Mobile Perfusion Assessment Tool as Alternative to LDI

Team 7 - Rohit Bhattacharya, Yvonne Jiang, Azwad Sabik

May 8, 2015

Abstract

Blood perfusion can be used as an indicator to assess the healing of a chronic wound. The current gold standard, LDI, is expensive and relatively inaccessible. In this paper we present an integrated software-and-hardware solution that is able to classify areas of the patient's hand as having low or high perfusion, using a standard phone video and infrared image as inputs. Our solution is able to produce accurate classifications for a two-class system distinguishing between high and medium perfusion from low perfusion however it provides poor classifications when attempting to classify a three-class system of low, medium, and high perfusion levels. We conclude that such a software solution might replace LDI in the foreseeable future but requires further testing and tuning for now.

1 Introduction

The goal of our project is to develop an integrated software-and-hardware tool that allows a clinician to use a mobile device to extract a usable metric that assesses local blood flow. Measures of local blood flow (perfusion) can help characterize healing of chronic wounds and assist physicians in developing appropriate treatment plans for patients. Currently laser doppler imaging (LDI) is the standard method of assessing perfusion, but it is often expensive and inaccessible. Consequentially, wound prognosis is often poor in quality and can necessitate skin grafts, amputation of the limb or even death.

The basis of our tool relies on the fact that skin color changes subtly when blood perfuses the tissue beneath it. This would normally be too

subtle of a change for the human eye to perceive, however using a recent technique developed at MIT [4], Eulerian Video Magnification, these color changes become more pronounced and quantifiable. Previous applications of this technique include using this skin color amplification to measure heart rate from a video of the face.

In addition to color amplified mobile phone captured video, thermal infrared imaging (via smartphone sensor attachment) is also studied as a usable metric, as blood flow has been shown to correlate with increased temperature [3]. These metrics are then used to teach a machine learning algorithm that will classify perfusion into relative bins. The details of the algorithm will be outlined in our methods sections below.

2 Methods

Our experimental protocol first and foremost relied on the collection of ground truth LDI data with corresponding iPhone captured video and thermal-infrared data (captured using the FLIR ONE iPhone attachment). All data was collected of the hands of subjects, as they were the easiest part of the body to stably and consistently take videos of. As access to laser doppler imaging was costly and time-limited (demonstrating the need for a more accessible solution), we mimicked different blood flow conditions during each data collection session so as to limit the number of test subjects needed as well as the number of sessions needed, but while maintaining variation of data.

To this end, we used a simple manual sphygmomanometer at a low (0 mm Hg), medium (60 mm Hg), and high pressure (120 mm Hg) to emulate good, medium, and poor perfusion respectively. Between each trial the arm was given time to rest and return to regular flow to ensure consistent conditions.

Each trial consisted of taking the pulse of the subject using a pulse oximeter, inflating the blood pressure cuff to the desired pressure, letting the flow stabilize for a minute, taking the thermal-infrared image, a ten second video, and then letting the LDI scan the hand. Following this, the collected data was analyzed and used to teach our algorithm using the techniques described in the following subsections.

2.1 Eulerian Video Magnification

The technique used to process the raw iPhone collected videos was Eulerian Video Magnification (EVM), an algorithm whose developers claim can "reveal subtle changes in the world." [4] Specifically, it was designed to amplify motion and color change in videos.

One of the proposed applications of this technique was to use the amplified color changes of the skin caused by underlying blood flow to determine a subject's heart rate [1]. We validated this application first before moving on to trying to apply it to the problem of perfusion assessment and found that while our results were not as accurate as the results described by the developers [4], they were significant enough that we felt it validated that particular application.

Thus using a similar approach, we applied the EVM algorithm to all of our input videos, with the same set of parameters that we found optimal for isolating heart beat, and then extracted metrics such as average intensity, change in intensity over time, and rate of changing intensity, on a section by section (each section being five by five pixels) basis. To focus on the most meaningful data points and consistently compare how they changed with increased sphygmomanometer pressure, we cropped each input video into five separate videos, each featuring a single fingertip. With heart rate, only the time between peak intensities was important, but for blood perfusion there are other significant factors at play. How we determined which metrics were the most related to perfusion will be discussed in the next section.

