BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 70(6), 2022, Article number: e143552
DOI: 10.24425/bpasts.2022.143552

Neuromorphic approach for breathing rate monitoring using data produced by FMCW radar

Krzysztof ŚLOT[1], **Piotr ŁUCZAK**[1]*, **and Sławomir HAUSMAN**[2]

[1] Institute of Applied Computer Science, Lodz University of Technology, Poland
[2] Institute of Electronics, Lodz University of Technology, Poland

Abstract. The paper introduces a neuromorphic computational approach for breathing rate monitoring of a single person observed using a Frequency-Modulated Continuous Wave radar. The architecture, aimed at implementation in analog hardware to ensure high energy efficiency and to provide system operation longevity, comprises two main functional modules. The first one is a data preprocessing unit aimed at the extraction of information relevant to the analysis objective, whereas the second one is a pre-trained recurrent neural regressor, which analyzes sequences of incoming samples and estimates the breathing rate. To ensure compatibility with neural processing and to achieve simplicity of underlying resources, several solutions were proposed for the data preprocessing module, which provides range-wise space segmentation, selection of a bin of interest (comprising the dominant motion activity), and delivery of data to regressor inputs. To implement these functions, we introduce an appropriate chirp frequency modulation scheme, apply a neuromorphic filtering procedure and use a Winner-Takes-All network for extracting information from the bin of interest. The architecture has been experimentally verified using a dataset of indoor recordings supplied with reference data from a Zephyr BioHarness device. We show that the proposed architecture is capable of making correct breathing rate estimates while being feasible for analog implementation. The mean squared regression error with respect to the Zephyr-produced reference values is approximately 3.3 breaths per minute (with a deviation of ± 0.27 in the 95% confidence interval) and the estimates are produced by a recurrent, GRU-based neural regressor, with a total of only 147 parameters.

Key words: neuromorphic architectures; analog computing; FMCW radars; vital parameter estimation.

1. INTRODUCTION

The advent of deep learning combined with rapid progress in computational hardware and sensor manufacturing technology enabled qualitative increase in intelligent data analysis of real-world data, delivering a variety of tools and systems that support human activities. One of the application domains that benefits from these advances is the development of systems for improving security and safety of living environments, which include monitoring aimed, for example, at unusual activity detection or supervision and care of children or elderly. Visual data is the predominant source of information on the environment and multiple systems that implement a variety of different objectives have been already successfully deployed. These include for example person recognition and tracking for surveillance and security purposes [1, 2], emotion and mood recognition for improving quality of human-computer interfacing or fatigue detection for monitoring of ability to perform critical tasks [3, 4]. Action recognition is another area of rapidly growing research interest, driven by a strive to develop truly situation-aware systems [5].

Unfortunately, the choice of visual input as the primary source of information on the environment is prone to several shortcomings. The first problem arises from data complexity: rich and variable image contents need to be extracted from rich and unpredictable background, which results in the necessity of employing large computational resources for successful accomplishment of posed tasks. Another shortcoming results from deterioration of visual data quality under reduced illumination levels, making monitoring useless when dark, foggy or under presence of heavy smog. These can be partially alleviated by applying sensors operating at longer wavelengths (near or far infrared), but this increases the cost of a system and still might not provide a viable problem solution [6]. Finally, from the point of view of health condition monitoring, visual data is of little use, as it can only provide indirect information, which is insufficient for efficient operation of target systems.

Therefore, other sources of information should be considered in developing effective systems aimed at health condition monitoring. A possible alternative is to use a low-power radar to sense indoor environment dynamics. Contactless monitoring of human vital parameters, such as breathing rate (BR), could save lives of people suffering from life-threatening conditions, e.g. sleep apnea or heart/lung diseases. In these cases, continuous, unobtrusive indoor patient surveillance would be beneficial as an alternative to devices attached to a subject's body, which compromise comfort or freedom of movement. Moreover, non-contact measurement of vital signs, as opposed to the contact methods, can reduce the spread of pathogenic microorganisms,

*e-mail: pluczak@iis.p.lodz.pl

Manuscript submitted 2022-03-21, revised 2022-09-13, initially accepted for publication 2022-10-22, published in December 2022.

such as viruses or bacteria. Information acquired by a remote radar could be processed to detect tiny body movements related to the activity of the human respiratory system and used for continuous health condition monitoring.

The presented paper is concerned with development of a computational architecture for contactless monitoring of the breathing rate in an indoor space, which exploits a Frequency-Modulated Continuous Wave (FMCW) radar as a source of input data. The proposed approach assumes two data processing stages: preprocessing, aimed at extraction of information relevant to the considered task, and analysis of the delivered data, which is expected to produce BR estimates. To implement both tasks, a neuromorphic architecture, aimed at its analog implementation in Applications-Specific Integrated Circuit (ASIC) has been developed. The main reason for focusing on analog algorithm implementation is a pursuit to achieve ultra-low power consumption, which is a critical requirement to ensure low-cost, longevity and reliability of the system. However, the algorithm is also suited for its possible energy-efficient digital implementation in e.g. FPGA devices. In such a case, an additional step of analog-to-digital conversion is required. Also, it implies less energy-efficient realization of elementary arithmetic operations; however, it offers data processing accuracy that is far superior to what can be achieved in analog hardware.

To attain the posed objective, we propose a BR estimation procedure that provides an alternative to algorithms commonly used in digital signal processing. The procedure includes a space-scanning scheme that enables minimization of data preprocessing hardware complexity, asynchronous detection and tracking of living objects that employs Winner-Takes-All (WTA) scheme and deep Recurrent Neural Network (RNN) based regression used for producing BR estimates. Majority of the data transformations involved in the proposed data analysis scheme utilize basic multiply-and-sum operations augmented with nonlinear transformations of resulting values, which are at the core of neuromorphic computation.

The proposed concept have been evaluated using a publicly available dataset of recordings [7] acquired using a 77 GHz Texas Instruments IWR1443 FMCW radar applied for monitoring of a single person present in an indoor environment. We show that the presented method is capable of estimating BR with over 80% accuracy, where simulations included a limited accuracy of the parameter representation that is inevitable when using analog technology. The complexity of the architecture is low: the hardware required for data preprocessing and data analysis comprises few dozens of analog memories and involves a few hundred parameters (weights), depending on the applied network model, making implementation of the algorithm feasible.

The paper has the following structure. A brief review of work in two relevant research fields – vital parameter monitoring and analog neuromorphic architectures, has been presented in Section 2. The proposed concept for BR estimation using a neuromorphic computational approach is elaborated on in Section 3 and the results of the experimental evaluation of the method have been provided in Section 4.

