IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 11 April 2023, accepted 29 April 2023, date of publication 8 May 2023, date of current version 14 June 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3274484

RESEARCH ARTICLE

Type 2 Diabetes Detection With Light CNN From Single Raw PPG Wave

SERENA ZANELLI[1], MOUNIM A. EL YACOUBI[2], (Member, IEEE), MAGID HALLAB[3], AND MEHDI AMMI[1], (Member, IEEE)

[1]Laboratoire Analyse, Géométrie et Applications (LAGA), University of Sorbonne Nord Paris, 93430 Villetaneuse, France
[2]SAMOVAR Telecom SudParis, Institut Polytechnique de Paris, 91120 Palaiseau, France
[3]Clinique Bizet, 75116 Paris, France

Corresponding author: Serena Zanelli (zanelli@math.univ-paris13.fr)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Data Protection Committee at the University Hospital of Nice under Cohort Register No. BS-004.

ABSTRACT Photoplethysmography (PPG) is a non-invasive and cost-efficient optical technique used to assess blood volume variations in the microcirculation. PPG technology is widely used in a variety of wearable sensors to investigate the cardiovascular system. Recent studies have demonstrated the utility of PPG analysis for carrying out large-scale screening to prevent and detect diabetes. However, most of these studies require feature extraction and/or several pre-processing steps. Over the past few years, the advent of deep learning has significantly impacted the analysis of biomedical signals. Despite their success in other fields, however, very few studies have focused on the application of deep learning to raw PPG signals for detecting diabetes. Existing studies have proposed large models trained on large amounts of data. In this paper, we present a Light CNN-based model for screening the presence of type 2 diabetes using a single raw pulse extracted from photoplethysmographic signals. In addition to the baseline architecture, we evaluate different model architectures that take as input age and biological sex or PPG handcrafted features. Furthermore, we apply transfer learning to all the tested architectures to evaluate the effectiveness of harnessing pre-trained models in detecting diabetes. We tested a model pre-trained on a general PPG shape dataset and another model pre-trained on a dataset containing hypertension PPG signals. Our model scored an AUC of 75.5 when trained with raw PPG waves, age, and biological sex without applying transfer learning, which is competitive with current state of the art.

INDEX TERMS Deep learning, photoplethysmography, screening, signal processing, type 2 diabetes, transfer learning.

I. INTRODUCTION

Worldwide, three main types of diabetes exist. Type 1 diabetes is a chronic condition in which the pancreas produces little or no insulin by itself [1]. Type 2 diabetes usually occurs in adults when the body becomes resistant to insulin or does not make enough insulin [2]. The last type is gestational diabetes, a transient disease that some women can develop during pregnancy [3]. In this study, we focused on type 2 diabetes, the most common one [4]. The development of new low-cost, non-invasive and light technologies that are able to detect the onset of diabetes could make a difference in diabetes large-scale prevention [5]. Type 2 diabetes (DT2) can lead to increased arterial stiffness, alteration in heart rate variability, increased blood viscosity and heart failure [6]. Cardiovascular activity can be assessed using many different technologies. However, most of them require a clinical environment or clinical expertise to be exploited. Photoplethysmography is a non-invasive technology that can be used to acquire cardiovascular pulse wave signals. This technology is attracting interest from industries and the scientific community thanks to its cost-effective technology and usability. Using the PPG signal to detect diabetes is less expensive compared to gold standard methods as the

The associate editor coordinating the review of this manuscript and approving it for publication was Chuan Li.

glycosylated hemoglobin test, and it does not require well equipped laboratory. Another advantage is that the PPG signal is acquirable through several devices that are already used in everyday life such as smartphones and wristbands. Devices for diabetes test, by contrast, need to be specifically bought and this usually becomes the case when there is a suspicion of diabetes. Recently, several studies have investigated the possibility of detecting the diabetes through PPG signals with traditional, machine learning or deep learning approaches [7]. Since the availability of data is one of the main issues in biomedical research [8], we have focused on searching for a method that could succeed without the need of a large amount of data. For this reason, we propose a light CNN-based model trained over a small dataset in terms of size and signal length. Our dataset is composed of 100 subjects (15% with diabetes) and consists of single PPG waves. Considering the size of our dataset, we have investigated the effect of transfer learning. We trained two models with two larger datasets: the first contains generic PPG waves from 500 subjects while the second dataset contains PPG waves from 181 hypertension (HT) and healthy subjects. The goal is to assess if utilizing the model weights and biases from a generic PPG shape classification model or from a specific pathology classification model can enhance the DT2 detection through transfer learning.

Our contributions can be summarized as follows:

- Firstly, we have designed a baseline CNN model for detecting DT2 using PPG signals. Our model takes as input much shorter PPG segments compared to the previous studies, as it operates with single pulses normalized to a length of 1 second. This segment length enables the model to operate with a variety of signal lengths without requiring a long PPG acquisition. Moreover, it allows us to obtain a larger dataset, since we can extract up to 15 waves from a single signal, with the additional benefit that the input layer size of the CNN model becomes much smaller.
- We propose a Light CNN trained over a small dataset. Our results are comparable to those reported by Avram et al. [9] that use much larger nets and larger datasets.
- Lastly, we utilize transfer learning to investigate the possible benefits of using firstly a pre-trained model over a larger dataset and then fine tuning it with a very small dataset containing DT2 PPG signals, and secondly to investigate the relation between hypertension and diabetes.

The rest of the paper is organized as follows: In Section II, we present related works. In Section III, we briefly outline the background of PPG and its correlation with diabetes. Then, we describe the datasets used, the extracted features, the model architecture, and the process we have followed to investigate the transfer learning effect. In Section IV, we present and discuss the results. Finally, in Section V, we conclude and discuss future work.

II. RELATED WORKS

Several approaches have recently been employed in detecting modifications in PPG waveforms [10], [11]. These approaches can be categorized into three types: traditional, machine learning-based, and deep learning-based. Bagus et al. [12] proposed a traditional approach in which the Poincaré plot of finger PPG signals was used to significantly discriminate ($p<0.05$) between healthy subjects (glycohemoglobin$<6.5\%$), diabetic subjects with good control ($6.5\%<$glycohemoglobin$<8\%$), and diabetic subjects with poor control (glycohemoglobin$<8\%$). Their dataset consisted of 22 healthy subjects, 23 diabetic subjects with good control, and 17 subjects with poor control. The length of the acquired signal was not specified. In [13], Pilt et al. proposed a new parameter, the PPGAI (PPG augmented index), to assess arterial stiffness in elderly patients. This work highlighted a significant difference ($p<0.0001$) between healthy and diabetic subjects. The dataset used in this study consisted of 24 healthy subjects and 20 diabetic subjects who were asked to acquire finger PPG signals for one minute. These studies demonstrated how PPG signal shape can be used to classify DT2 with traditional approaches. However, to achieve good classification performance, these methods sometimes involve several steps, such as filtering, feature extraction, and outlier detection. To simplify the preprocessing step, researchers began to utilize machine learning methods with basic PPG features and demographic data. Monte-Moreno et al. [14] used a random forest and gradient boosted model to detect diabetes with physiologically relevant PPG features and demographic features, achieving an area under the curve (AUC) of 0.7. The dataset in this study consisted of 1,170 subjects who were asked to acquire PPG signals twice for one minute. 29% of the subjects were labeled as diabetic based on their clinical records. In [15], Hettiarachchi used a Decision Tree model fitted with PPG features and physiological data to detect DT2 in the presence of hypertension (HT) and pre-hypertension. They used a public dataset composed of short PPG signals (2.1 seconds) acquired from 219 subjects with various pathologies [16]. The best AUC score was obtained when detecting DT2 in the absence of hypertension (9 healthy subjects, 9 diabetic subjects). They achieved an AUC of 0.83. In [17], Nirala et al. proposed a support vector machine model (SVM) to classify diabetic and healthy subjects. They selected the ten best morphological PPG features from a total of 37 total features, showing an AUC of 0.97. Their dataset consisted of 141 subjects with a five-minute-long recording each, with the diabetic subjects represented by 69% of the total dataset. Although the results are promising, feature extraction and identification of fiducial points can be heavily affected by signal quality. Thus, quality assessment is needed before applying this handcrafted feature extraction process. Over the past few years, the advent of deep learning has significantly impacted the analysis of biomedical signals [18]. Convolutional neural networks (CNN) are often used in PPG analysis due to their capability of auto-extracting discriminant features from the 1D PPG signal or

