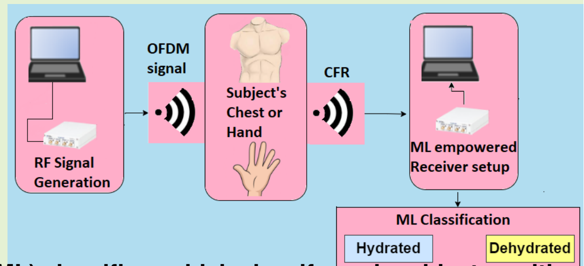
Buttar, H. M., Pervez, K., Ur Rahman, M. M., Mian, A. N., Riaz, K. and Abbasi, Q. H. (2023) Non-contact monitoring of dehydration using RF data collected off the chest and the hand. *IEEE Sensors Journal*, 24(3), pp. 3574-3582. (doi: 10.1109/JSEN.2023.3334590)

Non-Contact Monitoring of Dehydration using RF Data Collected off the Chest and the Hand

Hasan Mujtaba Buttar*, Kawish Pervez*, Muhammad Mahboob Ur Rahman, Adnan Noor Mian, Kashif Riaz, Qammer H. Abbasi

***Abstract*—We report a novel non-contact method for dehydration monitoring. We utilize a transmit software defined radio (SDR) that impinges a wideband radio frequency (RF) signal (of frequency 5.23 GHz) onto either the chest or the hand of a subject who sits nearby. Further, another SDR in the closed vicinity collects the reflected RF signals. The two SDRs exchange orthogonal frequency division multiplexing (OFDM) signal, whose individual subcarriers get modulated once it reflects off (passes through) the chest (the hand) of the subject. This way, the signal collected by the receive SDR consists of channel frequency response (CFR) that captures the variation in the blood osmolality due to dehydration. The received raw CFR data is then passed through a handful of machine learning (ML) classifiers which classify each subject as either hydrated or dehydrated. To train our ML classifiers, we have constructed our custom dataset by collecting data from 5 Muslim subjects who were fasting during the month of Ramadan. Specifically, we have implemented and tested the following ML classifiers: k-nearest neighbour, support vector machine, decision tree, ensemble classifier, and a neural network classifier. Among all the classifiers, the neural network classifier achieved the best classification accuracy, i.e., an accuracy of 93.8% (96.15%) for the proposed chest-based (hand-based) method. Compared to prior contact-based method where the reported accuracy is 97.83%, our proposed non-contact method provides slightly less accuracy than that of reported in the literature for contact-based method; nevertheless, the advantages of our non-contact dehydration method speak for themselves.**



***Index Terms*—dehydration, non-contact methods, RF-based methods, software-defined radio, covid19, machine learning.**

I. Introduction

A good sixty percent of the human body is composed of water, which is essential to many of the body's activities, including maintaining the body's temperature, transporting nutrients and oxygen to cells, lubricating joints, and eliminating waste products. Consuming sufficient water on a daily basis is necessary for preserving one's health and warding off a variety of diseases and adverse conditions [1]. Dehydration occurs when the body does not obtain enough water or when the body loses water through sweating and evaporation. When dehydration occurs, it throws off the natural equilibrium of the minerals and electrolytes found in the body. This could result in a variety of different health issues, ranging from quite harmless to life-threatening, depending on how much fluid is lost and what's causing it in the first place. Symptoms of mild dehydration include headache, dry mouth, thirst, dizziness, exhaustion, and dry and wrinkled skin [2], [3], [4]. In more extreme circumstances, dehydration can result in consequences such as kidney failure, convulsions, and even death. When the outside weather is hot and humid, then the dehydration could lead to heat exhaustion which could induce symptoms such as heavy perspiration, nausea, headache, and weakness. Heat exhaustion if not addressed quickly, could escalate to heatstroke, which is a life-threatening medical emergency that can cause damage to the brain, organ failure, and even death. Last but not the least, dehydration could also have some long-term adverse effects on the body, e.g., constipation, damage to the kidneys, and infections of the urinary tract, etc. [1].

In short, dehydration could have fatal implications if left untreated, thus, timely diagnosis of dehydration followed by imminent medical intervention is of utmost importance. For the elderly, and for the diabetic and diarrhea patients, it is especially important to track their hydration levels frequently [5]. But when it comes to the existing dehydration detection methods, they have their limitations as they are either invasive (e.g., blood sample based), or contact-based (e.g., pulse oximeter, smart watch based). Further, the existing methods are expensive, inconvenient and inconsistent, as discussed below.

H. M. Buttar, Kawish Pervez, M. M. U Rahman, and Kashif Riaz are with the Electrical Engineering Department, Information Technology University, Lahore 54000, Pakistan (e-mail: mahboob.rahman@itu.edu.pk).
A. N. Mian is with the Department of Computer Science, Information Technology University, Lahore 54000, Pakistan.
Kashif Riaz is also affiliated with the Division of Information and Computing Technology, College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar.
Q. H. Abbasi is with the James Watt School of Engineering, University of Glasgow, Glasgow, G12 8QQ, UK, and with Artificial Intelligence Research Center (AIRC), Ajman University, Ajman, UAE (e-mail: Qammer.Abbasi@glasgow.ac.uk).
* implies equal contribution by the authors.

Existing dehydration measures and the dilemma: Some of the most common methods for measuring hydration levels are: body mass change, total body water, serum and urine osmolality, plasma osmolality, urine specific gravity, and urine volume [6]–[10]. Another method that is sometimes considered as the ”gold standard” consists of a procedure whereby a subject ingests a known quantity of an isotope, which allows one to calculate its concentration in a bodily fluid in order to determine the body's total water content. Now, the dilemma. Though such ”gold standards” of hydration assessment are considered useful for sports science, medicine, or for creating reference standards, but since they necessitate extensive methodological control, they are not useful for tracking one's hydration status on daily basis during a training or competition [11]. In other words, none of aforementioned hydration measures has been demonstrated to be valid in all dehydration scenarios (i.e., lab and field) [12]. Last but not the least, many of the aforementioned hydration measures could be expensive, cumbersome, erroneous, and inconvenient (either invasive or contact-based). This calls for the innovative and preferably non-contact methods for dehydration monitoring, which is precisely the agenda of this work.

Contributions. This paper proposes an RF-based dehydration monitoring method that is non-invasive and contact-less, has high accuracy, allows continuous and seamless monitoring, is easy to use, and provides rapid results. Specifically, the key contributions of this work are as follows:

1) We propose a novel non-contact method called chest-based dehydration monitoring (CBDM) method. Under this method, the subject sits nearby an RF transceiver that impinges an OFDM signal onto the chest of the subject, while the receiver collects the signal reflected off the chest of the subject.

2) We propose a novel non-contact method called hand-based dehydration monitoring (HBDM) method. Under this method, the subject places his/her hand on a table and between two antennas such that the transmitted OFDM signal passes through the hand of the subject, and is subsequently collected by the receiver.