2. RELATED WORK

2.1. Monitoring vital parameters using FMCW radar

The general principle of any radar is sending a continuous wave (CW) or an amplitude-modulated or angle-modulated wave and then listening to its echo. Such radars can detect human chest movement caused by respiratory activity. One possibility is to use a CW Doppler radar. It can be used to detect chest velocity but not distance to the human subject, which, e.g., makes it difficult to distinguish between several subjects, even if they are in different range bins. An extensive overview of this approach can be found in [8]. Another approach is to emit swept frequency radio waves. One of the first studies in this area was presented in [9]. Many practical implementations have become feasible with the advent of low-cost Frequency-Modulated Continuous-Wave radar sensors.

In an FMCW radar the frequency of the transmitted signal changes over time, typically in a sweep across some preset bandwidth. A sawtooth function is the simplest and most often used frequency change pattern (Fig. 1). Due to the propagation delay Δt, a frequency difference Δf arises between the transmitted wave and the reflected one. The difference can be determined by mixing (multiplying) the two signals thus creating a new, intermediate frequency (baseband) signal.

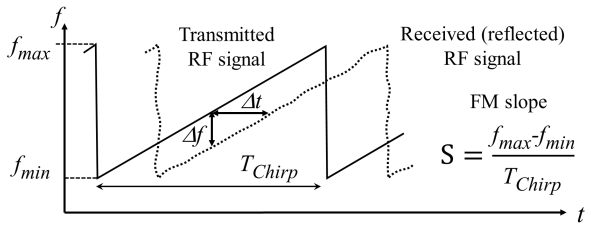

Fig. 1. Frequency-Modulated Continuous-Wave radar transmitted and reflected chirp

Blood circulation and breathing lead to both skin impedance variation and body movement. The body surface displacement has a higher influence on reflecting the signal than the impedance change of the skin, particularly for millimeter waves which cannot effectively penetrate more than 1–2 mm of the skin depth [10].

The distance R to the reflecting object can be determined by:

$$R = \frac{c\Delta t}{2} = \frac{c\Delta f}{2S}, \tag{1}$$

where Δf is the intermediate frequency, c is the speed of light in free space, and $S = df/dt$ is the slope of the FMCW frequency modulation rate. Additionally, if the reflecting object has a non-zero radial speed with respect to the radar receiving antenna, then the echo signal is subject to the Doppler effect, with a Doppler frequency f_D. Equation (1) and additionally the Doppler effect should be used for tracking fast-moving targets and large displacements, which is the case in automotive applications. Chest displacements caused by the respiratory activity, however, have amplitudes in the range of several millimeters and speed in the range of 10^{-3} m/s. In this case, an approach can

be adopted based on the determination of the received signal phase. This can be implemented using the complex-baseband architecture [11], which makes it possible to obtain the in-phase and quadrature components of the Intermediate Frequency (IF) signal. This mode of operation can be described by equations (2), (3) and (4) [12]:

$$s(t) = A \cdot \cos\left[2\pi f_{\min} t + \pi S t^2 + \phi(t)\right], \tag{2}$$

$$r(t) = A' \cdot \cos\left[2\pi f_c (t-\Delta t) + \pi S (t-\Delta t)^2 + \phi(t-\Delta t)\right], \tag{3}$$

$$y(t) = A'' \cdot e^{j(2\pi f_b t + \Phi(t) + \Delta\phi(t))}, \tag{4}$$

where $s(t)$ is the transmitted signal, $r(t)$ is the received signal, $y(t)$ is the low-pass filtered intermediate frequency complex signal, $f_b = R(t)(2S/c)$, $\Phi(t) = 2\pi f_c \Delta t + \pi S \Delta t^2$, and $\Delta\phi(t) = \phi(t) - \phi(t - 2R/c)$. The initial phase $\phi(t)$ of the transmitted signal is arbitrary. For simplicity, the calculation of the amplitudes A' and A'' is omitted and the influence of noise is neglected. The impact of frequency chirp nonlinearity, phase noise, and signal-to-noise ratio on FMCW radar accuracy can be found, e.g., in [13]. Conventionally, the displacement $\Delta R(t)$ of the chest can be found from (5) [14], where $Q(t)$ and $I(t)$ are the quadrature and in-phase components of the intermediate frequency signal $y(t)$:

$$\Delta R(t) = \arctan\left(\frac{Q(t)}{I(t)}\right)\frac{\lambda}{4\pi}. \tag{5}$$

If (5) is used, phase unwrapping is necessary to recover the correct motion information, since π phase discontinuity occurs in the arctangent function, when the signal trajectory crosses the boundary of two adjacent quadrants, which corresponds to the movement of $\lambda/4$. The presented procedure can be easily implemented by applying double-DFT (Digital Fourier Transform) procedure [15], where the first transformation locates objects (at f_b frequencies), whereas the second one tracks their motion ($\Delta R(t)$) from the phase of the complex slow-time signal.

2.2. Analog neuromorphic architectures

The most successful artificial intelligence approach to problem-solving involves deep neural models, which need to have capacity that is sufficient for matching the problem complexity. Therefore, a size of deep models can reach several or even hundreds of billion parameters [16]. As there is a little progress in incorporating prior knowledge into neural networks, neural architectures that are applied to even moderately complex problems tend to be huge and require specialized and powerful hardware, such as GPUs or GPU clusters, to train and execute. However, this power-hungry approach to problem-solving, which additionally generates heavy network traffic to deliver data and read out processing results, is not necessarily a future trend in AI, as energy efficiency recently becomes one of the most important system design criteria.

Intelligent data analysis utilizing deep neural networks that is to be performed on edge computing or embedded systems is clearly also subject to strict energy constraints. Among severe problems that prevent power-efficient operation of deep neural architectures in digital hardware there is for example, necessity to handle numerous data transfers between memory and processing units (referred to as von Neumann bottleneck), which applies not only to CPU-, but also, to GPU-based processing [17]. In a search to alleviate the aforementioned shortcomings, considerable amount of work has been done on developing neuromorphic, primarily, digital architectures that employ FPGA or digital ASICs [18]. However, it is analog computing that offers a real promise to drastically reduce power consumption, and at the same time, boost a computing speed. The main reason for this is to provide natural, physical means for execution of fundamental operations that underlie neural information processing pipeline, i.e. multiplication, accumulation and nonlinear data transformations. Therefore, development of analog neuromorphic modules that can augment conventional architectures with efficient execution of energy-critical operations, or even complete analog implementations of neural networks, becomes a focus of intense research that could alleviate shortcomings of conventional, digital architectures.

Research on analog implementations of various neural models has been carried out since the very beginning of the paradigm formulation and multiple analog architectures that implement simple feedforward and feedback networks have been reported [19–22]. The shift towards highly complex architectures necessary for implementing deep neural networks caused the corresponding modification of main targets in analog neural hardware development. The focus of research moved to enabling massively parallel execution of analog operations by introducing analog resistive (memristor-based) crossbars [23, 24] together with a variety of technological advances related to fabrication of non-volatile memories [25]. This in turn enabled analog implementations of more complex neural circuits or analog accelerators for deep convolutional networks [26], deep LSTM networks [27] as well as application-specific neuromorphic architectures [28]. Among these efforts, neuromorphic architectures aimed at radar signal processing have also been proposed [29].

