Sitting posture Analysis using CNN and RCNN

Nishitha R
Biomedical Engg. Department
Sathyabama Institute of Science and Technology
Chennai

T. Sudhakar*
Biomedical Engg. Department
Sathyabama Institute of Science and Technology, Chennai
drsudhakar35@gmail.com

Bethanney Janney J
Biomedical Engg. Department
Sathyabama Institute of Science and Technology, Chennai
bethanneyj@gmail.com

N. R. Krishnamoorthy
Dept. of ECE
Sathyabama Institute of Science & Technology , Chennai
moorthy26.82@gmail.com

Dhanalakshmi K
Dept. of Computer Science Engg
Sathyabama Institute of Science and Technology, Chennai
dhanalakshmi.cse@sathyabama.ac.in

Vigneshwaran S
M.Tech-Medical Instrumentation
Sathyabama Institute of Science and Technology, Chennai
vickyselvaraj20@gmail.com

2023 International Conference on Bio Signals, Images, and Instrumentation (ICBSII) | 979-8-3503-3817-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICBSII58188.2023.10181038

Abstract—**Based to statistics, about 80 percent of the population follows incorrect sitting posture. From the year 2020, due to the Covid-19 curfew, the traditional work office of many professionals turned into Work from home and remote office lifestyle. Many of the working professionals has no proper ergonomic workplace setup in their house leading to improper sitting position. Monitoring the sitting posture is needed to avoid chronic complication of MSD, neck, and spine related injuries. Many techniques were emerged for correcting and monitoring the posture. Neural Network based posture detection is being into the field of research as the come with more range of accuracy. This paper discusses about the various Neural Networks used and their accuracies.**

Keywords— MSD-musculoskeletal disorder, ergonomic workplace, convolution neural networks (CNN) and region based convolutional neural network (RCNN).

I. Introduction

As many of the people adopt for the work from home lifestyle the number of people affected by the spine and neck disorder has been increased to greater extend. The primary cause for the complication of disorder is that many professionals are not supported by the ergonomic workspace. All employees cannot make ergonomic workplace which result in spinal complications. In order to prevent the chronicity of the disorder, preventive posture monitoring and sitting posture correcting tools been emerging in the market. Both wearable and non-contact tools have been under research to make a preventive tool for the position monitoring. Many sensors been implemented in the posture management like Force resist sensor which is placed on the cushion of sitting to calculate the force acting on the sensors, Pressure sensor to detect the evenly distributed pressure on the chair. Also, some research works on the non-contact supportive tools so that the persons are free from sensors. One of the predominant non-contact tools is the use of Neural Network to monitor the posture of the person. Many research work has been carried out to extract the accurate posture detection.

II. Existing Methodologies

Research for sitting posture detection and analysis have been increasing for past couple of years. Predominantly the posture detection is carried out either by wearable sensors or by the Machine learning methods. Many research work stated the use of sensors like Flex sensors [1], flexible array pressure sensor [12], FRS and Sitting pressure sensors (SPS) [10], [2] will help in measuring the pressure and/or force exerted over the cushion while sitting. Some articles state the combinational use of sensors and Neural network for obtaining more accuracies. Technologies like IoT [2], [4] are used alone with sensors and machine learning for storing the data on cloud. By using sensors, the changes in pressure and/or force are monitored so that the persons are alerted when the values exceed the threshold in any of the sensors being used. In order to remove the noise and artifacts during sensor recording, they are supported by various filters like Kalman filter [2]. The monitoring and value analysis are made ease by creating a supporting android application [2]. Some research uses k-Nearest neighbor's, support vector machine, random forest, decision tree [10] for experimental analysis along with raspberry pi. Fully conventional neural network to implement the embeddable human posture recognition system [5] are used and proved that the accuracy is achieved at 98.99%. The use of Kinect V2 [17], [6] for classification of human body joints like using the Open Pose library. For performing the image classification either MATLAB [17] or python are used. In a research paper, integration of SVM, CNN and LSTM [19] are made to study about the posture of a person [9]. 3-D RCNN [20] with spatiotemporal process is used for multi-stage technique to recognize the position of a person. Neural network is emerging in this domain as they produce high range of accuracy with greater classification. Optical neural network [22] has random shift wavelet pattern for easy posture recognition.