The raw data collected by the receiver due to both (CBDM and HBDM) methods consists of channel frequency response (CFR) that is fed to multiple machine learning (ML) and deep learning (DL) classifiers which eventually determine whether a person is hydrated or dehydrated. *To the best of our knowledge, this is the first work that reports a non-contact method for dehydration monitoring.*

Rationale. The proposed CBDM and HBDM methods rely upon the following to infer dehydration related information from the data collected off the chest and the hand of the subject: i) Dehydration results in reduced blood volume and increased blood viscosity which in turn increases the heart rate and lessens the force of the blood against the walls of the arteries. ii) OFDM signal, being a wideband signal, helps in sensing for dehydration. That is, each OFDM subcarrier captures unique signatures of dehydration due to frequency, phase and amplitude modulation of the subcarrier reflected off the human body. Thus, the reflected signal which is eventually translated to channel frequency response (CFR) captures the variation in the blood osmolality (water content in the blood) due to dehydration, across the OFDM subcarriers. iii) Human body, particularly the tissues containing water, absorbs the microwave frequencies (i.e., Wi-Fi, 4G/5G signals). Thus, a hydrated person tends to absorb more RF energy and reflect lesser energy compared to a dehydrated person who will absorb lesser energy and reflect more energy. All three factors assist our ML and DL classifiers in achieving high classification accuracy.

Outline. The rest of this paper is organized as follows. Section II discusses the related work. Section III provides a compact discussion of the apparatus/equipment that provides the scaffolding for our proposed non-contact dehydration monitoring method. Section IV provides further details about the software and hardware setup used for data collection, specifics of each of the two proposed experiments (chest-based, and hand-based), as well as the data acquisition protocol implemented in order to construct our custom HCDDM-RF-5 dataset. Section V talks about the training and testing of various ML classifiers on our custom dataset, and provides a detailed performance analysis. Section VI concludes.

II. Related work

The literature on dehydration monitoring is scarce, but could be broadly classified into three categories: i) invasive methods, ii) non-invasive but contact-based methods, iii) non-contact methods. First kind of methods (i.e., invasive methods) which examine blood or urine samples in order to determine the plasma and urine osmolality (and are considered as gold standard) have already been discussed in section I. Further, to the best of our knowledge, there exists no work for third kind of methods (i.e., non-contact methods) for dehydration monitoring in the open literature. Therefore, we summarize the related work on second kind of methods (i.e., non-invasive methods) only.

A. Non-invasive methods for dehydration monitoring

The non-invasive methods for dehydration monitoring typically employ wearable sensors (e.g., oximeters, smart watch, smart wrist-bands, etc.) that capture the photoplethysmography (PPG) and electrodermal activity (EDA) signals and pass them through various ML algorithms in order to infer the dehydration status of a subject.

For example, [13] collects both the EDA and the PPG data from 17 subjects and feeds it to a range of ML algorithms in order to detect mild dehydration by exploiting the autonomic response to cognitive stress (induced by means of Stroop test). In [14], authors collect EDA data from 16 subjects for three different body postures (sitting, standing, and walking), and pass it to a hybrid Bi-LSTM neural network in order to classify the hydration level of a subject into one of the three different states (hydrated, moderate dehydration, extreme dehydration). Authors of [15] utilize a miniature pulse oximeter to collect PPG data from 17 dehydrated patients admitted in emergency of a tertiary care hospital. They then extract multiple features from the acquired PPG data using the variable frequency complex demodulation algorithm, feed them to a support vector machine classifier, and report an accuracy of 67.91%.

[16] collects the EDA data, skin temperature, heart rate and body mass index from 16 participants while they undergo a workout/physical activity known as circuit training. It then feeds this data to an empirically derived formula in order to quantify fluid loss (dehydration) caused by physical activity. In [17], authors developed a real-time Android-based tool called "monitoring my dehydration" that utilizes the EDA data to learn the dehydration level of a person using machine learning techniques. They did experimental evaluation of their tool by feeding it real-world data from five users, obtaining an accuracy of 84.5%. In [18], authors collect EDA data using BITalino kit from 5 subjects for three different activities by the subjects (sitting, standing, laying down), feed their data to various ML classifiers to solve the binary classification problem of dehydration detection, and report best classification accuracy of 91.3% using the random forests ML classifer. In [19] authors collect EDA data from several subjects under different conditions (sitting, standing), feed it to several ML classifiers to solve the binary classification problem of dehydration detection, and report a maximum accuracy of 87.78% using the simple KNN classifier. Finally, [20] takes a rather different approach, and utilizes a leg skin microbiome data from 63 female subjects in order to accurately predict their skin hydration levels and several other important bio-markers.

Before we conclude this section, it is imperative to have a quick discussion about the rise of non-contact methods for remote health sensing in the post-covid19 era.

B. Non-contact methods for health sensing

A vast majority of the literature on non-contact monitoring methods aims solely at monitoring of body vitals and respiratory system performance, in the post-covid19 era [21]–[24]. Non-contact health sensing methods for estimation of body vitals could be categorized into following four categories [25]: 1) Camera-based sensing (that uses the periodic change in skin colour of the face) [26], [27]. 2) Radar-based sensing (that utilizes the traditional radar principles of range and Doppler) [28], [29]. 3) Wi-Fi-based sensing (that uses the data collected off the reflections from the human subjects) [30], [29]. 4) Software-defined radio (SDR)-based sensing: (that utilizes the signals reflected off the human body) [31], [32].

Note that the proposed non-contact CBDM method and HBDM method both do SDR-based sensing for dehydration monitoring. To the best of authors' knowledge, non-contact monitoring of dehydration has not been reported in the open literature, to date.

III. The Apparatus for Non-Contact Dehydration Monitoring

The proposed non-contact system for dehydration monitoring is basically an RF transceiver that consists of two workstations, each connected with a software-defined radio (SDR) by means of a USB 3.0 port (see Fig. 1). Specifically, the SDR devices used for experiments are Universal Software Radio Peripheral (USRP) model B210[1]. Each SDR communicates with other by means of a directional horn antenna. We use MATLAB R2021a to program both the transmit and receive USRP SDRs. Specifically, the transmit SDR sends an orthogonal frequency division multiplexing (OFDM) signal with quadrature phase shift keying (QPSK) modulation on each subcarrier, while the receive SDR receives it and processes it.