its 2D representations [19]. In [20], Srinivasan and Foroozan proposed a CNN model (VGGNet architecture) that utilizes the scalogram obtained from 30-second finger PPG signals to detect DT2. The used dataset is a subset of the MIMIC-III dataset composed of 808 subjects (58% with type I and type II diabetes). A total of 592 subjects presented HT, of which 290 had DT2 as well (62% of diabetics also had HT). In addition to the baseline model, they tested different architectures with several different inputs such as HT (yes/no), age, biological sex, and heart rate. The model scored an AUC of 0.83 when trained with the PPG scalogram, age, biological sex, and HT classification. However, the main limitation of this approach is the difficulty of using it in portable devices such as smartphones and smartwatches. The idea of using raw PPG signals as input to CNN models to classify DT2 is not yet widely explored. Avram et al. [9] proposed a CNN model to classify the prevalence of diabetes. They used three different datasets: the first with 53,870 subjects (7% with diabetes), and the second and third to further validate the model, with 7,806 (8.7% with diabetes) and 181 (21% with diabetes) subjects, respectively. The proposed model takes a 21.3-second-long raw PPG signal as input. The PPG signal is defined as a sequence of PPG waves. The model achieved an averaged AUC of 0.69 over the three different datasets using a CNN output threshold of 0.427. Furthermore, they proposed a fusion approach where the CNN outputs and demographic data are taken as input by a logistic regression model to classify whether or not a PPG signal belongs to a diabetic subject. In this case, the AUC score increased up to 0.83. The proposed CNN model is a large network composed of 18 convolutional layers and requires a long PPG measure to detect DT2. Additionally, the model has been trained on a large amount of data, which is not always available in practice. The advantages and disadvantages of the cited studies are reported in Table 1.

III. MATERIAL AND METHODS

In this section, we describe the used datasets and the models we implemented to classify the presence of diabetes. Figure 4 shows an overview of the followed process.

A. PHOTOPLETHYSMOGRAPHY AND DIABETES

Photoplethysmography (PPG) is an optical measurement that assesses blood volume changes in the peripheral circulation. A photoplethysmograph is composed of a light emitting diode (LED) and a photo-detector (PD) [21]. The light emitted from the diode is partially transmitted through the tissues, the arteries and arterioles and then detected by the photo-detector, as shown in Figure 1. The PPG signal therefore contains information about the hemodynamic state of the subject [22].

As previously mentioned, diabetes is strongly correlated with cardiovascular diseases (CVDs). Although the directionality of the relationship between arterial stiffness and diabetes remains unknown, it has been confirmed that diabetes and increased arterial stiffness are linked [23]. It has also been confirmed that blood viscosity increases with blood glucose [24]. Heart rate variability (HRV) has been shown to be lower in all age groups of diabetic patients compared to healthy subjects [25]. Additionally, even in the early stages of diabetes mellitus, ECG alterations such as sinus tachycardia, changes in heart rate variability, and left ventricular hypertrophy, may be observed [26]. Arterial wall stiffness, blood viscosity, and changes in heart polarization/depolarization directly affect PPG wave shapes. As a result, diabetes can potentially be detected by analyzing the PPG pulse shape through various methods that estimate the aforementioned parameters [13], [27]. However, the differences between diabetic and non-diabetic PPG waves are sometimes not obvious. The PPG signal is affected by several physiological modifications [28], and identifying one among them can be very challenging. Figure 2 shows three different PPG waves: in black, a PPG wave from a non-diabetic subject; in red dotted and red dashed dotted lines, two waves from diabetic subjects. The absence of the dicrotic notch in the dashed-dotted red wave is clearly observable, whereas in the dashed red wave, the typical modifications due to diabetes are not clearly visible. All subjects belong to the same age range.

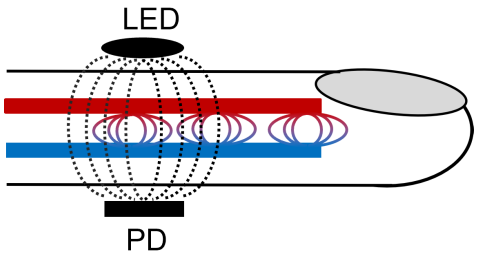

FIGURE 1. PPG transmitting technology with light emitting diode and photo-detector.

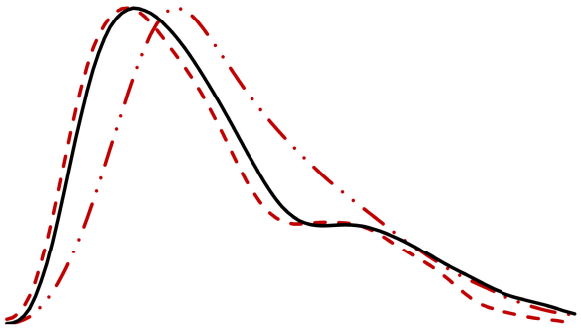

FIGURE 2. PPG waves from diabetic and non-diabetic subjects. The absence of the dicrotic notch in the dashed dotted red wave is observable, whereas the typical modifications due to diabetes are not clearly visible in the dashed red wave. The black line represents the PPG wave from a non-diabetic subject, the red dashed line represents the PPG signal from a diabetic subject, and the red dashed dotted line represents the PPG wave from a diabetic signal. All subjects are middle-aged (50-70 years old).

TABLE 1. Literature review summary.

Ref	Objective	Data type[a]	Approach	Feature	Main Outcome	Advantages	Disadvantages
Bagus et al. [12]	healthy VS diabetic VS unhealthy diabetic	PPG[-] [22/23/17] In-house	Multiscale poincaré analysis	PPG amplitude	SSR **	The proposed index is able to significantly differentiate between the three groups	Require heavy pre-processing, no test set
Pilt et al. [13]	healthy VS diabetic	PPG[-] [24/20] In-house	Statistical analysis	PPG Augmentation Index	PPGAI***	One index is able to significantly differentiate between groups	Require heavy pre-processing, no test set
Moreno et al. [14]	healthy VS diabetic	PPG[1min] [340/830] In-house	RF GBoost	PPG features + physio data	Accuracy: 70% Sensitivity: 80% Specificity: 48%	Large dataset, used PPG diabetes relevant features	Feature based, require quality assessment and identification of fiducial points
Hettiarachchi et al. [15]	healthy VS diabetic	PPG[2sec] [83/52] Public	LDA	PPG features + physio data	Accuracy: 83% Sensitivity: - Specificity: -	Assess the influence of hypertension on diabetes	Feature based, require quality assessment and identification of fiducial points, only accuracy is presented
Nirala et al. [17]	healthy VS diabetic	PPG[pulse] [44/97] In-house	SVM	PPG signal, first and second derivative parameters + eigenvalues	Accuracy: 97.87% Sensitivity: 98.78% Specificity: 96.6%	Feature selection based of feature importance	Feature based, require quality assessment, dataset division is not clearly specified
Avram et al. [9]	healthy VS diabetic	PPG[21sec] [50,099/3,770] In-house	1D - CNN + Logistic Regression	Raw signals + physio data	Accuracy: -% Sensitivity: 75% Specificity: 65.4% AUC: 0.77	Detection through raw PPG signals, validated over three different datasets, large dataset	Trained over a large dataset, large network
Srinivasan et al. [20]	healthy VS diabetic	PPG[30sec] [467/341] Public	2D - CNN	PPG scalogram + physio data	Accuracy: 76.34% Sensitivity: 76.6% Specificity: 76.1% AUC: 0.83	Analysis in the frequency domain, validated over a large dataset	Feature based, 2D approach

[a]: type of signal [signal length], [subjects for each class], type of database. SSR: ratio between long and short variations, PPGAI: PPG augmentation index, RF: random forest, LDA: latent dirichlet allocation, GBoost: gradient boost, SVM: support vector machine, CNN: convolutional neural networks, AUC: area under the curve.** $p < 0.01$, *** $p < 0.001$.

B. DATASETS

In this study, we utilized three datasets. A PPG wave is defined as the portion of the PPG signal that contains all information related to one heartbeat. The primary dataset consists of type 2 diabetic PPG waves (DB_DT2). Two additional datasets were employed to assess the impact of transfer learning. The first one (DB_shape) comprises 12,000 PPG waves that have been classified as high and low quality using our CNN model proposed in [29], then segmented into pulses, and manually labeled into four different classes based on their shape, as suggested in [30]. The second dataset (DB_HT) is composed of hypertension (HT) PPG waves. DB_HT and DB_DT2 are composed of two separate datasets, one available online [16], and one collected at the University Hospital of Nice (cohort register referenced BS-004) using the pOpmétre device [31]. Since HT and DT2 frequently coexist in the same subject, individuals who had both pathologies were excluded from the datasets to better isolate the effect of each pathology on the PPG signal. Table 2 summarizes the details of the datasets used. The pOpmétre device [31] acquires finger photoplethysmography (PPG) signals using red and infrared LEDs, and the signals are sampled at 1 kHz. The public dataset [16] consists of finger PPG signals acquired at 1 kHz with infrared light. Firstly, we applied a slight filter to the signals to remove high-frequency noise that could affect the PPG shape. We then identified the minimum points before and after the systolic maximum to segment the signals into single pulses. The pulses were normalized in time and amplitude. To perform inter-subject assessment and avoid overfitting, pulses belonging to the same subject were used either in training or testing.

