BMC Artificial Intelligence

RESEARCH

Open Access

Hyperspectral reconstruction for mobile diabetic foot blood perfusion monitoring

Alexander Gherardi[1*], Wei Bo[1], Ahmet Demirbas[1], Ye Zhan[2] and Wenyao Xu[1]

Abstract

Background Blood Perfusion is a key factor in the development and healing of wounded tissues including Diabetic Foot Ulcers (DFU), a harmful chronic wound caused by diabetic neuropathy. Recent works have explored the use of hyperspectral imaging (HSI) to non-invasively quantify the quality of blood perfusion with high spatial resolution. Later works consider the use of hyperspectral reconstruction (HSR) to provide the same capability using unmodified commodity hardware, such as smartphone cameras, using computational methods to yield full hyperspectral images from RGB ones. However, these HSR perfusion systems require profiles for each camera they are used with and furthermore require radiometric calibration to account for environmental lighting conditions before each use.

Methods In this work we demonstrate MobiPerf which extracts oxygenation signals/images along with high fidelity remote PPG signals while overcoming these challenges. To eliminate the need for camera profiles, our system uses deep learning HSR models that have been shown to generalize well across different cameras. Then to overcome the need for reference image calibration, we utilize a custom algorithm *Calibration Free Skin Compensation Estimation*.

Results Evaluated under 5 different simulated lighting conditions from the CIE Standard Illuminates, our system maintains strong agreement with oxygenation images/signals extracted directly from HSI cameras. Our testing on in-the-wild RGB data from a publicly available dataset of diabetic foot ulcer images ($N \approx 6000$) shows an acute sensitivity to Ischemia conditions ($p < 1 \times 10^{-5}$) as well as a more limited sensitivity to infection complications. Along with a dataset of videos with contract PPG ($N = 56$) which shows rPPG performance on par or better than other state-of-the-art algorithms.

Conclusions Our results demonstrate that a HSR system can be used to monitor diabetic foot ulcers using just images/videos minimizing the need for procedures prior to or during use and with mobile hardware patients already have. We anticipate that in the future our advancements in HSR can be used for other smart health applications that relate to perfusion, and we anticipate that similar HSR based systems can be used to monitor other tissue parameters such as sweat concentrations.

Keywords Hyperspectral reconstruction, Physics informed deep learning, Mobile health, Diabetic foot ulcers

*Correspondence:
Alexander Gherardi
ajgherar@buffalo.edu
[1]Department of Computer Science and Engineering, University at Buffalo, Davis Hall, Buffalo, NY 14260, USA
[2]Linde Engineering Americas, 175 E Park Drive, Tonawanda, NY 14150, USA

Introduction

Diabetic foot ulcers (DFUs) are a serious complication of diabetes, affecting both the mobility and the overall health of patients. Chronically high blood glucose levels cause damage to the small blood vessels (capillaries) that supply oxygen and nutrients to the tissue within the appendage. This vascular damage often leads to diabetic neuropathy (peripheral nerve damage), a condition that reduces sensation in the feet, masking minor injuries. As a result, unnoticed small blisters worsen into ulcers under the continued pressure of daily activities. The compromised capillaries around these ulcers are stressed further during the Healing process, which depends on adequate blood perfusion for oxygen and nutrient supply. Insufficient perfusion delays Healing, increases infection risk, and complicates recovery. Notably, infections occur in about 50% of DFU cases, with nearly 20% progressing to foot amputation, severely diminishing quality of life and survival rates [1]. In the United States, an estimated 1.6 million people live with DFUs [1], many of whom face mobility challenges that limit access to specialist care. Therefore, home tools are essential for the early detection of ulcer complications, reducing infection rates and minimizing the risk of loss of limbs and mortality. On a larger scale early detection of ulcer complications and timely care can greatly alleviate the significant healthcare burden associated with diabetic foot conditions.

Clinical technologies for evaluating DFUs utilize advanced imaging techniques to assess tissue health and vascular function, each with distinct capabilities. Computed Tomography (CT) Angiograms deliver high-resolution images of the macrovascular system, detecting arterial blockages or stenosis [2], but require radiation exposure, contrast agents, and large equipment, making them unsuitable for routine use. Hyperspectral Imaging (HSI) is a non-invasive technology that captures detailed spectral information across 10–100 color bands in the visible and near-visible spectrum [3]. Near-Infrared Spectroscopy (NIRS) operates similarly, focusing on near-infrared wavelengths for deeper tissue penetration [4]. Both HSI and NIRS leverage the unique spectral signatures of key chromophores in skin tissue—hemoglobin (Hb) and oxyhemoglobin (HbO_2), and melanin—to assess tissue health. These chromophores absorb light at specific wavelengths, much like a tuning fork resonates at certain frequencies. By observing and analyzing reflected light across spectral bands, these technologies are used to estimate chromophore concentrations and generate high-resolution composition maps of tissue. While highly valuable for evaluating tissue oxygenation and vascular health, these methods are constrained by their high costs and limited scalability, underscoring the need for more accessible and versatile solutions for the early detection and management of DFUs (Fig. 1).

Mobile solutions utilizing RGB cameras and smartphone-based imaging applications are among the most widely explored approaches. These tools enable non-invasive and straightforward tracking of changes in ulcer size or surface characteristics [6], making them convenient mobile solutions for patients and healthcare providers. However, their reliance on readily observable visible changes, which often take weeks to manifest, significantly

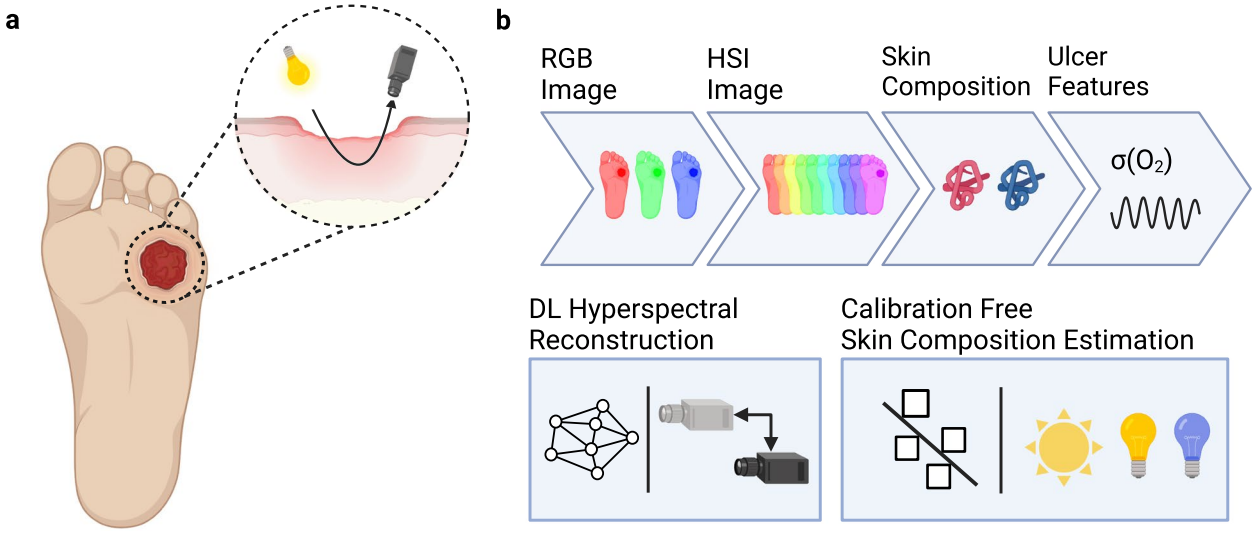

Fig. 1 Overview of MobiPerf, and pipeline. **a** Our system MobiPerf works by observing the precise color modulation applied to incident light reflected back by skin tissue using normal RGB cameras through hyperspectral reconstruction and custom algorithms/features. **b** Pipeline of MobiPerf (upper), RGB images taken with smartphones are converted to better defined and richer hyperspectral images using deep learning (DL-HSR), a novel algorithm calibration free skin composition estimation (CF-SCE) is then used to covert these hyperspatial images into a skin composition image which is then processed into robust physiological features like the sigmoid oxygenation image and the remote PPG signal. Our system operates across cameras and lighting conditions without requiring additional calibration/retraining (lower) [5]

limits their ability to do early detection. Emerging deep learning (DL) models integrated into mobile imaging systems offer automated classification of ulcers identifying issues like ischemia and infection from a single image taken with a smartphones camera. Despite their potential, these models face distrust (rightfully) for their "black box" nature, that is, they lack the explain ability of methods based on well understood principals which is very important when life-threatening conditions are in play as transparency is important for promoting patient autonomy and avoiding misdiagnosis as medical providers are untimely responsible for their practice [7].

Hyperspectral reconstruction (HSR) has recently emerged as a promising technology for tissue oxygenation and perfusion measurement. HSR transforms the limited and ill defined spectral bands of images captured by standard RGB cameras into the well defined images that a hyperspectral camera would produce, typically using an optimized linear transformation of some sort [8–12]. This offers a cost-effective, portable, and clear method for vitals imaging of high spatial resolution, enabling the estimation of tissue chromophore concentrations like hemoglobin (Hb) and oxyhemoglobin (HbO_2). Despite its promise, the practical implementation of HSR in mobile health applications faces a significant challenge. The first (intrinsic challenge) is the variability in smartphone imaging pipelines, which differ across devices in of lenses, sensors, illumination, and post-processing algorithms, especially color correction. These inconsistencies prohibit the use of HSR model parameters across different hardware configurations. Even if HSR is made to work in different configurations, there is still the extrinsic challenge that is the radiometric calibration process used to remove the effect of spatial unevenness and spectral variability in the lighting (such as from artificial lighting). Traditional HSI systems rely on per-use radiometric calibration with white and dark reference images [13], to standardize HSI pixels/images to a physically meaningful scale known as reflectance from where physical laws can be easily applied to infer the tissue composition. However, these processes are impractical for mobile health applications, where simplicity and ease of use are critical for widespread adoption.

In this paper, we introduce MobiPerf, a system that utilizes HSR to monitor blood perfusion, delivering physiological images and signals using only a standard smartphone while addressing both of the aforementioned challenges - the intrinsic and extrinsic challenges. Offering a cost-effective and portable solution for the monitoring of DFU vitals. Our goal is to empower patients to monitor their ulcers at home with an mobile app, receiving immediate insights into their condition. Looking ahead, we envision MobiPerf as a preventive health tool capable of detecting the conditions under which ulcers or blisters may form, potentially averting ulcers all together.

