View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by CURVE/open

Deep learning based photoplethysmography classification for peripheral arterial disease detection: a proof-of-concept study

Allen, J., Liu, H., Iqbal, S., Zheng, D. & Stansby, G.

Author post-print (accepted) deposited by Coventry University's Repository

Original citation & hyperlink:
Allen, J, Liu, H, Iqbal, S, Zheng, D & Stansby, G 2021, 'Deep learning based photoplethysmography classification for peripheral arterial disease detection: a proof-of-concept study', Physiological Measurement, vol. (In-Press), pp. (In-Press). https://dx.doi.org/10.1088/1361-6579/abf9f3

DOI 10.1088/1361-6579/abf9f3
ISSN 0967-3334
ESSN 1361-6579

Publisher: IOP Publishing

As the Version of Record of this article is going to be / has been published on a gold open access basis under a CC BY 3.0 licence, this Accepted Manuscript is available for reuse under a CC BY 3.0 licence immediately.

This document is the author's post-print version, incorporating any revisions agreed during the peer-review process. Some differences between the published version and this version may remain and you are advised to consult the published version if you wish to cite from it.

Physiological Measurement

IPEM Institute of Physics and Engineering in Medicine

ACCEPTED MANUSCRIPT • OPEN ACCESS

Deep learning based photoplethysmography classification for peripheral arterial disease detection: a proof-of-concept study

To cite this article before publication: John Allen *et al* 2021 *Physiol. Meas.* in press https://doi.org/10.1088/1361-6579/abf9f3

Manuscript version: Accepted Manuscript

Accepted Manuscript is "the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an 'Accepted Manuscript' watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors"

This Accepted Manuscript is **© 2021 Institute of Physics and Engineering in Medicine**.

As the Version of Record of this article is going to be / has been published on a gold open access basis under a CC BY 3.0 licence, this Accepted Manuscript is available for reuse under a CC BY 3.0 licence immediately.

Everyone is permitted to use all or part of the original content in this article, provided that they adhere to all the terms of the licence https://creativecommons.org/licences/by/3.0

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions may be required. All third party content is fully copyright protected and is not published on a gold open access basis under a CC BY licence, unless that is specifically stated in the figure caption in the Version of Record.

View the article online for updates and enhancements.

This content was downloaded from IP address 213.122.156.99 on 22/04/2021 at 08:12

Deep learning based photoplethysmography classification for peripheral arterial disease detection: a proof-of-concept study

John Allen[1,2,3], Haipeng Liu[2], Sadaf Iqbal[1,3], Dingchang Zheng[1,2], Gerard Stansby[4]

[1]Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne. UK.

[2]Research Centre for Intelligent Healthcare, Coventry University. UK.

[3]Northern Regional Medical Physics Department, Freeman Hospital, Newcastle upon Tyne. UK.

[4]Northern Vascular Centre, Freeman Hospital, Newcastle upon Tyne. UK.

Corresponding author (JA): John Allen (John.AllenVO@ieee.org)

Journal: **Physiological Measurement**

Manuscript version: 13th April 2021

Total number of words: ~3300 (Abstract <250)

Short running title: **Deep learning PPG detection of PAD**

Keywords: AI, Artery, Deep Learning, Peripheral Arterial Disease, Photoplethsymography, Pulse, Wavelet

Accepted Manuscript

Abstract

Objective: A proof-of-concept study to assess the potential of a Deep Learning (DL) based photoplethysmography PPG ('DLPPG') classification method to detect peripheral arterial disease (PAD) using toe PPG signals.

Approach: PPG spectrogram images derived from our previously published multi-site PPG datasets (214 participants; 31.3% legs with PAD by Ankle Brachial Pressure Index (ABPI)) were input into a pretrained 8-layer (5 Convolutional Layers + 3 Fully Connected Layers) AlexNet as tailored to the 2-class problem with transfer learning to fine tune the Convolutional Neural Network (CNN). k-fold random cross validation (CV) was performed [for k=5 and k=10], with each evaluated over k training / validation runs. Overall test sensitivity, specificity, accuracy, and Cohen's Kappa statistic with 95% confidence interval ranges were calculated and compared, as well as sensitivities in detecting mild-moderate (0.5≤ABPI<0.9) and major (ABPI<0.5) levels of PAD.

Main results: Cross validation with either k = 5 or 10 folds gave similar diagnostic performances. The overall test sensitivity was 86.6%, specificity 90.2% and accuracy 88.9% (Kappa: 0.76 [0.70-0.82]) (at k=5). The sensitivity to mild-moderate disease was 83.0% (75.5-88.9%) and to major disease was 100.0% (90.5-100.0%).

Significance: Substantial agreements have been demonstrated between the DL-based PPG classification technique and the ABPI PAD diagnostic reference. This novel automatic approach, requiring minimal pre-processing of the pulse waveforms before PPG trace classification, could offer significant benefits for the diagnosis of PAD in a variety of clinical settings where low-cost, portable and easy-to-use diagnostics are desirable.