III. Materials and Methods

A. Neural Network

Neural Network is a part of Machine Learning which makes the device to learn itself with the existing inputs. As the name implies, Neural network resembles the structure of Brain Neurons having many networks of weight layers such that the accuracy and the learning rate of the device is being improvised.

B. CNN

The Convolutional Neural Network is a part of Neural Networks that performs the mathematical convolution to train the network. This project uses a three-layer CNN. The CNN has feature map which generate both Part Confidence Maps and Part Affinity Field.

C. RCNN

Region based convolutional neural network is more derived version of neural network that concentrate deep on the region- based classification of the object in the image. RCNN is used because it mainly concentrates on the image classification based on region detection and object segmentation. Dataset are supported by MS-COCO dataset. Key point RCNN is slightly difference from Mask-RCNN.

D. Python

For coding and training the Neural networks, python is being used as python is one of the general-purpose, user-friendly programming language and is easily adopted by many non-programmers. Python along with libraries like TensorFlow, Scikit-learn, Keras, Pytorch can be used for Machine learning tasks.

E. Keras

Python written deep learning API which runs on the top of machine learning platform TensorFlow. This API is simple, flexible, and powerful used for research as it supports fast experimental procedure. Layers and models are the core data structures of Keras.

F. Gaussian filter

The gaussian filter is a type of low pass filter. This filter is used for reducing the noise and blurring in an input image. Here multidimensional gaussian filter is used. It targets to retain the edges so that key point extraction is made at ease.

G. Open Pose

OpenPose is a multi-person detecting real-time library can be used for detecting the human body key points. OpenPose is written in C++ an Caffe. OpenPose with python can support to create different operating systems and languages.

H. TensorFlow

An open-source platform for machine learning where the pose estimation can be achieved by merely estimating the key points and key joints output. It is a core library and has comprehensive, ecosystem of tools, community resources that helps the programmer to have state of art in development of Neural network.

I. TORCHVISION

Torchvision is a model that contains pretrained key point detection model which are built on top of the ResNet-50 FPN. FPN is a process of fusing features maps at multiple scales for preserving the information at multiple levels.

IV. PROPOSED METHODOLOGY

The sitting posture analysis using CNN and RCNN is a comparative work to analyse the accuracy and latency of the CNN and RCNN in detecting the posture of the person. The output of the network indicated the keypoints extracted from the given input and the latency for detection of the system. The postures that we considered to detecting the position are: a. Hunchback, b. Reclined, c. Straight, d. hand folds, e. kneeling, f. cross legs. This software detects the sitting position by getting the input from the lateral view of the person as we aim to calculate the posture by calculating the angle and degree between the keypoint of reference. The key points are extracted from the image by using OpenPose model. In this work 18 keypoints in the body are extracted and analysed.

Steps carried to recognize the sitting posture are:

- Gather the keypoint coordinates for different body parts.
- Calculate the sitting angle of the person.
- Getting the ear position
- Getting the hip position
- Calculating the angle between ear and hip

If angle > 110 → Reclined position; Elif angle > 70 → Hunchback position; Else → Straight position

Initially the software gets the input image and extract the keypoints from the person. This is achieved by constructing three python files: one for containing the whole architecture of the software, second containing the parameters essential to predict the keypoints, third having the functions to calculate the co-ordinates of keypoints.

Once the keypoints are extracted the system starts detecting the angle between the keypoints to analyse the deviation of spine from straight posture. Here we considered Hip and Ear keypoint since both the keypoints lay on same line when they are in straight position.

The above work has four programs:

- Posture detection in an input image using CNN.
- Posture detection in an input image using RCNN.
- Posture detection in an input video and real-time video input using CNN.
- Posture detection in an input video and real-time video input using RCNN.