2.2 Support Vector Machine

A support vector machine (SVM) is a supervised learning technique that is able to define a separating hyperplane between two distinct classes of data. An SVM must first be "trained" on hand-labelled data, from which it is able to identify patterns in the feature set that help it form the separating hyperplane. It can then be used to classify novel data using the features provided to it. For our purposes, we use SVMs to classify or bin areas of a patient's hand as having low, medium or high perfusion using features extracted from EVM processed videos of the hand.

2.2.1 Features

For the red, green and blue channels of the video some of the features we used include, rate of change of intensity, dominant frequency as given by a fast Fourier transform, average peak distance, and average distance between zero crossings. We also used the intensity from a gray-scale infrared image taken using the FLIR ONE attachment for the iPhone. A full list of these features can be found in our README. In addition to this, we used principal component analysis (PCA) to identify the most information rich features i.e. features that account for the most variability in our data. PCA is generally used on high dimensional feature sets to reduce computation time, and we thought it would be interesting to see how it would affect our SVM. In this paper, we use PCA to identify the 1, 10, 20, 30, 40, 50 and 60 (of a total of 64) of the most important features to train our SVM and analyze the tradeoff between accuracy and speed. We also created plots of the distribution of features across the whole training data set to enable us to manually validate how important a feature might be based on eyeballing the variation/similarity seen between class labels.

2.2.2 Multi-class SVM

SVMs are traditionally binary classifiers. In order to perform a low, medium and high classification as described above, we used a one-versus-rest [2] technique. This involved modelling three separate SVMs, one for each class. Each of these SVMs provides a score indicating the confidence with which an example can be said to belong to one class over the others. The class that scores the highest in this manner, is considered to be the final classification of this example.

2.2.3 Two-class SVM

Preliminary results from our multi-class SVM indicated to us that a 3-class system of low, medium, and high perfusion might be infeasible. As an alternative, we reduced the classification problem to one with two classes. This can be achieved by considering high and medium labels as belonging to one class and distinguishing them from low perfusion or alternatively considering low and medium labels as one class and distinguishing them from high perfusion. The results of both these methods are presented here.

2.2.4 EVM vs Non-EVM

Towards the end our project, we also decided to build an SVM trained using standard video data rather than pre-processing it with EVM. We were curious to see just how much of an impact EVM had on our classifier.

2.2.5 Training and Testing

As confirmed by the LDI ground truth, pressures of 0, 60, and 120 mm Hg do indeed correspond to low, medium and high perfusion. This allowed us to make the claim that any region of the hand in iPhone videos taken at these applied pressures can be labelled as low, medium or high accordingly. We used a combination of data obtained from Yvonne and Rohit's fingertips (as the fingers were regions that showed highest variability in perfusion between pressures) as training data. In addition to the tuning features used as talked about earlier, we also experimented with three different kernels - radial, linear, polynomial. We used 10-fold cross validation to assess the performance of our SVM.