Accepted Manuscript

1. Introduction

Peripheral arterial disease (PAD) of the lower limbs is a common form of widespread atherosclerosis and is associated with an increased risk of coronary artery disease and stroke (Abdulhannan *et al.*, 2012). PAD prevalence increases with age, especially in smokers and diabetics. Its diagnosis can be made on clinical grounds, but other conditions such as musculoskeletal, spinal disease and venous disease may also produce similar exercise induced symptoms. There are a range of tests to assess patients for possible PAD and this includes the widely used reference standard of an Ankle Brachial Pressure Index (ABPI) <0.9. The operation of ABPI measurement, however, typically takes 10-30 minutes even by specialist operators, may be inaccurate with vessel calcification and can cause discomfort and pain in some patients (Scott *et al.*, 2019).

Ideally a vascular screening technology for PAD should be low-cost, quick, reliable, repeatable, non-invasive, portable and simple to operate. One technique that has this potential is the low-cost optical pulse wave technology: photoplethysmography (PPG). PPG signals are derived from the changes in the blood volume in the microvascular bed of tissue, therefore can reflect on the properties of cardiovascular system in time and frequency domains (Allen, 2007). PPG signals can be measured at many different body sites (Perpetuini *et al.*, 2019; Chan *et al.*, 2019) although the toe site has been popular for developing methods to detect vascular disease. It is accepted that the peripheral PPG pulse wave usually becomes damped, delayed and diminished with increasing severity of PAD, which makes the PPG-based detection of PAD possible (Allen & Murray, 1993). In the recent works by our wider Newcastle research group, the PAD-

Accepted Manuscript

related changes in PPG waveform characteristics have been quantitatively investigated in subjects in different age groups and for PPG 'AC' as well as its lower frequency 'DC' components (Allen *et al.*, 2008; Allen *et al.*, 2005; Bentham *et al.*, 2018). The results indicate the potential value of using just bilateral great toe PPG measurements for low-cost and simple-to-do PAD detection. It has been suggested that the pulse wave from the great toe pad might be a better body site than ankle in detecting PAD, especially in "challenging populations" as those exhibiting arterial calcification (Herraiz-Adillo *et al.*, 2020).

As a promising approach towards large-scale healthcare services, the analysis of PPG signals based on deep learning (DL) has been recently applied to cardiovascular field for the detection of heart rate (Reiss *et al.*, 2019), hypertension (Liang *et al.*, 2018), arterial fibrillation (Cheng *et al.*, 2020). DL can accurately detect multiple PAD-related cardiovascular risks (e.g., hypertension, diabetes, cerebral infarction) (Panwar *et al.*, 2020). It has been suggested in two recent studies that DL-based PPG analysis could be useful in the detection of diseases of the arteries (Lee *et al.*, 2020; Panwar *et al.*, 2020). A DL-based analysis of the second derivative of the finger PPG trace showed a high accuracy in the prediction of the ABPI class for severity of arterial disease (accuracy 98.34%, for six ABPI classes covering mild through to severe disease, arterial hardening, and normal) (Lee *et al.*, 2020), although there could have been overtraining which compensated for known accuracy limitations with ABPI reference for PAD. Since PAD symptoms can manifest themselves early in the legs then it is considered appropriate to study the PPG measurements obtained more directly from the lower limbs. However, as far as we know there is a clear lack of DL-based studies on the detection of PAD based using PPG signals

Accepted Manuscript

collected from the toes. The aim of this proof-of-concept study was to assess the potential of a DL-based method for automatically detecting the presence (or absence) of PAD from the toe PPG measurements as well as its sensitivity to detecting higher and lower grade disease.

2. Methods

2.1 Measurements

The physiological measurements are described in detail in (Allen *et al.*, 2005; Allen *et al.*, 2008). In summary, they were performed in a warm temperature-controlled room (24 ± 1 °C), and at least 10 minutes were given for thermal acclimatization and relaxation. Measurements were made by a single operator (JA) with subjects in the supine position, firstly with the ABPI measurements for both legs as the 'gold standard' reference for the presence of PAD (i.e., <0.9, using highest of the right and left ankle systolic blood pressures), and secondly the bilateral PPG great toe pulse measurements from a multi-site PPG pulse measurement concept prototype system. The PPG measurement system simultaneously acquired bilateral toe pulses using electronically matched pairs of right and left pulse amplifiers of bandwidth (0.15–20 Hz, single pole high pass filtering) with data captured to computer for subsequent pulse wave analysis for at least 2.5 minutes (Allen & Hedley, 2019). Subject age, gender and height were also recorded.

2.2 Subjects

Accepted Manuscript

The two cohorts studied using the same protocol were combined to evaluate the DL-based PAD classifier and these sets are summarized in our 2005 pilot evaluation paper (Allen 2002, Allen *et al.*, 2005) and our 2008 prospective study paper (Allen *et al.*, 2008). In summary, they comprised a mixture of older subjects, with and without significant lower limb occlusive PAD. Subjects were excluded if they had an obvious cardiac arrhythmia, lower limb amputation, vasculitis, significant movement artefact (for example due to limb tremor), skin problems (e.g., cuts or bruising at a measurement site or photosensitive skin) and Raynaud's phenomenon. We did not specifically exclude those with hypertension, diabetes, hypercholesterolaemia, chronic renal failure or ischaemic heart disease as these are all relevant to the atherosclerotic phenotype. All subjects gave their written informed consent. More recent further ethical approval was also granted for a further re-analysis of the anonymized sets of PPG waveforms (Newcastle University Ethics Committee, 7840/2020). The combined 2 studies gave a total of 218 potential subjects although a small number of these subjects (3) had elevated ABPIs (>1.4 in either leg) which were excluded owing to the risk of false negative results observed when ABPI is used in calcified blood vessels. A further subject was excluded as they were found to show a consistent cardiac arrhythmia on their pulse traces. In total therefore, PPG toe pulses from 214 subjects were available for subsequent DL classification and study, comprising 134 subjects having normal arteries in both legs and 80 subjects having significant vascular disease in at least one leg (i.e., 37.4% of the total by subject; 31.3% of legs having PAD) by ABPI classification.