A. CNN Posture detection

Initially the image is passed through baseline CNN network to extract the feature maps of input. The feature map is processed in 3 stage CNN to obtain are Part Confidence Maps and Part Affinity Field. Greedy bipartite matching algorithm are used to get Part Confidence Maps (PCA) and Part Affinity Field (PAF). Confidence Map is a 2D representation that a particular body part can be located in pixel. PAF is a set of 2D vector field which encodes location and orientation of limbs of people. It encodes pairwise connections between body parts.

Three steps carried out in CNN are:

- To predict the Part Affinity Fields from the Feature maps of base network
- Use the output PAF from previous layers to refine the prediction of Confidence maps.
- Confidence map and Part Affinity Field are passed into the greedy algorithm.
- The loss function is used to calculate the loss between PCM and PAF to the ground truth maps and fields.

Proposed CNN model consists of two convolutional layers, followed by pooling layer. CNN network accepts image with pixel size of 28x28x3, and finally results in the 14x14x3 size features signal after passing through the two set of convolution layer with Relu activation function and max pooling layer. The input posture image will be processed by two convolution layers with 16 filters; each has kernel size of 3 in the first set and has kernel size of 4 in the second set of convolution layer. Further the extracted features will be down sampled to 14x14x3 image size using two max pooling layers. The confusion matrix of CNN model is shown in the figure 1. The accuracy of the model is evaluated for six different posture and overall accuracy of the CNN is obtained as 92.5%

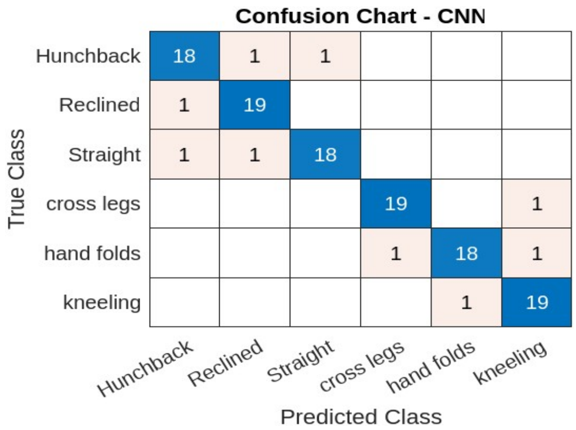

Fig.1. Confusion matrix of CNN Model

B. *CNN Realtime posture detection*

The CNN is coded for detecting the live input video streaming on the web camera attached to the laptop/PC and to process the recorded video input for detecting the posture of the person. The Torchvision is used for the dataset with gaussian filter to remove the noises in the input. If the posture is not able to detect or if the camera is not in the lateral position to the person, then the system responds with a notification of “unable to calculate the angle”.

C. *RCNN Posture detection*

Keypoint RCNN have same functionality as Mask-RCNN, they differ by output size and the way keypoints encoding in the keypoint mask. Keypoint RCNN has one-hot encoding a keypoint of the detected object instead of whole mask. Class-wise output feature map from Mask-RCNN: Here the feature map has two channels person and the background class. Pretrained keypoint detection model built on top of the ResNet– 50 FPN backbone. FPN is a fusion feature maps at different scales so that the information is preserved. The confusion matrix of RCNN model is shown in the figure 2. The accuracy of the model is evaluated for six different posture and overall accuracy of the RCNN is obtained as 95%

Faster RCNN has numerous layers where the Region-proposal- layer predicts the locations of N number of objects in feature maps. These regions are individually passed to ROI-Pooling layer which resizes the feature maps by quantising size grid to fixed size grid and getting the max-values to place in the fixed grid. A Fully Connected (FC) layer follows ROI-pooling layer. Mask-RCNN has some modification of layers, like ROI-Align layer is used instead of ROI-Pooling layer and the output of ROI-Align is passed to Mask-RCNN head. COCO dataset has 80 classes for detection and segmentation, the annotations are offered for person class. Torchvision is particularly trained to identify key points in a person.

