

## Article

# A Resource Constrained Neural Network for the Design of Embedded Human Posture Recognition Systems

Gian Domenico Licciardo , Alessandro Russo , Alessandro Naddeo , Nicola Cappetti ,  
Luigi Di Benedetto , Alfredo Rubino  and Rosalba Liguori 

Department of Industrial Engineering, University of Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano, Italy; alerusso@unisa.it (A.R.); anaddeo@unisa.it (A.N.); ncappetti@unisa.it (N.C.); ldibenedetto@unisa.it (L.D.B.); arubino@unisa.it (A.R.); rliguori@unisa.it (R.L.)

\* Correspondence: gdlicciardo@unisa.it

**Abstract:** A custom HW design of a Fully Convolutional Neural Network (FCN) is presented in this paper to implement an embeddable Human Posture Recognition (HPR) system capable of very high accuracy both for laying and sitting posture recognition. The FCN exploits a new base-2 quantization scheme for weight and binarized activations to meet the optimal trade-off between low power dissipation, a very reduced set of instantiated physical resources and state-of-the-art accuracy to classify human postures. By using a limited number of pressure sensors only, the optimized HW implementation allows keeping the computation close to the data sources according to the edge computing paradigm and enables the design of embedded HP systems. The FCN can be simply reconfigured to be used for laying and sitting posture recognition. Tested on a public dataset for in-bed posture classification, the proposed FCN obtains a mean accuracy value of 96.77% to recognize 17 different postures, while a small custom dataset has been used for training and testing for sitting posture recognition, where the FCN achieves 98.88% accuracy to recognize eight positions. The FCN has been prototyped on a Xilinx Artix 7 FPGA where it exhibits a dynamic power dissipation lower than 11 mW and 7 mW for laying and sitting posture recognition, respectively, and a maximum operation frequency of 47.64 MHz and 26.6 MHz, corresponding to an Output Data Rate (ODR) of the sensors of 16.50 kHz and 9.13 kHz, respectively. Furthermore, synthesis results with a CMOS 130 nm technology have been reported, to give an estimation about the possibility of an in-sensor circuitual implementation.

**Keywords:** posture recognition; human behavior; Convolutional Neural Network; low-power digital design



check for  
updates

**Citation:** Licciardo, G.D.; Russo, A.; Naddeo A.; Cappetti, N.; Di Benedetto, L.; Rubino, A.; Liguori, R. Tiny Convolution Neural Network for the Design of Embedded Human Posture Recognition Systems. *Appl. Sci.* **2021**, *11*, 4752. <https://doi.org/10.3390/app11114752>

Academic Editor: Redha Taïar

Received: 27 April 2021

Accepted: 19 May 2021

Published: 21 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The monitoring and interpretation of static and dynamic behavior of the human body are very attractive for a number of applications ranging from biomedical to industrial and automotive [1,2]. Although the capability to classify human postures can be considered a specific subset of the Human Activity Recognition (HAR) [3–5], it requires specific technological solutions, very different from HAR, and is very important in peculiar application fields. To improve the quality of life, for example, which can be significantly compromised by prolonged poor sitting and laying postures [6–8], causing serious health problems such as pressure ulcers, cervical and back diseases, and complex muscle and skeletal deformations. Professional vehicle drivers, like taxi/truck/farm tractor drivers, often suffer from Muscular-Skeletal Diseases (MSD) due to long time sitting [8,9]. Furthermore, automotive applications have been pushed by updated safety protocols of autonomous vehicles, which require that the driver postures must be monitored not only to verify their readiness to take over the control in warning situations, but also for perceived (dis)comfort [10]. Human posture while seated is one of the main parameters affecting safety and health of a sitting

person [11]; it has been demonstrated that posture changes and macro/micro movements are the first indicator of increasing discomfort or pain in time [12]. Thus, human posture monitoring can help in understanding what is happening to the driver/passengers and to apply countermeasures for reducing stress, discomfort and consequent errors while driving. Pressure at the interface between the seat and human body is widely used as a good indicator to evaluate the perceived (dis)comfort and to identify movements and postures on the seat [13].

Human Posture Recognition (HPR) has been recently implemented by using Machine Learning (ML) approaches [14,15] in conjunction with mechanical or image sensors [16–18]. In particular, recent papers deal with use of machine learning and AI for posture tracking and recognition in order to improve car drivers' safety and car-occupants' comfort. In Lin et al. [19], computer vision technology has been used to recognize the head and neck posture in order to detect drivers' drowsiness or sleepiness. Since 2010, in Wachs et al. [20], Neural Networks have been used for parts-based object detection of human body parts, both inside and outside the car. The purpose of this research was focused on techniques for identifying drivers' or pedestrians' postures for safety reasons. The need to acquire and recognize a posture with a contactless or embedded system has become one of the studied topics for future vehicle development. In 2017, Loeb et al. [21] used Kinect™ in order to track the human segments of car occupants (especially very young occupants) in order to recognize occupant posture and improve safety in case of an accident. In 2020, Zhao et al. [22] proposed a head pose estimation method based on deep learning applied to images to demonstrate that the head pose could be used as the basis for distraction detection.

Although image-based HPR systems have been favored by the advancements of image processing methods of recent years, they introduce important issues related to the privacy of people captured by the camera, the poor performances when the subject is partially occluded, as well as the cost of the systems. On the other hand, pressure sensors not only allow the design of much more compact systems which can be embedded in specific supports like chairs and beds, but they are also able to capture small body deformations more accurately than cameras [23]. Indeed, for the study of the posture of the human body on a mattress, a pressure pad between body and mattress is often used whose data is analyzed in different ways [17,24] to prevent the formation of bedsores in long-term patients. The same approach [25] can be used in the case of the sitting posture since contact pressure is the only way to support body weight and influence posture and comfort [26]. However, recent works have demonstrated that the accuracy of such HPR systems strongly depends on the careful distribution of the sensors in specific key-points of the chair, bed or any other kind of support equipped with the system, in order to acquire data from as many body parts as possible and avoid an excessive number of sensors [27]. Therefore, the reliability of pressure-based HPR systems appears much too dependent on the shapes of the specific supports. Moreover, despite of the reduced number of sensors, the overall extension of the system is not negligible, considering the processing unit and the connections between this and the sensors. On the other hand, other solutions exhibit a computational load scarcely compatible with embedded systems with autonomous power supplies [3–5].