To address the challenges of implementing HSR in mobile health, we adopt a novel approach that overcomes the two key limitations. First, instead of relying on traditional HSR techniques such as Wiener filters which are inflexible, we employ large, and robust HSR models based on the MST++ architecture, which have been shown to work moderately well with unknown smartphone cameras[14]. This allows us to bypass the need for complex model retraining and ensures broader applicability across a range of devices solving the intrinsic challenge. Second, we introduce a brightness-tolerant method for determining the relative concentrations of hemoglobin (Hb) and oxyhemoglobin (HbO_2) in skin tissue. This method works when the hyperspectral image has not been normalized using reference images, bypassing the need for radiometric calibration, thus enhancing the usability of HSI/HSR technology for mobile health applications solving the extrinsic challenge. Using these concentrations, we derive a custom oxygenation feature and remote photoplethysmography (rPPG) signals that offers insights into both tissue oxygenation and cardiac conditions [15], which are common comorbidities in diabetic patients, thereby affording MobiPerf broader diagnostic capabilities.

We validate our approach by testing the derived custom oxygenation feature and rPPG signal with publicly available ground truth hyperspectral videos (N=10) [16] and a dataset of RGB videos with ground truth rPPG data (N=56) [17]. Finally, we assess MobiPerf's sensitivity to DFUs in a real-world dataset of Diabetic Foot and Leg Ulcer images ($N \approx 6,000$) [18], taken under varying lighting conditions and using different cameras — something not feasible with previous HSR approaches. We observe a significant distribution shift in the sigmoid oxygenation values between ulcers with dangerous ischemic complication and those without, demonstrating the potential of MobiPerf in assessing DFU severity.

Results

We evaluated the performance of our system using three publicly available datasets. The first dataset, captured with a hyperspectral camera, involves subjects whose blood flow and oxygenation levels were modulated using a blood-pressure cuff [16]. This dataset was used to assess the accuracy of our system's reconstructed oxygenation and pulse intensity images against ground-truth hyperspectral measurements. The second dataset, a widely recognized rPPG benchmark comprising 56 videos, includes ground-truth PPG signals recorded during various tasks [17]. This dataset allowed us to validate the performance of our extracted rPPG signals and benchmark it against state-of-the-art approaches. To evaluate the applicability of our system for DFUs management, we utilized a

third dataset consisting of nearly 6,000 diabetic foot ulcer images [18], annotated with labels for ischemia and infection. These labels were used to analyze the distribution of oxygenation values under different pathological conditions, demonstrating the system's biomarkers relations to the conditions.

Pressure cuff study

To validate the oxygenation imaging capabilities of our system, we utilized the SPECTRALPACA dataset [16], which contains 10 videos recorded under controlled conditions. Blood flow to participants' hands was modulated using a pressure cuff applied to the arm, with pressures ranging from 0 mmHg to 160 mmHg. Adequate intervals were provided between pressure changes to allow blood flow and oxygenation levels to return to baseline. A neutral-spectrum xenon Light source illuminated the hand, ensuring uniform spectral conditions, and the recordings were free of stray illumination. The videos were captured using a 16-band hyperspectral camera with narrowband channels where each channels peak sensitivity Lies within the range 465 nm to 641 nm.

Our validation process starts with the preprocessing of ground truth hyperspectral data. The raw hyperspectral data were Transformed from the 16 original bands into 31 evenly spaced non-overlapping bands between 460 nm and 640 nm. This was achieved by calculating weighted sums of contributions from each raw spectral channel using the provided camera filter sensitivity vectors. Each output band was adjusted by dividing by the sum of the corresponding weights to account for variations in camera sensitivity. The same procedure was applied to the white calibration frame included in the dataset. The unnormalized 31-band images were then spectrally and spatially normalized by dividing them by the 31-band white reference image, producing reflectance images. Then, the reflectance images were also converted into three-channel RGB images by applying a similar weighted sum of spectral contributions procedure using the camera filter sensitivity function of a Samsung Galaxy smartphone [19], rescaled to align with the 460–640 nm range. The resultant RGB images were then max-normalized on a frame-by-frame basis to ensure consistency. For HSR reconstruction, the RGB images were processed using a retrained MST++ model (MST_{640} in Tables 2 and 3) in accordance with Deep learning hyperspectral reconstruction section. The training scenario for the model matchs the preprocessing done for both ground truth HSI and RGB images in this section. The ground truth HSI images and reconstructed HSR images were then run through the procedure in Calibration-free skin composition estimation section to extract the skin composition images in Fig. 2.

Both the reconstructed and ground truth skin composition videos were then processed as outlined in Sigmoid oxygenation section to obtain a video of sigmoid oxygenation images for each. The videos were temporarily down-sampled 90x in order to speed up all the previous steps. The performance of MobiPerf was understood to be strong if there was high agreement with the ground truth processed hyperspectral video. The agreement was measured both temporally and spatially. Pre-specified small ROI's were extracted from each frame in both videos and the mean of those ROI's frame-wise were used to construct a oxy signal, a moving average was applied to smoothen signals.

To determine the temporal agreement between the ground truth and reconstructed oxygenation signal we used common signal metrics and corresponding visualizations. Both signals were rescaled between 0 and 1 by performing a z-standardization $z = (x - \mu)/(4\sigma) + 1/2$. Then the Mean Absolute Error was taken between two signals for each signal pair in the dataset, Pearson's R was then also taken between the signals. Correspondingly we plotted the scatter plot between the two signals and a line of best fit for each subject. The summarized metrics, selected signal plots, and scatter plots are shown in Fig. 3. Metrics clearly show strong agreement between ground truth and reconstructed oxy signals with an overall average signal MAE of 0.122 (90% CI: 0.103, 0.141) and corresponding average Pearson's R of 0.865 (90% CI: 0.779, 0.919) MAE and R metrics computed signal-wise then averaged across all subjects/ROIs, moderate between-ROI variance in performance is apparent in the metrics as well. This agreement is also reflected visually in the scatter plots where a moderate-strong Linear fit can be clearly established in all but 1 subject.

To quantify the spatial agreement, the first oxy image from each subject video was taken and image metrics were applied between the ground truth and reconstructed ones. Images were segmented using the Segment Anything Model (SAM) [20]. The segments are z-standardized in the same way as the oxy signals (mean/std calculation ignored the background). We then crop the images tightly around the hands and zero-out the background to allow for PSNR and SSIM to be computed between images as well. To visualize the spatial agreement we selected a subject from the dataset and plotted the image pairs for each applied pressure level, we also plotted corresponding ridge plots with the distribution of oxy within each segment for both the ground truth and reconstructed oxy image series. All these metrics and visualizations are shown in Fig. 2. The metrics show a strong association between the image pairs with mean SSIM of 0.762 (90% CI: 0.741, 0.783) and PSNR/20 of 0.505 (90% CI: 0.493, 0.517) spatial performance was

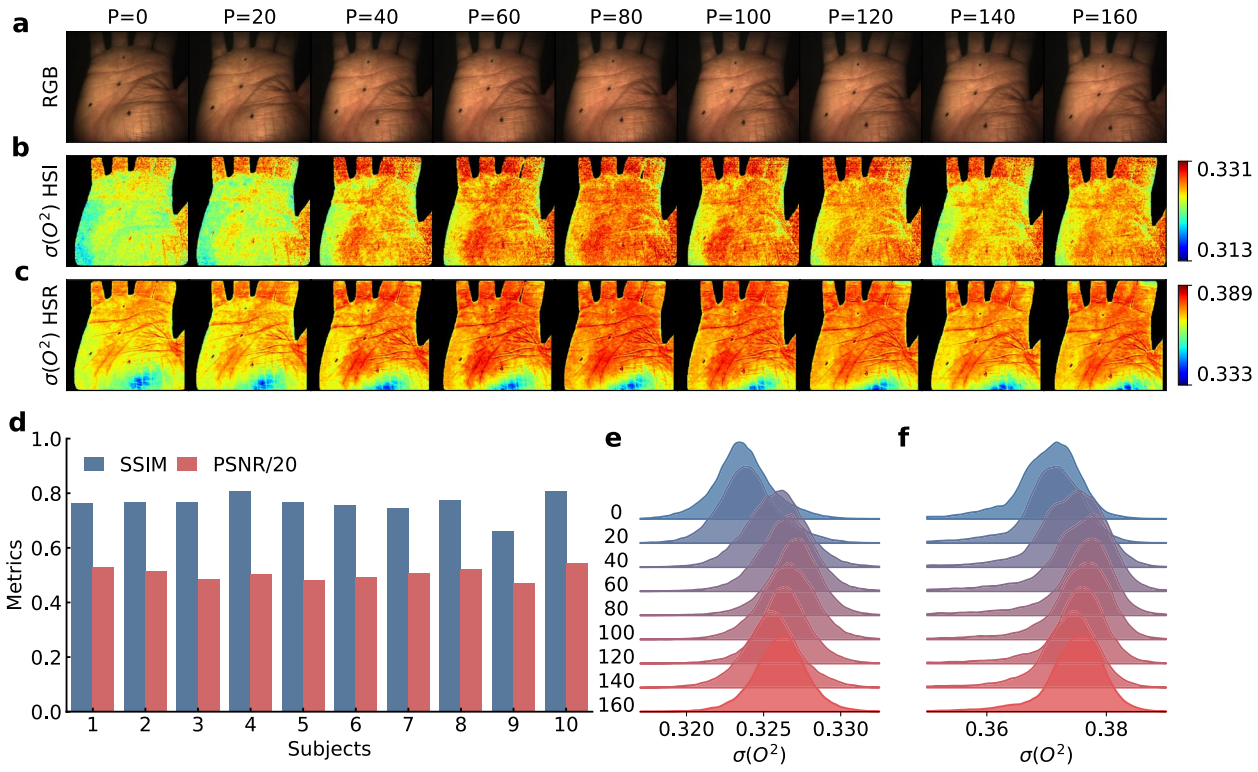

Fig. 2 Representatives sigmoid oxygenation image sequences from HSI, HSR and image metrics testing the agreement of HSI/HSR for first frames. **a** RGB images under various pressures generated by taking channels [11, 15, 3] from corresponding HSI 16ch raw images and enhancing contrast using CLAHE. **b** Sequence of oxygenation images computed from HSI images using MobiPerf (w/o reconstruction) taken under increasing pressures. Stress can be seen as pressure increases centered on the palms diagonal (/) and gradually expanding outwards. **c** Corresponding sequence of oxygenation images computed using MobiPerf from RGB smartphone images (generated from HSI images). Stress can again be seen centered on the palms diagonal (/) and gradually expanding outwards. **d** Metrics computed by comparing the sigmoid oxygenation images generated using MobiPerf from HSI and HSR images for the first image in each subject's video. Agreement can be seen to be strong with a mean SSIM 0.762 (90% CI: 0.741, 0.783) and mean PSNR/20 approx 0.505 (90% CI: 0.493, 0.517) dB. **e**, **f** Ridge plots generated for the sigmoid oxygenation image sequences from HSI and HSR respectively by performing kernel ridge regression for all pixels within the tissue segments. Strong agreement can again be seen visually

highly consistent as well with less inter-subject variability then the temporal metrics. Visually the two image series showed high agreement when viewed across different pressure levels (including in the ridge plot) with the center diagonal of the palm gradually becoming more stressed first followed by its surrounding areas in both sets, a artifact can be seen around the lower palm on the reconstructed images likely due to specular reflections.

