), an appropriate lens and IR lighting source
- camera tripod system on the bed to mount the device
- server connected with the edge device with a TCP/IP network, so that it can gather the data from the edge devices
- web interface for system administration

- a messaging subsystem to alert the nursing staff that a patient’s position has not changed for too long and a corrective action needs to be taken.

The camera of the edge device takes a shot often, and the neural network on the edge device recognizes the patient's position and sends the information about it to the server. This signal is processed on the server - a graph of the patient positions is formed in the admin panel and, if a decubitus protocol is violated, a message is sent to alert the personnel in charge.

The system is designed to respect medical confidentiality and patient privacy. Therefore, the processing of photos takes place directly on the edge devices, the photos are not stored anywhere and are not transferred to the server.

Neural network model is executed directly on the edge device running Android and is packaged into a Java app. For that purpose the model is converted to the tflite format via TensorFlow Lite library [26]. The resulting model size is approximately 35MB.

E. User interface

A web-based user interface is available to the hospital staff. In the administrator mode, both the network training functionality (including the management of image labelling) and the patient monitoring interfaces are available.

The administrator user interface provides access to the list of devices installed. By default, each edge device recognizes the position once in a specified period. The frequency can be adjusted on the device itself. Each device can be associated with a hospital name, room and a patient ID.

The details of the patient monitoring protocol, including the time the patient should stay in the same position for the system to issue an alert to the nursing staff, can also be adjusted using the administrator user interface.

For each specific device, a responsible person can see the graph of the patient's position change for the selected interval - day, week, month, or a custom time interval. This helps to keep track of stats.

III. Pre-clinical test results

Patients that were not included into the validation or training sets were included into the test set. The test consisted of the following number of pictures of patients in the classes:

- Lying on the back 1129
- On the left side 157
- On the right side 99
- Half-sitting 42
- Undefined 76
- Empty bed 35

Our model achieved 94.5% accuracy on the test set. Figure 2 below displays the confusion matrix for the target classes.

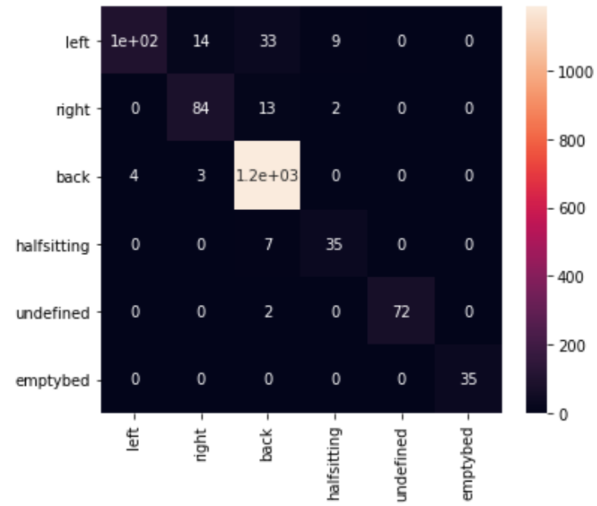

Figure 2. Test set confusion matrix.

IV. Discussion

With the advent of COVID-19, the intensive care units are facing new challenges and adopting new guidelines and protocols. Prone positioning have been found to significantly improve ventilation in patients with COVID-19 associated respiratory failure [27]. Thus, our nearest goal is to extend the dataset and achieve comparable accuracy for the prone position.

The intensity of pressure ulcer prevention depends on the severity of microcirculation disorders and skin temperature [28]. Some protocols take these parameters into account, and one of the future directions is to integrate these externally available parameters into the monitoring system.

Still, our main achievement is still the reformulation of the posture detection problem for the immobilized patients in the intensive care units as an object detection/image classification problem and the use of recent deep learning techniques. The above approach allowed us to achieve 94.5% accuracy on a pre-clinical test classifying 4 postures using imagery from an off-the-shelf camera and edge processing, which is a 60% improvement over the result previously known in literature. This in turn allowed us to build a ready for the clinical trials system based on inexpensive off-the-shelf cameras.

References

[1] J. McBride and A. Richardson, “A critical care network pressure ulcer prevention quality improvement project.,” Nurs. Crit. Care, vol. 21, no. 6, pp. 343–350, Nov. 2016, doi: 10.1111/nicc.12174.

[2] S. R. Nussbaum et al., “An Economic Evaluation of the Impact, Cost, and Medicare Policy Implications of Chronic Nonhealing Wounds.,” Value Heal. J. Int. Soc. Pharmacoeconomics Outcomes Res., vol. 21, no. 1, pp. 27–32, Jan. 2018, doi: 10.1016/j.jval.2017.07.007.

[3] Preventing Pressure Ulcers in Hospitals. Content last reviewed October 2014. Agency for Healthcare Research and Quality, Rockville, MD. https://www.ahrq.gov/patient-safety/settings/hospital/resource/pressureulcer/tool/index.html

[4] W. V. Padula and B. A. Delarmente, “The national cost of hospital-acquired pressure injuries in the United States,” Int. Wound J., 2019, doi: 10.1111/iwj.13071.

[5] M. Chang et al., “In-bed patient motion and pose analysis using depth videos for pressure ulcer prevention,” in 2017 IEEE International

Conference on Image Processing (ICIP), 2017, pp. 4118–4122, doi: 10.1109/ICIP.2017.8297057.

