Using the Remotion Plus Platform to Predict Emotion and Attention State from Smartphone Behavioral Data in Remote Usability Tests

By

Dinithi Silva-Sassaman

Electrical Engineering (BSEE) and Mathematics (BSc), University of Texas at Arlington, 2016

Thesis

Submitted in partial fulfillment of the requirements for the Degree of Master of Science in the Graduate Program of Biomedical Engineering at Brown University

PROVIDENCE, RHODE ISLAND

February, 2021

This thesis by Dinithi Silva-Sassaman is accepted in its present form by the Center for Biomedical Engineering as satisfying the thesis requirements for the degree of Master of Science

Date December 7, 2020 Jeff Huang

Dr. Jeff Huang, Advisor

Recommended to the Graduate Council

Date 12/8/20

Dr. Joseph Crisco, Reader

Date 12-9-20

Ian Gonsher, Reader

Approved by the Graduate Council

Date 12/9/20 Marissa Gray

Dr. Marissa Gray, Biomedical Engineering Master's Program Director

Acknowledgements

I would like to take this chance to thank everybody who helped me on this journey.

I would especially like to thank my advisor Dr. Jeff Huang. I still remember the first time when I asked if I could work with him on a project. Jeff was very welcoming and trusted me with a big project which later became Remotion Plus. This opportunity means so much to me.

Next, I would like to thank my thesis committee, Ian Gonsher and Dr. Joseph Crisco, who always made time for me and gave me very good insights to improve my research.

Next, I would like to thank my dear friend Jing Qian who helped me a lot with the paper submission process by giving very important feedback and reviewing it. Needless to say this was very handy for wrapping up the research and my thesis.

Next, my loving parents and my brother. Thank you for all of your support and encouragement.

Finally, my dearest Paul, thank you so much for listening to all my research ideas and being there, for encouraging me when I felt like it was very difficult to proceed, and for proofreading my work.

Table of Content

Table of Figures

Table of Tables

Chapter 1: Introduction

Smartphone-based interactions have become a crucial part of our lives. The number of smartphone users has surpassed three billion. China, India, and the United States are the countries with the highest number of smartphone users [1]. With this in mind, smartphones have opened up a whole new research field for understanding human emotions and attention.

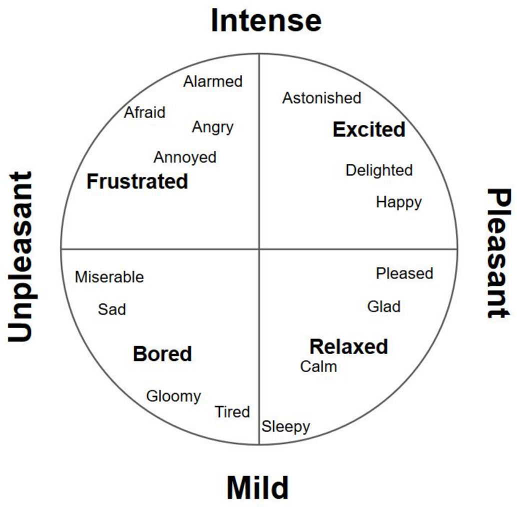
Human emotions and attention are very important in decision making and developing UX interfaces [2]. Previous work to recognize human emotions has been done using facial expressions [3,4], voice patterns [5,6] and by analysing context [7]. However, with the development of smartphone technology within the past few decades, mobile phones have advanced dramatically with different sensors including cameras, microphones, movement trackers, location trackers, and ambient lighting sensors becoming standardized in most devices. Even though there are a lot of benefits to collecting data from these readily available sensors, it can be difficult to calculate user cognitive state from them.

With advancements in biometric technology, it is now possible to use capacitive touch screens to measure touching dynamics including position and pressure. However, pressure is an inferred measurement from the touching dynamics calculated by measuring the surface area covered by the finger [8, 9] as capacitive touch sensors cannot measure pressure directly. This could lead to precision and proximity errors. Moreover, capacitive touchscreens are not usable by everybody, including people with conditions like zombie fingers [10]. Capacitive touchscreens work by producing an electric flow when the user keeps their finger on the screen.

However, people with zombie fingers [10] have thick calluses on their fingertips, and the dead skin does not conduct electricity. This is a common condition among carpenters and guitar players among others. Additional examples of people who have problems when using a capacitive touchscreen include people with long fingernails and people wearing gloves. However, if the users are able to use a capacitive touchscreen, recent research studies have found that users are more effective in choosing targets correctly across various screen sizes if the screen has capacitive touch capabilities [11].

There are additional problems when using the screen on the front of the device as both a touch input and a visual output including fat finger issues [12] and occlusion [13] problems. With all of this in mind, research has shown that there are benefits to using the back of the phone as an alternative touch input.

This research introduces Remotion Plus, a platform for analyzing smartphone users' emotion and attention levels while they are using smartphone apps remotely. The smartphone can remotely stream data to the Remotion Plus computer app. The computer app has the ability to annotate four qualified emotion states and three qualified attention levels. The emotions selected for this platform are excited, relaxed, frustrated and bored. These are taken from the emotion circumplex model [14] and each emotion is from a different quadrant to eliminate overlap of the emotions. The available attention levels are low, mid and high.



Figure(1) - Emotional Circumplex by Russel et al.

The Remotion Plus computer program is also capable of full playback of the remote recordings including microphone audio, screen visuals, device orientation, and rear finger placement and pressure.

In addition to the mobile and computer platform, this research consists of a Tensorflow Keras deep learning model consisting of dense and 2D convolution layers to predict user emotion and attention levels automatically. Using just the orientation and pressure data, the model's prediction accuracy for attention and emotion consistently reached greater than 80%. Two human analysts with UI/UX and behavioural data analysis experience were also asked to predict the emotion and attention level of actors given the same data. On average, their scored accuracy was between 20% and 40%, with an inter-rater reliability between the analysts of 50%.

In brief, this research contributes Remotion Plus, a system dedicated to remote usability testing for UI/UX researchers, and the Remotion Model, which has the ability to predict emotion and attention state with an automated machine learning model. More details on the Remotion Plus platform can be found here.