In this paper, a custom HW design of a tiny Fully Convolutional Network (FCN) is presented to implement a HPR system, which better combines high recognition accuracy and low-energy and low-area requirements with respect to the existent literature, in order to extend the application range. Indeed, the FCN achieves state-of-the-art recognition accuracy both for laying and sitting postures, by exploiting only pressure sensors grouped in a small area close to the FCN, according to the edge-computing paradigm [28,29], without any particular distribution strategy. The FCN implements an end-to-end classification by exploiting a base-2 quantization scheme for weights and binarized activations [30,31] to meet the optimal trade-off between high recognition accuracy, the number of mapped physical resources and low power consumption [32,33]. The FCN achieves an average

accuracy of 96.77% and 98.88% to classify laying and sitting postures, respectively. The main advantages of the proposed system over the existent literature can be summarized in the following points:

- the capability of the FCN to achieve high recognition accuracy by monitoring only the footprint of the human body in a limited space region covered by a reduced number of pressure sensors.
- Any sensor placement strategy is unnecessary, namely the system reliability is not dependent on the specific support.
- FCN can be easily reconfigured to different applications. Case studies are presented on laying and sitting postures recognition.
- FCN provides end-to-end classification by using a quantization scheme that overcomes binarized and ternary counterparts in terms of accuracy and meets the optimal trade-off between accuracy and employed physical resources for HW implementation.

Implemented on a small Xilinx Artix 7 FPGA, FCN dissipates 10.40 mW dynamic power and achieves a maximum operation frequency of 26.6 MHz, corresponding to sensors with Output Data Rate (ODR) of 9.13 kHz, when used for laying posture recognition. When used for sitting posture recognition, the FCN is reconfigured to use less physical resources and achieves 6.88 mW dynamic power dissipation and a maximum operation frequency of 47.64 MHz, compatible with a sensor Output Data Rate (ODR) of 16.50 kHz, which is very important for critical applications requiring a continuous monitoring and a real-time action in an emergency. In order to explore the possibility to embed the proposed accelerator for in-sensor circuitry, analysis with a conservative TSMC LP-HVT CMOS 130 nm technology has been done, which is compatible with that of the glue logic in modern MEMS. Synthesis results, by using the Cadence toolchain return a power dissipation of 425  $\mu$ W/MHz and 1.7 mW, respectively, at the maximum operating frequency of 40 MHz, and an area occupation 1.78 mm<sup>2</sup> when the FCN is configured for laying posture recognition, which support in real-time ODR up to 8.7 kHz.

The remainder of the paper is organized as follows: Section 2 describes the proposed models; design choice and architecture of the HW accelerator are discussed in Section 3; implementation results are presented in Section 4; comparisons with the state-of-the-art are discussed in Section 5; Section 6 concludes the paper.

## 2. The Proposed System and the Underlying Model

The HPR system has been designed according to the scheme in Figure 1. The FCN processes data from commercial pressure sensors and classifies them in a number of classes depending on the specific application. As a case study, a Medilogic<sup>®</sup> Seat Pressure carpet has been used for sitting posture classification of 8 classes, while data from a public dataset [34], obtained with a quite similar acquisition system, have been used for lying posture classification of 17 classes. The FCN can be reconfigured for the two applications by easily adapting the input and output layers to the different number of input sensors and output classes, respectively.

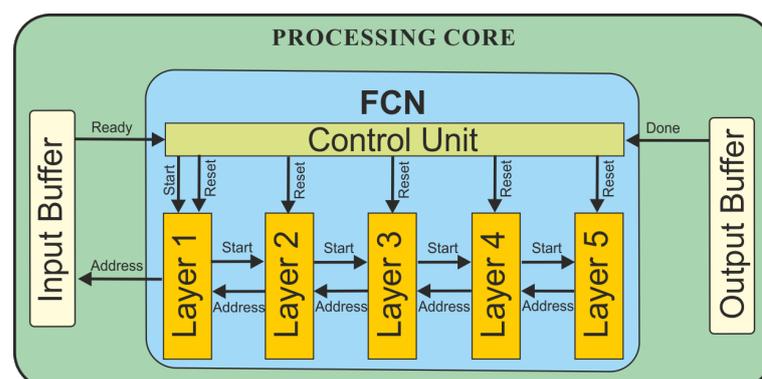
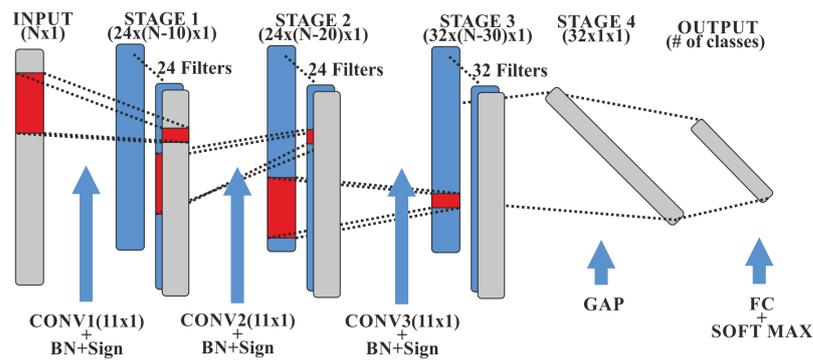


Figure 1. Scheme of the HPR system.