3. THE PROPOSED CONCEPT

An objective of the presented research was to develop a computational architecture for breathing rate estimation in FMCW radar signals that is feasible to be implemented in analog ASIC hardware. The considered vital parameter can be estimated from a series of observations of object micro-displacements (respiration-related chest movements) in subsequent time instants. Information on these displacements is embedded in phase shifts of harmonic components that are present in radar-generated data in consecutive sweeps. The conventional, digital signal processing-based approach to solving the posed problem, which involves double-DFT procedure, requires numerous data transfers that need to be executed between memory and the CPU, making it energy-inefficient.

The proposed concept offers a low-power alternative to accomplish the task: it gets solved using a computational architecture of topology that can be directly mapped onto hardware. The architecture of the proposed algorithm (depicted in Fig. 2),

comprises two main functional modules. An objective of the first one, which we refer to as *Vital Activity Information Extractor* (abbreviated by **VAIE**), is to distill only these components from radar-produced data that are related to an object that exhibits motion. This information is then subject to analysis in the second functional module, which we refer to as *Vital Activity Analyzer* (**VAA**), where the breathing rate gets estimated based on a sequence of data samples provided by the extractor.

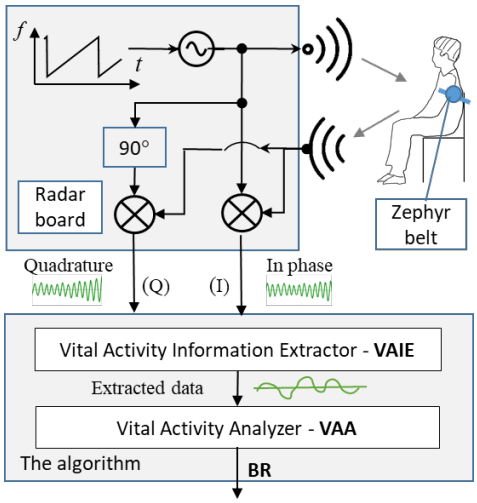

Fig. 2. Breathing rate monitoring setup (Zephyr data are used in training phase only)

Both modules of the proposed architecture have neuromorphic structures. The main building blocks of the VAIE module involve dot product calculations, scalar multiplications and additions, as well as the Winner-Takes-All unit. The VAA module is a pretrained Recurrent Neural Network. As all operations underlying data processing of both modules are analog, the proposed data analysis pipeline is well-suited for its analog VLSI implementation.

3.1. Vital Activity Information Extractor

The assumed objective of processing information produced by an FMCW radar is to provide online estimates of breathing rates for a single person inside a closed space, such as e.g. a room. For typical, furnished spaces, one can expect very rich structure of information present in data produced for each chirp, comprising several harmonic components corresponding to multiple objects present within such spaces. An essence of conventional analysis of such data is to use the first DFT to perform range-wise segmentation of space (testing for presence of subsequent intermediate frequencies), followed by the second DFT, performed only for frequencies where objects were found. A possible analog equivalent to the first procedure is to use a bank of analog bandpass filters, operating in parallel, and extracting information from different distance ranges. The resulting space segmentation is highly desirable, as it significantly simplifies data to be analyzed by discarding information related to objects located elsewhere. However, such an approach implies necessity of constructing multiple data analysis paths, which would significantly increase architectural complexity of a data preprocessing system. Also, one should search for alternative solutions to conventional analog filtering, which involves a use of bandpass LC filters (or their active equivalents) to provide compatibility with the assumed neuromorphic circuit structure.

To overcome these two problems we propose an alternative way of extracting information related to vital activity from radar-produced data, which combines architectural simplicity with neural processing-friendly approach. The proposed procedure (Fig. 3) involves space segmentation and selection of a spatial bin, which is likely to contain a 'living' object, and results in producing information that is passed to the VAA module. The proposed space segmentation utilizes only a single, fixed bandpass filter and a single data processing path. In addition, the proposed bin selection method enables seamless tracking of an object that moves among different range bins.

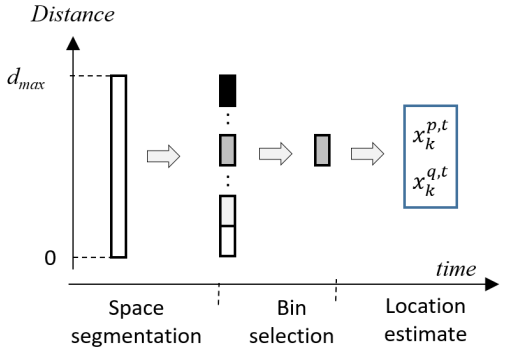

Fig. 3. Procedure for extracting information on moving object location from a single frame produced by an FMCW radar, involving space segmentation and selection of a bin with motion

Both operations involved in data processing in VAIE are detailed in the following subsections.

3.1.1. Space segmentation unit

The proposed idea to eliminate necessity of using multiple, parallel data processing paths that handle information from subsequent range bins, is to utilize possibility of modifying chirp parameters (offered by several FMCW radars present on the market) and to introduce time-multiplexed scheme for algorithm execution. It is observed that by varying a slope of chirp frequency modulation, the same intermediate frequency represents objects located at different distances:

$$\Delta f = \frac{2SR}{c}, \tag{6}$$

where the same notation as in equation (1) is used.

It follows that one can map different distance ranges to the same frequency band by appropriate modifications of a chirp slope S. Adoption of an appropriate scheme for chirp frequency modulation slope (e.g., a linear decrease for all chirps produced within each frame produced by a radar, as shown in Fig. 4) enables partitioning of the monitored space into distance intervals by means of just a single bandpass filter. As a result, instead of using a set of parallel signal processing paths, we propose

a single path, where information from subsequent spatial bins, extracted using chirps with the decreasing frequency modulation slopes and a single bandpass filter, will be processed in subsequent time slots.