Background

Emotion Based Recognition Techniques

Previously, research has been done to understand smartphone-based emotion recognition. This can be mainly divided into two main sections:

- smartphone-based recognition techniques
- touch-based recognition techniques

Smartphone Based Emotion Recognition Techniques

One example of emotion recognition with smartphones is EmotionSense [15] where they introduced a smartphone platform that performs audio-based emotion recognition. Next, Pielot et al. introduced a smartphone app for Android called Borapp which infers boredom from mobile phone usage [16]. In this research, they collected data from users in two ways: always collect data, and collect data when the phone is in use. After collecting data, they performed machine learning techniques on the data features they gathered. Similarly, LiKamWa et al. created MoodScope which infers the smartphone user's mood by analyzing data usage from various applications and data usage patterns such as phone calls, usage of applications, SMS, email, location, and web browsing [17]. These researchers asked the participants to annotate the data by entering their mood four times a day.

Gosh et al. analyzed typing characteristics when using smartphones [8]. Among these characteristics were parameters such as typing speed, the number of mistakes made while typing, and special character usage while typing. They also recorded finger pressure data and had their 24 participants self report their emotional state from four options: happy, sad, stressed, and relaxed [8]. Next, they used this data to train a multi-task learning based neural network to predict emotions. According to Gosh et al., the Neural Network reached an average accuracy of 84%. This accuracy is very similar to Remotion Plus' results, but Remotion Plus can also make predictions even when the user is not typing.

Touch Based Emotion Recognition Techniques

Gao et al. predicted their subjects' emotions using finger stroke data points collected from participants who played twenty sessions of the Fruit Ninja game [9]. Afterwards, the participants were asked to fill out a self-assessment questionnaire which was used to annotate the data. Similar to this research, Remotion Plus also gives participants a questionnaire. However, to combat losing data due to forgetfulness, all of the ground truth data was gathered while participants were engaged in the task and immediately after each session. Remotion Plus encouraged participants to think aloud as they were performing various tasks, and this audio data was recorded with their smartphone's microphone. The screen's video was recorded as well. Finally, participants were asked to fill out the questionnaire after each session to define the ground truth. While this occurred, the previously recorded audio and video was played back to remind the participants of any possible omissions.

Furthermore, Gao et al. measures finger pressures with a capacitive touchscreen. A capacitive touch sensor deduces finger pressures by assuming that an increase in contact area size means that the amount of pressure has also increased. This can lead to inaccuracies of data and precision as this assumption highly depends on how the user touches the screen. For instance, a touch from a vertical finger has a significantly different shape profile than a horizontal finger's touch. Simply changing the alignment of the finger while keeping consistent pressure on the capacitive touchscreen would trick the device into thinking the amount of pressure has changed when it has not. Remotion Plus is able to avoid this problem by using a conductive pressure pad that can directly measure the amount of pressure at different locations, separating the touch profile from the pressure readings.

Moving on, Wac et al. [18] developed an Android app to evaluate swipe, scroll, and text input data to evaluate users' stress states. The users were asked to enter data several times a day under different conditions. However, the app did not gather any intrusive privacy data such as their call or texting log or their browsing data. Similarly, Remotion Plus did not access any intrusive privacy data from the participants as Remotion Plus recorded data only when participants had manually activated recording. This does mean that Remotion Plus did not record data multiple times a day from participants but instead only during dedicated sessions.

Finally, Lee et al. [19] used sensors and typing features from Android smartphone users when they used social networks to predict emotions. The study was conducted with one male participant in order to log the emotional state of his daily life. The gathered device state features were divided according to behavioral patterns like shaking of the device, speed of typing, and user context. According to Lee et al., there is a correlation between typing speed and emotions. However, their study design only involved one participant which limits their system's diversity.

All the work above analyzes finger stroke data. However, there are limitations when gathering finger stroke data on the front of the screen from users who cannot interact with capacitive touchscreens. This includes people who wear gloves, have long fingernails, or have the condition "zombie finger" [10]. The latter condition is common among people who work with their hands and form calluses on their fingers, like guitarists and carpenters. Calluses are made of dry dead skin, and thus cannot conduct electricity. As capacitive touchscreens are looking for electrical conductivity, this makes it difficult for anyone who cannot touch live skin to the screen to perform finger-stroke activities. However, as previously stated, Remotion Plus uses a conductive pressure pad on the back of the phone, which eliminates all the above problems with the added benefit of focusing only on finger placement on the rear of the smartphone.

Attention Based Recognition Techniques

Paletta et al. introduces an eye-tracking system that provides the ability to detect where the user visually focuses on the smartphone [20]. This is later converted to a heatmap which shows where on the screen the user paid the most attention. However, this system faces one of the most common issues with computer vision techniques, where very bright or dark settings makes it incredibly difficult to gather the gaze positions. Next, Pagliari et al. introduced a system to detect smartphone users' attention levels by monitoring power management patterns and classifying them using a Convolutional Neural Network [21]. This system is limited by the fact that it cannot recognize the intermediary attention level, only distinct on/off states. It is further limited by the fact that even if the screen is active, it doesn't necessarily mean the user is paying any attention. For instance, a user might be mindlessly scrolling through a news feed without actually paying any attention.

Back-of-Device Phone Interaction

Smartphones mainly let users touch the screen on the front of the device to interact with it. Even though this is a very successful method, there are limitations. For instance, it can be challenging to precisely touch a position on the screen for people with “fat fingers” [12] or “zombie fingers” [10]. Furthermore, when the screen combines both visualization and interaction into a single interface there can be an occlusion problem [13]. As an alternative, Wigdor et al. reported that many smartphone users prefer touching the back of the phone as it reduces the occlusion problems and boosts the ability to make multi-finger inputs.

The study conducted by Le et al. gave three different tasks to their users: typing using the device’s virtual keyboard, watching a video, and reading text. In order to understand the user's finger positions [22], they intermittently had the user hold still, compared the finger positions to a color coded grid on the back of the phone, and then manually recorded the data. This is an unnatural way to interact with a device and is prone to human error. However, Remotion Plus uses a rear conductive pressure pad that automatically calculates positions and pressures which increases accuracy. Furthermore, it lets the user use the phone without intermittently freezing, being distracted, or being conscious about holding the phone.

BackXpress is another example of rear phone finger tracking but with real time interaction [23]. In this study, users were able to interact with applications by putting different amounts of pressures on the back of the phone in addition to the normal inferred screen pressures. The rear touch area was purposefully limited to specific subsections like left vs right or individual fingers such as the ring, middle, and index finger, and the users had to keep track of which section they were touching to correctly interact. Contrariwise, Remotion Plus users didn't have to consciously keep track of their finger positions and they were able to use the phone as they normally would.

Lastly, Infinitouch [24] also captures finger positions and pressures in real time across the back of the phone. They attached a second capacitive touchscreen to the back of the phone so they could capture capacitive touch elements. However, once again, capacitive touch infers finger pressure from finger touch surface area which is less precise. Additionally, as rear finger touches usually consist of more than just the fingertips, this could significantly distort the pressure readings. Remotion Plus' pressure pad can measure pressure directly at multiple locations at once, but the position data is at a lower resolution than most commercially available capacitive touchscreens.