### 2.1. The FCN Model

The FCN is schematized in Figure 2. It is composed by 3 convolutional (CONV) layers, a Global Average Pooling (GAP) layer and a dense fully convolutional layer. In order to reduce the number of physical resources for the HW implementation of the network, all the weights of the CONV layers have been quantized. As it will be shown in the next section, the conventional binary and ternary quantization schemes have given an unacceptable low accuracy for laying posture recognition. Therefore, a quantization scheme has been introduced, which exploits weights from the set  $\{-2, -1, 0, +1, +2\}$  in place of  $\{-1, 0, +1\}$  of ternary and  $\{-1, +1\}$  of binarized neural networks, selected according to the following criteria:



**Figure 2.** Model of the FCN. All the stages have been quantized with a 3Bit quantization scheme, taking values in the set  $\{-2, -1, 0, 1, 2\}$ . A 16-bits fixed point format is assumed as input.

$$f(x) = \begin{cases} -2 & \text{if } x < -1.5 \\ -1 & \text{if } -1.5 \leq x \leq -0.01 \\ 0 & \text{if } |x| < 0.01 \\ +1 & \text{if } 0.01 \leq x \leq 1.5 \\ +2 & \text{if } x > 1.5 \end{cases} \quad (1)$$

Additionally, all the activations have been binarized, and the activation functions have been reduced to [4]:

$$y = \text{sign}(x) = \begin{cases} -1 & \text{if } x < 0 \\ +1 & \text{if } x \geq 0 \end{cases} \quad (2)$$

The advantages of this choice with respect to a full precision implementation can be roughly estimated in reducing memory requirements by about 1 order of magnitude (a factor of 32/3 for quantized weights and 32 for binarized activations) since weights are coded with 3 bits and activations with 1 bit; Multiply-Accumulate (MAC) operations, typically required to implement the convolutions, are simplified into Shift-Accumulate (SAC) operations for the absence of floating point (FP) multiplications and the consequent reduction of about 2 orders of magnitude in the number of FPGA LUTs. Each CONV layer in Figure 2 is followed by a Batch Normalization (BN) layer, where, as schematized in Figure 3 with more details, each sample is scaled by a factor  $\sigma$  and subtracted by the mean value  $\mu$ , defined during the training and stored in a devoted memory.

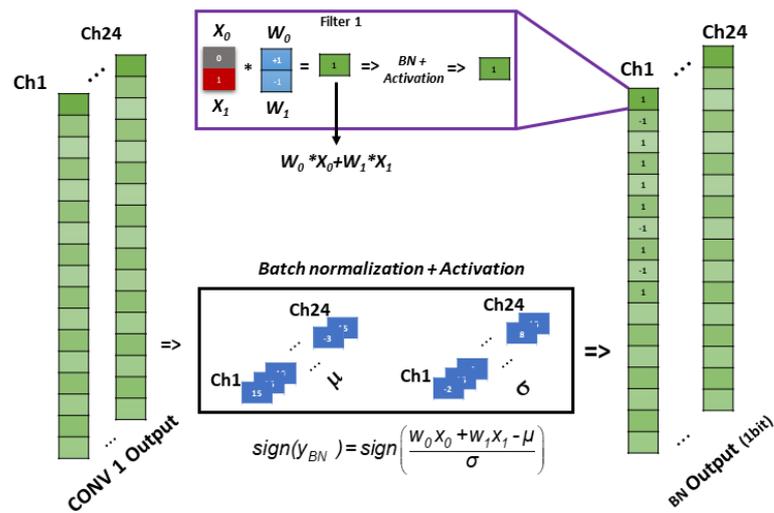


Figure 3. Operations of the BN and Activation functions.

No padding has been used. The fourth stage is made up of a Global Average Pooling (GAP), which is very robust to translations of the inputs and enables the Class Activation Map (CAM), which, together with the SoftMax, provides the final classification with less resources than typical dense layers of CNNs [35]. The GAP also reduces the dimensions of the network with respect to a conventional MaxPool since it transforms N-30 inputs to 1 and reduces the complexity of the following stage. The last stage is composed of a Fully Connected (FC) layer and a SoftMax classifier. The output of this last stage represents the probability of belonging to each output class, therefore the number of units of the fully connected corresponds to the number of considered classes. As shown in Table 1, the computational complexity of each layer in terms of required math operations and memory requirements changes depending on the specific application. When the FCN is used for laying posture recognition, it receives 108 input samples coded with 12 bits, representing a snapshot of the posture. The first and second CONV stages are composed by 24 one-dimensional filters of length 11, and the third by 32 filters. Depending on the number of classes to be considered, the output layer produces 17 and 8 values for laying and sitting postures, respectively.

Table 1. Complexity of the proposed FCN.

Layer	N° Parameters	Operations per Windows	N° of Bits
CONV 1	11 × 24	SAC: (* N – 10) × 11 × 24	11 × 24 × 3
NORM 1	2 × 24	ADD: (* N – 10) × 24	24 × 5 + 24
CONV 2	11 × 24 × 24	SAC: (* N – 20) × 11 × 24 × 24	11 × 24 × 24 × 3
NORM 2	2 × 24	ADD: (* N – 20) × 24	24 × 10 + 24
CONV 3	11 × 24 × 32	SAC: (* N – 30) × 11 × 24 × 32	11 × 24 × 32 × 3
NORM 3	2 × 32	ADD: (* N – 30) × 32	32 × 10 + 32
GAP	0	ADD: (* N – 30) × 32	0
SOFTMAX	32 × Classes **	SAC: 32 × Classes **	32 × Classes × 3

\* n = 108 is the dimension of the input window of the laying dataset. n = 56 is the dimension of the input window of the sitting dataset. \*\* Classes = 17 for the laying dataset. Classes = 8 for the sitting dataset.

### 2.2. FCN Training and Accuracy Results

Keras and Larq tools have been used to describe the FCN model. In order to prove the performance of the system in two of the most interesting contexts for HPR in biomedical and industrial application fields, two datasets have been employed for lying and sitting posture recognition, respectively. The public PmatData dataset in Table 2 has been specifically designed for in-bed posture classification [35]. The pressure data have been collected by using a Vista Medical FSA SoftFlex 2048, equipped with 2048 1 inch<sup>2</sup> pressure sensors

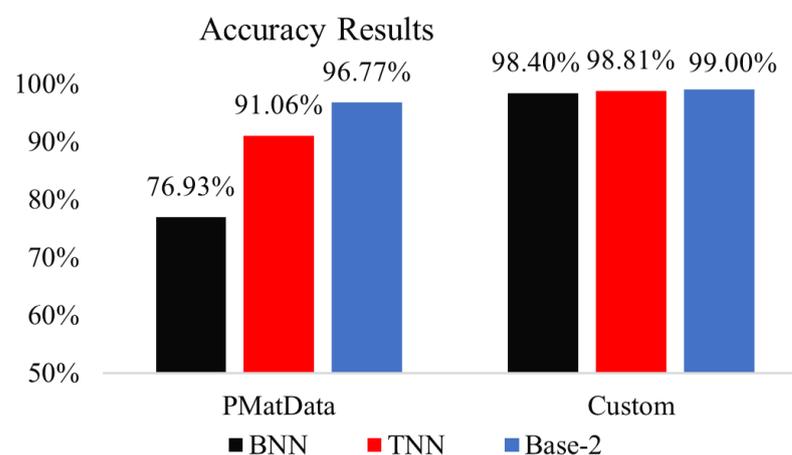
placed on a  $32 \times 64$  grid. The sensors provide output values coded with 12 bits and normalized in the range  $[0, 1]$ .