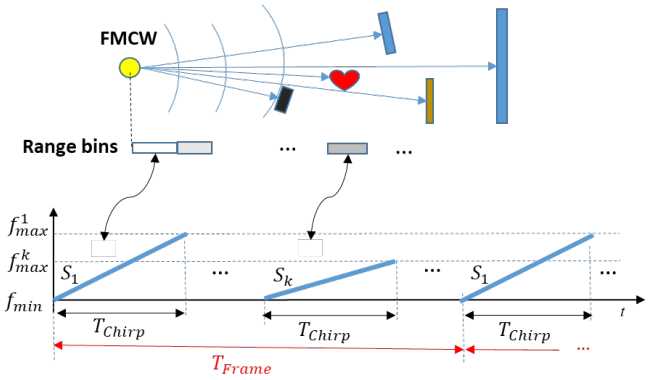

Fig. 4. The proposed chirp profile scheme that enables time-multiplexed scanning of a space monitored by FMCW radar (bottom) and range bins corresponding to chirps of selected slopes (top)

Using the adopted frequency modulation scheme, subsequent chirps map different spatial bins onto a fixed frequency band. This frequency band is a passband of a filter that extracts information on bin contents. We propose to implement this filter in a way that is compatible with a neural approach to data processing, by executing dot products that involve filter impulse response and data provided by a radar in a form of In-Phase and Quadrature components. Let us assume that impulse response of the adopted fixed bandpass filter is represented by a vector of coefficients $\mathbf{h} = [h_1 ... h_n]$ and data produced by a radar that emits a chirp k, generated for a frame t, are given by vectors $\mathbf{p}^t_k = [p^t_{k,1} ... p^t_{k,n}]$ for In-Phase component and $\mathbf{q}^t_k = [q^t_{k,1} ... q^t_{k,n}]$ for the Quadrature one (Fig. 5). Signal filtering in frequency domain is equivalent to convolution performed in time domain. Since discrete convolution is made up of dot products, and dot product is a way of assessing similarity of its two arguments, computation of the following expressions:

$$x^{p,t}_k = \left(\mathbf{p}^t_k\right)^T \mathbf{h} = \sum_{i=1}^{n} p^t_{k,i} h_i \tag{7}$$

and:

$$x^{q,t}_k = \left(\mathbf{q}^t_k\right)^T \mathbf{h} = \sum_{i=1}^{n} q^t_{k,i} h_i \tag{8}$$

produces information on how input vectors $\mathbf{p}^t_k$ and $\mathbf{q}^t_k$ fit a filter response $\mathbf{h}$. As the quadrature component $\mathbf{q}^t_k$ is a phase-shifted version of $\mathbf{q}^t_k$, the expressions (7) and (8) provide information that is equivalent to Fourier-domain filtering of Fourier-transformed $\mathbf{p}^t_k$, i.e., they could be considered the real and imaginary parts of the vector-filtering result. It follows that information provided by the pair $\{x^{p,t}_k, x^{q,t}_k\}$ enables evaluation of both the filtering outcome magnitude (i.e. to what extent frequencies in the assumed filter pass band are present in data-produced vectors) but also provides basis for phase evaluation (if only

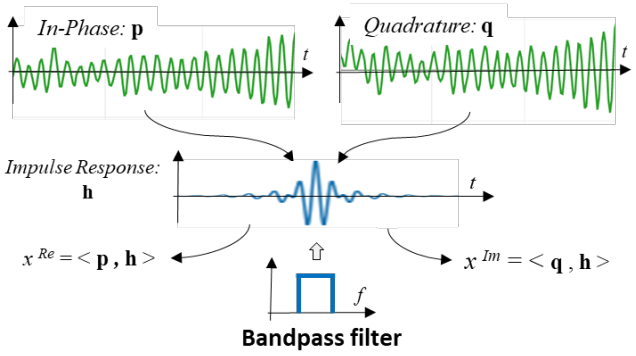

Fig. 5. The proposed range-filtering scheme employing dot products between In-Phase and Quadrature components of radar-produced data (represented by vectors **p** and **q**) and filter impulse response (**h**)

a single harmonic component is present within the considered frequency range).

The values $\{x^{p,t}_k, x^{q,t}_k\}$ produced by the filtering procedure should be forwarded for further analysis, i.e. to the VAA module, only for this range bin k (covered by a k-th chirp of a current frame t), which contains a moving object (a person), whereas the values produced for all remaining range-bins (chirps) should be discarded. To achieve this goal, we propose a procedure that involves detection of a spatial bin with motion activity and selection of the corresponding filtering outcome to be outputted from the VAIE module at a frame t.

3.1.2. Selection of the frame processing result

An objective of the proposed procedure is to produce a pair of values $\{x^{p,t}_k, x^{q,t}_k\}$, which represent a result of processing of a frame t, and which contain information on a moving object, located at some k-th range bin. The assumed 'liveness' attribute is object motion (we clearly simplify the problem, although the issue of handling motion of mechanical devices, such as fans or motion caused by airflow, will be addressed later). Therefore, the first part of the proposed procedure (see Fig. 6) is detection of a bin, where motion is present. Clearly, to detect a change, one needs to memorize results of processing of at least the preceding frame, i.e. all $\{x^{p,t-1}_k, x^{q,t-1}_k\}$ pairs, for $k = 1, \ldots, N_c$.

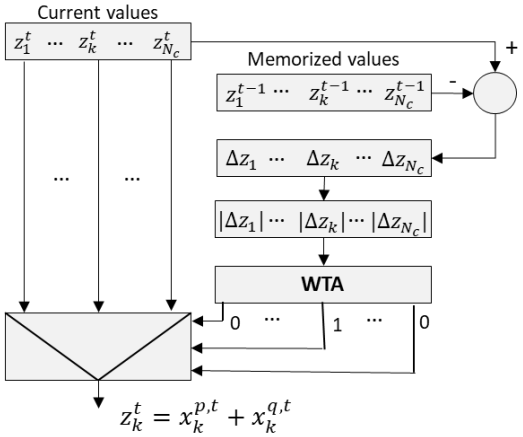

Fig. 6. Derivation of frame processing result: z^p_k denotes a complex number, composed of a real part $x^{p,t}_k$ and imaginary part $x^{q,t}_k$

The key observation that enables the presented procedure is that harmonic components produced by an FMCW radar that result from wave reflections from still objects are the same for every frame. As the distance to an object remains the same, so is a time delay for the arriving reflected wave, which results in the same frequency difference between outgoing and incoming components and in the same phase shift between the two. Therefore, information on still objects can be eliminated by simple subtraction of vectors $\{\mathbf{p}^{t-1}, \mathbf{q}^{t-1}\}$ and $\{\mathbf{p}^{t}, \mathbf{q}^{t}\}$ produced by the radar in subsequent frames. As the same filter impulse response vector $\mathbf{h}$ is used in dot products with both data vectors, the same effect of component cancellation for still object contributions can be obtained by subtracting dot product results. Therefore, to assess a change caused by motion that occurs in spatial bins, one needs to calculate a difference between dot products computed for all chirps in two subsequent frames:

$$\Delta z_k^t = z_k^t - z_k^{t-1}, \tag{9}$$

where z denotes a complex representation of the two products, i.e.:

$$z_k^t = x_k^{p,t} + j x_k^{q,t} \tag{10}$$

and where j is an imaginary unit.

If no motion is present, one should expect harmonic components that represent still objects to cancel out, leaving only non-zero values of (10) for spatial bins (chirps) with motion. A magnitude of the difference (10), i.e.:

$$\|\Delta z_k^t\| = \|(x_k^{p,t} - x_k^{p,t-1}) + j\left(x_k^{q,t-1} - x_k^{q,t}\right)\| \tag{11}$$

can be seen as a measure of amount of motion present in a k-th range bin. As monitoring of a single person only is assumed, results from a single frequency bin should get selected for final analysis. To accomplish this objective, a Winner-Takes-All module, that operates on magnitudes of subtraction results (11) has been used. The module has a 'neural' structure – it is composed of simple linear neurons with lateral inhibitory connections (Fig. 7). Each neuron is fed with input that represents motion energy estimates, provided by (11). After transient decays, WTA module produces a one-hot encoding of inputs, where the 'high' state indicates a spatial bin, where motion has been present. This 'high' output is then used as a selector of a channel to be passed through an output analog multiplexer.