Between the spatial and temporal analysis we can see that (in the case of ideal lighting) there is a strong agreement between the ground truth sigmoid oxygenation features and the reconstructed ones both visually and quantitatively.

Lighting variability

The influence of lighting can be regarded as either spatial (variance in intensity across the spatial dimensions of the image) or spectral (variance of intensity across bands/wavelengths) or more complicated than that if there are multiple light sources in the scene. In this section we concern ourselfs with the performance of our system under different lighting spectra. We take a reflectance image as generated in the previous section and element-wise multiply each pixel by one of four lighting spectral efficiency functions from the CIE standard illuminates [21] and one from a smartphone flash [22] sampled at 31 bands (wavelengths $[400, 700]_{31}$ for reasons described in Deep learning hyperspectral reconstruction section) to apply the illumination's spectra to the image.

We then apply the same procedure from the previous section to generate corresponding RGB images and apply MobiPerf to those images ultimately generating the same pairs of signals/image sequences for each different illuminate under consideration as in the previous section. We then plotted the image sequence and computed signal metrics shown shown in Fig. 4 as well as plotted the lighting spectra for each tested illuminate. The performance is generally fairly consistent across different light sources with the image sequences exhibiting similar behavior under increasing pressures aside from the Incandescent sequence (A) which seams to have a softer texture. The Pearson's R metric reveals a moderate performance

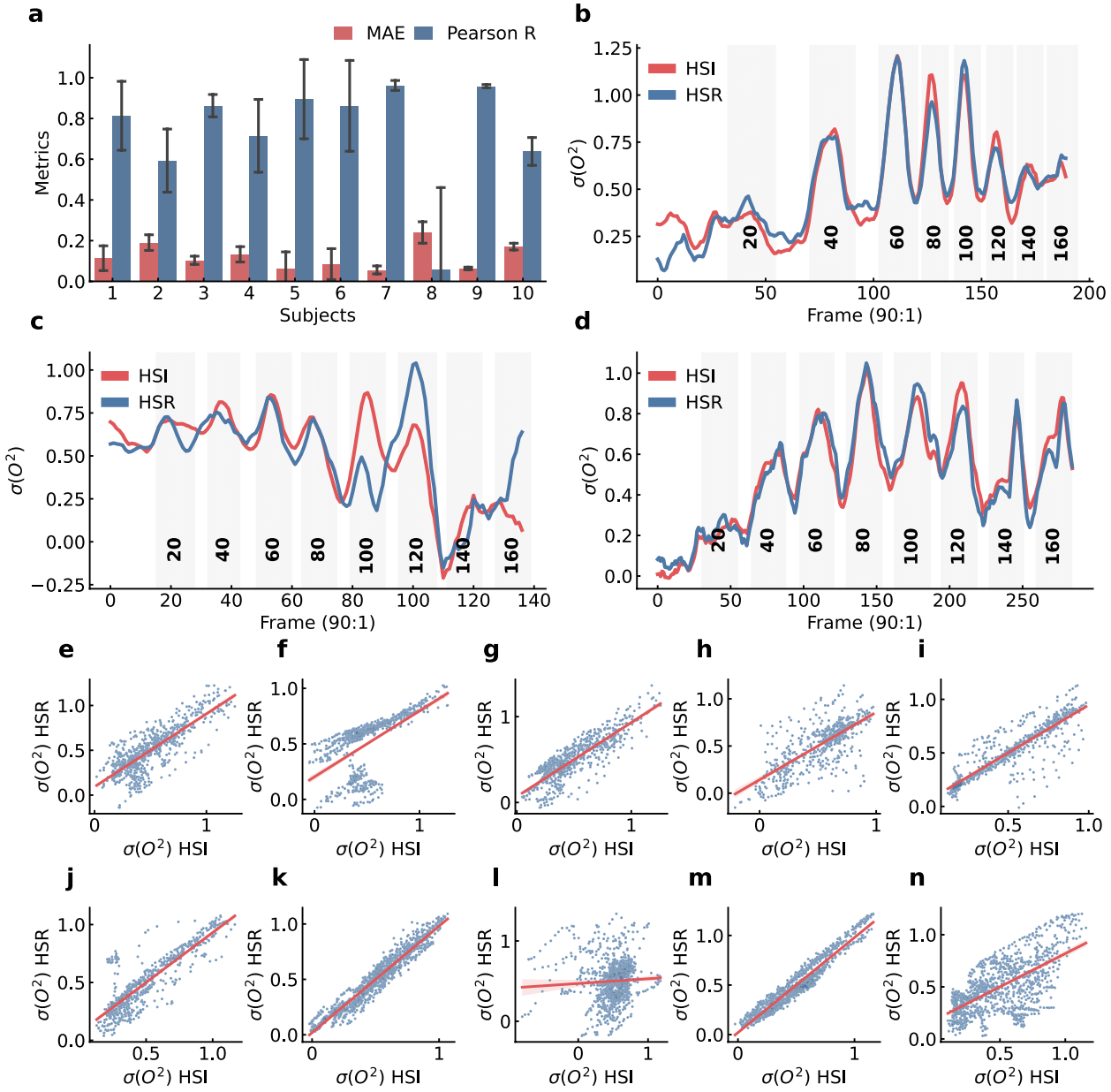

Fig. 3 Metrics from comparing oxygenation signals from HSI/HSR, representative signal pairs, scatter plots between signals from HSI/HSR. **a** Metrics computed by applying MobiPerf to HSI videos and to corresponding simulated smartphone RGB videos then extracting sigmoid oxygenation signals from 4 ROIs finally computing metrics between the extracted signal pairs. Subject-wise mean performance averages 0.865 (90% CI: 0.779, 0.919) for Pearson's R and 0.122 (90% CI: 0.103, 0.141) for MAE for most subjects. **b**, **c**, **d** Sigmoid oxygenation signal pairs (HSI/HSR) plotted together for 3 representative subjects/ROIs. Strong agreement can be seen in (**b**, **d**) whereas (**c**) demonstrates slightly weaker agreement. **e**-**n** Subject-wise scatter plots generated by pairing sigmoid oxygenation signals generated for all ROIs where (x=from HSI, y=from HSR). Agreement varies with (**l**) showing the weakest agreement by far (subject 8)

shortfall for the Flash 0.633 (90% CI: 0.442, 0.768), Florescent 2 0.696 (90% CI: 0.529, 0.811), and Incandescent Light 0.631 (90% CI: 0.441, 0.767) sources as well as a large increase in the variability of Pearson's R (across ROI/subject) in those cases as well. Overall we believe that the systems performance is consistent enough under different light sources to be usable and possibly explain to a degree the results of in-the-wild testing as in Remote photoplethysmography and Diabetic foot ulcers sections.

Reflectance reconstruction

Looking to address the performance shortfall found in Lighting variability section we apply the hyperspectral reconstruction models from Deep learning reflectance reconstruction section trained under different