**Table 2.** Characteristics of the datasets.

Dataset	PMatData:Laying Posture	Custom:Sitting Posture
Number of classes	17	8
Available classes	Supine (9 types), right, right (30°),right (60°), right fetus, left, left (30°), left (60°), left fetus.	Initial position, bent forward, rested back,bent left, legs up, right-bent thinker, straight legs, left-bent thinker
Training Set	15,232	6720
Test set	1692	1680

The dataset is composed of 17 postures listed in Table 2, sampled at 1 Hz and taken from 13 participants whose physical characteristics range in the intervals:  $[19, 34]$  years for ages,  $[170, 186]$  cm for heights and  $[63, 100]$  kg for weights.

The Sparse\_categorical\_crossentropy loss function has been used for training, set with 100 epochs, a batch size of 20 and a learning rate of  $5 \times 10^{-4}$ . Initial tests on the dataset showed that 2048 input samples provided for each acquisition are an unnecessary oversampling of the body footprint, which only increases complexity of the input layer without any evident advantages in terms of accuracy. The number of the inputs has been reduced to 108 by a downsampling of about 1:19, empirically determined as the best trade-off between the HW complexity of the input layer and the overall classification accuracy. The effects on the mean classification accuracy of the FCN of the binary (BNN) and ternary (TNN) quantization schemes for weights are shown in Figure 4 and compared with the Base-2 defined from Equation (1). Table 3 reports the main test results when a 10-fold cross validation has been used. The positions “Supine 1-4” are supine postures with different body attitudes: legs and arms more or less spread, cozy position, straddling left and right leg.

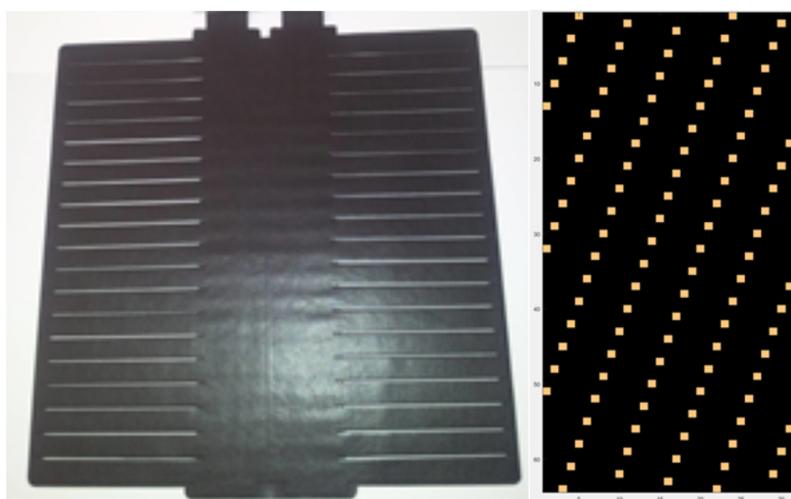


**Figure 4.** Comparisons between binary, BNN, ternary, TNN, and the proposed base-2 quantization in terms of accuracy. The models have been trained and tested on all the classes of the datasets. The accuracy has been estimated by using  $k$ -fold cross validation with  $k = 10$  for the PMatData and  $k = 5$  for the custom dataset.

In order to prove the effectiveness of the FCN to classify sitting postures, we built a custom dataset by using the Medilogic® Seat Pressure Measurement System in Figure 5 to train and test the FCN, considering that, based on our knowledge, there is no public dataset available for the purpose. The measurement system is composed by a carpet of 480 piezoresistive sensors distributed on a matrix of  $24 \times 20$  elements. The commercial

measurement system has been chosen to make reliable acquisitions, but the number of sensors is suitable for different applications and, also in this case, it is excessive for the FCN operations. Namely, a subsampling was applied to reduce the number of sensing elements to 56, which coincides with the number of inputs to the first layer. In Figure 5, also the sampling scheme is shown. Results of Figure 4 and Table 4 prove that binarization in the case of sitting posture recognition could be a sufficient quantization scheme.

In this paper, we chose to maintain the base-2 quantization for the better results in other applications. However, for fair comparisons between HW implementations, also binarized version of the FCN will be considered in the following. Since there are no studies on standard postures and all papers propose different approaches, usually depending on the chair examined and the type of analysis that is carried out, in Table 4, postures have been considered, the combination of which makes it possible to obtain plausible postures for many activities and for many types of chairs. However, the interaction with other objects such as a desk, armrests or steering wheel was not considered. The posture of a seated person involves the inclination of the trunk (supported, erect, inclined forward or sideways (left and right)).



**Figure 5.** Medilogic Seat Pressure Measurement System; Right: the imposed sampling scheme. The clear dots are the selected sensors.