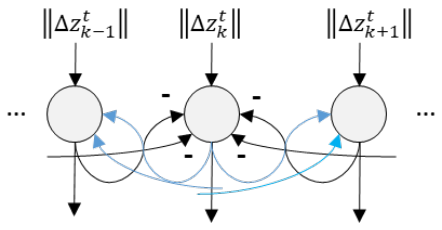

Fig. 7. A structure of WTA module for selecting the bin with motion

3.2. Vital Activity Analyzer

VAIE module produces for each frame t a pair of values that can be used to calculate a phase (arc tangent of a ratio of the 'imaginary' – $x_k^{q,t}$, and 'real' – $x_k^{p,t}$ components) of an intermediate frequency harmonic component, generated by a moving object. As the phase provides an accurate estimate of object location at a time instant t, reconstruction of phase evolution provides accurate information on small object displacements, enabling estimation of various motion parameters.

The proposed approach to analysis of a sequence of values produced by the VAIE module is to use a Recurrent Neural Network. Adoption of machine learning approach is justified by several reasons, which make application of any deterministic approach questionable. Firstly, the detected micro-motions are combinations of several simultaneously developing processes – breathing, pulse and person's motion, which makes the resulting waveform to be analyzed very complex. Secondly, exact motion patterns related to breathing are person-dependent and also, dependent on a type of activity, which additionally complicates their analysis. Finally, a neural network offers a convenient hardware platform for analog computing.

The VAA module is an RNN, which is fed with a sequence of two element vectors $\{x_k^{p,t}, x_k^{q,t}\}$ derived by VAIE for subsequent frames (we do not perform explicit phase extraction, as this is unnecessary and can introduce additional errors) and generates estimates for BR. We experimented with different RNN architectures ('vanilla' RNN, Gated Recurrent Units (GRU), and GRUs with Multilayer Perceptron (MLP) or convolutional post-processing) to find a structure that provides the best trade-off between analysis accuracy and complexity, as well as with different sequence lengths. To further account for limitations of the considered implementation method, we considered reduced parameter representation accuracy, as analog weights are inaccurate.

The proposed architecture requires simple control circuitry to handle data flow and distribution among memories that are used for storing intermediate processing results. Key computations of the proposed architecture: selection of a bin with motion and classification, are performed asynchronously.

4. EXPERIMENTAL EVALUATION

4.1. Radar configuration & dataset

The experimental evaluation of the proposed method was done using the IWR1443BOOST radar board combined with a DCA1000EVM data acquisition card from Texas Instruments. The board operates in the 76–81 GHz range and was configured to emit frames consisting of eleven 175 µs chirps every 250 ms. The chirp slopes have been chosen in a way to provide complete coverage of the assumed distance range. The configuration of the bandpass filter, where 3 dB passband starts at 126350 Hz and ends at 139 650 Hz, centered at 130 kHz, provides an operational range between approximately 1 m and 3 m. The exact chirp slopes and their corresponding distance ranges are shown in Table 1.

The labels (breathing rates) for the dataset have been acquired using the Zephyr BioHarness 3 device [30], which pro-

Table 1
Observation distances corresponding to each chirp for a bandpass filter centred at 130 kHz

Chirp slope	Minimum distance	Maximum distance
20.0 MHz/µs	0.95 m	1.05 m
18.0 MHz/µs	1.05 m	1.16 m
16.0 MHz/µs	1.18 m	1.31 m
14.5 MHz/µs	1.31 m	1.44 m
13.0 MHz/µs	1.46 m	1.61 m
11.5 MHz/µs	1.65 m	1.82 m
10.5 MHz/µs	1.80 m	2.00 m
9.5 MHz/µs	2.00 m	2.21 m
7.5 MHz/µs	2.53 m	2.79 m
6.5 MHz/µs	2.92 m	3.22 m

duces the reference information in two different forms: as raw data (instantaneous measurement results from a Zephyr-integrated tensometer), and as on-device calculated BR estimates. As no information is provided on the algorithm used by the device for BR evaluation, we adopted the tensometer raw data as a basis for computing the reference values. To align Zephyr-generated labels with radar-produced waveforms (Fig. 8), an internal device clock was synchronized with a radar controlling application, at the beginning of each experiment.

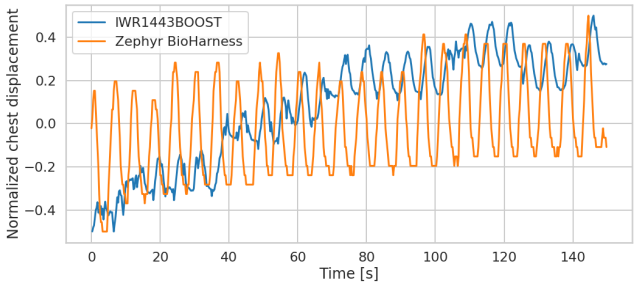

Fig. 8. Comparison of radar-produced chest motion estimation (blue trace) with Zephyr tensometer data (orange). Observe involuntary-motion induced drift in radar-extracted waveform

The acquisitions lasted between one and five minutes and featured a still human seated on a chair. This 'laboratory' setup, featuring only small subject's movements, provides the easiest context for BR estimation, however, it is plausible in real indoor monitoring scenarios, due to common long-lasting intervals of limited subject's mobility (sleep, TV-watching, working on a computer etc.). Moreover, as significant motion can be easily detected, this can be used to temporarily suspend the estimation procedure. The full dataset alongside a live viewer is available on the project website [7].

4.2. VAIE processing

A spectrum of a short snippet of a recording of a person located approximately 1 m from the radar is shown in Fig. 9. The top part of the figure shows the DFT spectrum derived for chirps with the same slope generated in consecutive frames. Two 'hot spots' visible at around 1 m and 3.4 m correspond to a person and a wall behind, respectively.

By stacking all chirps in a frame and observing the resulting spectra, as shown in the bottom part of Fig. 9, one can see how the apparent frequency of each object changes with the chirp slope. Reducing the slope shifts objects' frequency domain 'echoes' to match the predefined filter bandpass, indicated by the black lines (as it can be seen, the wall 'moves' from 0.46 MHz at 20 MHz/µs, to 0.15 MHz at 6.5 MHz/µs).