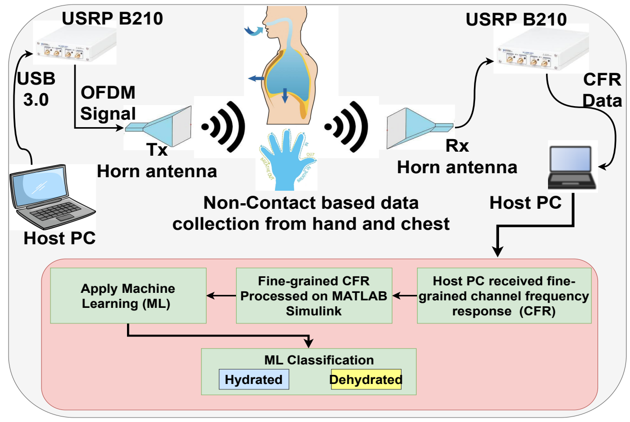


Fig. 1: The proposed non-contact method for dehydration monitoring: The apparatus consists of an SDR-based RF transceiver to collect radio data off the chest and the hand of the subject. The collected data is subsequently passed to various machine learning methods, which ultimately classify a subject either as hydrated or dehydrated.

Next, with the aim of non-contact dehydration monitoring, we design two distinct experiments. During the first experiment, the subject's chest is exposed to the OFDM signals, and thus, the receive SDR collects the signal reflected off the chest of the subject. We name this method as chest-based dehydration monitoring (CBDM) method. During the second experiment, the subject's hand is exposed to the OFDM signals, and thus, the receive SDR collects the signal that passes through the hand of the subject. We name this method as hand-based dehydration monitoring (HBDM) method[2].

IV. Acquisition of the HCDDM-RF-5 Dataset

This section provides sufficient details about the hardware and software setup used to construct the custom HCDDM-RF-5 dataset[3], our thoughtful data collection methodology (that helped us capture dehydration related data in a very controlled manner), as well as the intricate details of the two experiments performed in order to collect data for the two proposed (CBDM and HBDM) methods.

A. USRP SDRs based OFDM transceiver

OFDM Transmitter: For each OFDM frame, the random bits generator block creates pseudo-random data bits with a chunk size of 128 bits. The QPSK modulator block maps these bits

[1]The USRP B210 from National Instruments covers a wide frequency range (70 MHz to 6 GHz). It can process a wideband spectrum of up to 56 MHz in real time and sample at a high rate of up to 61.44 MS/s.

[2]This study was approved by the ethical institutional review board (EIRB) of Information Technology University, Lahore, Pakistan.

[3]The acronym HCDDM-RF-5 stands for **H**and and **C**hest **D**ata for **D**ehydration **M**onitoring via **R**adio **F**requency data collected from **5** subjects.

Parameter	Type/Value
Bits per OFDM frame	128
Bits per symbol	2
Coding scheme	Gray coding
Modulation scheme	QPSK
No. of OFDM subcarriers (N)	64
Data subcarriers	52
Pilot subcarriers	12
Size of FFT/IFFT	64 points
Size of cyclic prefix	16
Sampling rate	1000 samples/sec
Antenna type	directional horn
USRP B210 frequency range	70 MHz - 6 GHz
Centre frequency	5.23 GHz
Clock source & PPS source	Internal
Internal clock rate	200 MHz
Interpolation factor (at Tx)	250
Decimation factor (at Rx)	250
Transmitter gain (at Tx and Rx)	40 dB

TABLE I: Some parameters for the USRP-SDR-based OFDM transceiver used for non-contact monitoring of dehydration.

to (frequency domain) symbols which are then transformed into a time-domain signal by means of an inverse fast Fourier transform (IFFT) of size $N = 64$ points. Further, a cyclic prefix (CP) of size 16 samples is appended to each OFDM frame, making each OFDM frame 80 samples long. Gain of the transmit horn antenna is set to 40 dBi. Fig. 2(a) shows the Simulink flowgraph of USRP SDR based OFDM transmitter.

OFDM Receiver: After removing the CP from each OFDM frame, fast Fourier transform (FFT) is then used to transform the received time-domain OFDM samples into the equivalent frequency-domain OFDM symbol. Then, keeping in mind that the transmitted QPSK symbols on each sub-carrier are known to the OFDM receiver, the channel coefficient h_i for i-th sub-carrier could simply be computed as: $h_i = \frac{y_i}{x_i}$, where x_i,y_i are the transmitted and received QPSK symbol on i-th sub-carrier, respectively. This way, the raw channel frequency response (CFR) data $\mathbf{h} = [h_1, \cdots, h_N]^T$ is collected by the OFDM transmitter, which is to be utilized later by the ML algorithms in order to classify the status of each subject as either hydrated or dehydrated. Fig. 2(b) shows the Simulink flowgraph of USRP SDR based OFDM transmitter.

Table I provides a quick summary of setting of various relevant parameters of transmit and receive USRP SDRs.

B. Data Acquisition for the HCDDM-RF-5 dataset

The custom HCDDM-RF-5 dataset was constructed by collecting data from five volunteers during the month of Ramadan (between March 23rd, 2023 and April 21st, 2023). Ramadan is an Islamic holy month during which devout Muslims observe a strict fast from sunrise till sunset. That is, while they are fasting, Muslims refrain from eating and drinking from sunrise till sunset. We took advantage of this unique opportunity in order to collect dehydration related data from five devout Muslims who had been fasting during this month. Among five subjects, two were males (aged 28, 62 years), and three were females (aged 21, 26, 61 years). For each fasting subject, we collected data twice, once for each class label (hydrated and dehydrated) in order to construct a balanced dataset. Specifically, first episode of data collection took place about 30 minutes before the sunset when the subject was deemed to be maximally dehydrated (thus, this data belongs to the first/dehydrated class). Subsequently, the second episode of the data collection took place an hour after the sunset, after the subject had finished eating and drinking after breaking the fast (thus, this data belongs to the second/hydrated class). For each subject, we conducted two kinds of experiments where we exposed the subject's chest and hand to the RF signals, respectively. Some more pertinent details about data collection for our proposed CBDM and HBDM methods are given below.