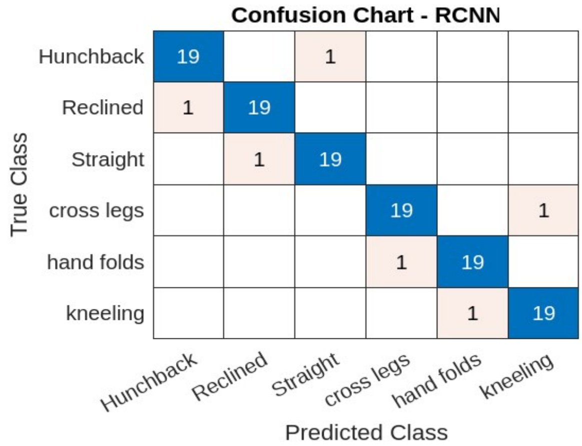

Fig.2. Confusion matrix of RCNN Model

D. *RCNN Realtime Posture detection*

RCNN is coded for analysing the live streaming video and recorded video input. The real time input is taken and processed with dataset from Torchvision, and the noise are removed by using the gaussian filter. If the camera is in wrong position or if the lateral view of the subject is not properly detected, then the system comes with a alert message of “Unable to calculate the angle”, later which the position can be altered.

V. Results and Discussion

The comparison of CNN and RCNN output time for same set of images are represented in Table I and Table II.

TABLE I. Results for CNN

Input (Image)	Expected Result (Position)	Actual Result (Position)	Time in Seconds
Sample 1	Hunchback	Hunchback	56.30753
Sample 2	Hunchback	Hunchback	61.99616
Sample 3	Hunchback	Hunchback	72.53529
Sample 4	Reclined	Reclined	52.14701
Sample 5	Reclined	Reclined	70.73289
Sample 6	Reclined	Reclined	64.77853
Sample 7	Straight	Straight	82.54892
Sample 8	Straight	Straight	69.10063
Sample 9	Straight	Straight	69.49605

TABLE II. Results for RCNN

Input (Image)	Expected Result (Position)	Actual Result (Position)	Time in Seconds
Sample 1	Hunchback	Hunchback	9.21453
Sample 2	Hunchback	Hunchback	9.80878
Sample 3	Hunchback	Hunchback	9.56103
Sample 4	Reclined	Reclined	8.6486
Sample 5	Reclined	Reclined	8.51954
Sample 6	Reclined	Reclined	8.89083
Sample 7	Straight	Straight	8.1843
Sample 8	Straight	Straight	8.41329
Sample 9	Straight	Straight	9.71633

Results for CNN, RCNN are shown in Table. III and Table. IV. Based on the result, the CNN has more latency period in detecting the posture than the time taken by the RCNN network. The key point detection for the image and output of the posture detection are as follows:

TABLE III. OUTPUT FOR CNN

Position	Key point detected Output Image	Output of CNN
Hunchback		processing time1 is 58.31735 secs processing time2 is 153.17488 secs Person is in Hunchback position! Incorrect Spine Posture
Straight		To enable them in other operations, rebuild TensorFlow er flags. processing time1 is 69.10063 secs processing time2 is 151.47656 secs Person is in Stright position! Correct Spine Posture
Reclined		processing time1 is 52.14701 secs processing time2 is 94.57633 secs Person is in Reclined position! Incorrect Spine Posture

TABLE IV. OUTPUT FOR RCNN

Position	Key point detected output image	Output of RCNN
Hunchback		start processing... C:\Users\91988\anaconda3\lib\site-packages\torch\fun the indexing argument. (Triggered internally at ..\ return _VF.meshgrid(tensors, **kwargs) # type: ig processing time2 is 9.56103 Degree is: 50 Person is in Hunchback position! Incorrect Spine Posture
Straight		C:\Users\91988\anaconda3\lib\site-packages\torch\func pass the indexing argument. (Triggered internally at return _VF.meshgrid(tensors, **kwargs) # type: ign processing time2 is 8.41329 Angle deviated from correct posture: [illegible] rad Degree is: 78 Person is in Straight position! Correct Spine Posture
Reclined		start processing... :\Users\91988\anaconda3\lib\site-package he indexing argument. (Triggered interna return _VF.meshgrid(tensors, **kwargs) processing time2 is 8.64860 Degree is: 114 Person is in Reclined position! Incorrect Spine Posture