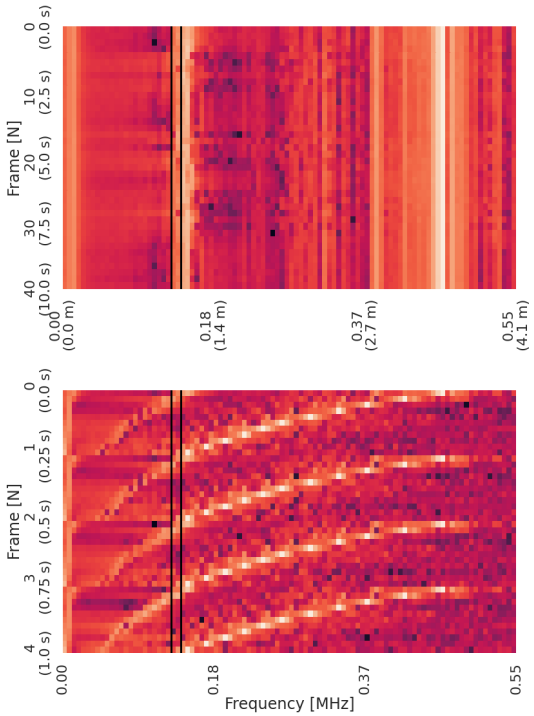

Fig. 9. Spectrogram (top) 10-second observation of 20 MHz/µs chirp; (bottom) 1-second observation of all chirps in the frame. Black lines mark the edges of the band pass filter

Results of WTA operation on differential data are presented in Fig. 10. Non-zero magnitudes of signal differences (left part of the figure), derived using (11), appear only for a moving object (a sitting person), with the vast majority of peaks being located in the range covered by the 18 MHz/µs chirp, albeit with small occasional flips into the neighbouring range bins. The wall, which would appear in the 6.5 MHz/µs bin, is not visible due to no motion, thus showing the effectiveness of the proposed approach in the removal of inanimate objects. The right part of the figure shows the result of the WTA operation, which is used for selecting a bin with data to be analyzed in the VAA module.

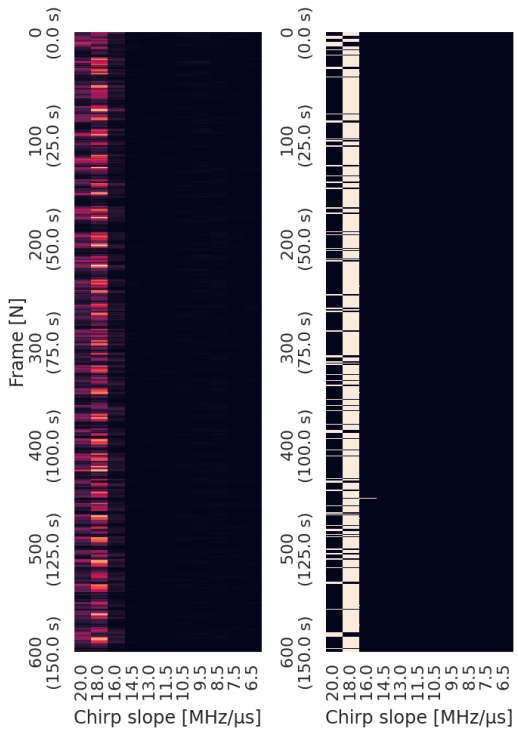

Fig. 10. Normalized magnitude of differential signals (left) and the corresponding WTA-processing result (right)

4.3. Neural network architectures

Three small architectures were proposed for solving the problem of breathing rate prediction (Fig. 11). Through an initial grid search, the optimal width of all hidden layers was determined to be 4. It should be noted that due to technology-implied constraints on feasible network size, several better models that were found during the search were deemed too large to implement.

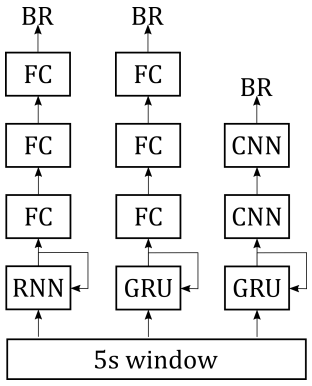

Fig. 11. Evaluated model architectures

The first and the second of the considered architectures combined recurrent units: simple, 'vanilla' ones in the former case, and gated recurrent units [31] in the latter case, with a three-layer dense modules with ReLu [32] activation (except for the output layer). While adoption of Gated Recurrent Units (GRU) approximately doubles the number of parameters, it was hypothesised that presence of additional information retention mechanisms would improve the performance. The third GRU-based architecture was aimed to simplify the second one by replacing dense layers with a couple of convolutional layers.

4.4. Neural network training

All models were trained for 300 epochs (Fig. 12) using 5-second windows of the radar signal (20 discrete samples at 4 Hz sampling rate). The Adam optimizer with a cosine annealing scheduler operating between 0.025 and 0.01 starting with a 60 epoch warm-up was applied. All models were trained with 16-bit floating point precision and subsequently quantized to 6-bit precision equivalent values. Additionally, all weights were L2 regularized and clipped to $[-1, 1]$ range.

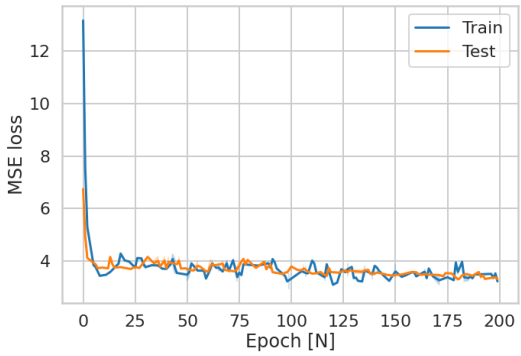

Fig. 12. Loss over the course of training (0.6 smoothing applied)

4.5. Neural network performance

The performance comparison of the evaluated models without quantization can be seen in Fig. 13 and Table 2. As can be observed, the best performance is achieved by the model that combines the two-layered GRU network with the convolutional layers. With only 287 parameters, this model is feasible for analogue implementation.

The addition of quantization resulted in comparable results to the non-quantized model, as shown in Fig. 14 and Table 3. In

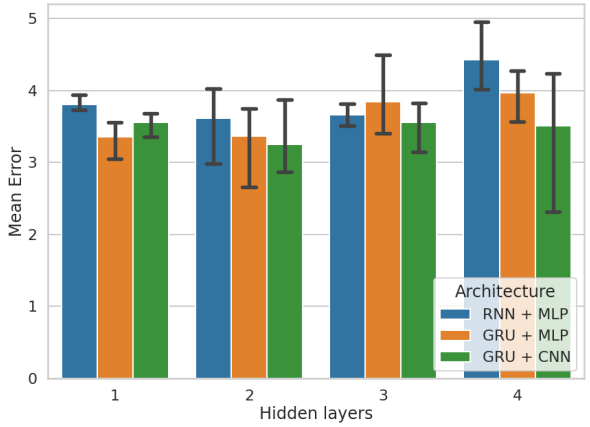

Fig. 13. Performance of the considered network models without quantization on the test set with 95% confidence intervals marked