Posture and Grip in Smartphone Use

With the increased flexibility of smartphone usage, users can use smartphones as they are walking, sitting down, or while taking public transportation. With this in mind, people hold their device depending on the situation or depending on the tasks they are engaged in. According to [25] there are three main ways of holding a phone

- One handed
- Cradled
- Two Handed

In a one handed hold, the user holds the phone with one hand and touches the screen with the same hand. A cradled hold involves holding the phone with both hands but only touching the screen with one hand. Two handed holds require the user to be holding the phone with two hands and touching the screen with both hands.

These categories can be further divided into subcategories based on the placement of the thumb and the size of the screen.

According to Yoo et al. [26] in one handed holds, rear placement of the index and the middle finger are prominent, and the grip shifts towards the top edge as the phone size increases.

Next, two handed mode can be divided into two categories:

- vertical split in landscape mode

- horizontal split in portrait mode

The two handed horizontal split is not common practice [25]. However, the actors in the Remotion Plus study used one handed and dual handed holds in both portrait and landscape modes when performing various tasks.

There is additional prior work on understanding grip positions from motions. For instance, Gripsense [27], uses motion data to understand grip position and the orientation of the device. Furthermore, the original Remotion [28] is a system which can understand grip positions. However, these systems cannot automatically predict emotion and attention state. Remotion Plus is able to use the current orientation of the smartphone to train models to make these predictions.

References

[1] O'Dea, S. 2020. “Smartphone Users by Country 2019.” Statista. February 28, 2020. https://www.statista.com/statistics/748053/worldwide-top-countries-smartphone-users/.

[2] Agarwal, Anshu, and Andrew Meyer. "Beyond usability: evaluating emotional response as an integral part of the user experience." In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*, pp. 2919-2930. 2009.

[3] Ekman, Paul. "Facial expression and emotion." *American psychologist* 48, no. 4 (1993): 384.

[4] Friesen, E., and Paul Ekman. "Facial action coding system: a technique for the measurement of facial movement." *Palo Alto* 3 (1978).

[5] Nicholson, Joy, Kazuhiko Takahashi, and Ryohei Nakatsu. "Emotion recognition in speech using neural networks." *Neural computing & applications* 9, no. 4 (2000): 290-296.

[6] Schuller, Björn, Gerhard Rigoll, and Manfred Lang. "Hidden Markov model-based speech emotion recognition." In *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03).*, vol. 2, pp. II-1. IEEE, 2003.

[7] Alm, Cecilia Ovesdotter, Dan Roth, and Richard Sproat. "Emotions from text: machine learning for text-based emotion prediction." In *Proceedings of human*

language technology conference and conference on empirical methods in natural language processing, pp. 579-586. 2005.

[8] Ghosh, Surjya, Shivam Goenka, Niloy Ganguly, Bivas Mitra, and Pradipta De. "Representation Learning for Emotion Recognition from Smartphone Keyboard Interactions." In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pp. 704-710. IEEE, 2019

[9] Gao, Yuan, Nadia Bianchi-Berthouze, and Hongying Meng. "What does touch tell us about emotions in touchscreen-based gameplay?." *ACM Transactions on Computer-Human Interaction (TOCHI)* 19, no. 4 (2012): 1-30.

[10] Raymond, Chris. 2015. Zombie Finger and Touchscreens - Consumer Reports. June 2, 2015. https://www.consumerreports.org/cro/news/2015/06/zombie-finger-and-touchscreens/index.htm.

[11] Baudisch, Patrick, and Gerry Chu. "Back-of-device interaction allows creating very small touch devices." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1923-1932. 2009.

[12] Siek, Katie A., Yvonne Rogers, and Kay H. Connelly. "Fat finger worries: how older and younger users physically interact with PDAs." In *IFIP Conference on Human-Computer Interaction*, pp. 267-280. Springer, Berlin, Heidelberg, 2005.

[13] Wigdor, Daniel, Clifton Forlines, Patrick Baudisch, John Barnwell, and Chia Shen. "Lucid touch: a see-through mobile device." In *Proceedings of the 20th annual ACM symposium on User interface software and technology*, pp. 269-278. 2007.

[14] Russel, J. "Three dimensions of emotion." *J Pers Soc Psychol* 9, no. 39 (1980): 1161-1178.

[15] Rachuri, Kiran K., Mirco Musolesi, Cecilia Mascolo, Peter J. Rentfrow, Chris Longworth, and Andrius Aucinas. "EmotionSense: a mobile phones based adaptive platform for experimental social psychology research." In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, pp. 281-290. 2010.

[16] Pielot, Martin, Tilman Dingler, Jose San Pedro, and Nuria Oliver. "When attention is not scarce-detecting boredom from mobile phone usage." In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, pp. 825-836. 2015.

[17] LiKamWa, Robert, Yunxin Liu, Nicholas D. Lane, and Lin Zhong. "Moodscope: Building a mood sensor from smartphone usage patterns." In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, pp. 389-402. 2013.

[18] Ciman, Matteo, Katarzyna Wac, and Ombretta Gaggi. "iSenseStress: Assessing stress through human-smartphone interaction analysis." In *2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, pp. 84-91. IEEE, 2015.

[19] Lee, Hosub, Young Sang Choi, Sunjae Lee, and I. P. Park. "Towards unobtrusive emotion recognition for affective social communication." In *2012 IEEE Consumer Communications and Networking Conference (CCNC)*, pp. 260-264. IEEE, 2012.

[20] Paletta, Lucas, Helmut Neuschmied, Michael Schwarz, Gerald Lodron, Martin Pszeida, Patrick Luley, Stefan Ladstätter, Stephanie M. Deutsch, Jan Bobeth, and Manfred Tscheligi. "Attention in mobile interactions: gaze recovery for large scale studies." In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, pp. 1717-1722. 2014.

[21] Pagliari, Daniele Jahier, Matteo Ansaldi, Enrico Macii, and Massimo Poncino. "CNN-Based Camera-less User Attention Detection for Smartphone Power Management." In *2019 IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED)*, pp. 1-6. IEEE, 2019.

[22] Le, Huy Viet, Sven Mayer, Katrin Wolf, and Niels Henze. "Finger placement and hand grasp during smartphone interaction." In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 2576-2584. 2016.