2.3 PPG Pulse analysis

Accepted Manuscript

2.3.1 Signal pre-processing

The raw PPG signals for right and left great toes over 2.5 minutes were pre-processed using detrending by removing their mean levels, and then normalizing to unit variance. No pre-filtering or denoising stages were employed. A continuous wavelet transform (CWT, 'amor' i.e., Gabor wavelet MATLAB, MathWorks Inc., version 2020a) was applied and a spectrogram image produced for each great toe measurement, i.e., for both right and left legs separately, and covering the 2.5 minute period ($\log_{10}$ y-axis: 0.0312 to 8 Hz to capture the dynamics of the PPG signal over time, Figure 1). The intensity for the colour bar of each CWT spectrogram was auto scaled, the image then saved in *.png file format. This gave 428 files (by ABPI reference standard: 292 / 136 legs without / with PAD, in which 94 legs with mild-moderate disease and 40 legs with major disease) for subsequent DL training and cross validation testing.

2.3.2 Deep learning convolutional neural network and training

Convolutional Neural Network (CNN) is a class of DL networks that is designed to learn features from the input data using its multiple Convolutional Layers. The DL analysis was performed using MATLAB Deep Learning Toolbox software (The Mathworks Inc., Natick, Massachusetts, United States) on a standard computing platform. AlexNet (http://alexlenail.me/NN-SVG/AlexNet.html) was used which is a pre-trained CNN that has been pre-trained on a database named ImageNet consisting of more than 1.2 million images belonging to one thousand classes (Krizhevsky *et al.*, 2012). AlexNet is well known (Krizhevsky *et al.*, 2012; Lu *et al.*, 2019; Wang *et al.*, 2019). Briefly, its structure has 8 layers: five Convolutional Layers (CL1 to 5) and three Fully Connected Layers (FCL6-8), with Rectified

Accepted Manuscript

Linear Unit (ReLU) as the activation function. The final FCL8 layer was tailored to the 2-class problem (Figure 2) (Wang *et al.*, 2019). Transfer learning was applied for fine tuning of specific layers of the CNN. Treating each leg separately, each great toe PPG spectrogram image was rescaled to the AlexNet CNN input layer structure size of 227x227 pixels, with 3-colour channels of red, green and blue. The 2 CNN outputs gave the classification outcome of 'no PAD' and 'PAD'. With the absence of fixed rules for deep learning parameter selection in the literature then initial experiments were carried out within the Matlab Deep Learning Toolbox development environment to set the key parameters for AlexNet-based PAD / PPG application. The learning rate and the momentum term parameters were initially and briefly explored in the ranges of 0.00001 to 0.01 and 0.4 to 0.9, respectively, with learning considered to be demonstrated when there was a trajectory towards an acceptable classification performance (e.g., 75-100%) at each fold tested (using 30 epochs with mini-batch size 20), i.e., without the algorithm getting stuck at the low starting performance level. The choice of number of epochs depends on training and validation error. Keeping the number of epochs small can lead to underfitting while making this number high can lead to overfitting of the model. In our study, the validation error trajectory approached a minimum at 30 epochs and hence this number was selected. Subsequently, a learning rate of 0.001 and momentum term of 0.5 were selected for the k-fold cross validation testing as these showed learning across all folds and carried out with reasonable speed. To aid the learning and generalization then pixel shift and image augmentation were also incorporated into the training stage.

Accepted Manuscript

2.3.3 Consistency with k-fold cross validation

As there is no standardization in the choice of k for k-fold random cross validation we experimented with two different values (k=5 and k=10), with k training / validation runs, as applied in existing studies on DL-based PPG signal analysis (Dall'Olio *et al.*, 2020; He *et al.*, 2016). For each training run the image sets were randomly split into a training set and a validation set, with transfer learning from the PPG spectrogram images. At each training run the confusion matrix was updated, and on the completion of all runs the overall test sensitivity, specificity and accuracy (with 95% confidence interval ranges) were calculated. The process was repeated for k set to 10 enabling a simple comparison of the classification performance obtained for a k-fold value of 5.