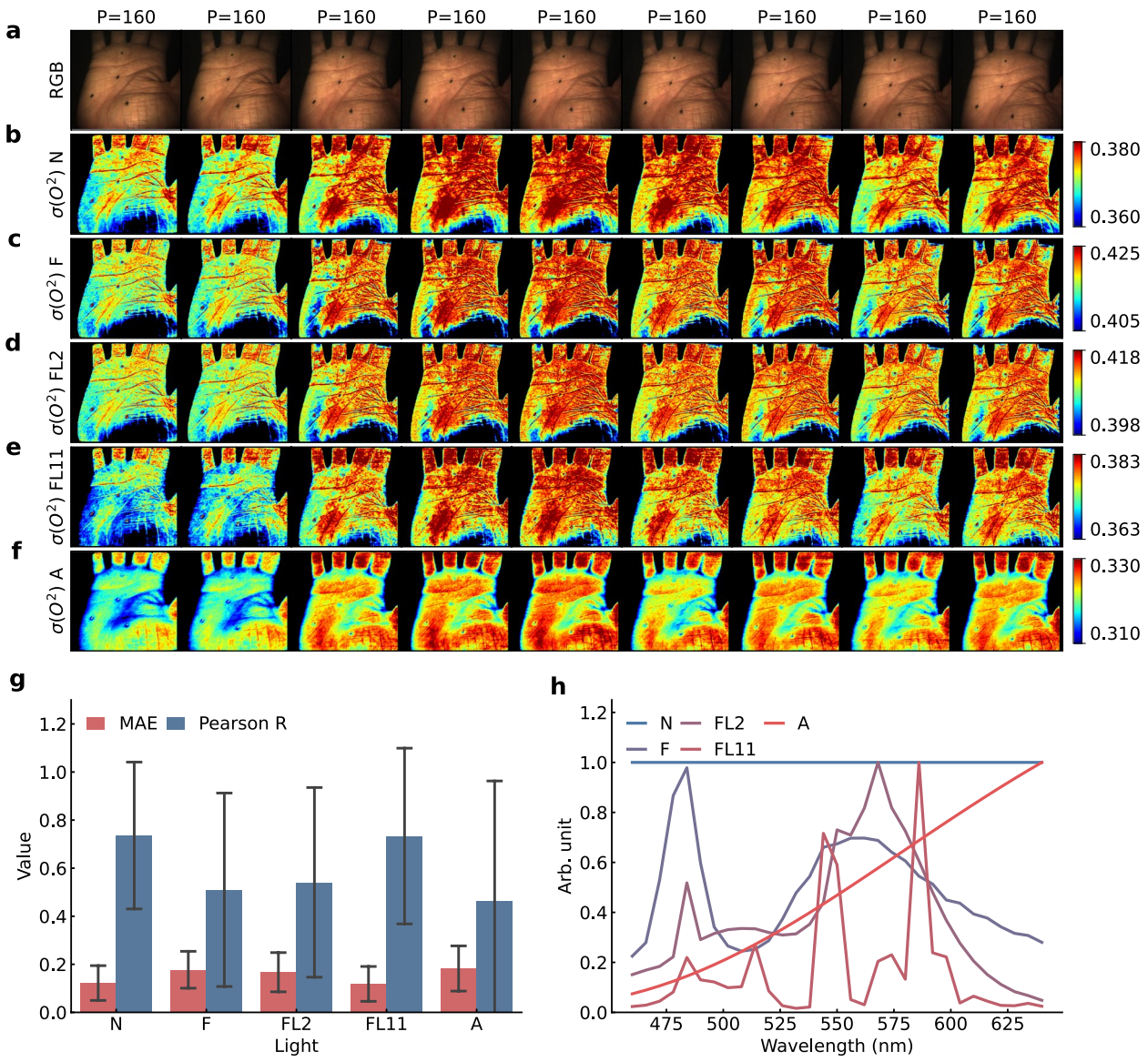

Fig. 4 Sigmoid oxygenation image sequences from MobiPerf (from HSR) and metrics under various illumination spectra. **a** RGB images under various pressures generated by taking channels [11, 15, 3] from corresponding HSI 16ch raw images and enhancing contrast using CLAHE. **b**-**f** Sigmoid Oxygenation image sequences computed by applying MobiPerf to smartphone RGB images (generated from HSI images) taken under increasing pressure levels and various lighting spectra. Images sequences taken under non-neutral lighting [F, FL2, FL11, A] can be seen to be visually similar to those taken under neutral [N] lighting. **g** Metrics computed by applying MobiPerf to HSI videos and to corresponding smartphone RGB videos (generated from HSI images) and comparing extracted sigmoid oxygenation signals from predefined ROIs for each lighting condition. Moderate performance degradation can be seen for some lighting conditions [F, FL2, A]. **h** Lighting Spectral Efficiency Functions for each tested lighting condition

objectives in order to see which one is the most consistent/performant focusing our attention on the case of reflectance reconstruction ($\mathrm{RGB_L} \rightarrow \mathrm{HSI_N}$) where both hyperspectral reconstruction and radiometric calibration are done using deep learning looking to see if it provides a performance gain over just hyperspectral reconstruction ($\mathrm{RGB_L} \rightarrow \mathrm{HSI_L}$). In order to do so we run the same procedure as in Lighting variability section replacing the MST_{640} model with the three models $\mathrm{RGB_L} \rightarrow \mathrm{HSI_L}$, $\mathrm{RGB_N} \rightarrow \mathrm{HSI_N}$, $\mathrm{RGB_L} \rightarrow \mathrm{HSI_N}$ in turn. The signal results were aggregated (with the results for MST_{640}) and are shown in Fig. 5.

It can be clearly seen that the two objectives that reconstruct reflectance images ($\mathrm{RGB_N} \rightarrow \mathrm{HSI_N}$, $\mathrm{RGB_L} \rightarrow \mathrm{HSI_N}$) regardless if the RGB image was expected to be taken/generated under neutral lighting conditions or not surprisingly have much better performance then traditional hyperspectral reconstruction ($\mathrm{RGB_L} \rightarrow \mathrm{HSI_L}$) with performance generally having less or comparable variation within and across illumination types. It can also be seen that $\mathrm{RGB_L} \rightarrow \mathrm{HSI_L}$ and

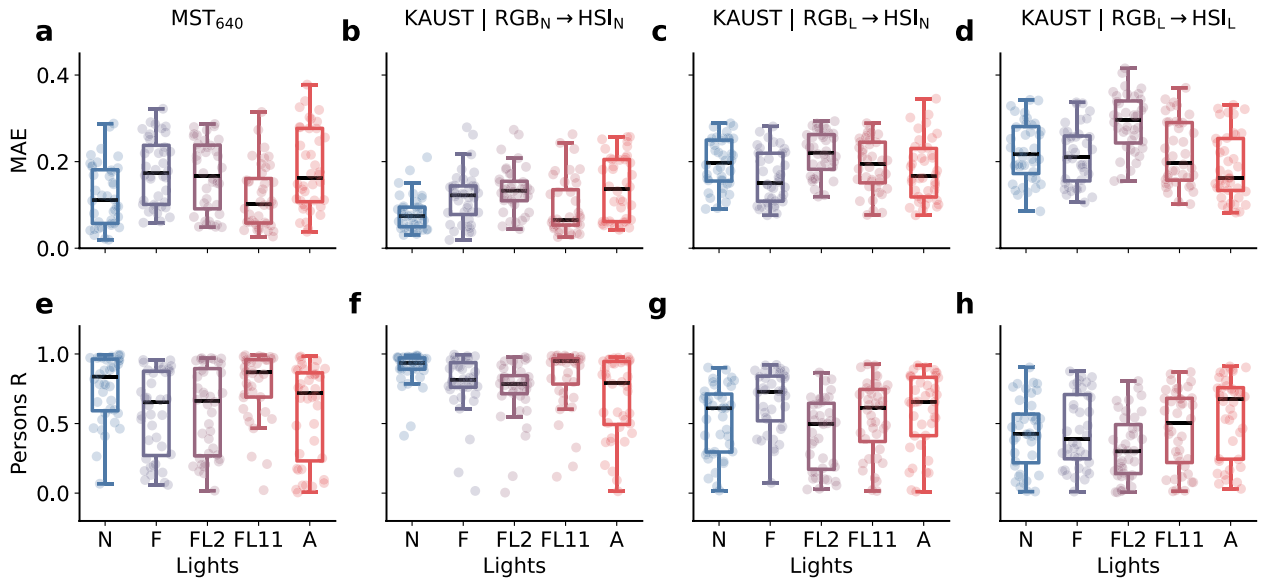

Fig. 5 Performance of various hyperspectral reconstruction models under various lighting spectra. **a**-**h** Metrics computed by applying MobiPerf to HSI videos and to corresponding smartphone RGB videos (generated from HSI images) and comparing extracted sigmoid oxygenation signals from predefined ROIs for each lighting condition. **a**, **b** metrics for the traditional hyperspectral reconstruction model used in majority of the testing included for reference only, (**b**, **f**) metrics for model that maps neutrally illuminated RGB images to neutrally illuminated hyperspectral images, (**c**, **g**) metrics for model that maps non-neutrally illuminated RGB image to neutral hyperspectral image (reflectance reconstruction), (**d**, **h**) model that maps non-neutrally illuminated RGB images to non-neutrally illuminated hyperspectral images (traditional hyperspectral reconstruction). Models that map to neutrally illuminated hyperspectral images ($\mathrm{RGB_N} \rightarrow \mathrm{HSI_N}$, $\mathrm{RGB_L} \rightarrow \mathrm{HSI_N}$) tend to have more consistent and higher performance than traditional hyperspectral reconstruction ($\mathrm{RGB_L} \rightarrow \mathrm{HSI_L}$)

MST_{640} have similar variation in performance especially within illumination types this can probably be attributed to both of these models being trained with similar objectives (see Table 3) although a direct comparison between these two models cant be made as they are not otherwise having the same parameters.

Remote photoplethysmography

To evaluate the remote photoplethysmography (rPPG) capability of MobiPerf, we utilized the UBFC-Phys dataset [17], which comprises recordings of 56 healthy participants engaged in three distinct tasks: resting (T1), speaking (T2), and performing arithmetic calculations (T3). For each participant, synchronized video data and reference PPG signals were collected, along with electrodermal activity and self-reported stress levels. The dataset was captured under unspecified artificial lighting conditions (spatially even, spectrally unknown), with videos compressed using JPEG and recorded at a resolution of 1024x1024 pixels. Participants were positioned approximately one meter from the camera during data collection, ensuring consistent imaging conditions. This setup provided a robust benchmark for validating the performance of MobiPerf's rPPG signal extraction under real-world conditions.

To preprocess the dataset, the reference PPG signals were first aligned to the sampling rate of the corresponding video frames using polyphase filtering. These signals were then filtered with a first-order Butterworth filter to suppress high-frequency noise and standardized by $z = (x - \mu)/(2\sigma)$, ensuring consistency in amplitude across recordings. Facial regions of interest (ROIs) were identified for each video frame using MediaPipe [23], which mapped predefined 3D facial landmarks to determine 2D polygonal segments corresponding to the forehead and the left and right cheeks. Each video frame, along with its associated ROI, was subsequently processed using the methodology described in Remote PPG section to extract the MobiPerf Hb and MobiPerf Motion-Compensated rPPG signals. This preprocessing pipeline ensured robust signal extraction and effective motion compensation throughout the dataset.

The video frames and corresponding ROI pairs were also utilized as inputs for other state-of-the-art algorithms to construct rPPG signals, each employing a distinct approach after deriving the channel-wise mean signal from the identified ROI. The Green algorithm [24] employed the simplest technique by directly using the mean signal from the green (G) channel of the ROI as the rPPG signal. In contrast, the Independent Component Analysis (ICA) method [25] applied independent component analysis to the 3-channel RGB mean signal extracted from the ROI, selecting the second independent component as the rPPG signal. The Plane Orthogonal to Skin (POS) algorithm [26], a method tailored for rPPG, adopted an analytic approach designed to leverage

the orthogonality of skin-tone variations for robust signal extraction. A standardized post-processing pipeline was applied to all generated signals. In instances where MediaPipe failed to detect an ROI—due to non-convergence or the face being out of frame—a forward-fill interpolation followed by backward-fill interpolation was applied to impute missing values in the ROI mean signals. Signals from each algorithm were passed through a 9th-order Butterworth bandpass filter with a frequency range of 0.7 Hz to 3.5 Hz to isolate the relevant cardiac frequency components. Subsequently, the signals were then standardized by $z = (x - \mu)/(2\sigma)$. This ensured data continuity and allowed consistent rPPG signal construction across all algorithms.

Following this, all signals, including the ground-truth PPG signal, were used for peak detection, identifying peaks with a prominence threshold of 0.5. The intervals between successive peaks, measured in elapsed samples, were calculated and converted to a pulse rate variability (PRV) signal by dividing by the sampling rate. The PRV signal was further transformed into a heart rate variability (HRV) signal using the relation $HRV = 60/PRV$. To refine the HRV signal, median filtering with a window size of 51 was applied, and the mean value of the resulting signal was computed to derive the final heart rate (HR) estimate. This reflects a typical methodology for extracting the Heart Rate in a robust manner.