[23] Corsten, Christian, Bjoern Daehlmann, Simon Voelker, and Jan Borchers. "BackXPress: using back-of-device finger pressure to augment touchscreen input on smartphones." In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 4654-4666. 2017.

[24] Le, Huy Viet, Sven Mayer, and Niels Henze. "InfiniTouch: Finger-Aware Interaction on Fully Touch Sensitive Smartphones." In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, pp. 779-792. 2018.

[25]Hoober, Steven. 2013. "Column: Mobile Matters." UXmatters. February 18, 2013. https://www.uxmatters.com/mt/archives/2013/02/how-do-users-really-hold-mobile-devices.php

[26] Yoo, Hyunjin, Jungwon Yoon, and Hyunsoo Ji. "Index finger zone: Study on touchable area expandability using thumb and index finger." In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, pp. 803-810. 2015.

[27] Goel, Mayank, Jacob Wobbrock, and Shwetak Patel. "GripSense: using built-in sensors to detect hand posture and pressure on commodity mobile phones." In *Proceedings of the 25th annual ACM symposium on User interface software and technology*, pp. 545-554. 2012.

[28] Qian, Jing, Arielle Chapin, Alexandra Papoutsaki, Fumeng Yang, Klaas Nelissen, and Jeff Huang. "Remotion: A Motion-Based Capture and Replay Platform of Mobile Device Interaction for Remote Usability Testing." *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, no. 2 (2018): 1-18.

Chapter 2: Methods

Remotion Plus System

The Remotion Plus platform was developed to assist researchers and developers in remotely exploring their smartphone users' emotion and attention levels while they are using smartphone apps. The system has 3 main parts

1. Remotion Plus smartphone app

2. Remotion Plus computer program

3. Remotion Plus conductive pressure pad

Remotion Smartphone App

The Remotion Plus smartphone app is built for Android smartphones. Users can download the app and start using it as a background app. This means the smartphone user can use their other apps as they normally would in their day to day lives, and the background app will keep recording and transmitting to the Remotion Plus computer program installed on a remote machine. The app has the ability to stream data such as microphone audio, screen video, and device orientation, as well as rear finger placement and pressure data collected from the pressure pad attached to the back of the phone. However, as this system is designed to be *non- intrusive*, data transmission only begins once the user manually flips a switch in the smartphone app and ends as soon as they turn the switch off.

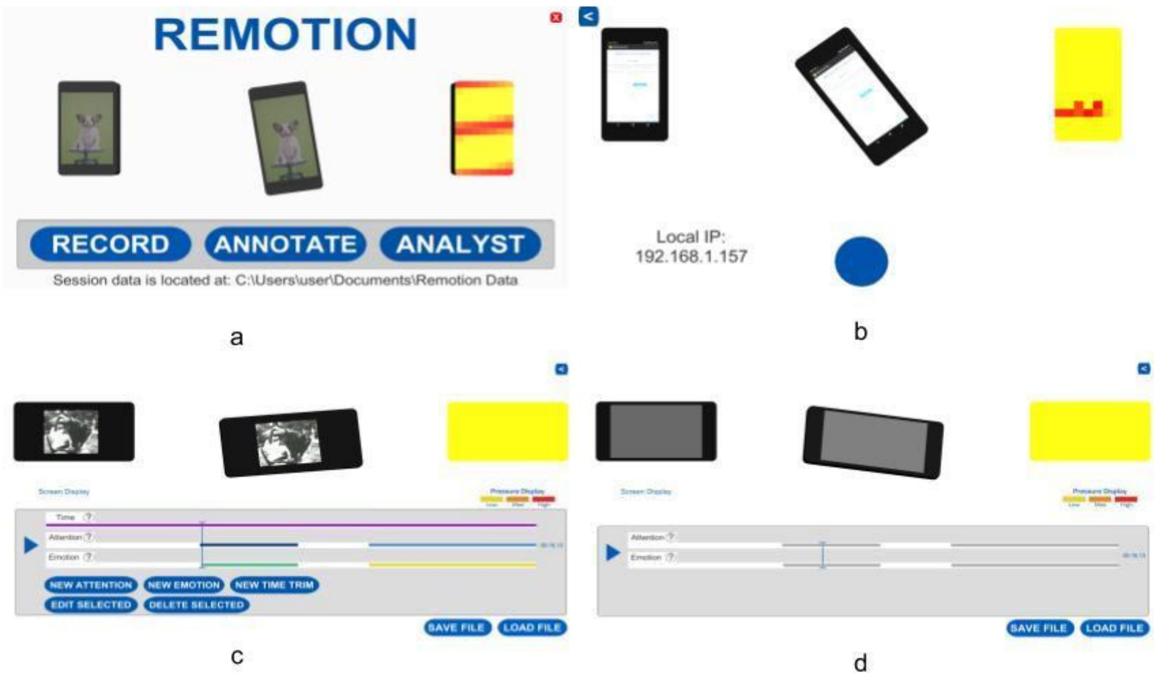
Remotion Computer Program

The Remotion computer program is compatible with Windows and Mac operating systems. It has three modes.

- Record - A researcher can view and record incoming data from the Remotion smartphone app.
- Annotate - A researcher can pause and play any recorded sessions with an intuitive interface that shows all of the recorded data in real time. Furthermore, it lets the researcher define ground truth time segments where the smartphone user was in a specific emotion or attention state according to the remote smartphone user's feedback. These segments are color coded with unique colors for emotion and attention state. The emotions are:
    - Excited - Green
    - Relaxed - Yellow
    - Bored - Orange
    - Frustration - Red

  and the attentions are:
    - Low Attention - Light Blue
    - Medium Attention - Darker Blue
    - High Attention: Dark Blue.

- Analyst - This mode is specifically made for performing prediction studies with hired analysts. Analysts can also play and pause the sessions and they have access to the orientation and pressure data. However, they don't have access to audio or video data as Remotion Plus is focusing on non-intrusive emotion and attention prediction. Similar to the annotate mode, analysts can color code emotion and attention time segments, but these segments are predefined from the ground truth annotations.



Figure(2) - The above figure shows the four different windows of the Remotion Desktop App.

a) Remotion Plus Main Menu b) Recording Screen c) Annotation Screen d) Analysis Screen

Pressure Pad

As Remotion Plus focuses on pressures from the rear of the device and most smartphones do not have rear pressure capabilities, a custom conductive pressure pad was made. The pressure pad consists of two leather pads with a sheet of Velostat in between. On one side of the velostate, there are 14 conductive copper foil strips forming rows. On the other side, there are an additional 7 columns of conductive copper foil. Each strip of conductive copper foil is attached to a wire which is connected to a multiplexer. All of the columns are connected to one multiplexer, and all of the rows are connected to another separate multiplexer. These are connected to an Arduino, and the data from the multiplexers are used to determine how much pressure is currently at each intersection point between the rows and columns, resulting in a grid composed of 98 data points. This is considerably lower resolution than the majority of capacitive touchscreens, but it does have the benefit of being able to directly and accurately detect pressure amounts instead of having to infer the data.