VI. CONCLUSION

The spinal posture of the person while working with computer have been analysed using the CNN and RCNN network. A comparative study has been plotted between the time taken by both the networks for detecting the posture. Based on the study for multiple images, it is clear that the RCNN have minimal latency for detecting the person position when compared to CNN (see Tab A and Tab B). This software supports to detect key points of many people, but posture analysis can be done only for one person at a time. This can be extended by iterating the detection part of the code for overall key points that are detected on each individual in the input. The research can further be extended by taking other type of Neural Networks and compare their accuracy and latency in detecting the posture.

ACKNOWLEDGMENT

I wish to acknowledge the help provided by Biomedical department and my college. I would like to show my gratitude to my guides for their guidance and support in the project phase.

REFERENCES

[1] Q. Hu, X. Tang and W. Tang, "A Smart Chair Sitting Posture Recognition System Using Flex Sensors and FPGA Implemented Artificial Neural Network," in IEEE Sensors Journal, vol. 20(14), pp. 8007-8016, 2020, doi: 10.1109/JSEN.2020.2980207.

[2] Arif Reza Anwary, Deniz Cetinkaya, Michael Vassallo, Hamid Bouchachia, "Smart-Cover: A real time sitting posture monitoring system," Sensors and Actuators A: Physical, vol. 317, pp.112451, 2021.

[3] R. Hasib, K. N. Khan, M. Yu and M. S. Khan, "Vision-based Human Posture Classification and Fall Detection using Convolutional Neural Network," 2021 International Conference on Artificial Intelligence (ICAI), Islamabad, Pakistan, pp. 74-79,2021. doi: 10.1109/ICAI52203.2021.9445263.

[4] Gupta, Jahnvi; Gupta, Nitin; Kumar, Mukesh; Duggal, Ritwik, "An Introduction to IoT Based Posture Detection and Underlying Sensor Technology," TechRxiv. Preprint, pp.1-21,2020. https://doi.org/10.36227/techrxiv.13466597.v1.

[5] Licciardo GD, Russo A, Naddeo A, Cappetti N, Di Benedetto L, Rubino A, Liguori R, "A Resource Constrained Neural Network for the Design of Embedded Human Posture Recognition Systems," Applied Sciences, vol. 11(11), 4752, 2021 https://doi.org/10.3390/app11114752.

[6] Zhang, Y., Zhu, T., Ning, H. et al. "Classroom student posture recognition based on an improved high-resolution network," J Wireless Com Network, vol. 140, 2021. https://doi.org/10.1186/s13638-021-02015-0.

[7] Piñero-Fuentes E, Canas-Moreno S, Rios-Navarro A, Domínguez-Morales M, Sevillano JL, Linares-Barranco A, "A Deep-Learning Based Posture Detection System for Preventing Telework-Related Musculoskeletal Disorders," Sensors, 21(15):5236, 2021, https://doi.org/10.3390/s21155236.

[8] U Ran, Cong Wang, Yao Xiao, Xuliang Gao, Zhiyuan Zhu, Bin Chen, "A portable sitting posture monitoring system based on a pressure sensor array and machine learning," Sensors and Actuators A: Physical, vol. 331, 112900, 2021, https://doi.org/10.1016/j.sna.2021.112900.

[9] S. Liaqat, K. Dashtipour, K. Arshad, K. Assaleh and N. Ramzan, "A Hybrid Posture Detection Framework: Integrating Machine Learning and Deep Neural Networks," in IEEE Sensors Journal, vol. 21, no. 7, pp. 9515-9522, 2021, doi: 10.1109/JSEN.2021.3055898.