2.3.5 Statistical analysis and diagnostic performance assessments

Clinical measurements were summarized using median and inter-quartile range [IQR, 25th percentile to 75th percentile], calculated using simple Microsoft Office Excel spreadsheet functions. Cross-tabulation of PAD status by ABPI and DL classifier was performed from which associated diagnostic test accuracy (DTA) measures of sensitivity, specificity, and diagnostic accuracy, were determined alongside associated 95% confidence intervals overall, as well as per severity classification. DTA statistical analyses were performed using SciStat.com online statistical software © 2020 MedCalc Software Ltd. The sensitivities in detecting overall disease (ABPI<0.9), mild-moderate disease (0.5$\leqslant$ABP<0.9) and major disease (ABPI <0.5) were calculated for the PPG spectrogram image sets. Cohen's Kappa statistic and 95% CI ranges were also calculated for the diagnostic performances (McHugh, 2012) on SPSS (SPSS 21.0 for Windows,

Accepted Manuscript

SPSS Inc., Chicago, IL), noting a Kappa value between 0.41 and 0.6 represents moderate agreement; between 0.61 and 0.8 represents substantial agreement; and 0.81 and 1.00 represents almost perfect agreement.

3. Results

Toe PPG pulse spectrogram images from a total of 214 subjects (137 males, 77 females) were classified using deep learning. The median [IQR] age was 64 [52 - 72] years, systolic blood pressure 142 [128 - 160] mmHg, and height 1.69 [1.60 - 1.70] m.

The overall diagnostic performances with AlexNet DL classification are shown in Tables 1 and 2, with all comparisons having substantial agreement (by Cohen's Kappa statistic, i.e., for agreement beyond chance) with ABPI. The tables also summarize the values (and 95% CI ranges) of sensitivity, specificity and accuracy for k=5 (and 10) over the k repeat training / validation runs.

For k=5 the overall test sensitivity was 86.6 (95% CI: 80.6-91.3) %, specificity 90.2 (86.2-93.3) %, accuracy 88.9 (85.7-91.6) %, and Kappa 0.76 [0.70-0.82]). The test sensitivity was clearly higher overall for the legs with higher grade PAD, with 83.0 (75.5-88.9) % for mild-moderate disease vs 100.0 (90.5-100.0) % for major disease, giving Cohen's Kappa values of 0.72 (0.65-0.79) and 0.67 (0.56-0.78), respectively.

For k=10 the results were similar to the k=5 results, with overall test sensitivity 82.4 (95% CI: 74.8-88.5) %, specificity 89.0 (84.8-92.3) % and accuracy 86.9 (83.3-90.0) %, and Kappa 0.70 [0.63-0.77]). As with k=5, the test sensitivity was

Accepted Manuscript

clearly higher overall for the legs with higher grade PAD, with 77.3 (67.7-85.2) % for mild-moderate disease vs 97.1 (84.7-99.9) % for major disease, giving Cohen's Kappa values of 0.64 (0.55-0.73) and 0.61 (0.50-0.73), respectively.

4. Discussion

An innovative DL-based PPG Classifier ('DLPPG') approach using pre-trained Convolutional Neural Network AlexNet with fine tuning by transfer learning to diagnose PAD from toe PPG measurements has been evaluated in this study, with performance assessed in this proof-of-concept study to show substantial agreements overall with the ABPI vascular reference for PAD.

Currently, PAD diagnostics depends on mainly on ABPI measurement as the first-line test. In the 2016 American Heart Association (AHA) / American College of Cardiology (ACC) Guideline on PAD, the technological advancement for simplified diagnosis of PAD and critical limb ischemia is listed as a major research direction (Gerhard-Herman *et al.*, 2017). The proposed DL-based analysis of the toe PPG signal provides the possibility for quick assessment and follow-up over time, and low-cost management of PAD. The novel DL-based approach assessed in this study has advantages over previously reported approaches (e.g., analysis involving fiducial waveform characteristics) (Karimian *et al.*, 2017) in PPG pulse analysis, with some earlier methods (Allen 2000; Allen *et al*., 2005; Allen *et al*., 2008) being resource intensive in pre-processing as they required the manual checking by an expert operator on a beat by beat basis to mark and exclude noisy pulses, and noting that derived pulse timing features and normalised pulse shapes are dependent on digital filter types which can be

Accepted Manuscript

difficult to replicate for devices working on-line i.e., in real time. In contrast, this new approach based on deep learning required minimal data pre-processing and had no denoising stage.

The manual extraction of physiological parameters from time-frequency spectra is cumbersome, highly parametrised, and tailored to specific scenarios, whereas the DL-based PPG analysis is appropriate for the large-scale application on bespoke measurement devices (Wilkes *et al.* 2015; Reiss *et al.*, 2019) as well as centralised cloud-based signal diagnostics. Ultimately, such pulse assessments could improve the management of patients with PAD by giving accessible and timely feedback to the patient and their doctors to help them reduce their associated risk factors for cardiovascular disease.

The AHA/ACC Guideline also highlighted the need for the improvement of clinical classification systems for PAD that incorporate symptoms, anatomic factors, and patient-specific risk factors and can be used to predict clinical outcome and optimize treatment approach (Gerhard-Herman *et al.*, 2017). Our results showed substantial accuracy, sensitivity, and Cohen's Kappa values for legs with mild-moderate PAD as well as for major PAD, although the detection sensitivity was lower for lower grade disease. These classification performances appear similar to the results of our earlier studies based on specific characteristics of the PPG toe pulse, e.g., using careful manual extraction of normalised shape, risetime and amplitude features (Allen *et. al.* 2005; Allen *et. al.* 2008) for overall accuracy and differences with PAD severity in the legs. Our new approach also includes information on the dynamics of the PPG, i.e., the 'DC' lower frequency components, in health and disease and this could well yield additional important

Accepted Manuscript

information that aids the diagnosis and understanding of the disease process, for example for diabetic patients with PAD where there is autonomic involvement and/or vascular calcification, and this should be further explored as an exciting way forward in vascular assessment.