TABLE 2. Databases description.

DB name	Subjects	Age	Pulses	Target class	Description
DB_shapes	300	50	12 000	-	4 wave shapes
DB_HT	181	55	917	61%	PPG waves, age, biological sex, PPG handcrafted features
DB_DT2	100	59	840	15%	PPG waves, age, biological sex, PPG handcrafted features

Target class %: percentage of pathological waves for the corresponding dataset. Age: mean age in years.

We used 70% of the available subjects to train and validate the model, and 30% of the subjects to test it. The same proportion of healthy and pathology subjects shown in Table 2 was

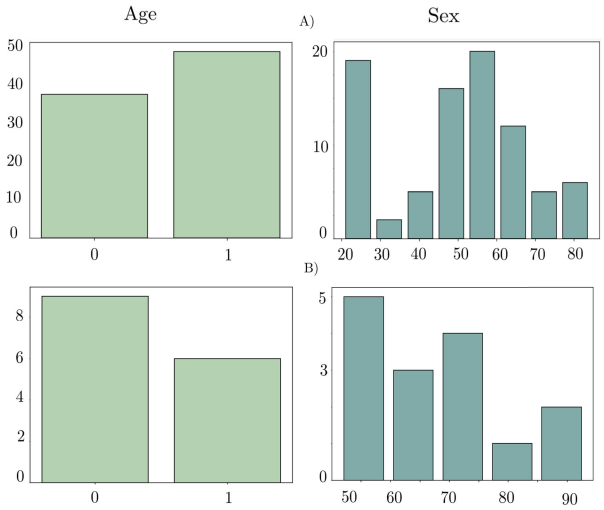

FIGURE 3. Age and biological sex histograms of A) non diabetic and B) diabetic subjects.

TABLE 3. PPG features extracted from the PPG waveforms.

Feature	Description
$\lambda 1$	first eigen value
$\lambda 2$	second eigen value
RT	rising time
RRD	rising time / descending time
Rarea	rising area / descending
Rslop	rising slope
Dslop	descending slope

maintained during the splitting. The presence of hypertension (HT) and type 2 diabetes (DT2) was clinically assessed, and no other pressure measurement or glucose analysis was performed.

C. FEATURES EXTRACTION

To assess the potential benefits of a handcrafted feature-based approach compared to a feature representation learning-based approach, we designed a convolutional neural network (CNN) that takes as input seven PPG features in addition to the PPG pulse wave. Table 3 presents a summary of the extracted features, which are injected after the convolutional layers, as shown in Tables 11, 12, and 13. We selected a significant subset of features presented in [17] that are commonly used to assess the PPG wave shape. We excluded derivative-based features because we did not apply any quality assessment to the HT and DT2 pulses, and the derivative is highly affected by noise in the signal, which could lead to errors. We also excluded features that required dicrotic notch identification since PPG waves often present multiple minima, and its detection is not straightforward. The selected features are extracted from a single PPG pulse.

D. CNN ARCHITECTURE

In this study, we have evaluated several different CNN architectures to detect DT2, with a total of nine models implemented for this purpose. We have also developed one model for classifying PPG waveform and one model detecting HT, both of which were considered for transfer learning. The baseline architecture consists of four convolutional layers, followed by n fully connected layers. The value of n changes during the fine-tuning process. Each CNN layer is followed by a max pooling layer, which produces a feature map containing the most prominent features of the previous feature map. A flattening layer is added after the last convolution to reduce the dimensions and pass the feature maps to the fully connected layer. We used ReLU activation functions inside the CNN layers and a sigmoid activation function for the classification layer. The network parameters were optimized using the Adam method. In the output layer, the loss function was categorical cross-entropy with class weights to mitigate the uneven class distributions' issue; Autonomio Talos [32] was used for learning rate optimization.

E. TRANSFER LEARNING

We implemented several models to classify whether a PPG wave belongs to a diabetic subject or not. Firstly, we trained a model on the DB_shape dataset to perform classification according to the four PPG shapes proposed by [30]. We then used the weights and biases of this shape model to pre-train the HT model. Additionally, we employed the same shape model weights and biases to pre-train three different models to detect diabetes. The HT weights and biases were also used to pre-train three DT2 models that differ in their required inputs. The first model is trained to classify the presence of DT2 based only on PPG pulse waves, the second requires PPG pulse waves along with age and biological sex, while the third model utilizes PPG pulse waves and PPG-based features. Each of the proposed models was tested with three different architectures. During the transfer learning process, we trained the models and unfroze the layers starting from the last layer to the input layer. In each new simulation, a new layer was unfrozen, and the net was trained on the selected dataset to fine-tune the model. To compare the benefits of transfer learning, we implemented the same architectures for DT2 detection without using any pre-trained models. A summary of all the implemented architectures is presented in Table 11, Table 12, and Table 13.

IV. RESULTS AND DISCUSSION

In this section, we present and discuss the results obtained for detecting diabetes with the proposed models. To evaluate the overall performance of the models, we computed several indices which are described below. The equations for these indices are displayed in Table 4. The sensitivity (also called recall) measures the ability to correctly identify patients with the disease, while the specificity (also called true

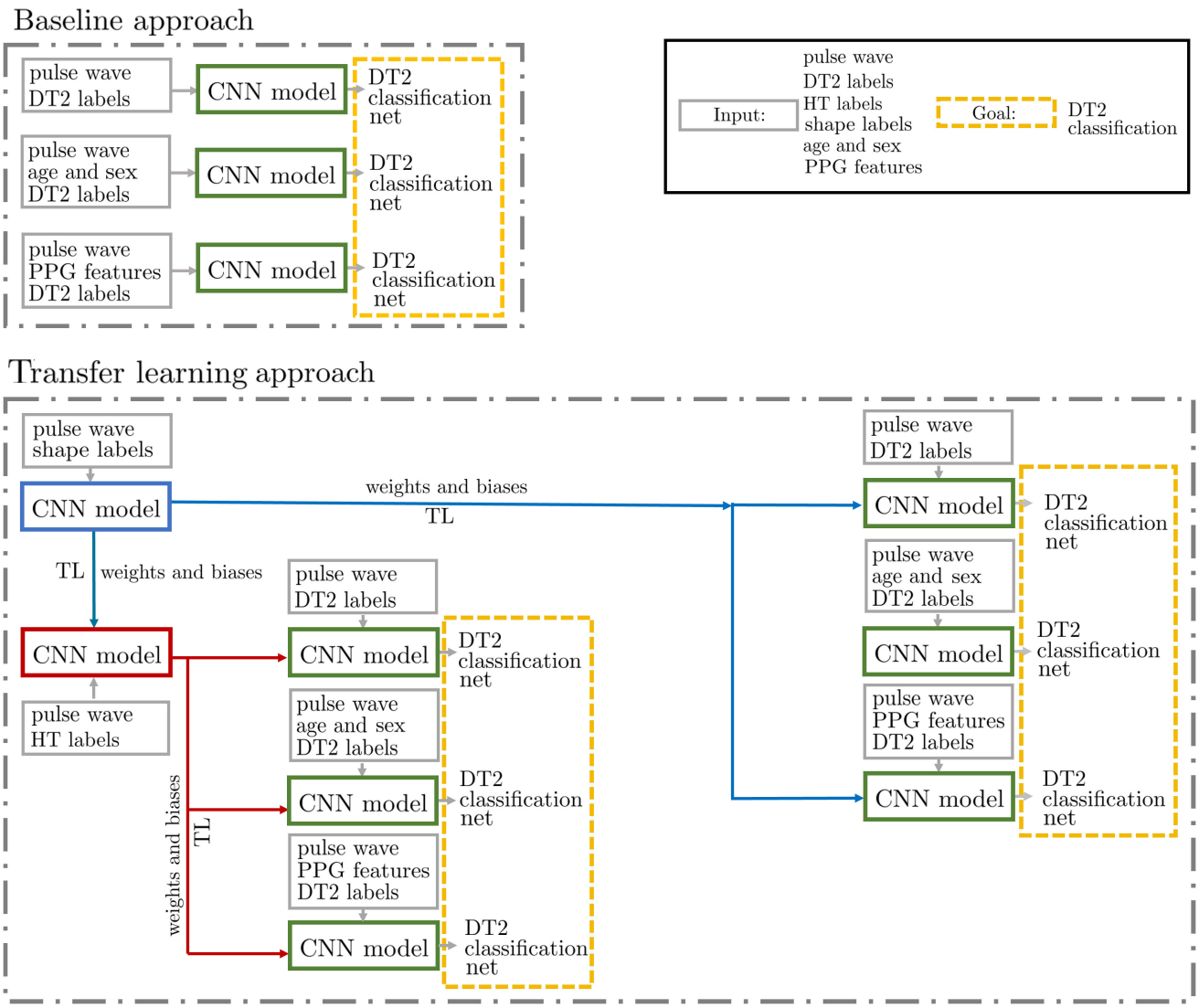

FIGURE 4. Overview of the proposed flow to assess the transfer learning efficacy in detecting diabetes. TL: transfer learning. HT: hypertension. DT2: type 2 diabetes. CNN: convolutiona neural network.

negative rate) measures the ability to correctly identify people without the disease. The negative predictive value (NPV) is the ratio of subjects who were truly diagnosed as negative to all those who had negative test results (including patients who were incorrectly diagnosed as healthy). It predicts how likely it is for someone to truly be healthy in case of a negative test result.