**Table 3.** Accuracy results of the FCN for all the laying postures.

Class	Precision	Recall	F1-Score	Support
Supine	1.00	0.95	0.97	100
Right	0.92	1.00	0.96	102
Left	0.96	0.97	0.96	96
Right 30° (1 wedge)	0.99	1.00	1.00	114
Right 60° (2 wedges)	1.00	1.00	1.00	96
Left 30° (1 wedge)	1.00	0.95	0.97	99
Left 60° (2 wedges)	0.97	0.98	0.97	88
Supine 1	0.94	1.00	0.97	102
Supine 2	0.98	0.99	0.99	120
Supine 3	0.99	1.00	0.99	88
Supine 4	0.99	0.98	0.99	117
Supine 5	1.00	0.99	0.99	98
Right Fetus	1.00	0.96	0.98	90
Left Fetus	1.00	0.98	0.99	99
Supine (30°)	0.97	0.97	0.97	101
Supine (45°)	0.97	0.89	0.93	88
Supine (60°)	0.94	1.00	0.97	94

**Table 4.** Accuracy results of the FCN for all the sitting postures.

Class	Precision	Recall	F1-Score	Support
Initial position	0.99	0.99	0.99	839
Bent forward	0.99	0.98	0.98	128
Rested back	1	0.98	0.99	123
Bent left	0.98	1	0.99	125
Left-bent thinker	1	1	1	116
Right-bent thinker	1	0.98	0.99	114
Legs up	0.98	0.99	0.99	120
Straight legs	1	1	1	115

Since the weight of the trunk and head can be partially discharged by placing the arms on a desk or armrests, the position of the thinker which foresees the elbow on the knees has been considered; the legs can be rested on or raised in the case of subjects of small stature. In total, 8400 samples have been acquired. Training has been done with a  $k$ -fold cross validation with  $k = 5$  on 6720 samples and tested on a sub-set of 1680 samples.

### 3. System Design

The HW architecture of the FCN follows the scheme in Figure 1. It is composed of 5 sequential layers and a simple control logic, which initiates and terminates the processing and initializes the memories embedded into the layers. The layer operations are the same of the model in Figure 2: the first three are CONV layers, followed by the GAP and a dense FC layer. The last layer differs from the model because the SoftMax classifier has been substituted by a simple prediction stage, which only selects the maximum output values from the FC layer. This has been possible since the combination of GAP and FC layers calculates the scores of the output classes as:

$$\begin{aligned}
 \text{Score}(C) &= \sum_{i=1}^{32} W_{i,C} \times X_i = \sum_{i=1}^{32} W_{i,C} \times \frac{\sum_{k=1}^{N-30} b_k}{N-30} = \\
 &= \frac{1}{N-30} \left( \sum_{i=1}^{32} W_{i,C} \times \left( \sum_{k=1}^{N-30} b_k \right) \right)
 \end{aligned} \tag{3}$$

where  $C$  is the class,  $W_{i,C}$  are the weights for each class and  $b_k$  are the binarized inputs to the GAP and the other quantities can be taken from Table 1. Given the linearity of Equation (3), and the only interest toward the class with the highest score, which can be obtained by a trivial comparison between the scores of all the classes, not only the SoftMax but also the division in Equation (3) can be avoided for a more compact HW implementation of the FCN, without loss of accuracy. FCN is fully synchronous, namely the layers exchange data after a fixed number of cycles. All the layers share the architecture in Figure 6, composed by the control unit (CU), the memories to store the kernel coefficients and the outputs of the BN and Operational (OP) Block whose dimensions can be obtained by Table 1 once  $N$  and the number of classes have been selected. The OPBlocks of CONV and FC layers are schematized in Figure 7. Thanks to the base-2 quantization, the OP block is free of multipliers and it has been essentially reduced to SAC operators, which are in turn composed by multiplexers and shifters.

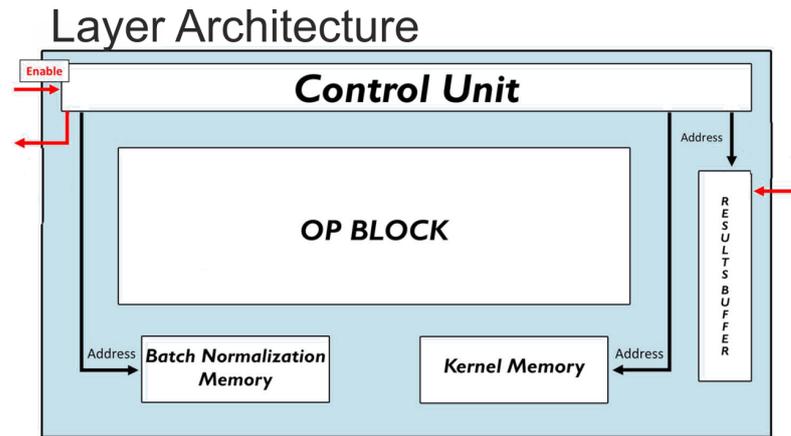


Figure 6. Block scheme of the layers.

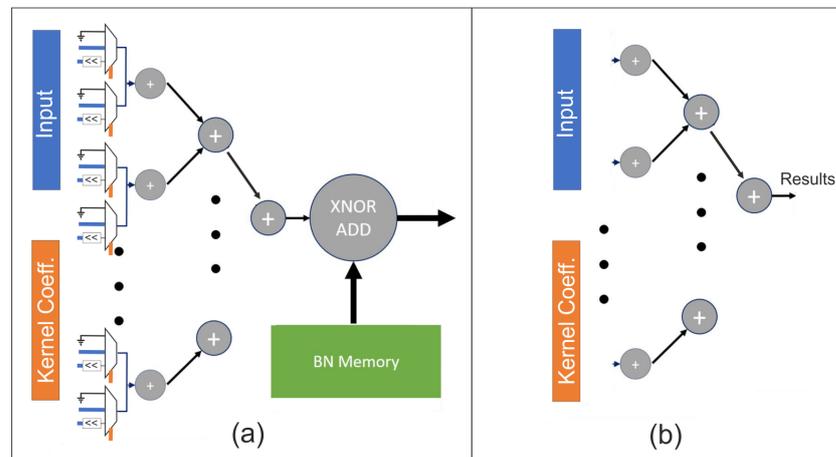


Figure 7. (a): Architecture of the OPBlock of Figure 3 in the case of CONV and FC Layers. Mux selects “0”, the input or the left-shifted input to emulate the products with weights. Mux with Adder tree compose the SAC operator. (b): PopCount tree for the OPBlock of Figure 3 in the case of GAP layer.