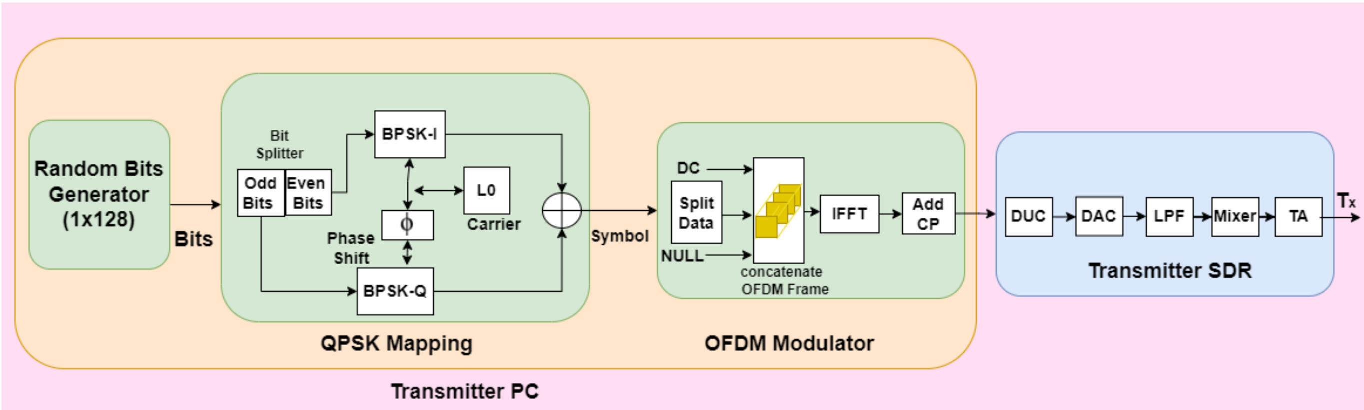
Data collection for the proposed CBDM method: During data acquisition for the proposed CBDM method, each participant sat on a chair that was about 80 cm away from the pair of directional horn antennas that pointed towards the chest of the subject (see Fig. 3). As described before, the transmit horn antenna impinged an OFDM signal onto the chest of the subject, while the receive horn antenna gathered the signal reflected off the subject's chest. During each experiment session, the subject sat still in order to avoid motion-induced artefacts in the data being gathered. Each single experiment session lasted for 30 seconds. For each subject, we conducted five experiment sessions before the sunset (to capture the raw CFR data for dehydrated class) and five experiment sessions after the sunset (to capture the raw CFR data for the hydrated class). This way, we were able to collect $30 \times 5 = 150$ seconds worth of data for each class (for a given subject), and thus, $150 \times 2 = 300$ seconds worth of data per subject. Ultimately, for 5 subjects, this led to a total dataset size of $300 \times 5 = 1500$ seconds (or, 25 minutes) of raw CFR data (that corresponds to a total of $5 \times 5 \times 2 = 50$ experiment sessions).

Data collection for the proposed HBDM method: During data acquisition for the proposed HBDM method, each participant sat on a chair that was about 60 cm away from the pair of directional horn antennas facing each other, and placed his/her hand on the table between the two antennas (see Fig. 4). Again, the transmit horn antenna impinged an OFDM signal onto the hand of the subject, while the receive horn antenna gathered the signal passed through the subject's hand. During each experiment session, the subject sat still in order to avoid motion-induced artefacts in the data being gathered. The rest of the details of data acquisition for the proposed HBDM method are the same as before. That is, for each subject, we conducted five experiment sessions before the sunset (to capture the raw CFR data for dehydrated class) and five experiment sessions after the sunset (to capture the raw CFR data for the hydrated class). This way, for 5 subjects, we acquired a dataset that consisted of $300 \times 5 = 1500$ seconds (or, 25 minutes) of raw CFR data (that corresponds to a total of $5 \times 5 \times 2 = 50$ experiment sessions).

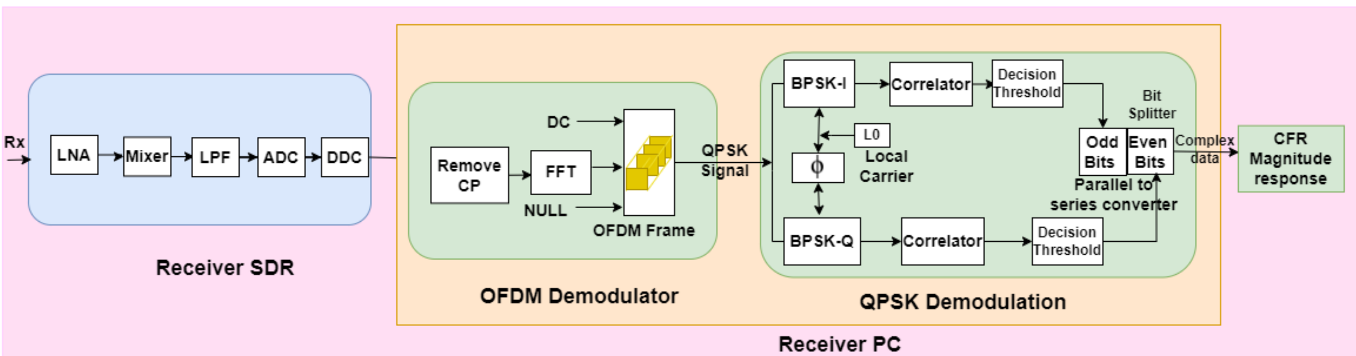
In short, combining the two smaller datasets due to CBDM method and HBDM method together, the custom HCDDM-RF-5 dataset consists of a total of 50 minutes of raw CFR data that corresponds to a total of 100 experiment sessions.

V. Training and Testing of Machine & Deep Learning Classifiers

For the binary classification problem (hydrated/dehydrated) under consideration, we train and test the following ML



(a) The transmitter flowgraph



(b) The receiver flowgraph

Fig. 2: The Simulink flowcharts of the USRP-SDR based OFDM transmitter and receiver.

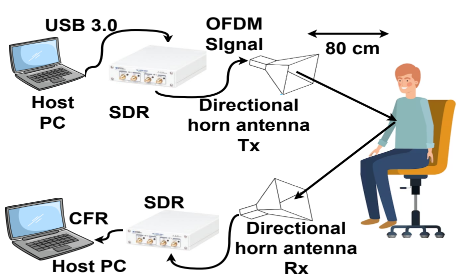


Fig. 3: Experimental setup of the proposed CBDM method.

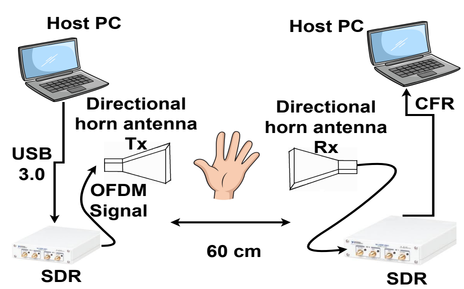


Fig. 4: Experimental setup of the proposed HBDM method.

classifiers and their variants: k-nearest neighbours (KNN), support vector machine (SVM), decision tree (DT), ensemble classifier. In addition, we also implement a neural network and its variants. Specifically, the neural network (NN) that we have implemented is a feedforward neural network (FFNN). The narrow, medium and wide NN variants are all FFNN with a single hidden layer, but with 10, 25, 100 neurons, respectively. Further, the bi-layered NN (tri-layered NN) has two (three) hidden layers with 10 neurons in each hidden layer.

Subsequently, we provide detailed performance analysis and comparison of all the ML and DL classifiers implemented.