The positive predictive value (PPV, also called precision) is the ratio of subjects truly diagnosed as positive to all those who had positive test results (including healthy subjects who were incorrectly diagnosed as pathological). This score gives a measure of how likely it is for someone to truly be pathological, in case of a positive test result. The F1 score is the harmonic mean of the PPV and sensitivity. The AUC represents the degree or measure of separability, indicating how well the model can distinguish between the classes. Since the model classifies a single pulse, we computed the performance scores in three different ways: *recording level*, where the scores are computed independently for each pulse; *subject's average*, where the CNN outputs corresponding to pulses belonging to the same subject are averaged; and

TABLE 4. Performance parameters equation.

Parameter	Equation
Sensitivity	TP / (TP+FN)
Specificity	TN / (TN+FP)
NPV	TN / (TN+FN)
PPV	TP / (TP+FP)
F1 score	2 (PPV*sensitivity) / (PPV + sensitivity)
AUC	$\int_0^1$TPR d(FPR)

TP: true positive; FN: false negative; TN: true negative; FN: false negative; PPV: positive predictive value; NPV: negative predictive value; TPR: true positive rate; FPR: false positive rate.

TABLE 5. Results without transfer learning.

DT2						
Test scores						
Score type	**PPV**	**NPV**	**Sensitivity**	**Specificity**	**AUC**	**f1**
Recording level	14.12	88.08	34.29	69.96	52.12	20.00
Subject's averaged	28.57	90.91	50.00	80.00	65.00	36.36
Majority voting	25.00	94.12	75.00	64.00	69.50	37.50
DT2[Age\Biological sex]						
Test scores						
Score type	**PPV**	**NPV**	**Sensitivity**	**Specificity**	**AUC**	**f1**
Recording level	15.00	88.07	25.71	79.01	52.36	18.95
Subject's averaged	25.00	88.00	25.00	88.00	56.50	25.00
Majority voting	33.33	95.00	75.00	76.00	75.50	46.15
DT2[Feature]						
Test scores						
Score type	**PPV**	**NPV**	**Sensitivity**	**Specificity**	**AUC**	**f1**
Recording level	16.79	91.49	65.71	53.09	59.40	26.74
Subject's averaged	25.00	94.12	75.00	64.00	69.50	37.50
Majority voting	18.18	100.00	100.00	28.00	64.00	30.77

PPV: positive predicted value, NPV: negative predicted value, AUC: area under the curve, f1: f1 score.

TABLE 6. Results with transfer learning from shape model.

DT2						
Test scores						
Score type	**PPV**	**NPV**	**Sensitivity**	**Specificity**	**AUC**	**f1**
Recording level	5.75	84.29	14.29	66.26	40.27	8.20
Subject's averaged	20.00	87.50	25.00	84.00	54.50	22.22
Majority voting	14.29	86.36	25.00	76.00	50.50	18.18
DT2[Age\Biological sex]						
Test scores						
Score type	**PPV**	**NPV**	**Sensitivity**	**Specificity**	**AUC**	**f1**
Recording level	14.46	88.21	34.29	70.78	52.53	20.34
Subject's averaged	20.00	87.50	25.00	84.00	54.50	22.22
Majority voting	22.22	90.00	50.00	72.00	61.00	30.77
DT2[Feature]						
Test scores						
Score type	**PPV**	**NPV**	**Sensitivity**	**Specificity**	**AUC**	**f1**
Recording level	17.69	91.89	65.71	55.97	60.84	27.88
Subject's averaged	16.67	88.24	50.00	60.00	55.00	25.00
Majority voting	22.22	100.00	100.00	44.00	72.00	36.36

PPV: positive predicted value, NPV: negative predicted value, AUC: area under the curve, f1: f1 score.

majority voting, where the class attributed to a subject is given by the majority class obtained by the totality of his pulses. Tables 5, 6, and 7 present the performance results obtained without transfer learning, with transfer learning from the shape CNN model, and with transfer learning from the HT model, respectively. The three best models (PPG wave, PPG wave and age/biological sex, PPG wave and PPG features) were selected based on their performances on the validation set. Figure 5 shows the performances of the three selected models. When applying transfer learning from the HT model, the best performances with all three models were reached by training only the fully connected layers without de-freezing the convolutional block.

When applying transfer learning from the HT model, the best performances with all three models were reached by training only the fully connected layers without de-freezing the convolutional block. In contrast, when applying transfer learning from the shape model, DT2[age/biological sex] and DT2[features] scored the best results when the convolutional block was also fine tuned, while the DT2 model performed better with all layers pre-trained and not de-freezed.

For DT2 detection with PPG waves, the best performances were achieved by the CNN model trained over the DB_DT2 without transfer learning using majority voting. It achieved a specificity of 64%, sensitivity of 75%, and an AUC of 69.5. Transfer learning from shape resulted in the worst performance, while transfer learning from HT scored lower results than the baseline. The best model had 41,316 parameters and, considering double precision, occupies around 2MB of space.

A. DT2 DETECTION WITH PPG WAVES, AGE, AND BIOLOGICAL SEX

The best performances were obtained by the CNN model trained over the DB_DT2 with age and biological sex without transfer learning, with majority voting achieving a specificity of 76% and sensitivity of 75%, and an AUC of 75.5.

Transfer learning from the shape and HT models scored similar and lower results compared to the baseline model. The number of parameters of the best model is 41,184. Taking into consideration double precision, the model occupies around 2MB of space.

B. DT2 DETECTION WITH PPG WAVES AND PPG FEATURES

The best performance was obtained by the CNN model trained over the DB_DT2 with PPG features using transfer learning from the shape model with majority voting, achieving a specificity of 44% and a sensitivity of 100% with an AUC of 72. Transfer learning from HT scored the worst results, while the model trained without transfer learning scored similar results compared to the best one. The number of parameters of the best model is 41,184, and taking into consideration double precision, the model occupies around 2MB of space.