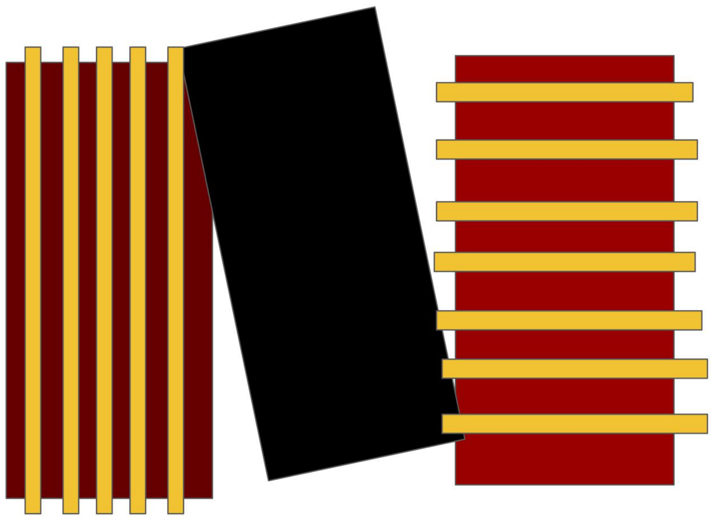


Figure (3) - An exploded view of the two leather pads with copper foil pasted on and the Velostat layer in the middle.

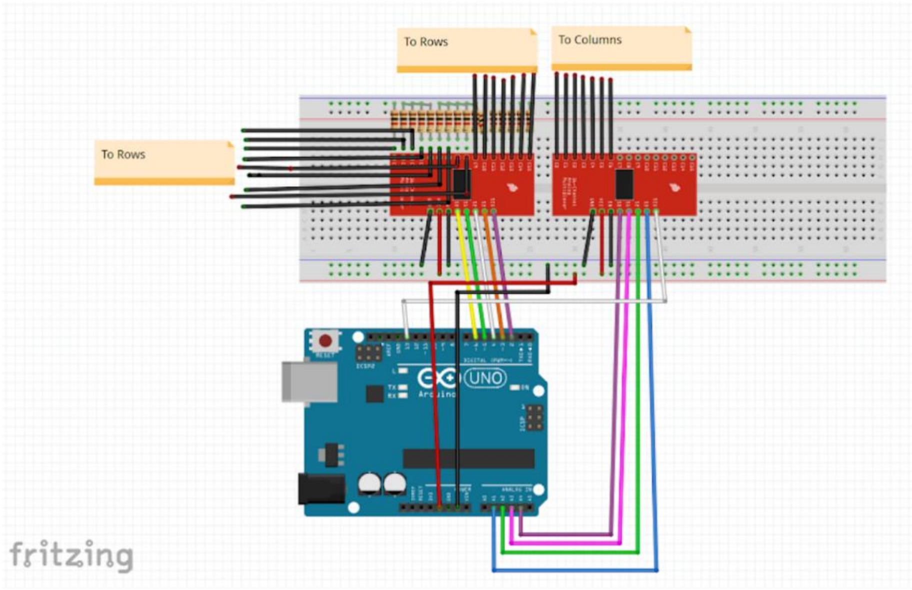


Figure (4) - A basic circuit diagram for the pressure pad to sense the finger pressures and positions

The figure below shows an example of pressure changes on the pad when different touches are applied. Bright yellow represents no pressure, light yellow represents low pressures, orange represents medium pressures, and red represents higher pressures. This is a continuous color scale, and incoming pressures are assigned a color based on linear interpolation inside the range.

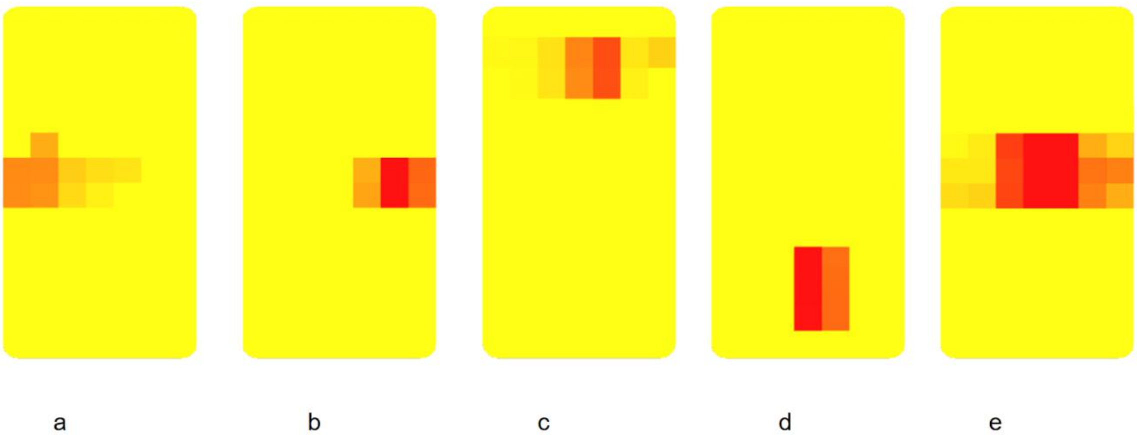


Figure (5) - a) left, b) right, c) up, d) down, and e) middle sides. The values change with increases in pressure, with yellow indicating no touches and red indicating very strong touches.

Communication

In order to gather the pressures and finger placement from the back of the phone, the Remotion Plus phone app communicates with the Pressure Pad's Arduino microcontroller via an OTG cable. The app gathers pressure data from the Arduino as well as microphone audio, screen video, and device orientation data from the smartphone and transmits it to the remote Remotion Plus computer program via UDP. The audio sampling rate is 16000Hz and the video resolution is one third of the smartphone screen's resolution. During this research, a Google Pixel 2 with a resolution of 1920x1080 was used. The video sampling only updated when the screen updated - fast for moving content like a video or a game, and not at all for static content like a picture or text. The orientation sampling rate was near 60Hz.

Design Considerations

When creating Remotion Plus, there were multiple usability design choices to consider. In this section, design choices are explained based upon prior research, the data types chosen for different cognitive states prediction, and system design.