A. Data Pre-processing & Training of Machine Learning Classifiers

Data Pre-processing: We utilised a low-pass filter and a Savitzky-Golay filter to denoise the CFR extracted from the received OFDM signal, for all the experiment sessions (for both CBDM and HBDM methods). We inspected the whole data manually and removed artifacts where found. Further, recall that we gathered data from five subjects, and for each subject, we conducted five experiment sessions each of duration 30 seconds. Thus, for each class (hydrated/dehydrated), we accu-

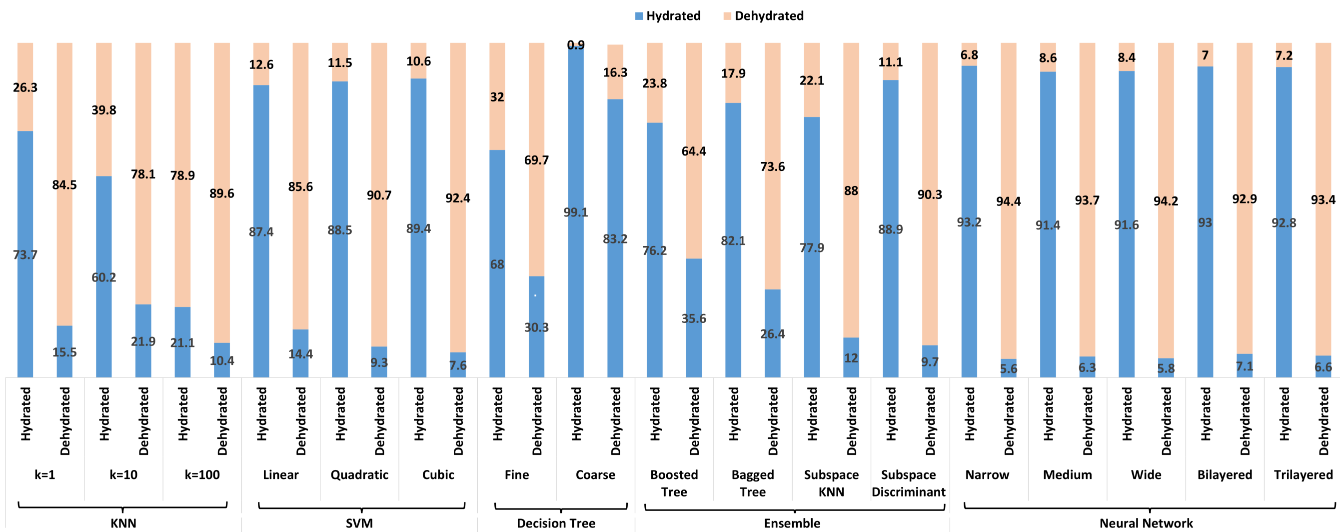


Fig. 5: Confusion matrix of each of KNN, SVM, DT, Ensemble classifiers, and NN for the proposed CBDM method.

mulated raw CFR data which consists of 750,000 samples, i.e., 5 subjects * 5 experiments/subjects * 30 second/experiment * 1000 samples/sec. We partitioned this data into segments, with each segment consisting of 3000 samples. Thus, through this meticulous approach, we successfully compiled 250 examples for each of the "hydrated" and "dehydrated" classes.

Training & validation of ML classifiers: The Matlab's classification learner app was used to train the following ML classifiers: k-nearest neighbour (KNN), support vector machine (SVM), decision tree (DT), ensemble classifier, and neural network. All the classifiers were trained on both labelled datasets (corresponding to the CBDM method and the HBDM method). The K-fold cross-validation strategy was used for validation in order to prevent the over-fitting issue, keeping in mind the small size of our dataset.

B. Performance metrics

Each classifier's performance is quantified in terms of accuracy, given as:

$$\text{Accuracy} = \frac{\text{Correct prediction}}{\text{Total observations}} \times 100 \quad (1)$$

$$\text{Accuracy} = \frac{T_n + T_p}{T_n + T_p + F_n + F_p} \times 100 \quad (2)$$

where T_n represents a true negative, T_p represents a true positive, F_n represents a false negative, and F_p represents a false positive. In addition, we also do a performance comparison of the various ML algorithms by means of a confusion matrix.

C. Performance of proposed CBDM method

We begin with performance analysis of the KNN classifier for three distinct values of k, i.e., k=1, k=10, and k=100 (where k is the number of neighbours used to calculate the distance to the new data point). We learn that the fine KNN (k=1) achieves an accuracy of 79.1%, medium KNN (k=10) achieves an accuracy of 69.2%, while the coarse KNN (k=100) achieves a very low accuracy of 55.3% (see Fig. 5 that displays the detailed confusion matrix).

Next, we do a performance comparison of the SVM classifier (with linear, quadratic, and cubic kernels), we note that the linear SVM achieves an overall accuracy of 86.5%, quadratic SVM achieves an overall accuracy of 89.6%, while the cubic SVM achieves an overall accuracy of 90.9%. Next, we focus on the decision tree classifier, and note that it has the lowest accuracy of all. That is, the fine tree (despite its many leaves and despite its ability to differentiate between classes precisely) achieved an accuracy of 68.8% only, while the coarse tree achieved a very low accuracy of 58.0% only. Next in line is the ensemble classifier (a mixture of many classifiers) that is typically implemented with the aim to boost classification accuracy. We observe the following: the ensemble Boosted tree has an overall accuracy of 70.3%, the ensemble Bagged tree has an accuracy of 77.9%, the ensemble subspace KNN has an accuracy of 82.9%, while the ensemble subspace discriminant has an accuracy of 89.6%. Finally, the neural network (NN) classifier. Each variant of the NN classifier is a fully-connected feedforward network. After each fully connected layer, the Relu activation function is applied, except the last year where softmax activation function is used. We observe that all the different variants of the NN classifier outperform the other ML classifiers. Specifically, the narrow variant of the NN achieves an accuracy of 93.8%, the medium NN achieves an accuracy of 92.5%, the wide NN achieves an accuracy of 92.9%, the bi-layered variant of NN achieves an accuracy of 93%, while the tri-layered variant of the NN achieves an accuracy of 93.1%.

Fig. 6 provides an alternate way of comparing the overall accuracy of all the five ML classifiers and their variants. We note that, for the proposed CBDM method, the neural network classifier (with the narrow neural network) achieves the highest accuracy, which is 93.8%.

Last but not the least, we also implemented an unsupervised learning method, namely, k-means clustering method,

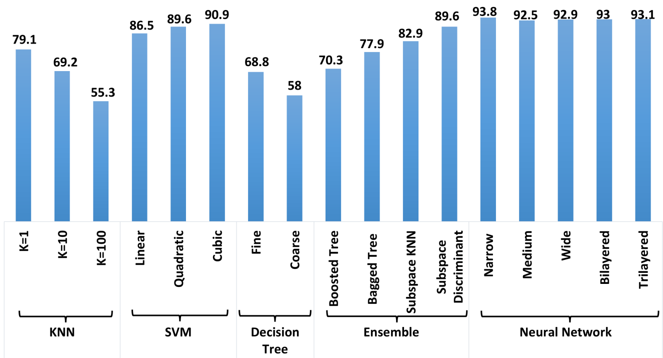


Fig. 6: Accuracy (in percentage) of all the classifiers for the proposed CBDM method.

by stripping-off the labels from both classes. We achieved an accuracy of 88.88% for CBDM, which is quite satisfying and attests to the fact that non-contact monitoring of dehydration is indeed feasible, and that, the two classes are indeed separable.