Muxes are used to select 0, the input value or the left-shifted input value to emulate the product between inputs and quantized weights, according to Equation (1). The CONV1 layer differs from the other two CONV layers for the number of the operators, considering the different number of filters between the first three layers, as well as for the dimensions of the input data which depend on the acquisition system, and that in our case is of 12 bits. The remaining parts are the same of Figure 7. BNs and activations are simply implemented by a XNOR gate considering that:

$$sign(Y_{BN}) = sign\left(\frac{ConvResult - \mu}{\sigma}\right) = sign(ConvResult - \mu) \text{ XNOR } sign(\sigma) \quad (4)$$

where  $\mu$  and  $\sigma$  are calculated during training and stored in a proper memory. Considering that activations are binarized, the GAP scheme is a simplified version of the one in Figure 7a. It is composed by a popcount and a memory buffer for results as reported in Figure 7b, where signed additions are calculated as in Figure 8.

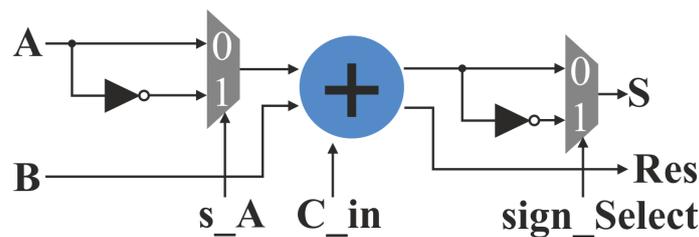


Figure 8. Adder for binarized additions.

Quantization means that logical and arithmetic operators are very compact and require few physical resources to be implemented. Therefore, resources are largely dependent on the memories for bias, activations and weights. Although, memory requirements should depend on the specific configuration, sitting or laying posture recognition in our cases, the amount of instantiated HW resources for the FCN is lower-bounded by the larger laying posture configuration, as results from the data in Table 1. Memory requirements of the CONV layers, in this configuration, are 0.11 kB, 2.35 kB and 3.14 kB for the first, second and third layer, respectively. The entire architecture requires 5.80 kB.

#### 4. Synthesis and Implementation Results

The proposed design has been implemented on a small Digilent CMOD A35T, equipped with a Xilinx Artix-7 (xc7a35tfgg484-1) FPGA by using the Vivado IDE suite. The dimensions of the resulting systems are defined by the sensor carpet considering that the FPGA board is about  $2 \times 8$  cm, therefore the overall dimensions of the systems are very compact.

The most interesting results of the FPGA implementation results are reported in Table 5 for both sitting and laying posture recognition configurations. For sitting posture recognition, our design requires 10,983 LUTs and 8424 FFs, in turn, 15,802 LUTs and 12,287 FFs are required for laying posture recognition. A resource reduction of about 14% both in the number of LUTs and FFs is obtained by imposing the use of BRAMs and DSPs in the synthesis tool. Our choice to impose the absence of these hard macros in the synthesis of our design is due to provide implementation results as much as independent from the specific FPGA topologies, which could significantly differ in the numbers and capabilities of the embedded macros. It is worthwhile to underline that a consistent reduction of the mapped resources, roughly about 66%, could be obtained by exploiting the similarities between the architectures of the layers, and implementing an iterative topology around a superset of a single CONV layer. However, considering that the unrolled architecture also fits well the small FPGA used for tests, the proposed implementation returns a much higher speed performance. In particular, the proposed design achieves a maximum operation frequency of 26.6 MHz and 47.64 MHz for laying and sitting posture configuration, respectively. Considering that the unrolled configuration completes the processing of an input set in 2920 and 2880 clock cycles, respectively, for the two configurations, sensors with 9.13 kHz and 16.50 kHz Output Data Rate (ODR) are supported, (processing time of 109  $\mu$ s and 60  $\mu$ s). Although for some applications, like sleep monitoring, the above ODRs could be unnecessarily high, other critical applications, like driver monitoring and situation awareness, take real advantages from our design choices. The proposed system meets state-of-the-art performance also in terms of power dissipation, which is a very relevant parameter considering the large number of possible embedded applications of HPR systems. At the maximum operation frequency the FCN dissipates 10.4 mW and 6.88 mW for lying and sitting posture recognition configurations, respectively. Namely, 391  $\mu$ W/MHz and 144.4  $\mu$ W/MHz. Considering that conventional human activity recognition systems operates at 50 Hz [4,5], a dynamic power dissipation less than 1 mW can be considered, which is not sensed by the Xilinx tool since it is much lower than the 70 mW of the quiescent power dissipation of the employed FPGA.

**Table 5.** FPGA implementation results of the of proposed FCN.

	Proposed Sitting	Proposed Laying
Neural Network	FCN	FCN
Network Complexity	3 CONV+GAP+FC	3 CONV+GAP+FC
N° Classes	8	17
Mean Accuracy	98.81%	96.77%
Target Platform	Artix-7	Artix-7
N° sensors	56	108
Sensing area [mm]	300 × 200 carpet	320 × 200 carpet
Dynamic Power [ $\mu$ W/MHz]	144	391
Dyn.Power @MaxFreq [mW]	10.4	6.88
Tot. Power @OpFreq [mW]	72	72
N° LUTs	15,802	10,983
N° FFs	12,287	8424
N° DSPs	0	0
N° BRAMs	0	0
Max Freq. [MHz]	47.64	26.6
Max Sensor ODR [kHz]	16.50	9.13
Delay @ Max Freq [ $\mu$ s]	60	109

## 5. Comparison with the Literature

Comparison with the current literature is not a trivial operation because there are no other systems which work well with both laying and sitting postures recognition. Moreover, custom HW implementations are tailored for a specific application or for a specific support, obviously, and require less physical resources than are proposed. This happens, for example, with the sitting posture recognition in [27], specifically designed to use six flexible sensors applied at the armrests, backrest and seat of a chair, in conjunction with a very compact two-layer Artificial NN (ANN) to classify seven sitting positions, representing the state-of-the-art solution for this specific problem. However, although the Artificial NN in [27] requires less physical resources than the FCN, (755 slice reg., 1822 FFs and 649 LUTs), it consumes more power (7.33 mW) with a much higher processing time of 267.5  $\mu$ s to classify one posture less than the proposed one with an average accuracy of 97.43%.