Orientation Data

As previous research found, there is a correlation between accelerometer data and emotions [1,2]. However, there are limitations in gathering data this way as users had to actively move in order for these sorts of systems to be able to predict emotions. These systems are less helpful if the user is sitting, lying in bed, or standing still.

While Remotion Plus uses this data, it is with the goal of measuring the relative orientation of the device compared to the user which captures additional data when the user is not actively moving. As most smartphone phones have a gyroscope and an accelerometer, this research uses these two modalities to provide the quaternion orientation to determine which direction the user is facing.

UDP Communication

Since the Remotion Plus smartphone app streams high volumes of data, the UDP protocol was used to make the data transmission faster. It is also currently the most common protocol for sending video and audio.

Pressure Pad

This section will be very concise as there was an in-depth analysis of the pressure pad in a prior section. The dimensions of the pressure pad used in this study was 45.7 mm x 69.7 mm.

Data Collection

This section presents the data collection methodologies used in this research. This was divided into three categories:

- Pre-study data collection
- Pilot study
- Primary study

Pre-Study Data Collection

In this section of the research, participants were hired via university mailing lists with the goal of gathering smartphone user data. This research refers to pre-study participants as actors.

The recruited actors' ages ranged from 19-27 years. Before the data collection began, they were presented with a consent form. After completing the form, the actors were asked to place the android phone on the pressure pad, pick both items up, and then use the smartphone as they would in their day to day lives.

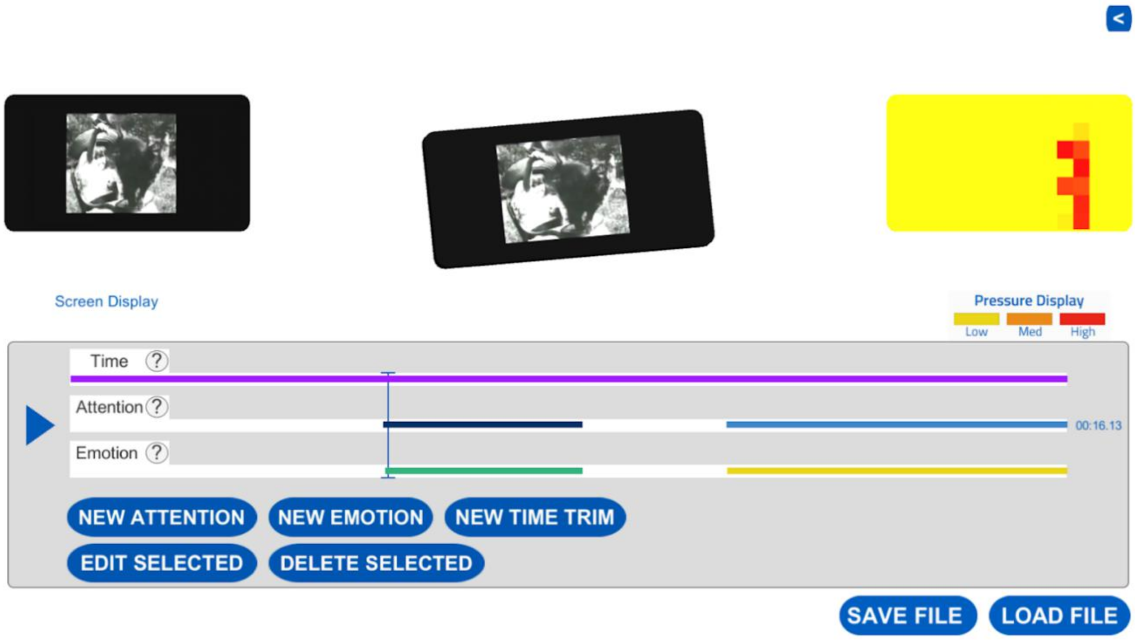
Actors were asked to perform three tasks. These ranged from web activities to entertainment and covered a wide range of orientation and back of device handedness modalities [3,4].

- Shopping App
- Labyrinth Game
- Dual Stick Shooter Game

a b c

Figure(6) - In pre-data collection, 3 tasks were used. a) eBay app b) Maze game c) Dual Stick Shooter game

Each app session was approximately one minute long and actors were highly encouraged to think aloud as they were engaging with the app to improve emotion and attention state recall. However, it was observed that if the actor had to concentrate more on the task, they often did not think aloud. After each session, data was annotated using the Remotion Plus computer program. The recorded data was played back, and actors were asked unbiased questions about their emotion and attention levels while they were engaging in the tasks. They were also asked similar questions from different angles, which gave the actor the room to expand the response. These responses were used to create color coded time segments in the session where the actor was in a specific emotion or attention state. This was then used as the ground truth for the pilot test and the primary study. This research focused only on four unique emotions, which are happy, relaxed, bored, and frustrated, and three unique attention levels, which are low, mid, and high. Once the task was completed, actors were given an exit survey about the session which asked what they were thinking when annotating the data. Finally, they were compensated with $15.



Figure(7) - Screenshot of the Remotion Plus Desktop app annotate window where actor ground truths were defined.

Pilot Test

The main reason to conduct the pilot test was to check the Remotion Plus platform's efficiency and to look for any unknown bugs that occur when using the system. The secondary reason was to determine if the participant analysts could understand the usage instructions properly.

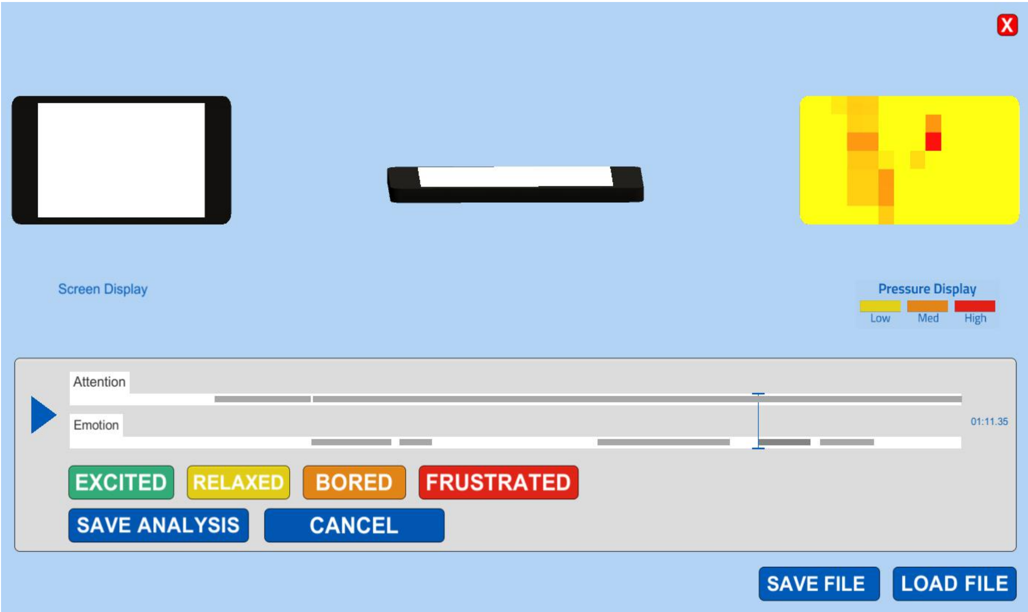
For this part of the research, five analysts were hired via email lists. They were provided with a version of the Remotion Plus software that was limited to just the Analyst section as well as the data from all of the sessions of one anonymous actor. They were presented with a consent form to fill out before beginning to reannotate the actor data. The process was 20 minutes long, and they were compensated with $15 after the session.

