Spring Midterm Progress Report

Core Body Temperature Estimation to Detect Ebola Virus Disease

CS 463, Spring 2018, Group 34

Claude Maimon, Brian Lee Huang, and Bianca Beauchamp

Abstract

The end goal of this project is to end up with a research paper. The paper should outline the problem that the project is trying to solve, the steps taken to solve the problem and how successfully we were at solving the problem. The paper should also allow the project to be continued if someone chooses to. The whole process should be explained in detail allowing whoever wants to continue the project to continue without any problem. The main body of this research paper will be about the program that we develop to predict a person's core body temperature. The program should first be able to extract data from a thermal image. The data of the image should come from the top half, focusing on the head. It will then interpret the data to create a mathematical model that uses the temperature of a person's skin as data and analyzes that information and predicts what their core body temperature is. A high accuracy rate is not strictly required as that is not the point of the project, the goal of the is to determine if this method will be effective to detect whether a person is symptomatic with Ebola. A high accuracy rate is a good indicator that a mathematical model is a good way to predict, where a low accuracy rate indicates that we should look for an alternative method.

CONTENTS

1 INTRODUCTION

Our project is being created for Bill Smart for doctors without borders. The idea behind it is to prevent the spread of Ebola by eliminating the need for doctors to check patients by hand, additionally potentially sick patients will be quarantined to potentially minimize the spread between sick and well patients. To do this we are using a thermal camera and look at the skin temperature. Then we are estimating core body temperature by creating a model and then determine whether or not the patient is symptomatic with a fever or not.

2 IMAGE PROCESSING, CLAUDE MAIMON

2.1 Recap and Goals

My part of the project is the image processing. I am responsible for getting the camera to take a picture and import that image into my code. Moreover, I am also responsible for processing the images and setting a range of temperature that the code uses. The goal of my part is to produce one temperature from the image and send it into the model part.

2.2 Current Progress

Overall the image processing part is done. The first thing that our code do is connect to the camera and make it take a picture on a key press. Bill's graduate student Chris helped us with this code. On a key press, the code takes about 30 images of the person and selects 5 of them. We do that because the camera can produce garbage and we don't want to use those frames. In order for this part to work the person has to be standing 2 meters away from the camera and his head must be under a specific height.

Right now we're using a key press to take an image because the mechanical team didn't install the movement sensors in the camera structure. In the future, a graduate student will implement those sensors into the code and the key press will not be required.

After taking the pictures our code transfer the images into CSV files. Each value in the CSV holds the temperature (in Celsius) of that specific pixel. After that our code finds the minimum value in the file and subtract that value from all the values in the CSV file. We do that to try and help with the camera calibration. After that, the code goes through the file values that are stored in an array and it calculates the mean values of the head temperatures. It does that by first finding the head. The way we find the head in the image is the following: we go through the values in the two dimensional array that holds the CSV values. Once we find 5 values in a row that are in the right range our code decides that that's where the head begins. After finding the head the code finds the average of all the head values that are in the right range. The code then produces a mean value that will be used in the model.

2.3 Problems and Solutions

We ran into many problems with the camera over this project. It started with an uncertainty about the camera, we didn't even know if we are going to get it in time. It's an $11,000 camera and it took a while to get that budget approved. At the beginning of winter term, we received a few images to work with and a few CSV

files but still didn't get access to the camera. Once we got access to the camera we spent week trying to get it to work. The camera actually returned wrong temperature values when used. When a human's skin is about 35 degrees Celsius, the camera would return values as high as 60 degrees Celsius. We tried many things to fix it. We checked if the error is constant, but it want's. We tried it with hot and cold reference points and it still didn't help. We tried calibrating it with more images, which helped a little but not enough for the camera to be right. We spent about 3 weeks on this with little progress on the temperature readings. At the end of winter term we worked with Bill's graduate student Chris on calibrating the camera. He tried to calibrate it with his code but that also wasn't sufficient. We found that the background of the images change all the time with no reason. Chris suggested that we'll take the minimum value in the image and subtract it from all the values in the image. Doing so helped a little with the calibration of the camera but not in all case.