D. Performance of proposed HBDM method

Fig. 7 shows the confusion matrix of each of the five ML classifiers and their variants. Beginning with the performance analysis of the KNN classifier, we learn that the fine KNN achieves an accuracy of 82.3%, medium KNN achieves an accuracy of 71.3%, while the coarse KNN achieves a very low accuracy of 56.1%. Next SVM classifier with linear, quadratic, and cubic kernels, we note that the linear SVM achieves an overall accuracy of 71.1%, quadratic SVM achieves an overall accuracy of 89.2%, while the cubic SVM achieves an overall accuracy of 88.2%. Next, the decision tree classifier. We observe that once again it has the lowest accuracy of all. That is, the fine tree achieved an accuracy of 72.2% only, while the coarse tree achieved a very low accuracy of 61.4% only. Next, the ensemble classifier. We observe the following: the ensemble Boosted tree has an overall accuracy of 74.8%, while the ensemble Bagged tree has an accuracy of 79.7%. Finally, the neural network classifier. Once again, all the different variants of the NN classifier outperform the other ML classifiers. Specifically, the narrow variant of the NN achieves an accuracy of 94.7%, the medium NN achieves an accuracy of 96.15%, the wide NN achieves an accuracy of 95.15%, the bi-layered variant of NN achieves an accuracy of 92.35%, while the tri-layered variant of the NN achieves an accuracy of 94.2%.

Fig. 8 provides an alternate way of comparing the overall accuracy of all the five ML classifiers and their variants. We note that, for the proposed HBDM method, the neural network classifier (with the medium neural network) achieves the highest accuracy, which is 96.15%. Finally, we have implemented the k-means unsupervised learning method, reporting an accuracy of 94.44%, which demonstrates a highly satisfactory performance among unsupervised learning methods.

E. Performance comparison with the state-of-the-art

Finally, Table II compares the accuracy of the proposed non-contact CBDM and HBDM methods with the state-of-the-art methods which are all contact-based methods for dehydration monitoring. Compared to the state-of-the-art where the maximum reported accuracy is 97.83%, our proposed non-contact method provides slightly less accuracy (as we report a maximum accuracy of 96.15%); nevertheless, the advantages of our non-contact dehydration method speak for themselves. That is, our proposed method is non-invasive and contact-less, allows continuous and seamless monitoring, is easy to use, and provides rapid results with quite a higher accuracy.

F. Computational complexity & latency analysis

To train and test our ML and DL models, we used a laptop with the following specs: Intel core i7 3.6 GHz processor with 16 GB RAM. We now do the computational complexity and latency analysis of our ML and DL classifiers in terms of the following: computational complexity (in big-O notation), training time (seconds), and prediction speed (observations/sec).

- The complexity of KNN is O(N), and training time remains fixed (to 2467.1 seconds), regardless of the value of k. Here, N represents the size of the dataset (i.e., number of examples in the dataset). The prediction speeds for KNN with k values of 1, 10, and 100 are 38, 38, and 34 observations/sec, respectively.
- The complexity of SVM is O($N \times D^p$) where $p = 1, 2, 3$ for linear, quadratic, and cubic SVM, respectively. Here, D represents the dimension of segmented data. The training time increases from linear to quadratic and to cubic, as follows: 6723.2 seconds for the linear, 9977.6 seconds for the quadratic, and 14077 seconds for the cubic. The prediction speed for all SVM variants are 55 observations/sec for the linear, 110 observations/sec for the quadratic, and 130 observations/sec for the cubic.
- The complexity of DT is O($\log(N) \times D \times T$) where T represents the number of splits. The training time decreases from the fine variant, which takes 164.29 seconds to the coarse variant, it only requires 62.5 seconds. While the training time varies, the prediction speed remains constant at 9600 observations/sec.

The computational complexity and latency of the four variants of the ensemble method are as follows:

- The complexity of Boosted tree is O($N \times D \times C \times M$) where, M represents the number of boosting iterations, and C represents complexity of training a single DT. The Boosted tree exhibits a longer training time, specifically 8457.5 seconds, owing to its inherently sequential and iterative optimization process. The prediction speed for Boosted tree is 9100 observations/sec.
- The complexity of Bagged tree is O($N \times \log(N) \times D \times K$) where K represents the number of trees. Bagged tree requires a lot less time to train than Boosted tree, only 391.4 seconds, mainly because it uses parallel processing. The prediction speed is 7400 observations/sec.
- The complexity of Subspace KNN is O($N \times S \times D_i$) where, S represents the number of subspaces, and D_i is the dimensionality of a subspace. The training time of subspace KNN is 1438.9 seconds, which is shorter than KNN due to the dimensionality reduction of the

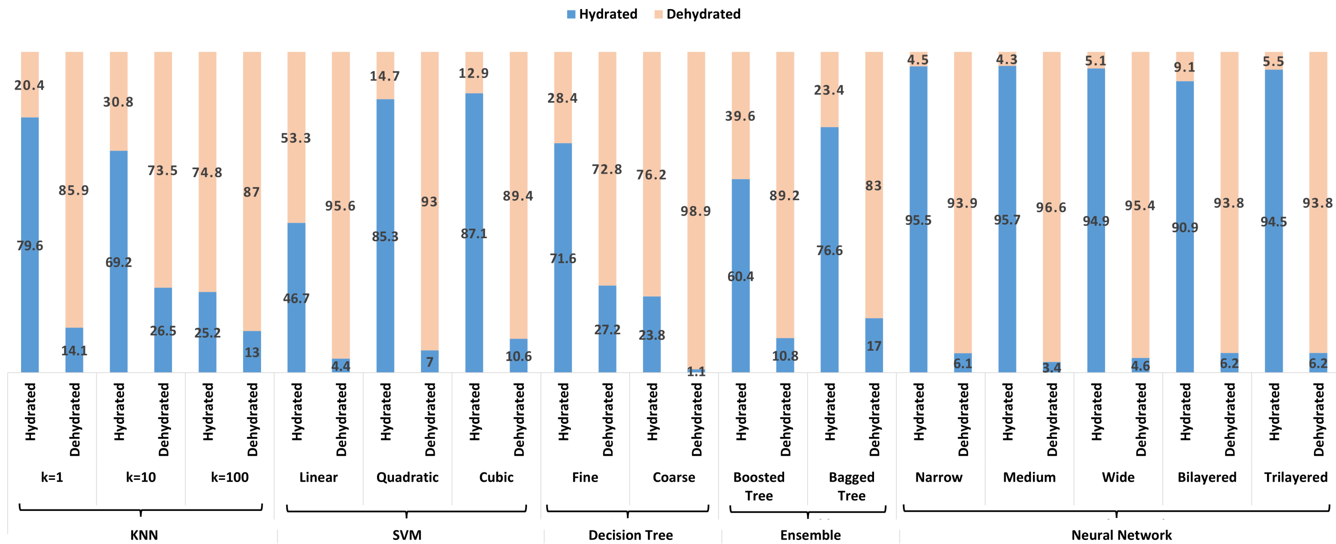


Fig. 7: Confusion matrix of each of KNN, SVM, DT, Ensemble classifiers, and NN for the proposed HBDM method.