Table 2
Performance and standard deviation of the considered network models without quantization. The lowest test error for each architecture was highlighted

Architecture	Hidden layers	Parameters	Mean train error	Mean test error
RNN + MLP	1	77	3.50 ± 0.41	3.81 ± 0.11
	2	117	3.54 ± 0.41	**3.62 ± 0.55**
	3	157	3.75 ± 0.35	3.66 ± 0.15
	4	197	3.78 ± 0.26	4.42 ± 0.48
GRU + MLP	1	141	3.04 ± 0.48	**3.36 ± 0.27**
	2	261	3.02 ± 1.09	3.37 ± 0.62
	3	381	2.48 ± 0.74	3.84 ± 0.58
	4	501	3.33 ± 0.92	3.97 ± 0.36
GRU + CNN	1	167	3.25 ± 0.28	3.56 ± 0.17
	2	287	2.71 ± 0.85	**3.26 ± 0.57**
	3	407	3.11 ± 0.93	3.56 ± 0.36
	4	527	3.69 ± 0.27	3.51 ± 1.04

this case, the best performance was also achieved by the combination of GRU and convolutional layers, albeit with an increased depth of the GRU. Given the significant size of this model, reaching 527 parameters, the second best configuration of GRU with fully connected layers, that achieved slightly worse performance while having only 141 parameters, could prove better suited for analogue implementation. It is also worth noting that this architecture yielded the most stable performance across all quantized network depths.

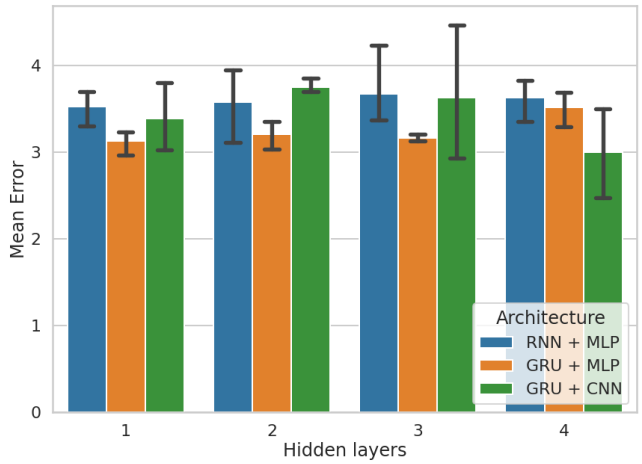

Fig. 14. Performance of the considered network models with quantization on the test set with 95% confidence intervals marked

While the average error of approximately 3 BPM does not constitute perfect performance, it is important to note that it is

Table 3
Performance of the considered models with quantization

Architecture	Hidden layers	Parameters	Mean train error	Mean test error
RNN + MLP	1	77	3.65 ± 0.20	**3.53 ± 0.20**
	2	117	3.71 ± 0.35	3.57 ± 0.42
	3	157	3.39 ± 1.03	3.67 ± 0.49
	4	197	3.88 ± 0.18	3.63 ± 0.25
GRU + MLP	1	141	3.15 ± 0.52	**3.13 ± 0.14**
	2	261	2.86 ± 0.23	3.21 ± 0.16
	3	381	2.81 ± 0.17	3.16 ± 0.04
	4	501	3.53 ± 0.46	3.52 ± 0.20
GRU + CNN	1	167	3.54 ± 0.38	3.39 ± 0.39
	2	287	3.89 ± 0.25	3.75 ± 0.09
	3	407	3.39 ± 0.52	3.63 ± 0.78
	4	527	2.89 ± 0.51	**3.00 ± 0.51**

comparable to accuracy of the reference device – Zephyr BioHarness, which reports the accuracy of ±3 BPM for stationary people and ±2 for a laboratory breathing emulator.

5. CONCLUSIONS

The presented paper shows feasibility of breathing rate estimation, based on data provided by a FMCW radar, using a neuromorphic architecture that can be implemented in VLSI. The target circuit, which is currently under development, is expected to offer ultra-low energy consumption and to provide data analysis accuracy comparable to conventional approaches to problem solution.

Although simulation results positively verify the proposed problem solution concept, further efforts are required to relax constraints adopted for the design, such as ability to monitor only a single subject (relevant for the adopted application scenario) or to appropriately respond to motion incurred by non-living objects. These are the forthcoming research directions, which require corresponding extensions to the dataset and introduction of additional, neural network-based data analysis units, operating on generated breathing rate estimates.

ACKNOWLEDGEMENTS

The Authors wish to thank Mr Przemysław Lewandowski and Mrs Magdalena Jaśkiewicz for their contributions at different stages of the concept development.

This work has been completed while the 2nd author was the Doctoral Candidate in the Interdisciplinary Doctoral School at the Lodz University of Technology, Poland.

This work has been done as a part of the AIR project, supported by the National Science Centre, Poland under CHIST-

ERA III, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 76897.

REFERENCES

[1] R.R. Varior, M. Haloi, and G. Wang, "Gated Siamese Convolutional Neural Network Architecture for Human Re-identification," in *Computer Vision – ECCV 2016*, ser. Lecture Notes in Computer Science, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Cham: Springer International Publishing, 2016, pp. 791–808, doi: 10.1007/978-3-319-46484-8_48.

[2] L. Ding, W. Fang, H. Luo, P.E.D. Love, B. Zhong, and X. Ouyang, "A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory," *Autom. Constr.*, vol. 86, pp. 118–124, Feb. 2018, doi: 10.1016/j.autcon.2017.11.002.

[3] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, and F. Babiloni, "Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness," *Neurosci. Biobehav. Rev.*, vol. 44, pp. 58–75, Jul. 2014, doi: 10.1016/j.neubiorev.2012.10.003.

[4] Y. Dong, Z. Hu, K. Uchimura, and N. Murayama, "Driver Inattention Monitoring System for Intelligent Vehicles: A Review," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 596–614, Jun. 2011, doi: 10.1109/TITS.2010.2092770.

[5] T.B. Moeslund, A. Hilton, and V. Krüger, "A survey of advances in vision-based human motion capture and analysis," *Comput. Vision Image Understanding*, vol. 104, no. 2, pp. 90–126, Nov. 2006, doi: 10.1016/j.cviu.2006.08.002.

[6] Z. Pan, G. Healey, M. Prasad, and B. Tromberg, "Face recognition in hyperspectral images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 12, pp. 1552–1560, Dec. 2003, doi: 10.1109/TPAMI.2003.125114810.1109/ TPAMI.2003.1251148.

[7] K. Ślot, P. Łuczak, P. Lewandowski, M. Jaśkiewicz, and A. Bielińska-Ślot, "CHISTERA AIR – FMCW radar recordings." [Online]. Available: https://chist-era-air.iis.p.lodz.pl/

[8] C. Li, V.M. Lubecke, O. Boric-Lubecke, and J. Lin, "A Review on Recent Advances in Doppler Radar Sensors for Noncontact Healthcare Monitoring," *IEEE Trans. Microwave Theory Tech.*, vol. 61, no. 5, pp. 2046–2060, May 2013, doi: 10.1109/TMTT.2013.2256924.