Other research groups have explored the use of DL-based PPG analysis in vascular assessments, including Lee *et al*. (2020) who used a convolutional long short term memory model (C-LSTM: with 5 convolutional layers, 5 pooling layers and one LSTM layer) to classify the second derivative of the finger PPG trace to predict the ABPI class i.e., severity of arterial disease (data from MIMIC-III database, Johnson *et al*., 2006). Their results are impressive and in approximately 1000 subjects attained an accuracy of greater than 98% (for six ABPI classes covering mild through to severe disease, arterial hardening, and normal). They did combine 2 DL models (CNN and LSTM) which could have improved performance, compared to the results from our single DL model. However, there may be an element of over-fitting which somehow compensates for the imperfect ABPI reference for PAD. Such results do show what can be possible with deep learning. It will be very interesting future step for us to look at both finger and toe PPG traces in our multi-site PPG measurements to see if classification performance can be boosted significantly.

A well-known aspect of PPG measurements is that movement artefact, e.g. movement of the limb during ambulation or even just whilst a patient is resting but talking, and changes in body posture can sometimes render a trace unusable (Carek *et al.*, 2020); Huthart *et al.*, 2020) and so a next step could be to test the DL-based classifier resilience to certain types of added noise and artefacts. We

Accepted Manuscript

would be optimistic that the CNN algorithm, with a max-pooling layer which is included in AlexNet, could enhance the algorithm robustness to noise, which has been proven in parallel studies on the DL-based analysis of electrocardiogram (Muhammed and Aravinth, 2019). The development and application of CNN and other DL-based methods could be investigated so that reliable evaluation of PAD can be made in poor quality signals (Waugh *et al.*, 2018).

Limitations and further work: A limitation of our proof-of-concept study is that the distribution of non-PAD, mild-moderate PAD, and major PAD cases was unbalanced. A future study should consider assessing unselected patients from primary care rather than specialist vascular centres. Due to the limitations on data, we also did not include other (patho-)physiological factors in the model, such as symptoms, history of heart disease, renal disease, and diabetes, etc. In future studies, we will recruit a new cohort of cases and controls and include a comparison with other PAD measures including the Edinburgh Claudication Questionnaire (ECQ), manual leg pulse palpation, toe systolic pressures, and/or duplex vascular ultrasound imaging (Kyle *et. al.*, 2020). Diabetic patients should also be studied further particularly the PPG 'AC' and 'DC' lower frequency components for with and without PAD cases (Bentham *et al.*, 2018). Other CNN architectures such as LeNet, GoogLeNet, and ResNet can also be tested to see if they can boost the efficiency and accuracy of the classifications, particularly in earlier milder cases of disease. Other pathophysiological parameters (e.g., symptoms, anatomic factors, and patient-specific risk factors) can be added to achieve more accurate classification of PAD cases. Finally, based on the Health Economic Modelling appraisal of the proposed concept technology, miniaturized

Accepted Manuscript

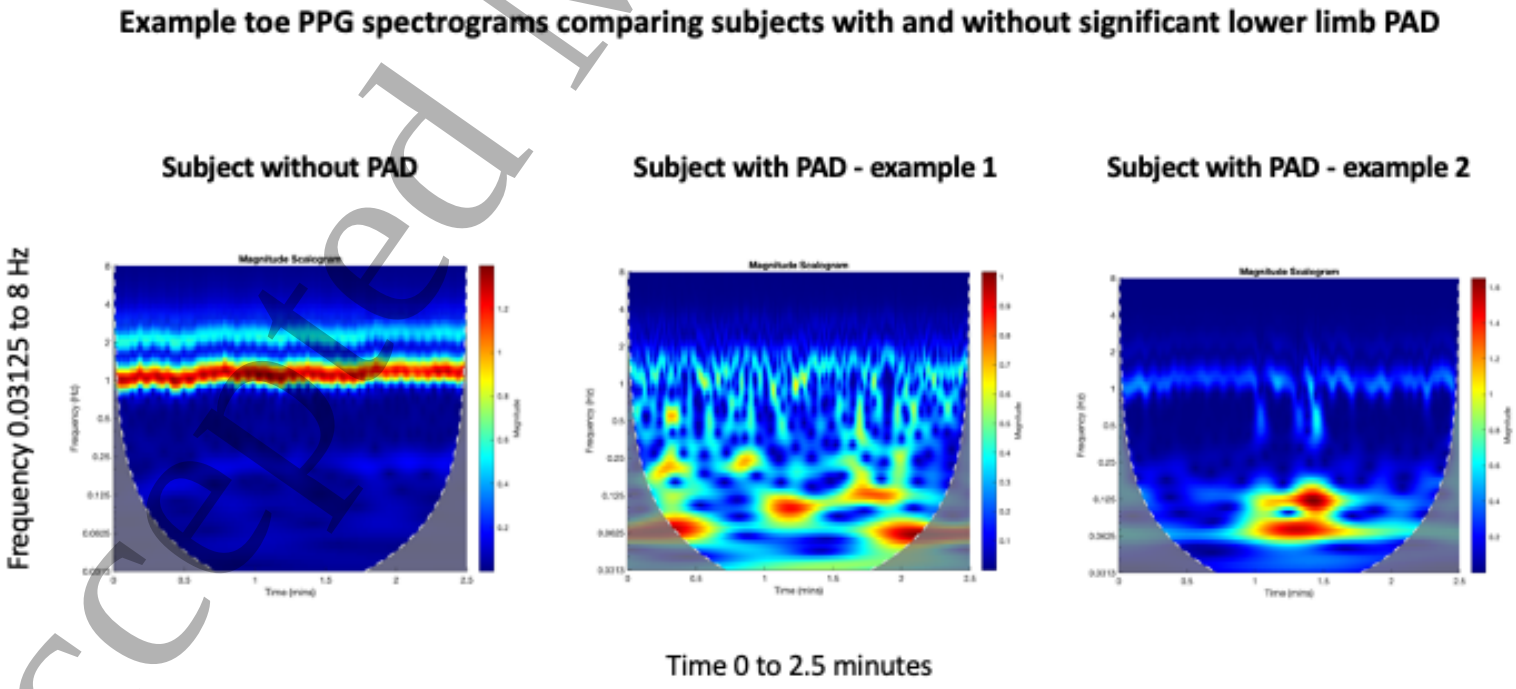
DL chips and cloud-based computing can be applied to achieve a use-friendly, cost-effective, and real-time system to improve healthcare services diagnosis, management, and treatment of PAD.