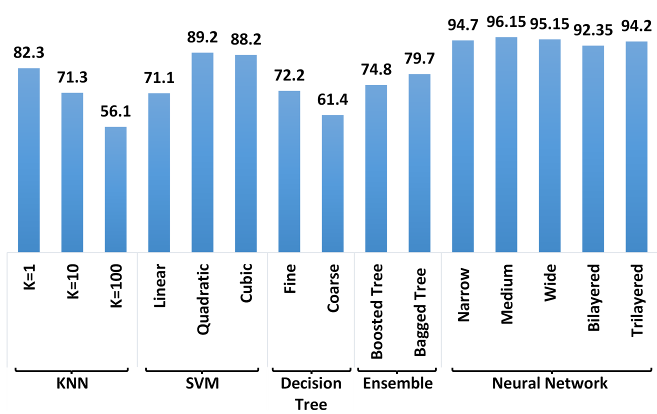


Fig. 8: Accuracy (in percentage) of all the ML classifiers for the proposed HBDM method.

subspace. The prediction speed for this classifier is 270 observations/sec.

- The complexity of Subspace discriminant is O($S\times$training complexity). The training time for this classifier is 370061 seconds, which is considerably higher due to the presence of multiple subspaces where the classifier computes discriminant for each subspace. The prediction speed is 3.5 observations/sec.

Among the four variants, the training time is minimum for Bagged tree and maximum for the subspace discriminant.

For the neural network, the computational complexity depends upon the number of hidden layers L, number of weights W' (neurons) in each layer, and number of data points N. For example, for backward propagation, the complexity of the different variants of the neural networks is as follows. The complexity of the narrow NN is: O($N \times W'$). The complexity of the medium NN is: O($N \times W'$). the complexity of the wide NN is: O($N \times W'$). The complexity of the bi-layered NN is: O($N \times W' \times L$). The complexity of the tri-layered NN is: O($N \times W' \times L$). Let $W' \times L = W$. Then, we can write the computational complexity of all the neural network variants (for backward propagation) in the compact form as follows: O($N \times W$). The training time for neural networks increases with an increase in the number of learnable parameters, specifically the number of hidden neurons. The training time and prediction speed for NN are as follows: the narrow NN takes 1767 seconds and a prediction speed is 8400 observations/sec, the medium NN takes 4316 seconds and a prediction speed is 7400 observations/sec, the wide NN takes 8630.3 seconds and prediction speed is 6400 observations/sec, the bi-layered NN takes 1685.8 seconds and a prediction speed is 8700 observations/sec, and the tri-layered NN takes 1694.1 seconds and prediction speed is 8300 observations/sec.

Work	Method	Accuracy
Liaqat et al. [14]	Contact-based	97.83%
Kulkarni et al. [17]	Contact-based	75.96%
Liaqat et al. [18]	Contact-based	91.53%
Rizwan et al [19]	Contact-based	85.63%
Carrieri et al. [20]	Contact-based	73.91 %
Our proposed CBDM method	Non-contact based	93.8%
Our proposed HBDM method	Non-contact based	96.15%

TABLE II: Accuracy comparison of our proposed CBDM and HBDM methods with the state-of-the-art.

In short, among all the ML and DL models implemented, decision tree (coarse) and Bagged tree (as ensemble method) have the lowest training times of 62.5 seconds and 391.4 seconds, decision tree has the highest prediction speed (of 9600 predictions/sec), while KNN has the least computational complexity.

VI. Conclusion

This work reported a novel non-contact method to monitor the dehydration status of a subject from a distance. Specifically, we utilized a pair of USRP SDRs whereby the transmit SDR impinged OFDM signals onto the chest or the hand of the subject, while the receive SDR collected the modulated signal reflected off the body of the subject. For the purpose of training our ML classifiers, we collected data from 5 Muslim subjects (before and after sunset) who were fasting during the month of Ramadan. We achieved a classification accuracy of 93.8% for the proposed CBDM method, and an accuracy of 96.15% for the proposed HBDM method.

Our proposed method is non-invasive and contact-less, has high accuracy, allows continuous and seamless monitoring, is easy to use, and provides rapid results. The anticipated beneficiaries of the proposed method include: sportsmen, athletes, elderly, diabetic and diarrhea patients, and labor working outdoors. Additionally, the proposed method may help realize smart mobile health (m-health) solutions that could be deployed in remote areas far away from the mega cities, in order to provide comprehensive health monitoring of the people living there.

This work opens up many exciting directions for the future work. For example, one could construct/acquire a more challenging dataset (unlike the current dataset that was obtained in a very controlled setting), and re-evaluate as well as fine-tune the performance of our ML and DL models further, in order to make them robust and amicable to the unseen data with potentially different distribution.

References

[1] B. M. Popkin, K. E. D'Anci, and I. H. Rosenberg, "Water, hydration, and health," *Nutrition reviews*, vol. 68, no. 8, pp. 439–458, 2010.

[2] A. M. El-Sharkawy, O. Sahota, and D. N. Lobo, "Acute and chronic effects of hydration status on health," *Nutrition reviews*, vol. 73, no. suppl_2, pp. 97–109, 2015.

[3] S. Liaqat, K. Dashtipour, A. Zahid, K. Arshad, S. Ullah Jan, K. Assaleh, and N. Ramzan, "A review and comparison of the state-of-the-art techniques for atrial fibrillation detection and skin hydration," *Frontiers in Communications and Networks*, vol. 2, p. 679502, 2021.

[4] D. Black, R. McCance, and W. Young, "A study of dehydration by means of balance experiments," *The Journal of physiology*, vol. 102, no. 4, p. 406, 1944.

[5] A. M. El-Sharkawy, P. Watson, K. R. Neal, O. Ljungqvist, R. J. Maughan, O. Sahota, and D. N. Lobo, "Hydration and outcome in older patients admitted to hospital (the hoop prospective cohort study)," *Age and ageing*, vol. 44, no. 6, pp. 943–947, 2015.

[6] L. E. Armstrong, "Assessing hydration status: the elusive gold standard," *Journal of the American College of Nutrition*, vol. 26, no. sup5, pp. 575S–584S, 2007.

[7] L. E. Armstrong, C. M. Maresh, J. W. Castellani, M. F. Bergeron, R. W. Kenefick, K. E. LaGasse, and D. Riebe, "Urinary indices of hydration status," *International Journal of Sport Nutrition and Exercise Metabolism*, vol. 4, no. 3, pp. 265–279, 1994.