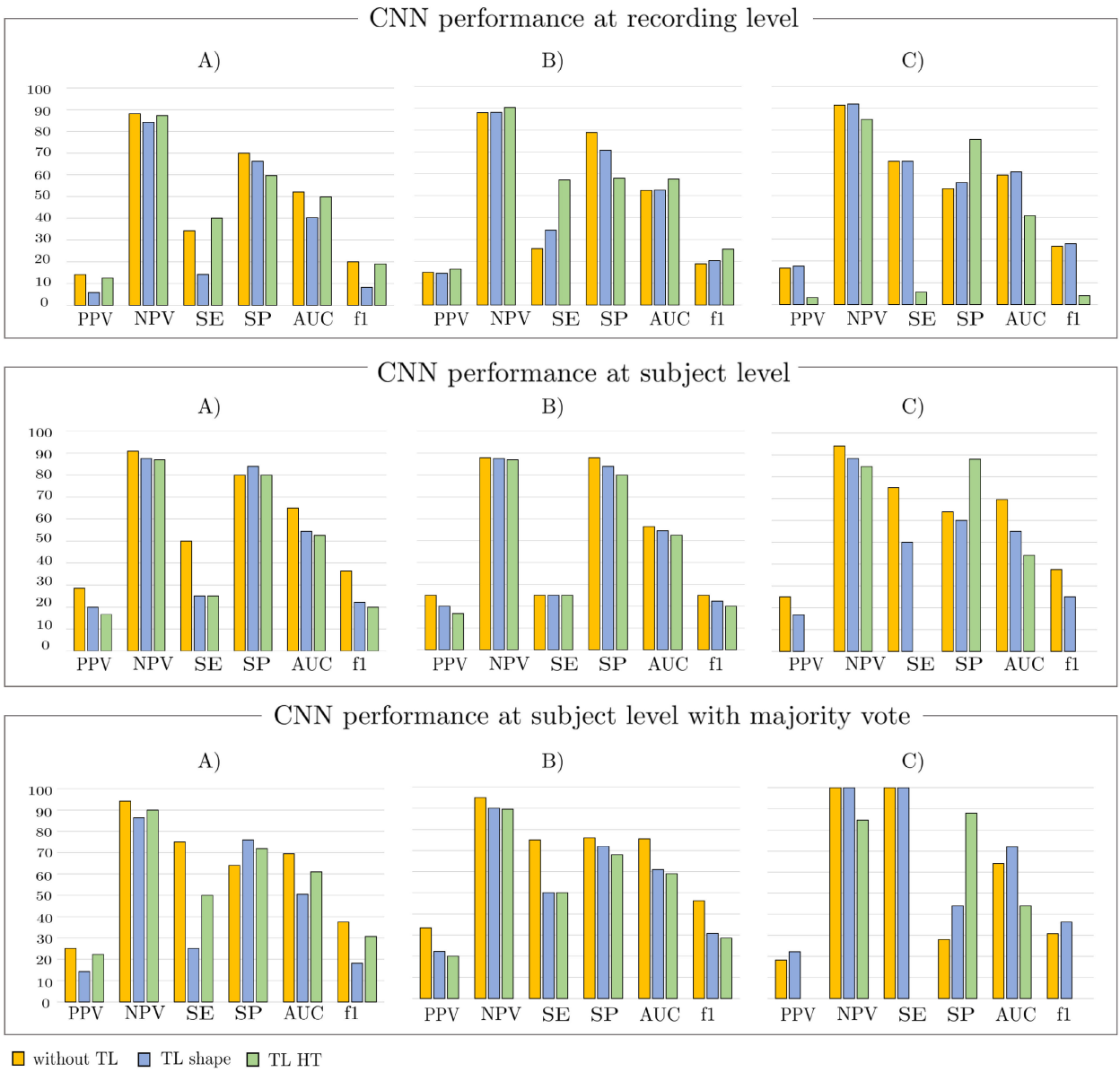

FIGURE 5. Visual representation of the models performances. On the columns, from the left to the right: A) performances of the models trained only with PPG waves, B) performances of the models trained with PPG waves, age and biological sex, C) performances of the models trained with PPG waves and handcrafted features respectively. From the top to the bottom: recording level, subject's averaged and majority voting. Yellow for the models without transfer learning, blue for the models with transfer learning from the shape model and green for transfer learning from the hypertension model (HT). PPV: positive predictive values; NPV: negative predictive values; SE: sensitivity; SP: specificity; AUC: area under the curve; f1: f1 score.

C. DISCUSSION

Our results suggest that PPG signals have the potential to detect diabetes. It is possible that with a larger dataset, we could obtain even better results. Thus, the proposed method could become competitive with traditional approaches, such as the HbA1c test, which requires well-equipped medical centers. A summary of the advantages and disadvantages of the database is shown in Table 8.

As presented in the previous section, the DT2 and DT2[age/biological sex] models obtained the best performances when trained without transfer learning (Table 10). The only model that benefited from transfer learning was the DT2[features] one.

Despite the fact that, when applying transfer learning from HT we reached the best performances defreezing only the final layers, highlighting the similarity between HT and DT2 PPG waves, the fact that it did not improve the results could be consistent with the results proposed by [15], who reported a lower AUC score in classifying DT2 in the presence of HT. It is not surprising that the best performances were obtained

TABLE 7. Results with transfer learning from HT model.

DT2						
Test scores						
Score type	**PPV**	**NPV**	**Sensitivity**	**Specificity**	**AUC**	**f1**
Recording level	12.50	87.35	40.00	59.67	49.84	19.05
Subject's averaged	16.67	86.96	25.00	80.00	52.50	20.00
Majority voting	22.22	90.00	50.00	72.00	61.00	30.77
DT2[Age\Biological sex]						
Test scores						
Score type	**PPV**	**NPV**	**Sensitivity**	**Specificity**	**AUC**	**f1**
Recording level	16.39	90.38	57.14	58.02	57.58	25.48
Subject's averaged	16.67	86.96	25.00	80.00	52.50	20.00
Majority voting	20.00	89.47	50.00	68.00	59.00	28.57
DT2[Feature]						
Test scores						
Score type	**PPV**	**NPV**	**Sensitivity**	**Specificity**	**AUC**	**f1**
Recording level	3.28	84.79	5.71	75.72	40.72	4.17
Subject's averaged	0.00	84.62	0.00	88.00	44.00	0.00
Majority voting	0.00	84.62	0.00	88.00	44.00	0.00

PPV: positive predicted value, NPV: negative predicted value, AUC: area under the curve, f1: f1 score.

TABLE 8. Database advantages and disadvantages.

Advantages	Disadvantages
Composed by two different devices	Unbalanced, small
Competitive results despite smaller size	Lack of information

TABLE 9. Best model architecture.

DT2[Age/Biological sex] without transfer learning	
option B	
Layers	Output shape
Convolution layer	(None,99,32)
MaxPooling	
Convolution layer	(None,50,32)
MaxPooling	
Convolution layer	(None,25,32)
MaxPooling	
Convolution layer	(None,13,32)
MaxPooling	
Flatten	(None,224)
Dense(30)	(None,30)
Concatenate	(None,32)
Dense(4)	(None,4)
Dense(2)	(None,2)

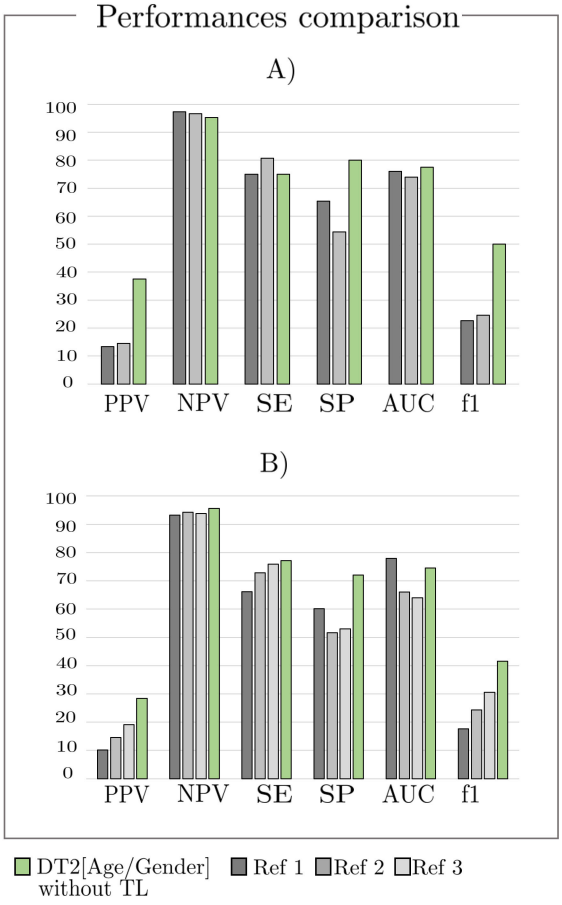

FIGURE 6. Performance of the proposed model with respect to the state of the art with modified CNN output threshold (0.427). From the top to the bottom: A) recording level and B) subject's averaged and majority voting. The reference [9] is represented with three different shades of grey while our model is represented in green. PPV: positive predictive values; NPV: negative predictive values; SE: sensitivity; SP: specificity; AUC: area under the curve; f1: f1 score.

TABLE 10. Best model performances and overview.

DT2[Age/Biological sex] without transfer learning	
Objective	healthy VS diabetic
Data type[a]	PPG[1 wave] - [85/15] - In-house and public
Approach	1D-CNN
Feature	Raw signals + physio data
Main Outcome	Specificity: 76% - Sensitivity: 75% AUC: 75.5%
Advantages	small dataset, two different devices, competitive results, detection through raw PPG signals
Disadvantages	Unbalanced and small dataset, lack of information

[a]: type of signal [signal length], [subjects for each class], type of database.
CNN: convolutional neural networks, AUC: area under the curve.

with the majority vote, since we did not apply any quality assessment to the PPG waves inside DB_DT2. In fact, the majority vote system allows the model to be more robust to noisy PPG waves, since it does not take into account the CNN score but only the final given class. Additionally, from a clinical point of view, the subject's averaged output and the majority voting are preferable to a single record analysis since the diagnosis cannot change in a short period of time. While age and biological sex enhanced the PPG classification, the PPG features as additional input did not improve the performances. This could be caused by the fact that the CNN model is already capable of extracting significant features

TABLE 11. Tested CNN architectures without transfer learning.