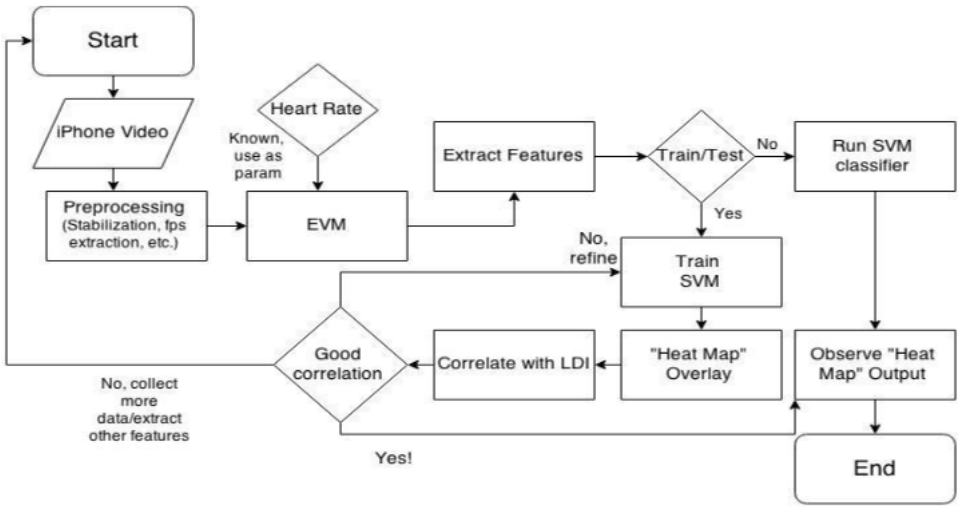


Figure 1: Development Pipeline

3 Results

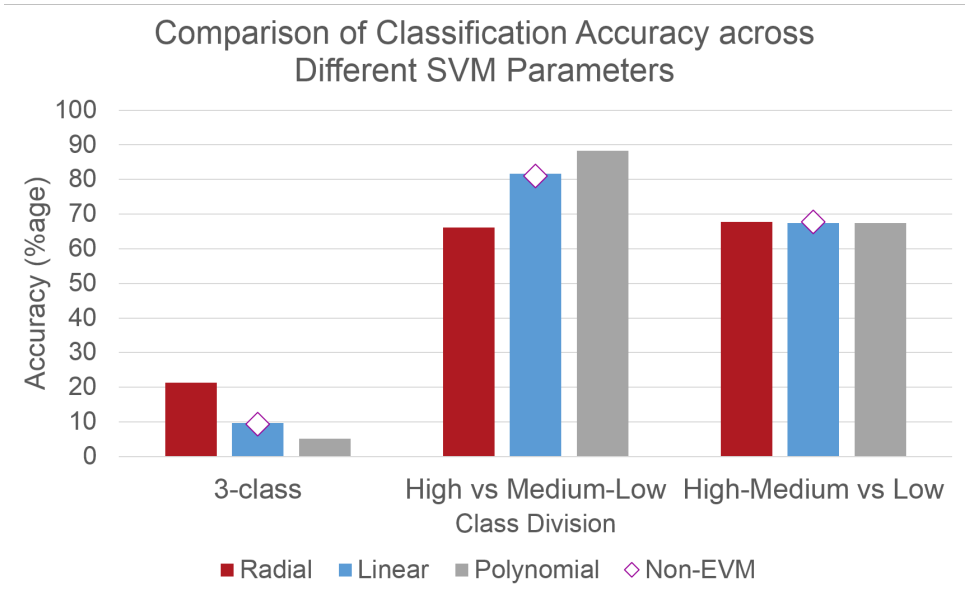


Figure 2: Results of running an SVM with 3 different kernels and 3 different class divisions. In addition, comparison with running one particular SVM kernel configuration on videos not processed using SVM

(a) 'High Perfusion' Hand (0 mmHg) (b) 'Medium Perfusion' Hand (60 mmHg)

Figure 3: Our trained SVM's output when run on the same hand under two different pressures (white - high perfusion, grey - medium/low perfusion)

3.1 Feature Selection

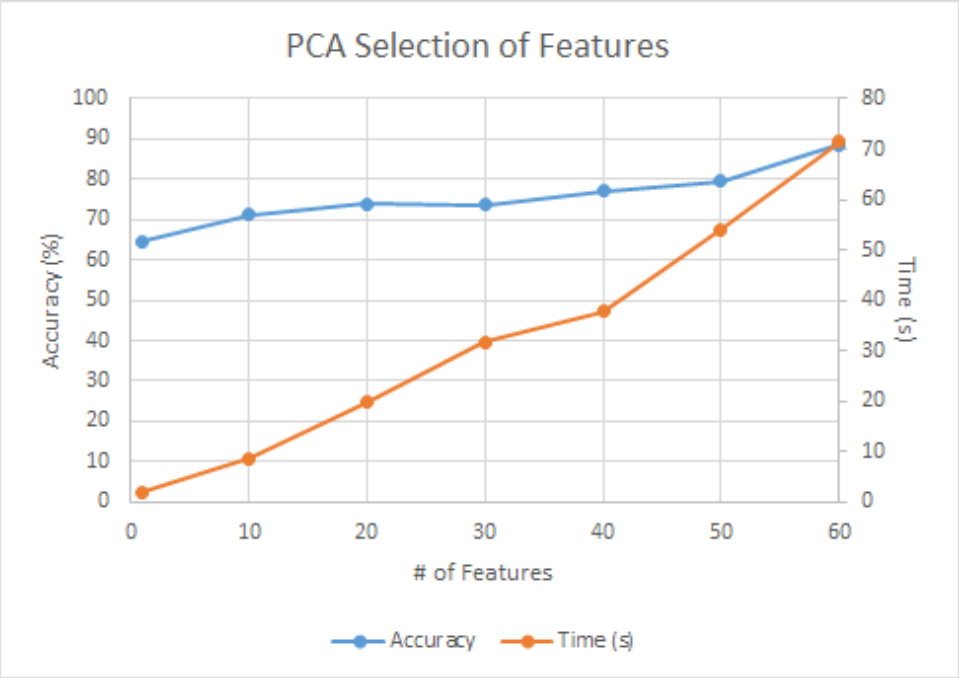
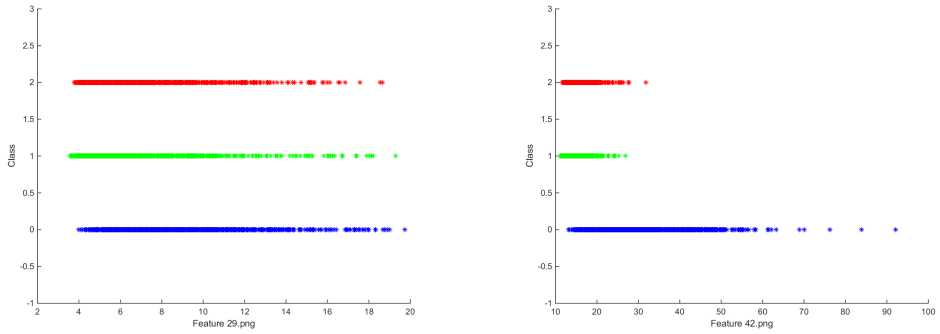