[9] J. Lin, "Noninvasive microwave measurement of respiration," *Proc. IEEE*, vol. 63, no. 10, pp. 1530–1530, Oct. 1975, doi: 10.1109/PROC.1975.9992.

[10] Ø. Aardal, Y. Paichard, S. Brovoll, T. Berger, T.S. Lande, and S.-E. Hamran, "Physical Working Principles of Medical Radar," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 4, pp. 1142–1149, Apr. 2013, doi: 10.1109/TBME.2012.2228263.

[11] K. Ramasubramanian, "Using a Complex-Baseband Architecture in FMCW Radar Systems," Texas Instruments, Tech. Rep., 2019. [Online]. Available: http://www.ti.com/lit/wp/spyy007/spyy007.pdf

[12] K. Yamamoto, K. Toyoda, and T. Ohtsuki, "Cnn-based respiration rate estimation in indoor environments via mimo fmcw radar," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6, doi: 10.1109/GLOBECOM38437.2019.9013951.

[13] S. Ayhan, S. Scherr, A. Bhutani, B. Fischbach, M. Pauli, and T. Zwick, "Impact of Frequency Ramp Nonlinearity, Phase Noise, and SNR on FMCW Radar Accuracy," *IEEE Trans. Microwave Theory Tech.*, vol. 64, no. 10, pp. 3290–3301, Oct. 2016, doi: 10.1109/TMTT.2016.2599165.

[14] G. Sacco, E. Pittella, E. Piuzzi, and S. Pisa, "A radar system for indoor human localization and breath monitoring," in *2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, Jun. 2018, pp. 1–6, doi: 10.1109/MeMeA.2018.8438759.

[15] S. Jardak, M.-S. Alouini, T. Kiuru, M. Metso, and S. Ahmed, "Compact mmWave FMCW radar: Implementation and performance analysis," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 34, no. 2, pp. 36–44, Feb. 2019, doi: 10.1109/maes.2019.180130.

[16] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, "Language Models are Unsupervised Multitask Learners," *OpenAI blog*, p. 24, 2019.

[17] R. Raina, A. Madhavan, and A.Y. Ng, "Large-scale deep unsupervised learning using graphics processors," in *Proceedings of the 26th Annual International Conference on Machine Learning*, ser. ICML'09. New York, NY, USA: Association for Computing Machinery, Jun. 2009, pp. 873–880, doi: 10.1145/1553374.1553486.

[18] V. Sze, Y.-H. Chen, T.-J. Yang, and J. S. Emer, "Efficient Processing of Deep Neural Networks: A Tutorial and Survey," *Proc. IEEE*, vol. 105, no. 12, pp. 2295–2329, Dec. 2017, doi: 10.1109/JPROC.2017.2761740.

[19] Holler, Tam, Castro, and Benson, "An electrically trainable artificial neural network (ETANN) with 10240 'floating gate' synapses," in *International 1989 Joint Conference on Neural Networks*, vol. 2, 1989, pp. 191–196, doi: 10.1109/IJCNN.1989.118698.

[20] R. Sarpeshkar, R.F. Lyon, and C. Mead, "A Low-Power Wide-Dynamic-Range Analog VLSI Cochlea," in *Neuromorphic Systems Engineering: Neural Networks in Silicon*, ser. The Springer International Series in Engineering and Computer Science, T.S. Lande, Ed. Boston, MA: Springer US, 1998, pp. 49–103, doi: 10.1007/978-0-585-28001-1_3.

[21] S. Draghici, "Neural networks in analog hardware – design and implementation issues," *Int. J. Neural Syst.*, vol. 10, no. 01, pp. 19–42, Feb. 2000, doi: 10.1142/S0129065700000041.

[22] R. Carmona, F. Jiménez-Garrido, R. Domínguez-Castro, S. Espejo, and A. Rodríguez-Vázquez, "CMOS realization of a 2-layer cnn universal machine chip," *Int. J. Neural Syst.*, vol. 13, no. 06, pp. 435–442, Dec. 2003, doi: 10.1142/S0129065703001716.

[23] T.P. Xiao, C.H. Bennett, B. Feinberg, S. Agarwal, and M.J. Marinella, "Analog architectures for neural network acceleration based on non-volatile memory," *Appl. Phys. Rev.*, vol. 7, no. 3, p. 031301, Sep. 2020, doi: 10.1063/1.5143815.

[24] M. Prezioso, F. Merrikh-Bayat, B.D. Hoskins, G.C. Adam, K.K. Likharev, and D.B. Strukov, "Training and operation of an integrated neuromorphic network based on metal-oxide memristors," *Nature*, vol. 521, no. 7550, pp. 61–64, May 2015, doi: 10.1038/nature14441.

[25] S. Ambrogio *et al.*, "Equivalent-accuracy accelerated neural-network training using analogue memory," *Nature*, vol. 558, no. 7708, pp. 60–67, Jun. 2018, doi: 10.1038/s41586-018-0180-5.

[26] M. Hu *et al.*, "Memristor-Based Analog Computation and Neural Network Classification with a Dot Product Engine," *Adv. Mater.*, vol. 30, no. 9, p. 1705914, 2018, doi: 10.1002/adma.201705914.

[27] C. Li *et al.*, "Long short-term memory networks in memristor crossbar arrays," *Nat. Mach. Intell.*, vol. 1, no. 1, pp. 49–57, Jan. 2019, doi: 10.1038/s42256-018-0001-4.

[28] S.-C. Liu, J.P. Strachan, and A. Basu, "Prospects for Analog Circuits in Deep Networks," *arXiv:2106.12444 [cs]*, Jun. 2021.

[29] J. López-Randulfe, T. Duswald, Z. Bing, and A. Knoll, "Spiking Neural Network for Fourier Transform and Object Detection for Automotive Radar," *Front. Neurorob.*, vol. 15, 2021, doi: 10.3389/fnbot.2021.688344.

[30] J. Hailstone and A.E. Kilding, "Reliability and Validity of the Zephyr™ BioHarness™ to Measure Respiratory Responses to Exercise," *Meas. Phys. Educ. Exercise Sci.*, vol. 15, no. 4, pp. 293–300, Oct. 2011, doi: 10.1080/1091367X.2011.615671.

[31] K. Cho, B. van Merrienboer, D. Bahdanau, and Y. Bengio, "On the Properties of Neural Machine Translation: Encoder-Decoder Approaches," *arXiv:1409.1259 [cs, stat]*, Oct. 2014.

[32] V. Nair and G.E. Hinton, "Rectified linear units improve Restricted Boltzmann machines," in *ICML 2010 – Proceedings, 27th International Conference on Machine Learning*, 2010, pp. 807–814.