We employed three evaluation metrics to assess the quality of the extracted rPPG signals. The Mean Absolute Error (MAE) of heart rate (HR) estimation, expressed in beats per minute $|\mathrm{BPM}|$, is a key metric for evaluating the signal's suitability for downstream applications such as pulse transit time analysis, which is commonly derived from PPG signals. For the overall signal quality, we used two complementary metrics: Pearson's correlation coefficient (R) and dynamic time warping (DTW) with a Euclidean distance metric. Pearson's R was calculated over short, non-overlapping windows, and the mean R-value across all windows was used as a measure of the signal's correlation quality. To account for possible phase shifts between signals, a time-delay adjustment was applied by shifting one signal by n samples ($n < F_s$, where F_s represents the sampling rate) and selecting the maximum correlation observed over the range of shifts. DTW is widely recognized for its robustness in comparing time series signals, particularly when dealing with variations in signal timing or phase. Previous studies have demonstrated its reliability in assessing the quality of rPPG signals [27]. This combination of metrics provides a comprehensive evaluation of signal integrity and its potential for physiological monitoring applications.

The results of the three evaluation metrics are presented in Fig. 6. For all algorithms performance is reduced greatly when motion is introduced as in tasks T2 and T3. Most notably for task 1 our motion compensated signal far outperformed other algorithms in BPM error 5.139 (90% CI: 3.336, 6.941) with better error skew/variance of 12.356 (90% CI: 10.270, 14.442) for task 2 and comparable performance 11.457 (90% CI: 9.197, 13.718) to other algorithms in task 3. Again in Pearson's R our motion compensated mobiperf signal far outperformed other algorithms in task 1 at 0.335 (90% CI: 0.285, 0.386) while maintaining comparable (low) performance in the other tasks. DTW error was fairly noninformative with all algorithms (aside for POS) being comparable to each other in most cases. Overall, the MobiPerf signals either significantly outperform competing approaches or achieve comparable performance, with notably lower error variation and improved error distribution skew in many cases. As with the other algorithms shown here some motion is clearly tolerable but too much motion clearly makes the signals not usable for BPM estimation and other downstream tasks.

Diabetic foot ulcers

We evaluated the applicability of our system to real-world diabetic foot ulcer images using the DFUC2021 dataset [18, 28–31], which comprises over 15,000 images of foot and leg ulcers. For our analysis, we utilized the provided training set containing approximately 6,000 images, each labeled with ulcer pathology. The labels specify whether an ulcer exhibits ischemia, infection, both conditions, or neither. The images were captured using one of three different cameras under varying, unspecified room lighting conditions, without the use of flash. While most images are close-ups of the ulcers, some include background elements, as observed upon inspection. These inherent variations in imaging conditions provide a realistic and challenging testbed for our system, enabling us to evaluate its robustness and performance under practical, real-world scenarios.

To evaluate the power of our sigmoid oxygenation feature for identifying foot ulcer images with different pathologies, each image in the training set was processed following the methodology outlined in Sigmoid oxygenation section to generate its corresponding sigmoid oxygenation map. To derive a comparable metric across all images, we computed the mean sigmoid oxygenation value over the entire image, resulting in a single representative oxygenation value for each ulcer image.

Figure 7 presents the distributions of the sigmoid oxygenation feature for different foot ulcer conditions—ischemia, infection, both (ischemia and infection), and none—using violin plots. The x-axis represents the value of the oxygenation feature $\sigma(O^2)$, while the y-axis corresponds to the ulcer conditions. Notably, Very clearly ischemia ulcers mean oxy value skews right heavily and less apparently infection ulcers skews right moderately.

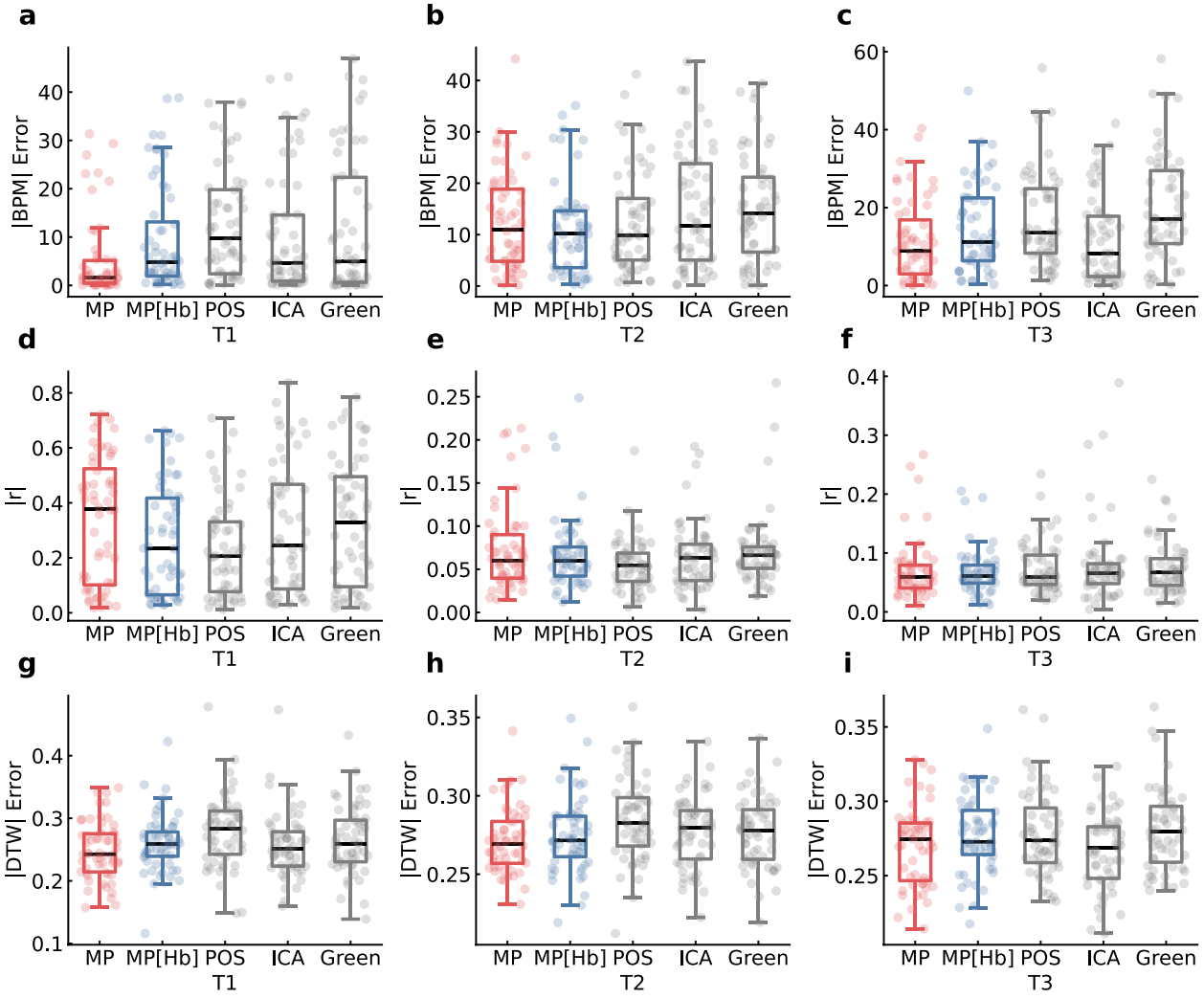

Fig. 6 Performance of MobiPerf rPPG signals vs signals generated by other common algorithms. **a**-**c**$\|\mathrm{BPM}\|$ error of rPPG signals extracted using various algorithms from the UBFC-Phys dataset on tasks T1 (little motion), T2 (some motion) and T3 (significant motion) respectively. MobiPerf motion-robust (MP) and MobiPerf hemoglobin-only (MP[Hb]) can be seen in red and blue with stronger or comparable performance to other algorithms in all three tasks. **d**-**f** Absolute Pearson's R again comparable or better performance can be seen from MP and MB[Hb]. **g**-**i** Dynamic Time Warping Error, comparable performance can be seen across the board

Empirically the distribution for both conditions looks Like the superposition of the distributions for ischemia and infection with twin peaks. On the other hand ulcers with no conditions demonstrated no skew or superposition behavior. We take it that this indicates that the sigmoid oxy values is related to the condition of the ulcer to verify this we employ 2-way ANOVA with the results shown in Table 1. As $p \ll 0.0001$ for both conditions we conclude that there is a statistically significant effect of condition on sigmoid oxy values obtained from our system.

Figure 7 also shows a few examples of processed images from our system on the DFCU2020 dataset [31] these images show the entire foot and clearly illustrate that the oxy images extract something unseen by the eye and related to ulcers, most notably the impacted oxy in the skin tissue around the ulcers.

Methodology

The goal of MobiPerf is to extract a comprehensive set of static and dynamic physiological features relevant to monitoring blood perfusion, using only ordinary images captured with off-the-shelf smartphones. As illustrated in the Fig. 1, the pipeline begins with the hyperspectral reconstruction phase, where an RGB image is captured using a standard smartphone. Using Deep Learning Hyperspectral Reconstruction (DL-HSR), the RGB image is converted into a 31-band hyperspectral image, closely resembling outputs from dedicated HSI cameras. These reconstructed hyperspectral images are then processed to estimate tissue composition through our innovative