DT2 without transfer learning					
option r			**option s**		
Layers	Output shape	Pre-trained	Layers	Output shape	Pre-trained
Convolution layer	(None,99,32)		Convolution layer	(None,99,32)	
MaxPooling			MaxPooling		
Convolution layer	(None,50,32)		Convolution layer	(None,50,32)	
MaxPooling			MaxPooling		
Convolution layer	(None,25,32)		Convolution layer	(None,25,32)	
MaxPooling			MaxPooling		
Convolution layer	(None,13,32)		Convolution layer	(None,13,32)	
MaxPooling			MaxPooling		
Flatten	(None,224)		Flatten	(None,224)	
Dense(30)	Dense(30)		Dense(30)	Dense(30)	
Dense(4)	(None,4)		Dense(8)	(None,8)	
Dense(4)	(None,2)		Dense(4)	(None,4)	
			Dense(4)	(None,2)	

DT2[Age/Biological sex] without transfer learning								
option A			**option B**			**option C**		
Layers	Output shape	Pre-trained	Layers	Output shape	Pre-trained	Layers	Output shape	Pre-trained
Convolution layer	(None,99,32)		Convolution layer	(None,99,32)		Convolution layer	(None,99,32)	
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,50,32)		Convolution layer	(None,50,32)		Convolution layer	(None,50,32)	
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,25,32)		Convolution layer	(None,25,32)		Convolution layer	(None,25,32)	
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,13,32)		Convolution layer	(None,13,32)		Convolution layer	(None,13,32)	
MaxPooling			MaxPooling			MaxPooling		
Flatten	(None,224)		Flatten	(None,224)		Flatten	(None,224)	
Concatenate	(None,226)		Dense(30)	Dense(30)		Dense(30)	Dense(30)	
Dense(30)	Dense(30)		Concatenate	(None,32)		Dense(4)	(None,4)	
Dense(4)	(None,4)		Dense(4)	(None,4)		Concatenate	(None,6)	
Dense(4)	(None,2)		Dense(4)	(None,2)		Dense(4)	(None,2)	

DT2[Features] without transfer learning								
option A			**option B**			**option C**		
Layers	Output shape	Pre-trained	Layers	Output shape	Pre-trained	Layers	Output shape	Pre-trained
Convolution layer	(None,99,32)		Convolution layer	(None,99,32)		Convolution layer	(None,99,32)	
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,50,32)		Convolution layer	(None,50,32)		Convolution layer	(None,50,32)	
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,25,32)		Convolution layer	(None,25,32)		Convolution layer	(None,25,32)	
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,13,32)		Convolution layer	(None,13,32)		Convolution layer	(None,13,32)	
MaxPooling			MaxPooling			MaxPooling		
Flatten	(None,224)		Flatten	(None,224)		Flatten	(None,224)	
Concatenate	(None,231)		Dense(30)	Dense(30)		Dense(30)	Dense(30)	
Dense(30)	Dense(30)		Concatenate	(None,37)		Dense(4)	(None,4)	
Dense(4)	(None,4)		Dense(4)	(None,4)		Concatenate	(None,11)	
Dense(4)	(None,2)		Dense(4)	(None,2)		Dense(4)	(None,2)	

Tested CNN architectures to detect diabetes from PPG signal (top), PPG signal, age and biological sex (middle) and PPG signal and handcrafted PPG features (bottom). The Pre-trained column shows the layers that have been pre-trained from the shape CNN based model and trained again against DT2 detection defreezing the layers from the last layer to the input layer.

from the PPG wave and using additional PPG features only replicates the information without adding additional relevant information to the input. We will explore new features and their relation with diabetes. In addition, we will study other architecture combination schemes between the convolutional features and the handcrafted features. Overall, the best performance was obtained by the CNN model trained with PPG waves, age, and biological sex without transfer learning, reaching an AUC of 75.5. When tested negative, a subject has a probability of 95% of truly being negative (NPV). When tested positive, the probability decreases to 33% (PPV). The model architecture is presented in Table 9. The model has much fewer parameters than the deep learning architectures considered in [9]. Our results are consistent with [33], which showed that low-complexity models perform comparably well or better than high-complexity models when dealing

TABLE 12. Tested CNN architectures with transfer learning from shape model.

DT2 from shape model

option x			option y			option z		
Layers	Output shape	Pre-trained	Layers	Output shape	Pre-trained	Layers	Output shape	Pre-trained
Convolution layer	(None,99,32)	x	Convolution layer	(None,99,32)	x	Convolution layer	(None,99,32)	x
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,50,32)	x	Convolution layer	(None,50,32)	x	Convolution layer	(None,50,32)	x
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,25,32)	x	Convolution layer	(None,25,32)	x	Convolution layer	(None,25,32)	x
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,13,32)	x	Convolution layer	(None,13,32)	x	Convolution layer	(None,13,32)	x
MaxPooling			MaxPooling			MaxPooling		
Flatten	(None,224)	x	Flatten	(None,224)	x	Flatten	(None,224)	x
Dense(30)	(None,30)	x	Dense(30)	(None,30)	x	Dense(30)	(None,30)	x
Dense(4)	(None,4)	x	Dense(4)	(None,4)		Dense(8)	(None,8)	
Dense(2)	(None,2)		Dense(2)	(None,2)		Dense(4)	(None,4)	
						Dense(2)	(None,2)	

DT2[Age/Biological sex] from shape model

option A			option B			option C		
Layers	Output shape	Pre-trained	Layers	Output shape	Pre-trained	Layers	Output shape	Pre-trained
Convolution layer	(None,99,32)	x	Convolution layer	(None,99,32)	x	Convolution layer	(None,99,32)	x
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,50,32)	x	Convolution layer	(None,50,32)	x	Convolution layer	(None,50,32)	x
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,25,32)	x	Convolution layer	(None,25,32)	x	Convolution layer	(None,25,32)	x
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,13,32)	x	Convolution layer	(None,13,32)	x	Convolution layer	(None,13,32)	x
MaxPooling			MaxPooling			MaxPooling		
Flatten	(None,224)		Flatten	(None,224)		Flatten	(None,224)	
Concatenate	(None,226)		Dense(30)	(None,30)	x	Dense(30)	(None,30)	x
Dense(30)	(None,30)		Concatenate	(None,32)		Dense(4)	(None,4)	x
Dense(4)	(None,4)		Dense(4)	(None,4)		Concatenate	(None,6)	
Dense(2)	(None,2)		Dense(2)	(None,2)		Dense(2)	(None,2)	

DT2[Features] from shape model

option A			option B			option C		
Layers	Output shape	Pre-trained	Layers	Output shape	Pre-trained	Layers	Output shape	Pre-trained
Convolution layer	(None,99,32)	x	Convolution layer	(None,99,32)	x	Convolution layer	(None,99,32)	x
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,50,32)	x	Convolution layer	(None,50,32)	x	Convolution layer	(None,50,32)	x
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,25,32)	x	Convolution layer	(None,25,32)	x	Convolution layer	(None,25,32)	x
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,13,32)	x	Convolution layer	(None,13,32)	x	Convolution layer	(None,13,32)	x
MaxPooling			MaxPooling			MaxPooling		
Flatten	(None,224)		Flatten	(None,224)		Flatten	(None,224)	
Concatenate	(None,231)		Dense(30)	(None,30)	x	Dense(30)	(None,30)	x
Dense(30)	(None,30)		Concatenate	(None,37)		Dense(4)	(None,4)	x
Dense(4)	(None,4)		Dense(4)	(None,4)		Concatenate	(None,11)	
Dense(2)	(None,2)		Dense(2)	(None,2)		Dense(2)	(None,2)	

Tested CNN architectures to detect diabetes from PPG signal (top), PPG signal, ange and biological sex (middle) and PPG signal and handcrafted PPG features (bottom). The Pre-trained column shows the layers that have been pre-trained from the shape CNN based model and trained again against DT2 detection defreezing the layers from the last layer to the input layer.

with small datasets. With small model architectures, a small dataset is likely to be sufficient for reliable training of the parameters without the need for transfer learning. Although there is certainly room for improvement, the presented results are comparable to the state of the art [9]. To better compare our results with [9], we additionally computed the performance analysis using the same CNN output threshold (0.427). The results are presented in Figure 6. In [9], the model has been tested on different datasets. The first reference (Ref1) is when the model is tested over the test dataset (11,313 subjects), the second (Ref 2) when tested on a contemporary cohort dataset (7,806 subjects), and the last one (Ref 3) is when the model is tested over a clinic cohort (181 subjects). Our test set is composed of 31 subjects. The reference results are represented in gray in the figure, while the performance of our model is represented in green. The performance

TABLE 13. Tested CNN architectures with transfer learning from hypertension model.