With reference to laying posture recognition, the proposed FCN obtains an average accuracy value of 96.77% to classify 17 laying postures in real-time exploiting 108 sensors, and a throughput of 9.13 kHz. The state-of-the-art in this case is represented by the recent work in [36], dealing with in-bed posture recognition, exploiting a microcontroller unit to implement a very complex ResNet composed by 17 CONV layers, two MaxPool and three FC layers, in order to obtain an average classification accuracy of 95.08% by using 1024 force sensitive resistor sensors.

In order to explore the possibility to embed the proposed accelerator for in-sensor circuitry, analysis with a conservative TSMC LP-HVT CMOS 130 nm technology has been done, which is compatible with that of the glue logic in modern MEMS. Synthesis results, by using the Cadence tool-chain of the larger configuration for lying posture recognition, return a power dissipation of 425  $\mu$ W/MHz and 1.7 mW, respectively, at the maximum operating frequency of 40 MHz, and an area occupation 1.78 mm<sup>2</sup> and which support in real-time an ODR up to 8.7 kHz. All the above results overcome the state-of-the-art for this kind of system.

## 6. Conclusions

In this work, a new FCN has been designed to implement HPR operations. The design exploits 2-base quantization schemes to limit the amount of mapped physical resources and achieves state-of-the-art performance in terms of power consumption and area occupation. The FCN has been tested with datasets for sitting and laying posture recognition. In both applications, state-of-the-art performance and an adaptation capability which largely

overcomes the existent solutions were demonstrated. The compactness of the design has also suggested a perspective ASIC implementation, encouraged by the synthesis results with a CMOS 130 nm technology. Future improvements will be aimed at the reduction of the overall area required by the sensor array, which could limit the application of the proposed system when used on small supports.

**Author Contributions:** All the authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Boulay, B.; Brémond, F.; Thonnat, M. Applying 3D human model in a posture recognition system. *Pattern Recognit. Lett.* **2006**, *27*, 1788–1796. [[CrossRef](#)]
- Ni, W.; Gao, Y.; Lucev, Z.; Pun, S.H.; Cifrek, M.; Vai, M.I.; Du, M. Human posture detection based on human body communication with multi-carriers modulation. In Proceedings of the 2016 39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 30 May–3 June 2016; pp. 273–276.
- De Vita, A.; Pau, D.; Di Benedetto, L.; Rubino, A.; Pétrot, F.; Licciardo, G.D. Low Power Tiny Binary Neural Network with improved accuracy in Human Recognition Systems. In Proceedings of the 2020 23rd Euromicro Conference on Digital System Design (DSD), Kranj, Slovenia, 26–28 August 2020; pp. 309–315.
- De Vita, A.; Russo, A.; Pau, D.; Di Benedetto, L.; Rubino, A.; Licciardo, G.D. A Partially Binarized Hybrid Neural Network System for Low-Power and Resource Constrained Human Activity Recognition. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2020**, *67*, 3893–3904. [[CrossRef](#)]
- De Vita, A.; Pau, D.; Parrella, C.; Di Benedetto, L.; Rubino, A.; Licciardo, G.D. Low-Power HWAccelerator for AI Edge-Computing in Human Activity Recognition Systems. In Proceedings of the 2020 2nd IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS), Genova, Italy 31 August–4 September 2020; pp. 291–295.
- Tattersall, R.; Walshaw, M. Posture and cystic fibrosis. *J. R. Soc. Med.* **2003**, *96*, 18.
- Grandjean, E.; Hünting, W. Ergonomics of posture—review of various problems of standing and sitting posture. *Appl. Ergon.* **1977**, *8*, 135–140. [[CrossRef](#)]
- O’Sullivan, K.; O’Dea, P.; Dankaerts, W.; O’Sullivan, P.; Clifford, A.; O’Sullivan, L. Neutral lumbar spine sitting posture in pain-free subjects. *Man. Ther.* **2010**, *15*, 557–561. [[CrossRef](#)] [[PubMed](#)]
- Grujicic, M.; Pandurangan, B.; Xie, X.; Gramopadhye, A.; Wagner, D.; Ozen, M. Musculoskeletal computational analysis of the influence of car-seat design/adjustments on long-distance driving fatigue. *Int. J. Ind. Ergon.* **2010**, *40*, 345–355. [[CrossRef](#)]
- Fiorillo, I.; Piro, S.; Anjani, S.; Smulders, M.; Song, Y.; Naddeo, A.; Vink, P. Future vehicles: The effect of seat configuration on posture and quality of conversation. *Ergonomics* **2019**, *62*, 1400–1414. [[CrossRef](#)]
- Cardoso, M.; McKinnon, C.; Viggiani, D.; Johnson, M.J.; Callaghan, J.P.; Albert, W.J. Biomechanical investigation of prolonged driving in an ergonomically designed truck seat prototype. *Ergonomics* **2018**, *61*, 367–380. [[CrossRef](#)]
- Fasulo, L.; Naddeo, A.; Cappetti, N. A study of classroom seat (dis) comfort: Relationships between body movements, center of pressure on the seat, and lower limbs’ sensations. *Appl. Ergon.* **2019**, *74*, 233–240. [[CrossRef](#)]
- Hiemstra-van Mastrigt, S.; Groenesteijn, L.; Vink, P.; Kuijt-Evers, L.F. Predicting passenger seat comfort and discomfort on the basis of human, context and seat characteristics: A literature review. *Ergonomics* **2017**, *60*, 889–911. [[CrossRef](#)] [[PubMed](#)]
- Wafaa, M.; Salih Abedi, E.A. Modified Deep Learning Method for Body Postures Recognition. *Int. J. Adv. Sci. Technol.* **2020**, *29*, 3830–3841.
- Han, S.H.; Kim, H.G.; Choi, H.J. Rehabilitation posture correction using deep neural network. In Proceedings of the 2017 IEEE international conference on big data and smart computing (BigComp), Jeju Island, Korea, 13–16 February 2017; pp. 400–402.
- Elforaici, M.E.A.; Chaaraoui, I.; Bouachir, W.; Ouakrim, Y.; Mezghani, N. Posture recognition using an RGB-D camera: Exploring 3D body modeling and deep learning approaches. In Proceedings of the 2018 IEEE life sciences conference (LSC), Montreal, QC, Canada, 28–30 October 2018; pp. 69–72.
- Matar, G.; Lina, J.M.; Kaddoum, G. Artificial neural network for in-bed posture classification using bed-sheet pressure sensors. *IEEE J. Biomed. Health Inform.* **2019**, *24*, 101–110. [[CrossRef](#)] [[PubMed](#)]
- Ren, W.; Ma, O.; Ji, H.; Liu, X. Human Posture Recognition Using a Hybrid of Fuzzy Logic and Machine Learning Approaches. *IEEE Access* **2020**, *8*, 135628–135639. [[CrossRef](#)]
- Lin, G.; Zhan, Z.; Peng, X.; Xu, H.; Fu, Y.; Jiang, L. *A Study of Driver’s Driving Concentration Based on Computer Vision Technology*; SAE Technical Paper; SAE International: Warrendale, PA, USA, 2020.
- Wachs, J.P.; Kölsch, M.; Goshorn, D. Human posture recognition for intelligent vehicles. *J. Real Time Image Process.* **2010**, *5*, 231–244. [[CrossRef](#)]
- Loeb, H.; Kim, J.; Arbogast, K.; Kuo, J.; Koppel, S.; Cross, S.; Charlton, J. Automated recognition of rear seat occupants’ head position using Kinect™ 3D point cloud. *J. Saf. Res.* **2017**, *63*, 135–143. [[CrossRef](#)] [[PubMed](#)]