5. Summary

We have demonstrated in this proof-of-concept study that, with only limited signal pre-conditioning, our DL-based PPG method can be configured reliably to detect PAD (as diagnosed by ABPI) from simple toe PPG measurements in a hospital vascular department setting. The sensitivity of the technique was higher for more severe disease (i.e., an ABPI<0.5). Further works are already in progress to help improve the classification algorithm and its resilience to noise and measurement artefacts. Ultimately, the trajectory for the project area will be to use *explainable AI* techniques to help understand the complex changes in PPG characteristics seen with atherosclerosis and in vascular ageing.

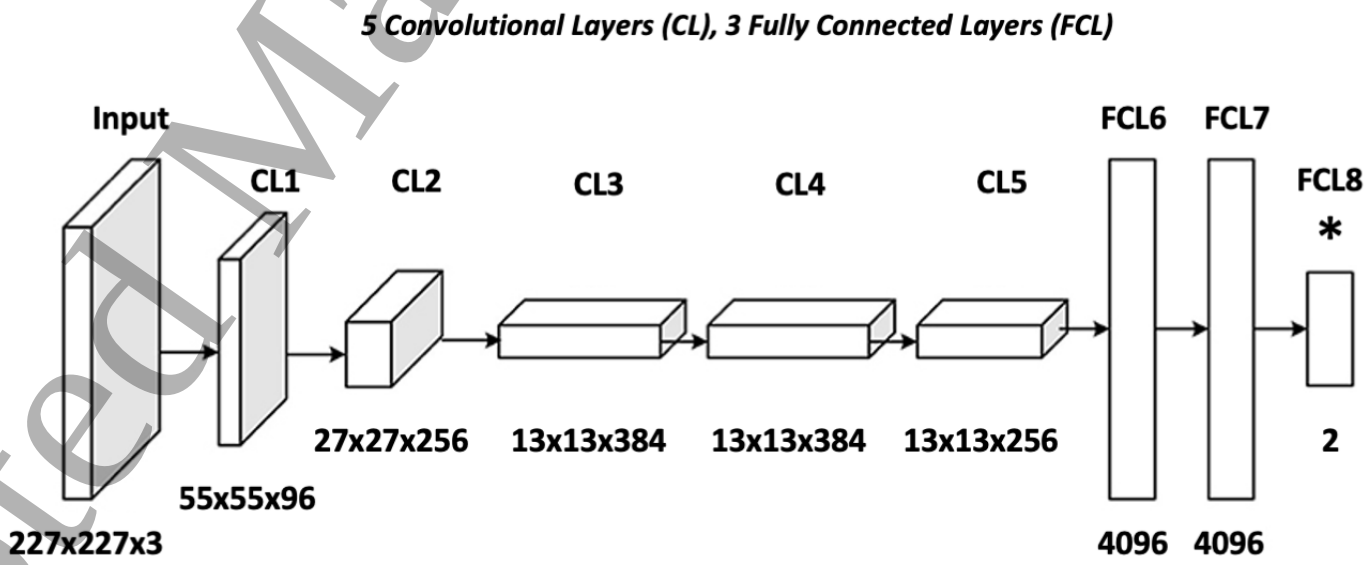
Accepted Manuscript

Figure 1: Example spectrograms comparing time-frequency characteristics over a period of 2.5 minutes for the raw unfiltered PPG signals from the great toe. Clear differences can be seen between these PAD and non-PAD examples, for example in pulse harmonics, variability over time as well as the "AC" to "DC" lower frequency components.



Accepted Manuscript

Figure 2: Overview of structure of AlexNet and its tailoring to PPG PAD classification.