DT2 from HT model

option x			option y			option z		
Layers	Output shape	Pre-trained	Layers	Output shape	Pre-trained	Layers	Output shape	Pre-trained
Convolution layer	(None,99,32)	x	Convolution layer	(None,99,32)	x	Convolution layer	(None,99,32)	x
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,50,32)	x	Convolution layer	(None,50,32)	x	Convolution layer	(None,50,32)	x
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,25,32)	x	Convolution layer	(None,25,32)	x	Convolution layer	(None,25,32)	x
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,13,32)	x	Convolution layer	(None,13,32)	x	Convolution layer	(None,13,32)	x
MaxPooling			MaxPooling			MaxPooling		
Flatten	(None,224)	x	Flatten	(None,224)	x	Flatten	(None,224)	x
Dense(30)	(None,30)	x	Dense(30)	(None,30)	x	Dense(30)	(None,30)	x
Dense(4)	(None,4)	x	Dense(4)	(None,4)		Dense(8)	(None,8)	
Dense(2)	(None,2)		Dense(2)	(None,2)		Dense(4)	(None,4)	
						Dense(2)	(None,2)	

DT2[Age/Biological sex] from HT model

option A			option B			option C		
Layers	Output shape	Pre-trained	Layers	Output shape	Pre-trained	Layers	Output shape	Pre-trained
Convolution layer	(None,99,32)	x	Convolution layer	(None,99,32)	x	Convolution layer	(None,99,32)	x
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,50,32)	x	Convolution layer	(None,50,32)	x	Convolution layer	(None,50,32)	x
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,25,32)	x	Convolution layer	(None,25,32)	x	Convolution layer	(None,25,32)	x
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,13,32)	x	Convolution layer	(None,13,32)	x	Convolution layer	(None,13,32)	x
MaxPooling			MaxPooling			MaxPooling		
Flatten	(None,224)		Flatten	(None,224)		Flatten	(None,224)	
Concatenate	(None,226)		Dense(30)	(None,30)	x	Dense(30)	(None,30)	x
Dense(30)	(None,30)		Concatenate	(None,32)		Dense(4)	(None,4)	x
Dense(4)	(None,4)		Dense(4)	(None,4)		Concatenate	(None,6)	
Dense(2)	(None,2)		Dense(2)	(None,2)		Dense(2)	(None,2)	

DT2[Features] from HT model

option A			option B			option C		
Layers	Output shape	Pre-trained	Layers	Output shape	Pre-trained	Layers	Output shape	Pre-trained
Convolution layer	(None,99,32)	x	Convolution layer	(None,99,32)	x	Convolution layer	(None,99,32)	x
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,50,32)	x	Convolution layer	(None,50,32)	x	Convolution layer	(None,50,32)	x
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,25,32)	x	Convolution layer	(None,25,32)	x	Convolution layer	(None,25,32)	x
MaxPooling			MaxPooling			MaxPooling		
Convolution layer	(None,13,32)	x	Convolution layer	(None,13,32)	x	Convolution layer	(None,13,32)	x
MaxPooling			MaxPooling			MaxPooling		
Flatten	(None,224)		Flatten	(None,224)		Flatten	(None,224)	
Concatenate	(None,231)		Dense(30)	(None,30)	x	Dense(30)	(None,30)	x
Dense(30)	(None,30)		Concatenate	(None,37)		Dense(4)	(None,4)	x
Dense(4)	(None,4)		Dense(4)	(None,4)		Concatenate	(None,11)	
Dense(2)	(None,2)		Dense(2)	(None,2)		Dense(2)	(None,2)	

Tested CNN architectures to detect diabetes from PPG signal (top), PPG signal, ange and biological sex (middle) and PPG signal and handcrafted PPG features (bottom). The Pre-trained column shows the layers that have been pre-trained from the shape CNN based model and trained again against DT2 detection defreezing the layers from the last layer to the input layer.

parameters have been computed also in this case with the recording level, subject's average, and majority voting approach. Our model scored a comparable NPV and sensitivity, while our PPV and specificity performances overcome the state of the art. We obtain a comparable AUC. To obtain a more reliable model, we should train our model over a more stratified dataset containing more specific information about the subjects' disease.

1) LIMITATIONS

Our study has some limitations. First, the diagnosis of diabetes was based solely on the subject's medical record, and no additional glucose measurements were taken. Second, information about the drug treatments followed by the patients was not included in the available databases. Third, the lack of information about medical history, such as cardiovascular accidents, biological sex transition, and drug treatment, could

lead to biased results due to mislabeled data, as discussed in previous studies [34], [35]. To compensate for the unbalanced classes, we used penalization weights in our analysis. However, a larger and more balanced database would be needed to obtain a clinical validation of the model.

V. CONCLUSION

In this study, we have presented a lightweight CNN model capable of obtaining competitive results compared to the state of the art in classifying diabetic PPG waves. The best performance was achieved when age and biological sex were used as input in addition to PPG waves, without transfer learning. The model achieved an AUC of 75.5 when calculated with majority voting. Embedding the proposed model into a portable device for large-scale screening could improve diabetes prevention and early treatment. Additionally, the proposed method only requires a single PPG pulse and no quality assessment. The chosen segment length makes the model capable of operating with a variety of signal lengths without requiring a long PPG acquisition. Age and biological sex are simple parameters that do not require any clinical knowledge and allow the model to be used by end-users without any assistance.

Future improvements will focus on obtaining a more balanced and complete database with glucose measurements and drug treatments and validating the model with a permutation test. We will also search for better PPG handcrafted features to enhance DT2 detection. Since we normalized the length of the PPG waves, one possible improvement is to inject this information, which is related to the heart rate, as an additional feature to the neural net. Further studies are needed to explore the possibility of using a fused approach between PPG waves, demographic data, and PPG features. Another interesting future improvement is to assess the vascular status of diabetic subjects through PPG signals.

ACKNOWLEDGMENT

Serena Zanelli collaborates with Axelife, a company that designs and develops PPG-based medical devices. Magid Hallab is the CEO of Axelife and has authored patents used by Axelife.

REFERENCES

[1] M. A. Atkinson, G. S. Eisenbarth, and A. W. Michels, "Type 1 diabetes," *Lancet*, vol. 383, no. 9911, pp. 69–82, 2014.

[2] S. Chatterjee, K. Khunti, and M. J. Davies, "Type 2 diabetes," *Lancet*, vol. 389, no. 10085, pp. 2239–2251, 2017.

[3] T. A. Buchanan and A. H. Xiang, "Gestational diabetes mellitus," *J. Clin. Invest.*, vol. 115, no. 3, pp. 485–491, 2005.

[4] A. Esteghamati, B. Larijani, M. H. Aghajani, F. Ghaemi, J. Kermanchi, A. Shahrami, M. Saadat, E. N. Esfahani, M. Ganji, S. Noshad, E. Khajeh, A. Ghajar, B. Heidari, M. Afarideh, J. I. Mechanick, and F. Ismail-Beigi, "Diabetes in Iran: Prospective analysis from first nationwide diabetes report of national program for prevention and control of diabetes (NPPCD-2016)," *Sci. Rep.*, vol. 7, no. 1, pp. 1–10, Oct. 2017.

[5] G. Fagherazzi and P. Ravaud, "Digital diabetes: Perspectives for diabetes prevention, management and research," *Diabetes Metabolism*, vol. 45, no. 4, pp. 322–329, Sep. 2019.

[6] A. N. Long and S. Dagogo-Jack, "Comorbidities of diabetes and hypertension: Mechanisms and approach to target organ protection," *J. Clin. Hypertension*, vol. 13, no. 4, pp. 244–251, Apr. 2011.

[7] S. Zanelli, M. Ammi, M. Hallab, and M. A. El Yacoubi, "Diabetes detection and management through photoplethysmographic and electrocardiographic signals analysis: A systematic review," *Sensors*, vol. 22, no. 13, p. 4890, Jun. 2022.

[8] N. R. Anderson, E. S. Lee, J. S. Brockenbrough, M. E. Minie, S. Fuller, J. Brinkley, and P. Tarczy-Hornoch, "Issues in biomedical research data management and analysis: Needs and barriers," *J. Amer. Med. Informat. Assoc.*, vol. 14, no. 4, pp. 478–488, Jul./Aug. 2007.

[9] R. Avram, G. Tison, P. Kuhar, G. Marcus, M. Pletcher, J. E. Olgin, and K. Aschbacher, "Predicting diabetes from photoplethysmography using deep learning," *J. Amer. College Cardiol.*, vol. 73, no. 9, p. 16, Mar. 2019.

[10] A. Prabha, J. Yadav, A. Rani, and V. Singh, "Design of intelligent diabetes mellitus detection system using hybrid feature selection based XGBoost classifier," *Comput. Biol. Med.*, vol. 136, Sep. 2021, Art. no. 104664.

[11] V. R. Reddy, A. D. Choudhury, S. Jayaraman, N. K. Thokala, P. Deshpande, and V. Kaliaperumal, "PerDMCS: Weighted fusion of PPG signal features for robust and efficient diabetes mellitus classification," in *Proc. 10th Int. Joint Conf. Biomed. Eng. Syst. Technol.*, 2017, pp. 553–560.