Figure 4: Results of experimenting with PCA and restricting the number of features used



(a) Median Red Channel Intensity (b) Max Red Channel Zero Crossing

Figure 5: Distribution of the Values of two features over the training data (0, 1, and 2 are the labels given to low, medium, and high respectively)

4 Discussion

4.1 Class Division

As can be seen in figure 2 extremely poor classification accuracy was achieved when using a 3-class system. The slightly higher accuracy achieved by the radial classifier suggests to us that the 3-class system feature space is not linearly separable and thus explains the poor performance of the linear and polynomial classifiers. When we use either of the 2-class systems described in the previous section, the accuracy drastically increases as seen again in figure 2. Thus, applying such a simplification to the problem, made for a much "nicer" feature space that was significantly easier to separate. The highest accuracies were achieved when using the 2-class system that used the same label for the medium and low perfusion classes. This is an encouraging result when considering the implications of our problem simplification. The 2-class system described above better describes a distinction between a normal patient (high perfusion) and a patient with a chronic wound and ill-healing underlying blood vessels (medium/low perfusion). The other 2-class system on the other hand, wherein the same label was given to high and medium perfusions, might simply describe the difference between a human being (high/medium perfusion) and an inanimate object (low perfusion). Results of an SVM trained under the former 2-class system (medium and low perfusion in one class) can be seen in figures 3a and 3b.

4.2 Kernel Selection

As mentioned earlier, the radial classifier performed better than the linear and polynomial ones for the 3-class system due to the feature space being linearly inseparable. For the 2-class systems however, the linear and polynomial classifiers were able to achieve the same or higher accuracies. In the case where high and medium perfusions were given the same label, the polynomial classifier obtained a significantly higher accuracy than the linear one. While this may seem like a favourable result, this may be caused by the polynomial classifier overfitting the data. Hence, we concluded that the linear classifier may be the most suitable and consistent for our purposes.

4.3 Feature Selection

As with any data set, there are features that seem to be more important than others. Upon manual examination of plots such as the ones seen in figures 5a and 5b, it is easy to see why - some features are just too similar across all three classes while some are similar between two classes but different from the third (the latter case is what makes classification accuracy better for 2-class systems). There were no features that were distinct among all three of our classes, which may be a reason why our 3-class classifier's performance is so poor. In addition to manual examination, we also used PCA to pick out important features in our data. Most of the features obtained through PCA's dimensionality reduction corresponded to features that we ourselves had identified through manual examination, which is an encouraging result. The results of our experiments with PCA (shown in figure 4) demonstrate the tradeoff between run time and accuracy. With our current feature set we found the run time to be acceptable even without PCA reduction; however, it is possible that in a less developed country computational power might not be up to par and figure 4 helps to identify a happy medium between speed and accuracy.