[6] T. Grimm, M. Martinez, A. Benz, and R. Stiefelhagen, “Sleep position classification from a depth camera using Bed Aligned Maps,” in 2016 23rd International Conference on Pattern Recognition (ICPR), 2016, pp. 319–324, doi: 10.1109/ICPR.2016.7899653

[7] Y. Li, Y. Lei, L. C. Chen, and Y. Hung, “Sleep posture classification with multi-stream CNN using vertical distance map,” in 2018 International Workshop on Advanced Image Technology (IWAIT), 2018, pp. 1–4, doi: 10.1109/IWAIT.2018.8369761

[8] M. Martinez, B. Schauerte, and R. Stiefelhagen, “‘BAM!’ depth-based body analysis in critical care,”, pp. 465-472. Springer, Berlin, Heidelberg, 2013 doi: 10.1007/978-3-642-40261-6_56.

[9] A. Yoshino and H. Nishimura, “Study of Posture Estimation System Using Infrared Camera BT - HCI International 2016 – Posters’ Extended Abstracts,” 2016, vol. 617, pp. 553–558

[10] F. Achilles, A. E. Ichim, H. Coskun, F. Tombari, S. Noachtar, and N. Navab, “Patient MoCap: Human pose estimation under blanket occlusion for hospital monitoring applications,” 2016. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9900, pp. 491–499. Springer, Cham (2016), doi: 10.1007/978-3-319-46720-7_57.

[11] W.-M. Liu, C.-L. Chen, L.-Y. Chang, S.-C. Pong, and H.-M. Chen, “Multimodal and Multispectral Imaging for Chronic Pressure Ulcer Assessment,” in Proceedings of the 2nd International Conference on Biomedical Signal and Image Processing, ser. ICBIP 2017. New York, NY, USA: ACM, 2017, pp. 47–52.

[12] Liu S., Ostadabbas S. (2019) Seeing Under the Cover: A Physics Guided Learning Approach for In-bed Pose Estimation. In: Shen D. et al. (eds) Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. MICCAI 2019. Lecture Notes in Computer Science, vol 11764. Springer, Cham. https://doi.org/10.1007/978-3-030-32239-7_27

[13] S. Liu, Y. Yin, and S. Ostadabbas, “In-Bed Pose Estimation: Deep Learning With Shallow Dataset.,” IEEE J. Transl. Eng. Heal. Med., vol. 7, p. 4900112, 2019, doi: 10.1109/JTEHM.2019.2892970.

[14] S. Liu and S. Ostadabbas, “A Vision-Based System for In-Bed Posture Tracking,” in 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), 2017, pp. 1373–1382, doi: 10.1109/ICCVW.2017.163.

[15] B. Sapp and B. Taskar, “MODEC: Multimodal decomposable models for human pose estimation,” 2013, doi: 10.1109/CVPR.2013.471.

[16] M. Andriluka, L. Pishchulin, P. Gehler, and B. Schiele, “2D human pose estimation: New benchmark and state of the art analysis,” 2014 in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2014, pp. 3686–3693, doi: 10.1109/CVPR.2014.471

[17] S. Johnson and M. Everingham, “Clustered pose and nonlinear appearance models for human pose estimation,” 2010, in Proc. BMVC, vol. 2, 2010, p. 5, doi: 10.5244/C.24.12.

[18] Ferrari, V., Marin, M., Zisserman, A.: “Progressive search space reduction for human pose estimation.” 2008, In: IEEE Conference on Computer Vision and Pattern Recognition.

[19] Liu, Shuangjun and Huang, Xiaofei and Fu, Nihang and Li, Cheng and Su, Zhongnan and Ostadabbas, Sarah, “Simultaneously-Collected Multimodal Lying Pose Dataset: Towards In-Bed Human Pose Monitoring under Adverse Vision Conditions”, arXiv:2008.08735

[20] M. Tan and Q. V. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” 2019 , arXiv:1905.11946

[21] M. Tan, R. Pang, and Q. V. Le, “EfficientDet: Scalable and efficient object detection,” 2020, doi: 10.1109/CVPR42600.2020.01079, arXiv:1911.09070

[22] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng. “TensorFlow: a system for large-scale machine learning.” In Proceedings of the 12th USENIX conference on Operating Systems Design and Implementation (OSDI'16). USENIX Association, USA, 265–283.

[23] T. Y. Lin et al., “Microsoft COCO: Common objects in context,” 2014, In ECCV, pages 740–755, 2014 doi: 10.1007/978-3-319-10602-1_48.

[24] Goyal P., Dollar P., Girshick R., Noordhuis P., Wesolowski L., Kyrola A., Tulloch A., Jia Y., and Kaiming He. “Accurate, large minibatch sgd: training imagenet in 1 hour.” arXiv preprint arXiv:1706.02677, 2017.

[25] Buslaev, A., Iglovikov, V.I., Khvedchenya, E., Parinov, A., Druzhinin, and M., Kalinin, A. A. "Albumentations: Fast and Flexible Image Augmentations." Information 11, no. 2: 125, 2020.

[26] https://www.tensorflow.org/lite

[27] Mittermaier, M., Pickerodt, P., Kurth, F., Jarcy, L., Sander, L., Uhrig, A., Garcia, C., Machleidt, F., Pergantis, P., Weber, S., Li, Y., Breitbart, A., Bremer, F., Knape, P., Dewey, M., Doellinger, F., Weber-Carstens, S., Slutsky, A., Kuebler, W., Müller-Redetzky, H., 2020. “Evaluation of PEEP and Prone Positioning in COVID-19 ARDS.” SSRN Electronic Journal. 10.2139/ssrn.3622409.

[28] Lyder C.H., Ayello E.A. Pressure Ulcers: A Patient Safety Issue. In: Hughes R.G., editor. Patient Safety and Quality: An Evidence-Based Handbook for Nurses. Rockville (MD): Agency for Healthcare Research and Quality (US); 2008 Apr. Chapter 12. Available from: https://www.ncbi.nlm.nih.gov/books/NBK2650/