* FCL8 Layer	Original AlexNet	Replaced by
23	FCL (1000), pre-trained weights & biases	FCL (2 classes), random initialization
24	Softmax Layer	Softmax Layer
25	Classification Layer (= 1000 classes)	Classification Layer (= 2 classes: 'no PAD' and 'PAD')

Accepted Manuscript

Tables

Tables 1a and b 5-fold random cross validation with Confusion Matrix from 5 training/validation runs and DTA performance for DLPPG vs ABPI-determined disease (Overall, Mild-Moderate, Major) on a leg basis. Cohen's Kappa statistic for agreement between 2 raters is also give. The 95% Confidence Intervals are shown in brackets.

(a)

Diagnosis by ABPI	DL-based PPG classification	
	No PAD	**PAD**
No PAD	**266**	29
Mild-Moderate	23	**112**
Major	0	**37**

(b)

	Sensitivity (%)	Specificity (%)	Accuracy (%)	Cohen's Kappa
Overall	**86.6 (80.6 – 91.3)**		**88.9 (85.7 – 91.6)**	**0.76 (0.70-0.82)**
Mild – Moderate	83.0 (75.5 – 88.9)	**90.2 (86.2 – 93.3)**	87.9 (84.5 – 90.8)	0.72 (0.65-0.79)
Major	100.0 (90.5 - 100.0)		91.3 (87.7 – 94.1)	0.67 (0.56-0.78)

Accepted Manuscript

Tables 2a and b 10-fold random cross validation with Confusion Matrix from 10 training/validation runs and DTA performance for DLPPG vs ABPI-determined disease (Overall, Mild-Moderate, Major) on a leg basis. Cohen's Kappa statistic for agreement between 2 raters is also give. The 95% Confidence Intervals are shown in brackets.

(a)

Diagnosis by ABPI	DL-based PPG classification	
	No PAD	**PAD**
No PAD	**258**	32
Mild-Moderate	22	**75**
Major	1	**33**

(b)

	Sensitivity (%)	Specificity (%)	Accuracy (%)	Cohen's Kappa
Overall	**82.4 (74.8 – 88.5)**		**86.9 (83.3 – 90.0)**	**0.70 (0.63-0.77)**
Mild – Moderate	77.3 (67.7 – 85.2)	**89.0 (84.8 - 92.3)**	86.1 (82.2 – 89.3)	0.64 (0.55-0.73)
Major	97.1 (84.7 – 99.9)		89.8 (86.0 – 92.9)	0.61 (0.50-0.73)

Accepted Manuscript

References

Abdulhannan P, Russell D A and Homer-Vanniasinkam S 2012 Peripheral arterial disease: a literature review *British Medical Bulletin* **104** 21-39

Allen J 2002 Measurement and analysis of multi-site photoplethysmographic pulse waveforms in health and arterial disease. Newcastle University, PhD thesis.

Allen J 2007 Photoplethysmography and its application in clinical physiological measurement *Physiological Measurement* **28** R1-R39

Allen J and Hedley S 2019 Simple photoplethysmography pulse encoding technique for communicating the detection of peripheral arterial disease—a proof of concept study *Physiological Measurement* **40** 08NT1

Allen J and Murray A 1993 Development of a neural network screening aid for diagnosing lower limb peripheral vascular disease from photoelectric plethysmography pulse waveforms *Physiological Measurement* **14** 13-22

Allen J, Oates C P, Lees T A and Murray A 2005 Photoplethysmography detection of lower limb peripheral arterial occlusive disease: a comparison of pulse timing, amplitude and shape characteristics *Physiological Measurement* **26** 811-21

Allen J, Overbeck K, Nath A F, Murray A and Stansby G 2008 A prospective comparison of bilateral photoplethysmography versus the ankle-brachial pressure index for detecting and quantifying lower limb peripheral arterial disease *Journal of Vascular Surgery* **47** 794-802

Bentham M, Stansby G and Allen J 2018 Innovative Multi-Site Photoplethysmography Analysis for Quantifying Pulse Amplitude and Timing Variability Characteristics in Peripheral Arterial Disease **6** 81

Carek A M, Jung H and Inan O T 2020 A Reflective Photoplethysmogram Array and Channel Selection Algorithm for Weighing Scale Based Blood Pressure Measurement *IEEE Sensors Journal* **20** 3849-58

Chan G, Cooper R, Hosanee M, Welykholowa K, Kyriacou P A, Zheng D, Allen J, Abbott D, Lovell N H, Fletcher R and Elgendi M 2019 Multi-Site Photoplethysmography Technology for Blood Pressure Assessment: Challenges and Recommendations **8** 1827

Cheng P, Chen Z, Li Q, Gong Q, Zhu J and Liang Y 2020 Atrial Fibrillation Identification With PPG Signals Using a Combination of Time-Frequency Analysis and Deep Learning *IEEE Access* **8** 172692-706

Dall'Olio L, Curti N, Remondini D, Safi Harb Y, Asselbergs F W, Castellani G and Uh H-W 2020 Prediction of vascular aging based on smartphone acquired PPG signals *Scientific Reports* **10** 19756