4.4 EVM vs Non-EVM

As discussed earlier, an important step in our pipeline was to process our videos using EVM. We believed that this would amplify the effects of the features that we feed to our SVM and improve classification accuracy. As seen in figure 2 however, this is clearly not the case. The results of our classification are very comparable for both non-EVM and EVM processed data. When we looked through plots similar to those seen in figures 5a and 5b, we concluded that while EVM is amplifying the absolute values of our data, it is not in fact, affecting the relative difference of the features between our class labels. Thus, our SVMs identify similar features in both sets of data as being important and result in similar classifications.

5 Conclusions

There are promising results for two class classification, especially between cases of high-to-medium perfusion versus low perfusion, but there is still

significant work needed to tune our algorithm for three-class classification or better.

Our experimentation with Principal Component Analysis (PCA) showed that it resulted in faster classification at the cost of slightly lower accuracies in general, as expected. It also showed that no one feature was enough to distinguish between perfusion levels.

We were, however, unable to show a significant advantage in using EVM to process the videos compared to just using the original video, which may be a favorable result, because it implies that the final product could potentially be less computationally intensive than previously projected.

6 Management Summary

Here we present a summary of the management of our project, including division of labor, goals we planned to achieve, and what we were actually able to achieve.

6.1 Management Plan

Azwad was in charge of perfusion metric extraction, input-to-classifier pipeline development, and statistical analysis.

Rohit was in charge of data collection, SVM design and implementation, and SVM performance assessment.

Yvonne was in charge of pre-processing of video and image data, front end development, and source code control.

6.2 Planned Deliverables

At the very minimum, we were to provide proof or disproof of EVM as a means of characterizing perfusion. We expected to get this done by 03/02 and move on to building a classifier that would utilize EVM data (if found to be feasible) to classify areas in a video as having low, medium or high perfusion. If EVM did not show promising results, we were to integrate a single point laser doppler system into our solution to provide additional data points that would aid our classifier. We projected that this would be completed by 04/12, leaving us with time to work towards our maximum

deliverable, which was to transfer our solution from MATLAB to a mobile platform and also perform some pre processing in order to stabilize the video.

6.3 Final Deliverables

Due to the inaccessibility of the LDI, there were significant delays in obtaining data that we could use for correlation with EVM. We also dropped the idea of using a single point laser doppler due to limited market availability and unexpectedly high costs associated with acquiring one, going against the principle of this being a more accessible alternative to LDI. Despite this, our work has culminated in a relatively streamlined process of video/IR data collection using a smart phone, followed by EVM processing and the final classification (performed by an SVM) of the region of interest, output on a computer.

6.4 Future Direction

The current version of our tool runs in a MATLAB environment, which made it easier for us to prototype and test ideas. Moving forward however, we hope to convert the code base from MATLAB to one that is executable on mobile platforms, making our solution truly mobile. We would also like to be able to include real chronic wound data into our analysis to see how well our hypothesis holds up in the real world.

6.5 Lessons Learned

Correlations obtained through image and video processing require very precise data collection or advanced image stabilization that might not always be feasible in a clinical setting. We also realized that having well documented code right from the start can save on a lot effort later on in the project. And finally, better communication with our mentors might have alleviated several key issues we were facing both in terms of dependencies and overall direction of the project at a much earlier stage.

7 Technical Appendix

Here is a link to our repository - https://bitbucket.org/yjiang23/cisperfusion

References

[1] M. DaneshiKohan and A. NasrAbadi. A temporal video-processing method to improve heart rate estimation. *Perfusion*, Dec 2014.

[2] S. Hua and Z. Sun. Support vector machine approach for protein subcellular localization prediction. *Bioinformatics*, 17(8):721–728, Aug 2001.

[3] A. Merla, L. Di Donato, G. L. Romani, M. Proietti, and F. Salsano. Comparison of thermal infrared and laser doppler imaging in the assessment of cutaneous tissue perfusion in scleroderma patients and healthy controls. *Int J Immunopathol Pharmacol*, 21(3):679–686, 2008.

[4] Hao-Yu Wu, Michael Rubinstein, Eugene Shih, John Guttag, Frédo Durand, and William T. Freeman. Eulerian video magnification for revealing subtle changes in the world. *ACM Transactions on Graphics (Proc. SIGGRAPH 2012)*, 31(4), 2012.