Labelling Task

For this part of the research, three analysts were hired via university email lists. Since the general toolkit was targeted at UI/UX researchers, the analysts employed were expected to have a UI/UX background experience. Furthermore, since they would predict behavioral data in this research, they were expected to have some experience with analyzing behavioral data. To make sure the analysts had these features, applicants were limited to students who had taken UI/UX courses at the university.

The labeling tasks happened remotely as remote usability testing is a major goal for the Remotion Plus platform. Analysts were first asked to fill out a consent form. They were then shown the ground truth emotion and attention time segments which were collected from the actors involved in the pre-data collection process, but the correct emotion or attention label was removed. Analysts were also not provided with any audio or video capture data from actor sessions, even though the Remotion Plus platform can handle playback of audio and video capture data. This artificial limitation was chosen because this research focuses on predicting user cognitive states in a non-intrusive way. Since the audio and video data wasn't available, analysts had to make educated guesses on the actors' emotion and attention states based strictly on the current device orientation or rear pressure data.

Analysts had to reannotate every session from every actor. Additionally, they had to do this three separate times while considering a limited set of data: just pressures, just orientation, or both pressures and orientation together. Each category took approximately five hours and about 15 hours all together. After the session, they



Figure(8) - Screenshot of the Remotion Plus computer program analyst window where analysts relabel actor data. It shows the grayed out baseline annotations.

were presented with an exit survey to understand more about their thought process when engaging in labelling decisions. By the end, one analyst dropped out due to the tediousness of the work. The remaining analysts were compensated with $15 per hour.

Remotion Model

In addition to the Remotion Plus Platform, a deep learning model was developed which automatically predicts an actor's emotions and attention levels.

The Remotion Plus Model's inputs are the device orientations and rear finger positions and pressures. Orientations are represented as a four float quaternion and the pressures are 98 floats representing each point on the 7x14 pressure pad grid.

As for outputs, the model can be trained to output either three different attention levels or four emotions states.

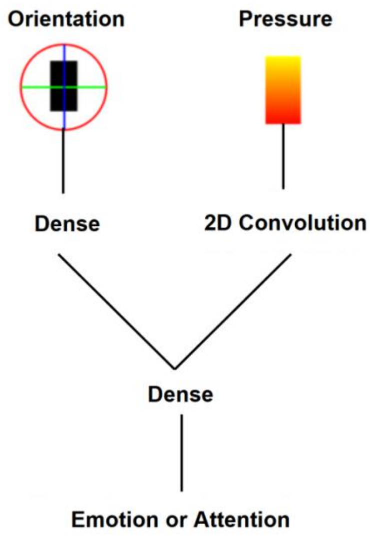
To test the Remotion Plus Model, we trained it on the ground truth data from the actors. However, before starting, some of the actor data was eliminated as their sessions were not annotated properly. In the end, the data from nine actors were selected. In order to confirm consistent results from the model, the nine actors' data was split into nine equal groups to perform kfold validation.

The model itself consists of 4 parts.

- First Part - The orientation quaternion data was processed. This part has an input size of 4, an output size of 4, and it contained four dense layers.
- Second Part - The goal here was to process pressure data. The input size is 98, but it is reshaped to a 7 by 14 array as it represents the 7 columns and 14 rows of the pressure pad. The model treats the pressure pad grid data like a grayscale image with each intersection point representing a pixel. However, as this "image" is very small, it only has a single 2D convolution layer and max pooling layer to keep from losing too much information. Afterwards, the max pooling output is flattened and sent through a dense layer to produce a new output with a size of 21.

- Third Part - The outputs from parts 1 and 2 are concatenated into a single output with a size of 25

- Final Output - The output from the third part are processed by four dense layers to produce the likelihood of each label for either emotions (excited, relaxed, bored, frustrated) or attentions (low, medium, high) depending on what was being trained for.



Figure(9) - Remotion Model in a nutshell

During the initial research phase, analysis on the actor data was first performed with regression analysis. However, the Remotion Plus model ended up providing higher accuracy than the regression analysis, so the final prediction analysis was performed using the Remotion Model. However, regression analysis did provide a baseline to understand which input types are more useful in predicting. It quickly became clear that orientation data was less useful than pressure data, and both pressure and orientation combined was the most useful of the bunch.

The Remotion Plus actor data was also evaluated with other existing models. The following models were selected to represent various types of machine learning and classification.

- Logistic Regression (LR) - Linear Classifier
- Random Forest (RF) - Ensemble Learning
- Support Vector Machines (SVM) - Radial Based Kernel

The Remotion Plus Model and Random Forest consistently produced the highest accuracy percentages. Specific accuracy values are shown later in Tables (1) - (6) in the results section.

References

[1]. Cui, Liqing, Shun Li, and Tingshao Zhu. "Emotion detection from natural walking." In *International Conference on Human Centered Computing*, pp. 23-33. Springer, Cham, 2016.

[2]. Olsen, Andreas Fsrøvig, and Jim Torresen. "Smartphone accelerometer data used for detecting human emotions." In *2016 3rd International Conference on Systems and Informatics (ICSAI)*, pp. 410-415. IEEE, 2016.