[12] H. Bagus, L. Gen-Min, Y. Chieh-Ming, C. Shiao-Chiang, and W. Hsien-Tsai, "Poincaré plot of fingertip photoplethysmogram pulse amplitude suitable to assess diabetes status," in *Proc. 4th Int. Conf. Electr. Eng., Comput. Sci. Informat. (EECSI)*, Sep. 2017.

[13] K. Pilt, K. Meigas, R. Ferenets, K. Temitski, and M. Viigimaa, "Photoplethysmographic signal waveform index for detection of increased arterial stiffness," *Physiolog. Meas.*, vol. 35, no. 10, pp. 2027–2036, Oct. 2014.

[14] E. M. Moreno, M. J. A. Lujan, M. T. Rusinol, P. J. Fernandez, P. N. Manrique, C. A. Trivino, M. P. Miquel, M. A. Rodriguez, and M. J. GonzalezBurguillos, "Type 2 diabetes screening test by means of a pulse oximeter," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 2, pp. 341–351, Feb. 2017.

[15] C. C. Hettiarachchi ansd C. Chitraranjan, "A machine learning approach to predict diabetes using short recorded photoplethysmography & physiological characteristics," in *Proc. 17th Conf. Artif. Intell. Med. (AIME)*. Poznan, Poland: Springer, Jun. 2019, pp. 322–327.

[16] Y. Liang, Z. Chen, G. Liu, and M. Elgendi, "A new, short-recorded photoplethysmogram dataset for blood pressure monitoring in China," *Sci. Data*, vol. 5, no. 1, pp. 1–7, Feb. 2018.

[17] N. Nirala, R. Periyasamy, B. K. Singh, and A. Kumar, "Detection of type-2 diabetes using characteristics of toe photoplethysmogram by applying support vector machine," *Biocybern. Biomed. Eng.*, vol. 39, no. 1, pp. 38–51, Jan. 2019.

[18] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya, "Deep learning for healthcare applications based on physiological signals: A review," *Comput. Methods Programs Biomed.*, vol. 161, pp. 1–13, Jul. 2018.

[19] F. Esgalhado, B. Fernandes, V. Vassilenko, A. Batista, and S. Russo, "The application of deep learning algorithms for PPG signal processing and classification," *Computers*, vol. 10, no. 12, p. 158, Nov. 2021.

[20] V. B. Srinivasan and F. Foroozan, "Deep learning based non-invasive diabetes predictor using photoplethysmography signals," in *Proc. 29th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2021, pp. 1256–1260.

[21] A. Kamal, J. Harness, G. Irving, and A. Mearns, "Skin photoplethysmography—A review," *Comput. Methods Programs Biomed.*, vol. 28, no. 4, pp. 257–269, 1989.

[22] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiolog. Meas.*, vol. 28, no. 3, pp. R1–R39, Mar. 2007.

[23] M. Zheng, X. Zhang, S. Chen, Y. Song, Q. Zhao, X. Gao, and S. Wu, "Arterial stiffness preceding diabetes: A longitudinal study," *Circulat. Res.*, vol. 127, no. 12, pp. 1491–1498, Dec. 2020.

[24] C. Irace, C. Carallo, F. Scavelli, M. S. De Franceschi, T. Esposito, and A. Gnasso, "Blood viscosity in subjects with normoglycemia and prediabetes," *Diabetes Care*, vol. 37, no. 2, pp. 488–492, Feb. 2014.

[25] S. Masaoka, A. Lev-Ran, L. Robert Hill, G. Vakil, and E. H. G. Hon, "Heart rate variability in diabetes: Relationship to age and duration of the disease," *Diabetes Care*, vol. 8, no. 1, pp. 64–68, Jan. 1985.

[26] S. Stern and S. Sclarowsky, "The ECG in diabetes mellitus," *Circulation*, vol. 120, no. 16, pp. 1633–1636, Oct. 2009.

[27] H. Njoum and P. A. Kyriacou, "Photoplethysmography for the assessment of haemorheology," *Sci. Rep.*, vol. 7, no. 1, pp. 1–11, May 2017.

[28] M. Nitzan and Z. Ovadia-Blechman, "Physical and physiological interpretations of the PPG signal," in *Photoplethysmography*, J. Allen and P. Kyriacou, Eds. New York, NY, USA: Academic, 2022, pp. 319–340.
[29] S. Zanelli, M. A. E. Yacoubi, M. Hallab, and M. Ammi, "Transfer learning of CNN-based signal quality assessment from clinical to non-clinical PPG signals," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Nov. 2021, pp. 902–905.
[30] T. R. Dawber, H. E. Thomas, and P. M. McNamara, "Characteristics of the dicrotic notch," *Angiology*, vol. 24, no. 4, pp. 244–255, 1973.
[31] H. Obeid, G. Soulat, E. Mousseaux, S. Laurent, N. Stergiopulos, P. Boutouyrie, and P. Segers, "Numerical assessment and comparison of pulse wave velocity methods aiming at measuring aortic stiffness," *Physiolog. Meas.*, vol. 38, no. 11, pp. 1953–1967, Oct. 2017.
[32] Talos. (2019). *Autonomio Talos Computer Software*. [Online]. Available: http://github.com/autonomio/talos
[33] L. Brigato and L. Iocchi, "A close look at deep learning with small data," in *Proc. 25th Int. Conf. Pattern Recognit. (ICPR)*, Jan. 2021, pp. 2490–2497.
[34] V. A. Cargill, "Commentary: Valuing the vulnerable—The important role of transgender communities in biomedical research," *Ethnicity Disease*, vol. 30, no. 2, pp. 247–250, Apr. 2020.
[35] M. A. Almarshad, M. S. Islam, S. Al-Ahmadi, and S. A. Bahammam, "Diagnostic features and potential applications of PPG signal in healthcare: A systematic review," *Healthcare*, vol. 10, no. 3, pp. 1–28, 2022.

SERENA ZANELLI received the B.S. and M.S. degrees in biomedical engineering from the Polytechnic of Milan, in 2019. She is currently pursuing the Ph.D. degree in developing machine learning and deep learning approaches to detect cardiovascular diseases through PPG analysis. Her Ph.D. degree is carried on in collaboration with Axelife, a French enterprise specialized in the analyses of PPG signals, the University of Sorbonne Nord Paris, and TelecomSud Paris. Her research interests include signal quality assessment, waveform analysis, arterial stiffening, and cardiovascular diseases. She is a member of VascAgeNet, an European COST Action to promote awareness about vascular ageing.

MOUNIM A. EL YACOUBI (Member, IEEE) received the Ph.D. degree from the University of Rennes, France, in 1996. He was with Service de Recherche Technique de la Poste (SRTP), France, from 1992 to 1996. He was a Visiting Scientist with the Centre for Pattern Recognition and Machine Intelligence (CENPARMI), Montreal, Canada, and an Associate Professor with the Catholic University of Parana, Curitiba, Brazil, from 1998 to 2000. From 2001 to 2008, he was a Senior Software Engineer with Parascript, Boulder, Colorado. At the SRTP and Parascript, he has developed handwriting recognition software for real-life automatic mail sorting, bank check reading, and form processing. Since June 2008, he has been a Professor with the SAMOVAR Telecom Sud-Paris, Institut Polytechnique de Paris. His research interests include artificial intelligence, data science, machine learning, modeling human user data, especially behavioral signals like handwriting, voice, gesture and activity recognition, biometrics, e-health, smart agriculture, and smart cities.

MAGID HALLAB received the degree in medical studies, in 1987, the master's degree in internal medicine from Angers University Hospital, in 1992, and the master's degree in statistical medical studies from the University of Jussieu-Paris. He has participated in numerous studies in the field of hypertension and diabetes, particularly in diabetics nephropathy. He is the first author of *The BMJ* and participating in more than 50 scientific articles. In 2010, he founded a company with a medical device based on signal processing and AI to measuring the vascular ageing.

MEHDI AMMI (Member, IEEE) received the degree in electronics engineering, in 2000, and the Ph.D. degree in robotics, in 2005. He is currently pursuing the H.D.R. degree in computer science. He is also a Professor with the University of Paris 8. He was the Leader or Co-Leader of several international working groups, such as IEEE TCH and EuroVR Haptic SIG. He was involved in more than 20 national and international projects and industrial collaborations. His research interests include the study of pervasive environments and their societal and industrial applications (e-health, smart building, and the factory of the future). He addresses this field with a multidisciplinary approach combining artificial intelligence (machine learning and neuroscience), mechatronic technologies (organic electronics and smart textiles), and human factors (psychology and behavioral analysis).