When Bill's graduate student sent us some new images we found that the "subtructing the min value" approach doesn't work on all images. In order to solve this we tried to normalize the values in the csv files but that approach didn't help as well. We chose to stick with just subtracting the minimum values from the csv files. That's the method that gave us best results so far.[1]

3 Model, Bianca Beauchamp

3.1 Recap and Goals

The skin temperature found from analyzing the pixels needs to be related to the measured core body temperature in order to produce an equation that represents the relationship between skin temperature and core body temperature. The skin temperature of a subject and the measured core body temperature of the same subject will be provided. These two pieces of data will be taken from many subjects and then used to create a model. The model will then be able to take just the summary statistics, which is the subjects estimated skin temperature, and produce the core body temperature. The model that will be used to start with is a linear regression. This was chosen as a starting point because it is the most basic model and complexity can be added if it is determined to be valuable. Once the model has trained on the set of data that has both the subjects estimated skin temperature and their measured core temperature, it needs to be tested on a new set of data to determine how accurate the model is. This will be done by providing the model with only the estimated skin temperature and comparing the core temperature the model calculates to the measured core temperature. To quantify this comparison it is best to calculate the absolute error between the calculated core temperature and the measured core temperature. Absolute error is the best way to calculate the accuracy of the model because it averages the size of each error and weighs each error the same. This will need to be done for a large set of data to get a good idea of how well the model is working. [2]

3.2 Current Progress

In terms of creating the model I have done some research on how to code a linear regression model. This research ended up being very complicated and well beyond my realm of understanding. But since there was difficulty meeting with both the TA and the client at the beginning of the term, I had to try and understand

it until I could get more guidance. Finally at a meeting with Bill I was able to ask him how I should go about creating the model. He sent me a link to a website of a python library called scikit-learn which has pre-made functions that create a linear regression for you. This was extremely helpful and I was able to read up on the library to understand how to properly use their functions. I have gotten started on this code and have it working where it will read in an array of fake skin temperature data and core temperature data from a file and output an equation to model the relationship between the two. I have tested my program by using the same data in excel to create a linear model and comparing the model that comes from my program to the one from excel to see how similar they are. At this point I am now waiting for real data from the camera to create a real model.

For the model testing I started to write some code that will take in all of the estimated temperatures as well as all of the corresponding measured temperatures and then calculate the absolute error. Right now I am using fake data for the portion as well because I don't have a real model yet. Even though the model can't be created and used yet due to the problem with the camera, I have started working on this code just to have it working so I don't have to worry about it later. Once the camera is working, data is collected and the model is created this code will be used and I will find out if our model is good or not. If it is not good I will have to go back and make changes to the model to make it better.

3.3 Remaining Work

Once there is real data from the camera and a model is created it will need to be evaluated by the model testing program. If the model has a high absolute error then I know that it is not a good model. To improve the model the first step that I will take is to add information from reference temperature sources as well as from a possible humidity sensor if the mechanical team ends up getting it. This will be the first step to create a better model. If the absolute error is too high with this additional information then I will try a different type of model. It is hard to say for sure at this time what model I will try next but I am thinking of trying a polynomial regression since it has a bit more complexity. If I get to this point I will use the information I have learned as well as our client and TA to make the decision of which model to try next. This process will most likely be a long one of trial and error but I hope that I can find a model that will have a reasonable absolute error such as 40 percent or lower.

3.4 Problems and Solutions

The main problem I am facing right now is the lack of data. Our camera is having a hard time calibrating and because of this all of the temperatures it senses are much higher than they should be. This is a problem because I can not truly work on my portion of this project until the camera is working and I have the data. A consequence of this problem I am going to face is that I will be limited on time when it comes to fiddling with the model to make it as accurate as possible. I know that this process will take a long time and it is possible that I will not be able to get a high accuracy especially if the camera takes a long time to fix. To solve this problem, I will do as much research before hand so that I have a good idea of how to work with these models.

4 Production Mode, Brian Huang

4.1 Recap and Goals

My part of the project is mostly the ending components and bringing the project together as a whole. I am responsible for the user interface, the production mode, and the evaluation. The production mode and the user interface both require having the all previous portions to be done to be created completely. However a simple skeleton can be created for each of these pieces, where the code for the image analysis and the model could be inserted. The evaluation is meant to determine how well the model is working and determining what threshold temperature should be used to determine what qualifies as a fever.

4.2 Current Progress

I have not started working on the production mode yet as it requires the previous portions of the project to be completed before this piece can be made. The production mode depends on the previous pieces of the project to be worked on and finished. This is the final working piece of the project. As a whole, it takes in a thermal images and then outputs an estimated temperature. Without the images analysis and the model this piece can not be finished.