[8] S. N. Cheuvront, B. R. Ely, R. W. Kenefick, and M. N. Sawka, "Biological variation and diagnostic accuracy of dehydration assessment markers," *The American journal of clinical nutrition*, vol. 92, no. 3, pp. 565–573, 2010.

[9] R. A. Oppliger, S. A. Magnes, L. A. Popowski, and C. V. Gisolfi, "Accuracy of urine specific gravity and osmolality as indicators of hydration status," *International journal of sport nutrition and exercise metabolism*, vol. 15, no. 3, pp. 236–251, 2005.

[10] L. A. Popowski, R. A. Oppliger, P. Lambert, R. F. Johnson, C. Gisolf *et al.*, "Blood and urinary measures of hydration status during progressive acute dehydration." *Medicine and science in sports and exercise*, vol. 33, no. 5, pp. 747–753, 2001.

[11] S. N. Cheuvront and M. N. Sawka, "Hydration assessment of athletes," *CHINESE JOURNAL OF SPORTS MEDICINE*, vol. 25, no. 2, p. 238, 2006.

[12] N. Kulkarni, C. Compton, J. Luna, and M. A. U. Alam, "Monitoring my dehydration: A non-invasive dehydration alert system using electrodermal activity," *arXiv preprint arXiv:2009.13626*, 2020.

[13] H. F. Posada-Quintero, N. Reljin, A. Moutran, D. Georgopalis, E. C.-H. Lee, G. E. Giersch, D. J. Casa, and K. H. Chon, "Mild dehydration identification using machine learning to assess autonomic responses to cognitive stress," *Nutrients*, vol. 12, no. 1, p. 42, 2019.

[14] S. Liaqat, K. Dashtipour, A. Rizwan, M. Usman, S. A. Shah, K. Arshad, K. Assaleh, and N. Ramzan, "Personalized wearable electrodermal sensing-based human skin hydration level detection for sports, health and wellbeing," *Scientific Reports*, vol. 12, no. 1, p. 3715, 2022.

[15] N. Reljin, Y. Malyuta, G. Zimmer, Y. Mendelson, D. J. Blehar, C. E. Darling, and K. H. Chon, "Automatic detection of dehydration using support vector machines," in *2018 14th Symposium on Neural Networks and Applications (NEUREL)*. IEEE, 2018, pp. 1–6.

[16] N. K. Suryadevara, S. C. Mukhopadhyay, and L. Barrack, "Towards a smart non-invasive fluid loss measurement system," *Journal of medical systems*, vol. 39, pp. 1–10, 2015.

[17] N. Kulkarni, C. Compton, J. Luna, and M. A. U. Alam, "A non-invasive context-aware dehydration alert system," in *Proceedings of the 22nd International Workshop on Mobile Computing Systems and Applications*, 2021, pp. 157–159.

[18] S. Liaqat, K. Dashtipour, K. Arshad, and N. Ramzan, "Non invasive skin hydration level detection using machine learning," *Electronics*, vol. 9, no. 7, p. 1086, 2020.

[19] A. Rizwan, N. A. Ali, A. Zoha, M. Ozturk, A. Alomainy, M. A. Imran, and Q. H. Abbasi, "Non-invasive hydration level estimation in human body using galvanic skin response," *IEEE Sensors Journal*, vol. 20, no. 9, pp. 4891–4900, 2020.

[20] A. P. Carrieri, N. Haiminen, S. Maudsley-Barton, L.-J. Gardiner, B. Murphy, A. Mayes, S. Paterson, S. Grimshaw, M. Winn, C. Shand *et al.*, "Explainable ai reveals key changes in skin microbiome associated with menopause, smoking, aging and skin hydration," *bioRxiv*, pp. 2020–07, 2020.

[21] A. Hafeez, S. Ahmad, S. A. Siddqui, M. Ahmad, and S. Mishra, "A review of covid-19 (coronavirus disease-2019) diagnosis, treatments and prevention," *Ejmo*, vol. 4, no. 2, pp. 116–125, 2020.

[22] K. Yatani and K. N. Truong, "Bodyscope: a wearable acoustic sensor for activity recognition," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, 2012, pp. 341–350.

[23] E. Ertin, N. Stohs, S. Kumar, A. Raij, M. Al'Absi, and S. Shah, "Autosense: unobtrusively wearable sensor suite for inferring the onset, causality, and consequences of stress in the field," in *Proceedings of the 9th ACM conference on embedded networked sensor systems*, 2011, pp. 274–287.

[24] U. Saeed, S. Y. Shah, A. Zahid, J. Ahmad, M. A. Imran, Q. H. Abbasi, and S. A. Shah, "Wireless channel modelling for identifying six types of respiratory patterns with sdr sensing and deep multilayer perceptron," *IEEE Sensors Journal*, vol. 21, no. 18, pp. 20 833–20 840, 2021.

[25] M. Rehman, R. A. Shah, M. B. Khan, S. A. Shah, N. A. AbuAli, X. Yang, A. Alomainy, M. A. Imran, and Q. H. Abbasi, "Improving machine learning classification accuracy for breathing abnormalities by enhancing dataset," *Sensors*, vol. 21, no. 20, p. 6750, 2021.

[26] I. Sato and M. Nakajima, "Non-contact breath motion monitor ing system in full automation," in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*. IEEE, 2006, pp. 3448–3451.

[27] A. Al-Naji, K. Gibson, S.-H. Lee, and J. Chahl, "Monitoring of cardiorespiratory signal: Principles of remote measurements and review of methods," *Ieee Access*, vol. 5, pp. 15 776–15 790, 2017.

[28] K. Van Loon, M. Breteler, L. Van Wolfwinkel, A. Rheineck Leyssius, S. Kossen, C. Kalkman, B. Van Zaane, and L. Peelen, "Wireless non-invasive continuous respiratory monitoring with fmcw radar: a clinical validation study," *Journal of clinical monitoring and computing*, vol. 30, pp. 797–805, 2016.

[29] C. Massaroni, A. Nicolo, M. Sacchetti, and E. Schena, "Contactless methods for measuring respiratory rate: A review," *IEEE Sensors Journal*, vol. 21, no. 11, pp. 12 821–12 839, 2020.

[30] K. Ali, M. Alloulah, F. Kawsar, and A. X. Liu, "On goodness of wifi based monitoring of vital signs in the wild," *arXiv preprint arXiv:2003.09386*, 2020.

[31] M. Rehman, R. A. Shah, M. B. Khan, N. A. AbuAli, S. A. Shah, X. Yang, A. Alomainy, M. A. Imran, and Q. H. Abbasi, "Rf sensing based breathing patterns detection leveraging usrp devices," *Sensors*, vol. 21, no. 11, p. 3855, 2021.

[32] K. Pervez, W. Aman, M. M. U. Rahman, M. W. Nawaz, and Q. H. Abbasi, "Hand-breathe: Non-contact monitoring of breathing abnormalities from hand palm," *IEEE Sensors Journal*, 2023.