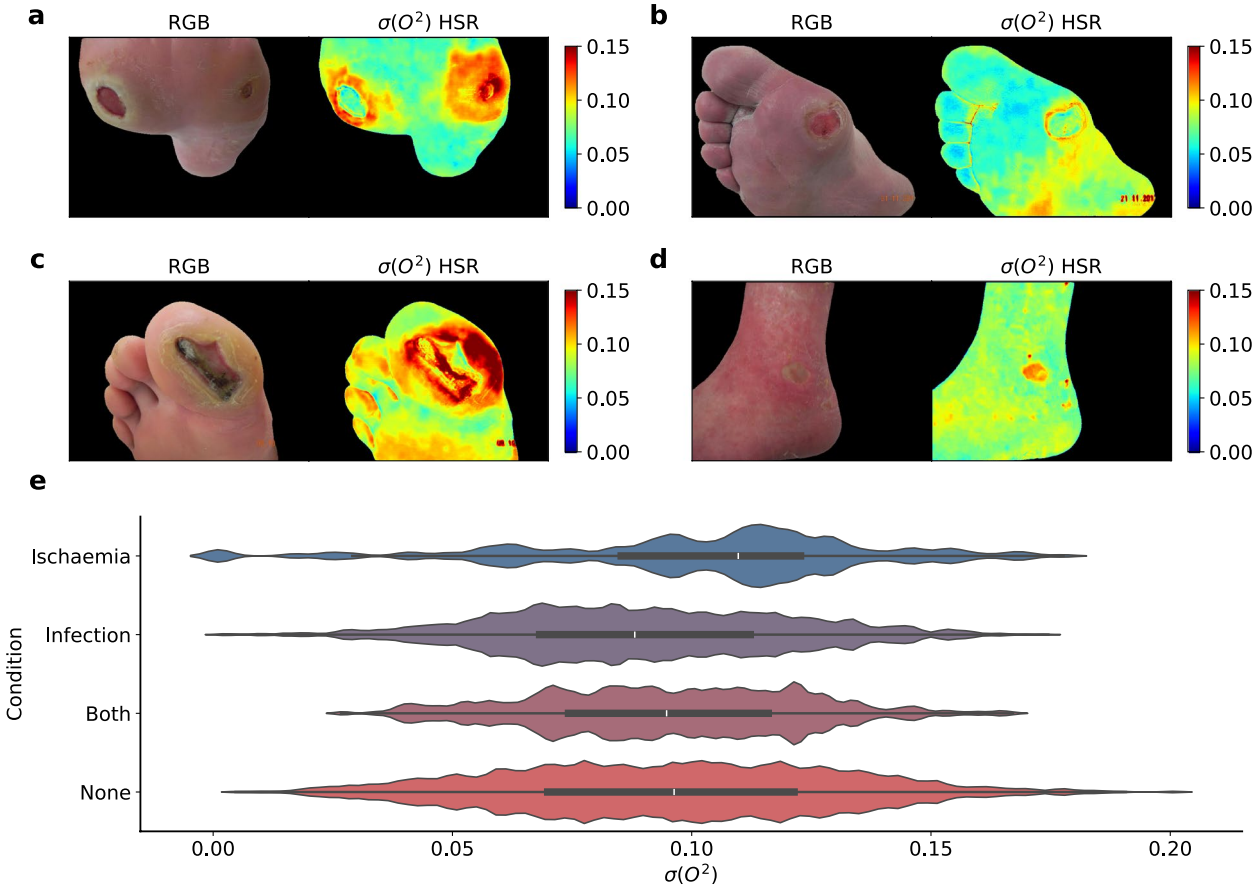

Fig. 7 Representative oxygenation images generated using MobiPerf from the DFCU2020 dataset, and per condition distributions of mean oxygenation over the DFCU2021 dataset. **a**-**d** representative sigmoid oxygenation and source RGB image pairs generated by MobiPerf from the DFCU2020 dataset around and/or within the ulcer higher values can be seen indicating poorer oxygenation something which can clearly not be seen from the RGB images. **e** violin plots of per condition distributions generated by taking the mean sigmoid oxygenation of each image in the DFCU2021 dataset. A large skew to the right can be seen for ulcers with ischemia, a small skew to the left can be seen for ulcers with infection. The distribution of images with Both conditions can be seen as a superposition of the individual-condition distributions with peaks that line up on left/right, whereas the distribution of images with no condition can be seen to have no discernible peaks or skew

Table 1 ANOVA table

Source	Sum Sq	df	F	PR(> F)
C(Ischaemia)	0.021260	1.0	20.301810	6.74×10^{-6}
C(Infection)	0.049680	1.0	47.440646	6.25×10^{-12}
Residual	6.232995	5952.0		

Calibration-free Skin Composition Estimation (CF-SCE) algorithm. In the subsequent Feature Derivation phase, the system extracts physiological information from the tissue composition data. A Sigmoid Oxygenation Image is generated to visualize tissue oxygenation levels. Additionally, by capturing a series of RGB images over time and processing them through the pipeline, the system derives a remote photoplethysmography (rPPG) signal for each frame by identifying and averaging a region of interest (ROI) than applying signal processing techniques, including alpha-tuning and frequency-domain filtering. This dynamic signal offers valuable insights into cardiovascular health, addressing a significant comorbidity commonly linked to diabetes.

Deep learning hyperspectral reconstruction

In this work, hyperspectral images are reconstructed using deep learning, leveraging the MST++ model [32], a state-of-the-art transformer-based model with a U-shaped encoder-decoder architecture. We train a selection of models with parameters summarized in Tables 2 and 3.

31 bands were used when sampling camera spectral response function (CSRFs), lighting spectral efficiency functions (LSEFs), and resampling HSI images (when

Table 2 Model training parameters general

Model	Dataset	RGB CSRF	Noise Injection
MST_{700}	NTIRE	DLSR	Yes
MST_{640}	NTIRE	Smartphone	Yes
$\mathrm{RGB_L} \rightarrow \mathrm{HSI_L}$	KAUST	Smartphone	No
$\mathrm{RGB_N} \rightarrow \mathrm{HSI_N}$	KAUST	Smartphone	No
$\mathrm{RGB_L} \rightarrow \mathrm{HSI_N}$	KAUST	Smartphone	No

Table 3 Model training parameters lighting and bands

Model	HSI CSRF Bands	RGB CSRF Bands	HSI lighting	RGB lighting
MST_{700}	$[400, 700]_{31}$	$[400, 700]_{31}$	Environment	Environment
MST_{640}	$[460, 640]_{31}$	$[400, 700]_{31}$	Environment	Environment
$\mathrm{RGB_L} \rightarrow \mathrm{HSI_L}$	$[460, 640]_{31}$	$[400, 700]_{31}$	Augmented	Augmented
$\mathrm{RGB_N} \rightarrow \mathrm{HSI_N}$	$[460, 640]_{31}$	$[400, 700]_{31}$	Neutral	Neutral
$\mathrm{RGB_L} \rightarrow \mathrm{HSI_N}$	$[460, 640]_{31}$	$[400, 700]_{31}$	Augmented	Neutral

applicable). Much of our simulated testing used HSI images resampled to $[460, 640]_{31}$ from $[400, 700]_{31}$ for the purpose outlined in Pressure cuff study section, so for models involved in that testing we also resample the training images to $[460, 640]_{31}$ these resampled images are also used to generate the corresponding training RGB images. All RGB CSRFs and LSEFs were sampled however at $[400, 700]_{31}$ so that any limitations associated with smartphone cameras are accurately represented to a extent.

Traditional deep learning hyperspectral reconstruction

For the majority of our testing both simulated and in-the-wild we use two models trained under similar parameters from the same data. The MST_{700} model which was used for testing on in-the-wild data like in Diabetic foot ulcers and Remote photoplethysmography sections and the MST_{640} was used for the testing in Pressure cuff study and Lighting variability sections. These models are both trained using the NITRE 2022 dataset [33] which is originally used in the MST++ paper and contains HSI images captured in indoor/outdoor environments. RGB images for training pairs were generated using a CSRF from a RGB camera and the corresponding HSI images to perform a weighted sum of spectral bands as in Pressure cuff study section. Shot noise was injected into the RGB images in order to make the generated images more realistic. These setups closely match up with the original setup in the MST++ paper.

Deep learning reflectance reconstruction

The influence of lighting is a concern for many systems dependent on hyperspectral imaging these systems typically operate using a specific type of hyperspectral image known as a reflectance image which is a image where both spatially and spectrally the lighting is completely neutral such that for each band/location in the image 1 indicates complete reflectivity and 0 indicates no reflectivity. The process of converting a hyperspectral image to a reflectance image is known as radiometric calibration. Here we attempt to train a model which essentially performs both hyperspectral reconstruction on a RGB image and then substantially also performs radiometric calibration on that image. We focus our attention on the limited case of one light source and spatially even lighting. To the best of our knowledge we are the first and only paper to apply reflectance reconstruction to whole images using deep learning. We believe this may be one of the first applications of whole-image reflectance reconstruction to a biomedical application.

We train 3 models under different objectives in order to test reflectance reconstruction ageist the task in Pressure cuff study section. In all three cases a RGB image is being mapped ($\rightarrow$) to a HSI one with the subscript N indicating a reflectance image (or RGB image generated from a reflectance image) and the subscript L indicating that a CIE Standard Illuminate has been applied. $\mathrm{RGB_L} \rightarrow \mathrm{HSI_L}$ can be said to preserve the effect of the illuminate in the RGB to the reconstructed HSI, $\mathrm{RGB_N} \rightarrow \mathrm{HSI_N}$ is said to map RGB images generated from reflectance ones back to reflectance images, and $\mathrm{RGB_L} \rightarrow \mathrm{HSI_N}$ performs reflectance reconstruction removing the effect of the illuminate. All three models are trained without noise injection and use the same bands as MST_{640}. The models were trained using the KAUST dataset [34] a dataset of reflectance images.

Calibration-free skin composition estimation

Hemoglobin and Oxyhemoglobin are known as chromophores meaning that within the skin along with melanin they are responsible for the overall color of the tissue. At each wavelength every chromophore absorbs a percentage of the light incident to it and reflects the rest in accordance to its absorption spectra. A very typical method for estimating the amount or concentration of each of these materials is matrix inversion which can be applied when

there are more camera channels then chromophores in the skin as this results in a matrix of sufficient rank (invertible). In this method a linear mapping (matrix) is constructed which combines both prior known and complete information about the camera sensitivity to different bands of light and the materials absorption of bands of light via a common physical law known as the Beer-Lambert law obtained by prior characterization of both. This method requires a known lighting source spectra as well in order to derive the matrix transformation as the lighting can be thought of as a modulation of either the camera or material spectra. We refer to this as the extrinsic challenge and set out to solve it in order to allow for different unknown lighting sources to be used with this common and well understood method.

In our Calibration-Free Skin Composition Estimation (CF-SCE) algorithm, we apply an approximation to the typical method for estimating skin composition from a hyperspectral pixel in order to avoid needing to do radiometric calibration (solving the extrinsic challenge). According to the Beer-Lambert law, the relationship between the transmitted or reflected spectra, $T(\lambda)$, the absorbed spectra, $A(\lambda)$, and the concentrations of the substances in the sample as follows.

$$A(\lambda) = -\ln T(\lambda) = \varepsilon(\lambda) \cdot c \cdot l. \tag{1}$$

Where λ denotes the wavelength of light, $\varepsilon(\lambda)$ is molar absorptivity (or extinction coefficient) at λ, c is the concentration of the absorbing substance and l is the path length.