Gerhard-Herman M D, Gornik H L, Barrett C, Barshes N R, Corriere M A, Drachman D E, Fleisher L A, Fowkes F G R, Hamburg N M, Kinlay S, Lookstein R, Misra S, Mureebe L, Olin J W, Patel R A G, Regensteiner J G, Schanzer A, Shishehbor M H, Stewart K J, Treat-Jacobson D, Walsh M E, Halperin J L, Levine G N, Al-Khatib S M, Birtcher K K, Bozkurt B, Brindis R G, Cigarroa J E, Curtis L H, Fleisher L A, Gentile F, Gidding S, Hlatky M A, Ikonomidis J, Joglar J, Pressler S J and Wijeysundera D N 2017 2016 AHA/ACC Guideline on the Management of Patients with Lower Extremity Peripheral Artery Disease: Executive Summary *Vascular Medicine* **22** NP1-NP43

He R, Huang Z, Ji L, Wu J, Li H and Zhang Z 2016 Beat-to-beat ambulatory blood pressure estimation based on random forest. In: *2016 IEEE 13th International Conference on Wearable and Implantable Body Sensor Networks (BSN),* pp 194-8

Accepted Manuscript

Herraiz-Adillo Á, Cavero-Redondo I, Álvarez-Bueno C, Pozuelo-Carrascosa D P and Solera-Martínez M 2020 The accuracy of toe brachial index and ankle brachial index in the diagnosis of lower limb peripheral arterial disease: A systematic review and meta-analysis *Atherosclerosis* **315** 81-92

Huthart S, Elgendi M, Zheng D, Stansby G, Allen J. Advancing PPG signal quality and know-how through knowledge translation- from experts to student. Frontiers in Digital Health – Health Informatics section. 2020. https://doi.org/10.3389/fdgth.2020.619692

Johnson A E W, et al. 2016 MIMIC-III, a freely accessible critical care database. *Scientific Data* **3** 160035.

Karimian N, Tehranipoor M and Forte D 2017 Non-fiducial PPG-based authentication for healthcare application. In: *2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI),* pp 429-32

Krizhevsky A, Sutskever I, Hinton GE 2012 Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems **25** 1097-1105

Kyle D, Boylan L, Wilson L, Haining S, Oates S, Sims A, Guri I, Allen J, Wilkes S, Stansby G 2020 Accuracy of Peripheral Arterial Disease Registers in UK General Practice: Case-Control Study *Angiology* https://doi.org/10.1177/2150132720946148

Lee J J, Heo J H, Han J H, Kim B R, Gwon H Y and Yoon Y R 2020 Prediction of Ankle Brachial Index with Photoplethysmography Using Convolutional Long Short Term Memory *Journal of Medical and Biological Engineering* **40** 282-91

Liang Y, Chen Z, Ward R and Elgendi M 2018 Photoplethysmography and Deep Learning: Enhancing Hypertension Risk Stratification **8** 101

Lu S, Lu Z & Zhang Y D 2019 Pathological brain detection based on AlexNet and transfer learning. *J Computational Science* **30** 41-47

McHugh M L J B m B m 2012 Interrater reliability: the kappa statistic **22** 276-82

Muhammed M A A and Aravinth J 2019 CNN based Off-the-Person ECG Biometrics. In: *2019 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET),* pp 217-21

Panwar M, Gautam A, Dutt R and Acharyya A 2020 CardioNet: Deep Learning Framework for Prediction of CVD Risk Factors. In: *2020 IEEE International Symposium on Circuits and Systems (ISCAS),* pp 1-5

Perpetuini D, Chiarelli A M, Maddiona L, Rinella S, Bianco F, Bucciarelli V, Gallina S, Perciavalle V, Vinciguerra V, Merla A and Fallica G 2019 Multi-Site Photoplethysmographic and Electrocardiographic System for Arterial Stiffness and Cardiovascular Status Assessment **19** 5570

Reiss A, Indlekofer I, Schmidt P and Van Laerhoven K 2019 Deep PPG: Large-Scale Heart Rate Estimation with Convolutional Neural Networks **19** 3079

Scott J, Lecouturier J, Rousseau N, Stansby G, Sims A, Wilson L and Allen J 2019 Nurses' and patients' experiences and preferences of the ankle-brachial pressure index and multi-site photoplethysmography for the diagnosis of peripheral arterial disease: A qualitative study *PLOS ONE* **14** e0224546

Vriens B, D'Abate F, Ozdemir B A, Fenner C, Maynard W, Budge J, Carradice D and Hinchliffe R J 2018 Clinical examination and non-invasive screening tests in the diagnosis of peripheral artery disease in people with diabetes-related foot ulceration *Diabetic Medicine* **35** 895-902

Accepted Manuscript

Wang S-H, Xie S, Chen X, Guttery DS, Tang C, Sun J and Zhang Y-D 2019 Alcoholism identification based on an AlexNet transfer learning model *Front. Psychiatry* **10** 205. doi: 10.3389/fpsyt.2019.00205

Waugh W, Allen J, Wightman J, Sims A J and Beale T A W 2018 Novel Signal Noise Reduction Method through Cluster Analysis, Applied to Photoplethysmography *Computational and Mathematical Methods in Medicine* 6812404

Wilkes S, Stansby G, Sims A, Haining S, Allen J 2015 Peripheral arterial disease: diagnostic challenges and how photoplethysmography may help. *British Journal of General Practice* **65 (635)**: 323-324 doi.org/10.3399/bjgp15X685489

Accepted Manuscript