22. Zhao, Z.; Xia, S.; Xu, X.; Zhang, L.; Yan, H.; Xu, Y.; Zhang, Z. Driver Distraction Detection Method Based on Continuous Head Pose Estimation. *Comput. Intell. Neurosci.* **2020**, *2020*, 9606908. [[CrossRef](#)]
23. Tlili, F.; Haddad, R.; Ouakrim, Y.; Bouallegue, R.; Mezghani, N. A Review on posture monitoring systems. In Proceedings of the 2018 International Conference on Smart Communications and Networking (SmartNets), Yasmine Hammamet, Tunisia, 16–17 November 2018; pp. 1–6.
24. Viriyavit, W.; Sornlertlamvanich, V.; Kongprawechnon, W.; Pongpaibool, P.; Isshiki, T. Neural network based bed posture classification enhanced by Bayesian approach. In Proceedings of the 2017 8th International Conference of Information and Communication Technology for Embedded Systems (IC-ICTES), Chonburi, Thailand, 7–9 May 2017; pp. 1–5.
25. Wang, J.; Hafidh, B.; Dong, H.; El Saddik, A. Sitting Posture Recognition Using a Spiking Neural Network. *IEEE Sens. J.* **2020**, *21*, 1779–1786. [[CrossRef](#)]
26. Cappetti, N.; Di Manso, E. Study of the relationships between articular moments, comfort and human posture on a chair. *Work A J. Prev. Assess. Rehabil.* **2021**, *68*, S59–S68.
27. Hu, Q.; Tang, X.; Tang, W. A smart chair sitting posture recognition system using flex sensors and FPGA implemented artificial neural network. *IEEE Sens. J.* **2020**, *20*, 8007–8016. [[CrossRef](#)]
28. Bianchi, V.; Bassoli, M.; Lombardo, G.; Fornacciari, P.; Mordonini, M.; De Munari, I. IoT wearable sensor and deep learning: An integrated approach for personalized human activity recognition in a smart home environment. *IEEE Internet Things J.* **2019**, *6*, 8553–8562. [[CrossRef](#)]
29. Rahimiazghadi, M.; Lammie, C.; Eshraghian, J.K.; Payvand, M.; Donati, E.; Linares-Barranco, B.; Indiveri, G. Hardware implementation of deep network accelerators towards healthcare and biomedical applications. *IEEE Trans. Biomed. Circuits Syst.* **2020**, *6*, 1138–1159. [[CrossRef](#)] [[PubMed](#)]
30. Long, X.; Zeng, X.; Ben, Z.; Zhou, D.; Zhang, M. A Novel Low-Bit Quantization Strategy for Compressing Deep Neural Networks. *Comput. Intell. Neurosci.* **2020**, *2020*, 7839064. [[CrossRef](#)] [[PubMed](#)]
31. Courbariaux, M.; Hubara, I.; Soudry, D.; El-Yaniv, R.; Bengio, Y. Binarized neural networks: Training deep neural networks with weights and activations constrained to+ 1 or-1. *arXiv* **2016**, arXiv:1602.02830.
32. Licciardo, G.D.; Cappetta, C.; Di Benedetto, L.; Rubino, A.; Liguori, R. Multiplier-less stream processor for 2D filtering in visual search applications. *IEEE Trans. Circuits Syst. Video Technol.* **2016**, *28*, 267–272. [[CrossRef](#)]
33. Licciardo, G.D.; Cappetta, C.; Di Benedetto, L.; Vigliar, M. Weighted partitioning for fast multiplierless multiple-constant convolution circuit. *IEEE Trans. Circuits Syst. II Express Briefs* **2016**, *64*, 66–70. [[CrossRef](#)]
34. Pouyan, M.B.; Birjandtalab, J.; Heydarzadeh, M.; Nourani, M.; Ostadabbas, S. A pressure map dataset for posture and subject analytics. In Proceedings of the 2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), Orlando, FL, USA, 16–19 February 2017; pp. 65–68.
35. Zhou, B.; Khosla, A.; Lapedriza, A.; Oliva, A.; Torralba, A. Learning deep features for discriminative localization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2921–2929.
36. Diao, H.; Chen, C.; Yuan, W.; Amara, A.; Tamura, T.; Fan, J.; Meng, L.; Liu, X.; Chen, W. Deep Residual Networks for Sleep Posture Recognition With Unobtrusive Miniature Scale Smart Mat System. *IEEE Trans. Biomed. Circuits Syst.* **2021**, *15*, 111–121. [[CrossRef](#)] [[PubMed](#)]