We note Here that the reflectance, by definition, is bounded between 0 and 1, and reflectance from the skin is generally not expected to fall below 0.05 (5%). Based on this, we propose our first approximation: the natural logarithm function, $\ln(x)$, can be approximated by a Linear function within the range of approximately 0.05 to 1.

$$A(\lambda) \approx -(a \cdot T(\lambda) + b). \tag{2}$$

Additionally, we assume that the skin primarily contains only the three chromophores of interest: hemoglobin, oxyhemogloban, and melanin (specifically eumelanin). And we simplify the analysis by assuming an optical path length of $l = 1$ (previous works have disregarded the path length as well [10]). Using the molar extinction coefficients [35, 36] for each chromophore ($\epsilon_{\mathrm{Hb}}, \epsilon_{\mathrm{HbO2}}, \epsilon_{\mathrm{M}}$), the reflected spectra can be expressed as a linear combination of the material concentrations as follows:

$$-(a \cdot T(\lambda) + b) \approx \epsilon_{\mathrm{Hb}}(\lambda) \cdot c_{\mathrm{Hb}} + \epsilon_{\mathrm{HbO2}}(\lambda) \cdot c_{\mathrm{HbO2}} + \epsilon_{\mathrm{M}}(\lambda) \cdot c_{\mathrm{M}}. \tag{3}$$

By enumerating the equation for the 31 spectral bands of wavelength λ outputted by the spectral reconstruction, we obtain a system of 31 linear equations. This system can be represented in matrix form with $\mathrm{E} = [\boldsymbol{\epsilon}_{\mathrm{Hb}}, \boldsymbol{\epsilon}_{\mathrm{HbO2}}, \boldsymbol{\epsilon}_{\mathrm{M}}]$, $\mathrm{c_t} = [c_{\mathrm{Hb}}, c_{\mathrm{HbO2}}, c_{\mathrm{M}}]$:

$$-a \cdot \mathrm{p} - 1b = \mathrm{Ec}. \tag{4}$$

$$\mathrm{c_t} = \left(\mathrm{E}^T\mathrm{E}\right)^{-1} \mathrm{E}^T(-a \cdot \mathrm{p} - 1b). \tag{5}$$

Where the vector p is an observation of $T(\lambda)$ for the 31 band wavelengths. We then lastly choose to eliminate the constants *a* and *b* as we note that for our features we only need to estimate a quantity that moves with the true concentration the scale/offset unimportant for that unlike in the case of the true physical values. As such, we can replace the term $(-a \cdot \mathrm{p} - 1b)$ with $(-\mathrm{p})$. Now all nonlinearity in *p* has been addressed/eliminated. We finally can compute the "pseudo" concentration vector c of the three chromophores.

$$\mathrm{c} = \left(\mathrm{E}^T\mathrm{E}\right)^{-1} \mathrm{E}^T(-\mathrm{p}). \tag{6}$$

$$\frac{\mathrm{dc_t}}{\mathrm{dc}} = a. \tag{7}$$

Since c is linear in p we can now normalize p as we see fit for each feature calculated based on c in the case of our oxygenation feature we chose to L1 normalize *p* in order to eliminate the influence of brightness variations throughout the image and to make images of different overall brightnesses comparable to each other. For rPPG extraction we chose not to normalize.

Feature derivation

Sigmoid oxygenation

We design a feature named sigmoid oxygenation, $\sigma(O_2)$, to measure the tissue oxygenation level based on the outputted concentration data from CF-SCE. For each pixel concentration vector *c* we apply the following relation to map it to a pixel sigmoid oxygenation value.

$$\sigma(O_2) = 1 - \sigma\left(\left|\frac{\mathrm{HbO_2}}{\mathrm{Hb}}\right|\right). \tag{8}$$

As mentioned before with CF-SCE the scale/offset of the Hb, HbO2 quantities is lost/unknown. So no subtraction/addition of these quantities is meaningful. So we chose just to take the ratio of the two quantities to capture the understanding that when HbO2 is increased relative to Hb higher oxygenation should follow. As a ratio is involved the quantity $\left|\frac{\mathrm{HbO_2}}{\mathrm{Hb}}\right|$ may experience asymptotic behavior which makes visualization/scaling difficult. So we apply the sigmoid function Here to bound the quantity between 0 and 1. Finally, we apply an inverse transform to this bounded item to make higher sigmoid

oxygenation values indicate critical (low) areas also to make visualization easier.

Remote PPG

The process of extracting rPPG signals in MobiPerf is as follows. Each video frame is first passed through the pipeline outlined in Deep learning hyperspectral reconstruction and Calibration-free skin composition estimation sections, generating the corresponding concentration images. For each frame, the region of interest (ROI) is defined based on the given 2D segment, and the mean concentration vector is computed by averaging the pixel-wise concentration vector values spatially within this region. This spatially-averaged vector serves as the basis for the signal extraction process.

To construct an rPPG signal, the hemoglobin (Hb) concentration from the mean concentration vector is selected for each frame, forming a time-series signal. Similarly, an alternative rPPG signal can be generated by considering the oxyhemoglobin (HbO2) concentration.

$$S_1(t) = \sum_{(i,j)\in G_t} C_{t,i,j,\mathrm{Hb}}. \tag{9}$$

$$S_2(t) = \sum_{(i,j)\in G_t} C_{t,i,j,\mathrm{HbO_2}}. \tag{10}$$

To further enhance signal robustness, we apply an alpha-tuning technique [37], which leverages the in-phase dynamics shared between the Hb and HbO2 signals. By assuming that noise components are largely out of phase, alpha-tuning amplifies the shared in-phase components while suppressing noise, resulting in a signal that is robust to dynamics caused by ballistocardiogram. The final enhanced output, termed the MobiPerf Motion-Compensated rPPG, provides a more robust representation of cardiac dynamics, enabling accurate physiological assessments.

$$\alpha = \frac{\sigma(S_1(t))}{\sigma(S_2(t))}. \tag{11}$$

$$\mathrm{rPPG}_m(t) = S_1(t) - \alpha \cdot S_2(t). \tag{12}$$

Discussion

In this paper, we presented MobiPerf, a dual-function wound monitoring system specifically designed for diabetic foot care. Utilizing standard RGB images captured with a smartphone, MobiPerf estimates skin composition parameters by reconstructing HSI images using advanced deep learning techniques. The system was rigorously validated on hyperspectral video datasets and "in-the-wild" image datasets captured under diverse conditions and using multiple types of cameras. To achieve its robust performance, MobiPerf introduces key innovations at every stage of its pipeline. For HSR, we employed state-of-the-art image-to-image MST++ model architecture, to ensure adaptability across various camera setups without requiring device-specific training. For skin composition estimation, we developed a novel brightness-tolerant methodology that eliminates the dependence on controlled lighting or radiometric calibration, thereby enhancing usability in real-world scenarios. Additionally, features such as the sigmoid oxygenation metric and robust rPPG signals were designed to provide diagnostic insight in diverse conditions. These innovations position MobiPerf as a cost-effective, portable, and accessible solution for diabetic foot ulcer monitoring. Beyond wound care, its systematic robustness improvements hold promise for advancing other mobile health technologies and improving outcomes for patients managing chronic conditions. In addition to the concerns we are about to explore future works can also look into quantifying the effect of skin tone here on the performance of MobiPerf. As well as considering how to make the system real time for video capabilities such as rPPG through the use of model quantization reducing needed channel and image resolution and the like.

Lighting variability

HSI is conducted with radiometric calibration when used for scientific purposes as reflectance images which are physically meaningful are needed (typically xenon lighting which is considered spectrally neutral due to its consistent emission of light across all wavelengths). Radiometric Calibration is typically done using reference white/black images placed in the scene to normalize the image as disentangling the spectra/amplitude of the Light source from the reflectance at each pixel is very difficult, as it entails blind-factoring two 31-element vectors from another 31-element vector at each pixel a under defined problem.

In practical mobile environments, however users rarely have objects laying around that would redly serve as a reference target for radiometric calibration and the procedure for calibrating can be regarded as straining/requires care to keep the environment static during subsequent image capture. In order to address this we take two different approaches the first approach is to mitigate the need for radiometric calibration entirely as we do in Calibration-free skin composition estimation section, and the second approach is to integrate radiometric calibration into the hyperspectral reconstruction process yielding reflectance reconstruction without reference targets as we do in Deep learning reflectance reconstruction section.

The results, presented in Lighting variability section, reveal that performance does degrade moderately under challenging lighting conditions when we are only using CF-SCE. And as we can see in Reflectance reconstruction section using reflectance reconstruction models is a way to address this performance shortfall as well as reduce performance variation generally. However since the shortfall is still only moderate and further since deep learning reflectance reconstruction is still very early (not been applied to multiple light sources, or uneven lighting), we chose not to use it for our in-the-wild testing although encourage readers to explore it further.

Despite the performance shortfall under challenging lighting, our "in-the-wild" evaluation on diabetic foot ulcer images (Diabetic foot ulcers section) shows that the system remains effective under non-ideal lighting conditions with the sensitivity of our sigmoid oxygenation feature still being evident even under fully un-constrained circumstances. Additionally, as highlighted in Remote photoplethysmography section, the derived rPPG signals retain strong diagnostic reliability, even with unknown lighting conditions. These findings underscore MobiPerf's robustness and adaptability in real-world environments, affirming its viability as a practical and portable solution for mobile health applications, particularly in diabetic foot care monitoring.

Camera variability

Previous studies have demonstrated that mobile phone cameras typically exhibit highly similar camera spectral response functions (CSRFs), which are defined by three vectors representing the sensitivity of each camera channel (R, G, and B) to bands of incident light [19]. This uniformity is in stark contrast to the substantial variation observed in DSLR spectral sensitivity functions, such as those present in the UBFC-Phys [17] and DFCU2021 [18] datasets used in our evaluations. Notably, the MST_{700} model employed by our system was trained using a DSLR spectral sensitivity function as well. This variation suggests that the evaluation conditions for our system, in terms of CSRF variability, may represent a more challenging scenario than applying the system to smartphones. Additional considerations however arise from the optical characteristics of mobile lenses, which may introduce chromatic aberrations due to size and cost constraints. These aberrations are intrinsic to the imaging pipeline and could influence the spectral accuracy of HSR. More extreme chromatic aberrations likely can cause significant hallucinations in hyperspectral reconstruction models which form the backbone of the work. Likewise so may highly unusual lighting artifacts or sources even. Both are likely to have a effect on estimation of the chromophores using CF-SCE even as it is designed to reduce the effect of unknown lighting.

Furthermore, while smartphones typically apply post-capture color corrections—such as dynamic white balancing and pre-applied color correction matrices to enhance image realism—these adjustments may introduce inconsistencies in HSR. A potential mitigation strategy involves capturing raw images via permissive camera APIs, where available, and preprocessing them to enforce a standard white balance. Although this approach has been explored in prior studies, further investigation is necessary to fully address the effects of color correction in mobile imaging pipelines.

Bacteria

In this work, we demonstrated that our system is sensitive to the differences between ulcers with ischemic conditions and those without any pathological indications, using only the mean sigmoid oxygenation value derived from hyperspectrally reconstructed images. However, the system showed a much more limited capability in distinguishing infections from normal images. Here we propose a possible extension of the work to improve the infection capability.