[10] Ahmad J, Sidén J, Andersson H, "A Proposal of Implementation of Sitting Posture Monitoring System for Wheelchair Utilizing Machine Learning Methods," Sensors, vol. 21(19), 6349, 2021, https://doi.org/10.3390/s21196349.

[11] Silvia Caggiari, Peter R. Worsley, Sarah L. Fryer, Joseph Mace and Dan L. Bader, "Detection Of Posture and Mobility in Individuals at Risk of Developing Pressure Ulce," Rs Elsevier Medical Engineering and Physics. Vol. 91, pp: 39–47, 2021.

[12] Cai W, Zhao D, Zhang M, Xu Y, Li Z, "Improved Self-Organizing Map-Based Unsupervised Learning Algorithm for Sitting Posture Recognition System," Sensors, vol. 21(18), pp. 6246, 2021, https://doi.org/10.3390/s21186246.

[13] Weili Ding, Bo Hu1, Han Liu, Xinming Wang and Xiangsheng Huang, "Human Posture Recognition Based on Multiple Features and Rule Learning," International Journal of Machine Learning and Cybernetics. vol. 11(11), pp:2529–2540, 2020.

[14] Luna-Perejón F, Montes-Sánchez JM, Durán-López L, Vazquez-Baeza A, Beasley-Bohórquez I, Sevillano-Ramos JL, "IoT Device for Sitting Posture Classification Using Artificial Neural Networks. Electronics," Sensors, vol. 10(15), 1825, 2021. https://doi.org/10.3390/electronics10151825.

[15] Wan Q, Zhao H, Li J, Xu P, "Hip Positioning and Sitting Posture Recognition Based on Human Sitting Pressure Image," Sensors, vol. 21(2), 426, 2021. https://doi.org/10.3390/s21020426

[16] T. -H. Tran, D. T. Nguyen and T. Phuong Nguyen, "Human Posture Classification from Multiple Viewpoints and Application for Fall Detection," 2020 IEEE Eighth International Conference on Communications and Electronics (ICCE), Phu Quoc Island, Vietnam, pp. 262-267, 2021. doi: 10.1109/ICCE48956.2021.9352140.

[17] Klishkovskaia T, Aksenov A, Sinitca A, Zamansky A, Markelov OA, Kaplun D, "Development of Classification Algorithms for the Detection of Postures Using Non-Marker-Based Motion Capture

Systems," Applied Sciences, vol. 10(11), 4028, 2020. https://doi.org/10.3390/app10114028.

[18] Wafaa M. Salih Abedi, Dr. Ibraheem Nadher and Dr. Ahmed T. Sadiq, "Modification Of Deep Learning Technique for Face Expressions and Body Postures Recognitions," International Journal of Advanced Science and Technology, vol. 29 (3). pp. 313-320, 2020.

[19] R. Gupta, D. Saini and S. Mishra, "Posture detection using Deep Learning for Time Series Data," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2020, pp. 740-744, doi: 10.1109/ICSSIT48917.2020.9214223.

[20] Wafaa M. Salih Abedi, Dr. Ibraheem Nadher and Dr. Ahmed T. Sadiq, "Modified Deep Learning Method for Body Postures Recognition," International Journal of Advanced Science and Technology. vol. 29, No. 2, (2020), pp. 3830 – 3841, 2020.

[21] Shumei Zhang And Victor Callaghan,"Real Time Human Posture Recognition Using an Adaptive Hybrid Classifier," International Journal of Machine Learning and Cybernetics, vol. 12, pp.489–499, 2020.

[22] H. Zhang, S. Shui, Y. Wu, Q. Yang and Y. Cai, "Posture Recognition Based on the Improved Optical Diffractive Neural Network," 2020 IEEE 3rd International Conference on Electronic Information and Communication Technology (ICEICT), Shenzhen, China, pp. 783-785, 2020, doi: 10.1109/ICEICT51264.2020.9334358.