For the evaluation portion of the project I have created a small program that creates a Receiver operating characteristic curve that can show how well our model is doing. What it does is it plots the true positive value against the false positive value at all different thresholds to see at what threshold we need to use to minimize false positives, or maximize true positives. For this project however, we want to minimize false negative values so I need to change my program to plot false negative against true negatives. This should be an easy change as I simply need to adjust what data the program uses to plot.

I have made very little progress in creating the user interface for the project. The plan for the user interface is to have it be in command line, and if we have time we could improve it to something more. So far the user interface is mostly in the image analysis portion, where you type in the program name and use command line arguments to select which images we want to use. The user interface will most likely be something similar to that, with maybe more options for different modes.

4.3 Remaining Work

We have the image analysis portion of this project complete, so I am able to fit that into the framework for the production mode when I finish it. We have still yet to make progress on the model portion of this project, but we are slowly figuring out how to progress. Our image analysis is complete, but we would like to test it against some thermal images. However, the thermal camera that is producing the images has some calibration issues, so we have recently been working on fixing the calibration on the camera increasing the amount of images that it calibrates on.

A evaluation method is complete, however I need to check my program to see if it is actually plotting correctly, and check if my data generator is creating data correctly. I also plan to modify my data generator to have a certain correctness, so I can see if my ROC curve program is actually working. I am also planning on

looking into different python libraries to see if there is something built into a library that I could check my program against. If there is a python library that generates an ROC curve I would consider using that instead of the one I wrote. I also plan on looking into more evaluation methods, so we are not just relying on one method.

Currently the user interface that we have is built into the image analysis portion of this project. It is in command line and takes command line arguments. However, we have recently found out that the software the camera is written in uses C++ and we may have to use it to take images with our program. We made need to create something more complicated that uses bash commands to redirect the output from the camera software to our program.

4.4 Problems and Solutions

The production mode depends on the previous pieces of the project to be worked on and finished. This is the final working piece of the project.As a whole, it takes in a thermal images and then outputs an estimated temperature. Without the images analysis and the model this piece can not be finished. I have been thinking about creating a simple framework for this piece of the project. I would simply create a skeleton code where I could simply insert the other pieces of the code I need. I could create some fake data and a fake model for my framework to use, so that it is able to compile and run on it's own. I just need to make sure that the inputs and outputs of each piece match with what they will actually be.

For the ROC curve to work properly I need real data to feed into it. I also need a lot of data to have a good estimate of what the threshold needs to be. The curve I have created also seems to be plotting backwards, so I need to check if my conditions for true positive and false positives are switched. The ROC curve also needs data to plot, so I created a small program that creates a data for me. It creates a list of patients and have a real temperature, a measured temperature and says whether or not they actually have a fever. However this does generates completely random data so it is not an effective method for testing my ROC curve. What I plan on doing is modifying my data generator to produce a set of data that has some percentage of correctness. That way I could test if my ROC curve is actually working or not.

The user interface does not really depend on the different portion of the project, but I would like to see how they all fit together before creating something. I want to avoid the problem of having created the user interface, and not having the correct type of inputs for the different pieces. The user interface works directly in conjunction with the production mode, so any progress on the production mode is also progress on the user interface. Both of these pieces depend on the other pieces to be completed, so this will be done towards the end.

5 Conclusion

This term good progress was made on the image processing software. However, there has been a major problem with the camera that has yet to be fixed. This camera being fixed is crucial to the success of the project. Another major issue was understanding how to come up with the mathematical model but this was resolved through communication. Since the rest of the project has been prepared, once the camera is fixed and data is collected it should be relatively straight forward to finish this project. If the camera can not be fixed in time and data can

not be collected fake data may be used in order to have something to show at expo. However, the main goal is to have a functional prototype of the model by the end of this term.

REFERENCES

[1] C. Maimon, "Claude's technical review," https://github.com/maimonc/Ebola-Virus-Project/blob/master/Documents/TechReview/Claude_Maimon_Tech_Review/Tech_Review_Claude_Maimon.pdf, 2017.

[2] B. Beauchamp, "Bianca's technical review," https://github.com/maimonc/Ebola-Virus-Project/blob/master/Documents/TechReview/Bianca_Beauchamp_Tech_Review/Technology_Review.pdf, 2018.