Previous studies have demonstrated the potential of spectrometer-based measurements to detect bacterial presence on wounds [38]. Specifically, spectrometer data were used to identify spectral signatures associated with common bacterial infections, such as Staphylococcus aureus and Escherichia coli, with these spectral signatures known a priori. The study explored whether hyperspectral imaging could similarly distinguish these infections using the same spectral markers. Building on this, we hypothesize that integrating these bacterial spectral signatures into our HSR framework could enhance its ability to detect infections. Future work should focus on incorporating these spectral features into the system and evaluating its performance in detecting and classifying infected wounds.

Acknowledgements

This study was funded by National Institute of Biomedical Imaging and Bioengineering 1R01EB035188-01. The funder played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript.

Authors' contributions

A.G Lead Conceptualization, Methodology, Software, Manuscript, Validation. W.B Preparing the manuscript, Analysis/Advisory. A.D Worked on software development. Y.Z Preparing the manuscript, Analysis/Advisory. W.X Lab PI advised all aspects of the project including Conceptualization, provided resources, and performed manuscript review.

Data availability

All data used to generate the results in this paper are from third party datasets. The SPECTRALPACA dataset is publicly available and can be accessed at Synapse Repository under the following DOI: https://doi.org/10.7303/syn51625685. The UBFC-Phys dataset is also publicly available and can be accessed at IEEE Dataport Repository under the following DOI: https://dx.doi.org/10.21227/5da0-7344. Access to the DFCU2021 and DFCU2020 datasets are restricted by their owners. Researchers interested in obtaining this dataset must submit

a request at the following website to the owner and agree to their terms and conditions: https://helward.mmu.ac.uk/STAFF/M.Yap/dataset.php.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 26 April 2025 / Revised: 28 August 2025 / Accepted: 3 September 2025
Published online: 22 September 2025

References

1. Voelker R. What are diabetic foot ulcers? JAMA. 2023;330(23):2314.
2. Tran V, Galvan B, Khemka S, Holder K, Ansari MM. Importance of using angiography for the early detection of chronic limb ischemia in diabetic foot wounds. Cureus. 2024;16(6):e61906.
3. Bhargava A, Sachdeva A, Sharma K, Alsharif MH, Uthansakul P, Uthansakul M. Hyperspectral imaging and its applications: a review. Heliyon. 2024;10(12):e33208.
4. Scheeren TWL, Schober P, Schwarte LA. Monitoring tissue oxygenation by near infrared spectroscopy (NIRS): background and current applications. J Clin Monit Comput. 2012;26(4):279–87.
5. Gherardi A. Created in BioRender. 2025. https://BioRender.com/u17g887.
6. Kuang B, Pena G, Szpak Z, Edwards S, Battersby R, Cowled P, et al. Assessment of a smartphone-based application for diabetic foot ulcer measurement. Wound Repair Regen. 2021;29(3):460–5.
7. Xu H, Shuttleworth KMJ. Medical artificial intelligence and the black box problem: a view based on the ethical principle of "do no harm". Intelligent Medicine. 2024;4(1):52–7.
8. He Q, Wang R. Hyperspectral imaging enabled by an unmodified smartphone for analyzing skin morphological features and monitoring hemodynamics. Biomed Opt Express. 2020;11(2):895. https://doi.org/10.1364/BOE.378470.
9. He Q, Li W, Shi Y, Yu Y, Geng W, Sun Z, et al. SpeCamX: mobile app that turns unmodified smartphones into multispectral imagers. Biomed Opt Express. 2023;14(9):4929. https://doi.org/10.1364/BOE.497602.
10. Nishidate I, Maeda T, Niizeki K, Aizu Y. Estimation of melanin and hemoglobin using spectral reflectance images reconstructed from a digital RGB image by the Wiener estimation method. Sensors. 2013;13(6):7902–15. https://doi.org/10.3390/s130607902.
11. Park SM, Visbal-Onufrak MA, Haque MM, Were MC, Naanyu V, Hasan MK, et al. Mhealth spectroscopy of blood hemoglobin with spectral super-resolution. Optica. 2020;7(6):563. https://doi.org/10.1364/OPTICA.390409.
12. Ji Y, Park SM, Kwon S, Leem JW, Nair VV, Tong Y, et al. Mhealth hyperspectral learning for instantaneous spatiospectral imaging of hemodynamics. PNAS Nexus. 2023. https://doi.org/10.1093/pnasnexus/pgad111.
13. Pillay R, Hardeberg JY, George S. Hyperspectral Calibration of Art: Acquisition and Calibration Workflows. arXiv. 2019. https://doi.org/10.48550/ARXIV.1903.04651.
14. Arad B, Ben-Shahar O, Timofte R, Van Gool L, Zhang L, Yang MH, et al. NTIRE 2018 Challenge on Spectral Reconstruction from RGB Images. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2018. pp. 1042–104209. https://doi.org/10.1109/CVPRW.2018.00138.
15. Cicek M, Buckley J, Pearson-Stuttard J, Gregg EW. Characterizing multimorbidity from type 2 diabetes. Endocrinol Metab Clin North Am. 2021;50(3):531–58. https://doi.org/10.1016/j.ecl.2021.05.012.
16. Ayala L, Mindroc-Filimon D, Rees M, Hübner M, Sellner J, Seidlitz S, et al. The SPECTRAL perfusion arm clamping dataset (SPECTRALPACA) for video-rate functional imaging of the skin. Sci Data. 2024. https://doi.org/10.1038/s41597-024-03307-y.
17. Meziati R, Benezeth Y, De Oliveira P, Chappé J, Yang F. UBFC-Phys. IEEE Dataport. 2021. https://doi.org/10.21227/5da0-7344.
18. Yap MH, Cassidy B, Pappachan JM, O'Shea C, Gillespie D, Reeves N. Analysis Towards Classification of Infection and Ischaemia of Diabetic Foot Ulcers. arXiv. 2021. https://doi.org/10.48550/ARXIV.2104.03068.
19. Tominaga S, Nishi S, Ohtera R. Measurement and estimation of spectral sensitivity functions for mobile phone cameras. Sensors. 2021;21(15):4985. https://doi.org/10.3390/s21154985.
20. Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L, et al.. Segment Anything. arXiv. 2023. https://doi.org/10.48550/ARXIV.2304.02643.
21. Cie. CIE 015:2018 Colorimetry, 4th Edition. Tech. Rep., International Commission on Illumination; 2018.
22. Vu BV, Lei R, Mohan C, Kourentzi K, Willson RC. Flash characterization of smartphones used in point-of-care diagnostics. Biosensors. 2022;12(12):1060.
23. Lugaresi C, Tang J, Nash H, McClanahan C, Uboweja E, Hays M, et al.. MediaPipe: A Framework for Building Perception Pipelines. arXiv. 2019. https://doi.org/10.48550/ARXIV.1906.08172.
24. Ontiveros RC, Elgendi M, Missale G, Menon C. Evaluating RGB channels in remote photoplethysmography: a comparative study with contact-based PPG. Front Physiol. 2023. https://doi.org/10.3389/fphys.2023.1296277.
25. Poh MZ, McDuff DJ, Picard RW. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. Opt Express. 2010;18(10):10762. https://doi.org/10.1364/oe.18.010762.
26. Wang W, den Brinker AC, Stuijk S, de Haan G. Algorithmic principles of remote PPG. IEEE Trans Biomed Eng. 2017;64(7):1479–91. https://doi.org/10.1109/TBME.2016.2609282.
27. Haugg F, Elgendi M, Menon C. Effectiveness of remote PPG construction methods: a preliminary analysis. Bioengineering. 2022;9(10):485. https://doi.org/10.3390/bioengineering9100485.
28. Goyal M, Reeves ND, Rajbhandari S, Ahmad N, Wang C, Yap MH. Recognition of ischaemia and infection in diabetic foot ulcers: dataset and techniques. Comput Biol Med. 2020;117:103616. https://doi.org/10.1016/j.compbiomed.2020.103616.
29. Goyal M, Reeves ND, Davison AK, Rajbhandari S, Spragg J, Yap MH. DFUNet: convolutional neural networks for diabetic foot ulcer classification. IEEE Trans Emerg Top Comput Intell. 2020;4(5):728–39. https://doi.org/10.1109/TETCI.2018.2866254.
30. Goyal M, Reeves ND, Rajbhandari S, Yap MH. Robust methods for real-time diabetic foot ulcer detection and localization on mobile devices. IEEE J Biomed Health Inform. 2019;23(4):1730–41. https://doi.org/10.1109/JBHI.2018.2868656.
31. Cassidy B, Reeves ND, Pappachan JM, Gillespie D, O'Shea C, Rajbhandari S, et al. The DFUC 2020 Dataset: Analysis Towards Diabetic Foot Ulcer Detection. touchRev Endocrinol. 2021;17(1):5. https://doi.org/10.17925/ee.2021.17.1.5.
32. Cai Y, Lin J, Lin Z, Wang H, Zhang Y, Pfister H, et al. Mst++: Multi-stage spectral-wise transformer for efficient spectral reconstruction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022. pp. 745–55.
33. Arad B, Timofte R, Yahel R, Morag N, Bernat A, Cai Y, et al. NTIRE 2022 Spectral Recovery Challenge and Data Set. In: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2022. pp. 862–80. https://doi.org/10.1109/CVPRW56347.2022.00102.
34. Li Y, Fu Q, Heidrich W, Multispectral illumination estimation using deep unrolling network. In: 2021 IEEE International Conference on Computer Vision(ICCV). 2021. pp. 1–8.
35. Prahl S. Tabulated molar extinction coefficient for hemoglobin in water. 1999. https://omlc.org/spectra/hemoglobin/summary.html.
36. Jacques S. Optical Absorption of Melanin. 1998. https://omlc.org/spectra/melanin.
37. de Haan G, Jeanne V. Robust pulse rate from chrominance-based rPPG. IEEE Trans Biomed Eng. 2013;60(10):2878–86. https://doi.org/10.1109/TBME.2013.2266196.
38. Poosapadi Arjunan S, Tint AN, Aliahmad B, Kumar DK, Shukla R, Miller J, et al. High-resolution spectral analysis accurately identifies the bacterial signature in infected chronic foot ulcers in people with diabetes. Int J Low Extrem Wounds. 2018;17(2):78–86. https://doi.org/10.1177/1534734618785844.

39. Gherardi AJ, Demirbas A, Xu W. SigmoidOxy: A Light-weight mobile perfusion tool for diabetic foot management. In: Proceedings of the 30th Annual International Conference on Mobile Computing and Networking. ACM MobiCom '24. New York: Association for Computing Machinery; 2024. pp. 2290–6. https://doi.org/10.1145/3636534.3698119.

Publisher's Note