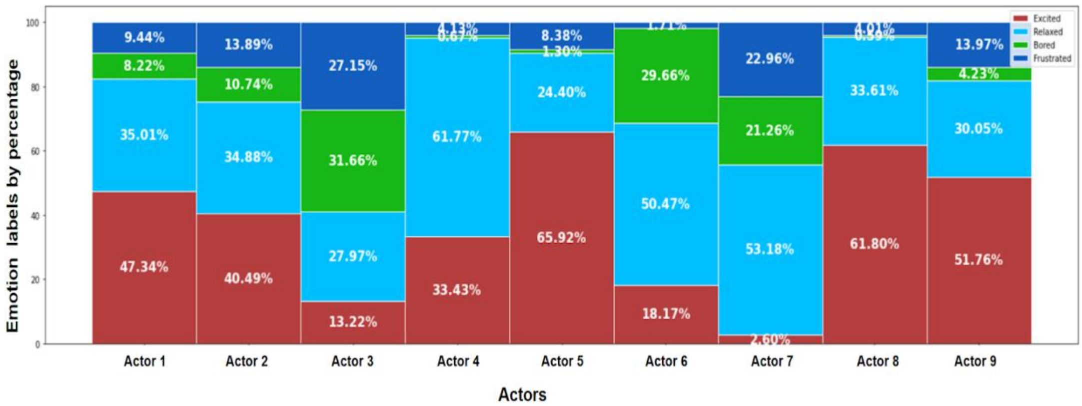
[3] Coursaris, Constantinos K., and Dan J. Kim. "A meta-analytical review of empirical mobile usability studies." Journal of usability studies 6, no. 3 (2011): 117-171.

[4] Do, Trinh Minh Tri, Jan Blom, and Daniel Gatica-Perez. "Smartphone usage in the wild: a large-scale analysis of applications and context." In *Proceedings of the 13th international conference on multimodal interfaces*, pp. 353-360. 2011.

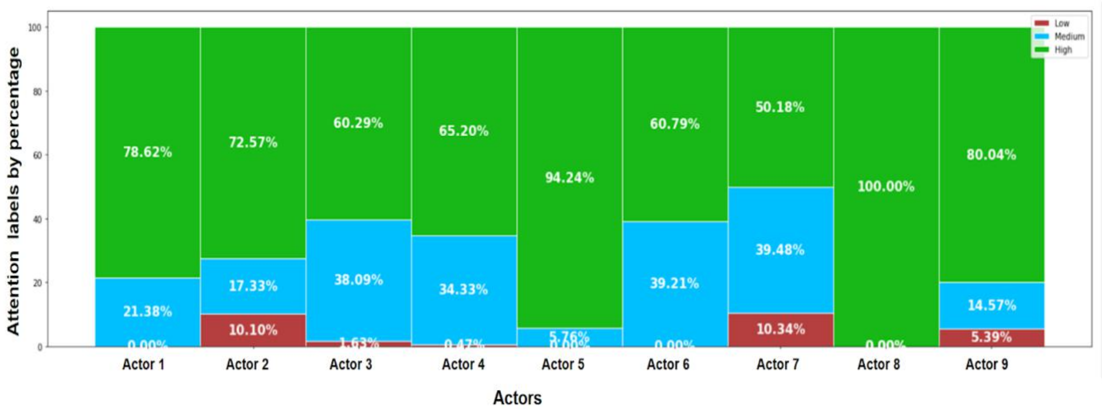
Chapter 3: Results

Pre- Data Collection

A summary of the distribution of actor emotion and attention levels are shown below. As certain labels are more common than others, there is a chance this may have distorted the results. However, the Remotion Plus model does consistently produce higher accuracy results than if the user were to always pick the label with the highest likelihood, which bodes well for its trustworthiness.



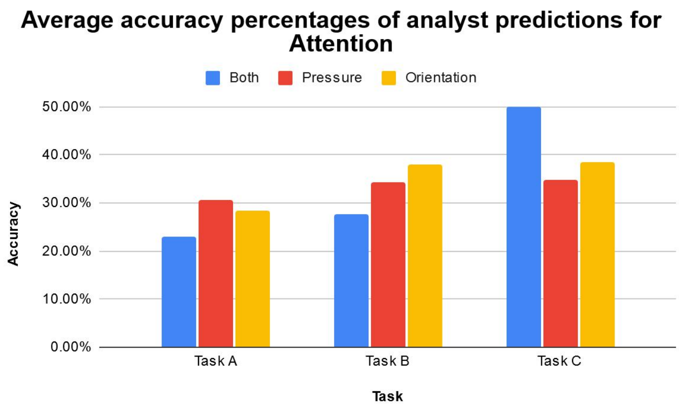
Figure(10) - Distribution of the Emotion labels over different participants. Each bin of the histogram is divided into 4 sections as 4 emotions are considered in Remotion Plus.



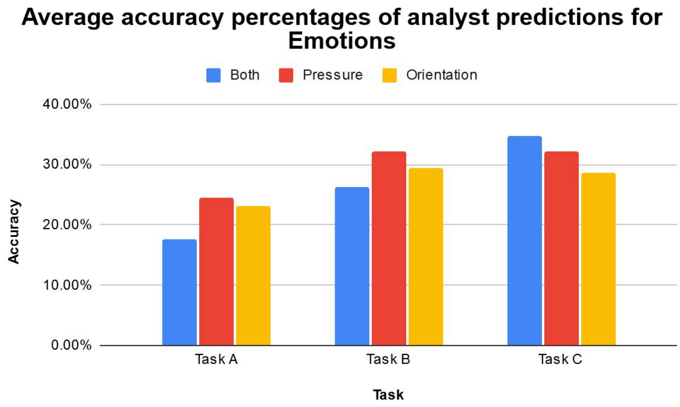
Figure(11) - Distribution of the Attention labels over different participants. Each bin of the histogram is divided into 3 sections as 3 attention levels are considered in Remotion Plus.

Labelling Task

A summary of the results from the analysts labeling task are shown below. Each column represents an averaged accuracy across all analysts for each combination of task, input type, and output type.



Figure(12) - Figure above shows averaged analyst accuracies for attention for three tasks



Figure(13) - Figure above shows averaged analyst accuracies for emotions for three tasks

Attention

Analyst 1 corresponds dramatic movements, changes in orientation, and pressure locations with higher attention. Similarly, they correspond lower attention with little to no movements and changes. Analyst 2 took a similar approach to Analyst 1, assuming that the rate of change corresponded to the amount of attention.

Analyst 1 and Analyst 2 reported that it was most helpful to see both orientation and pressure at the same time, followed by just pressure, and then just orientation. Both analysts reported that when they could see both orientation and pressure, it gave them more context when deciding attention levels. Furthermore, analyst 1 said that they did not find orientation helpful at all in deciding about attention levels. Analyst 2 also reported that being able to see both types of data were helpful because they could cross-check among data types. They said “A lot of pressure in one area might indicate high attention to that area, but when paired with no orientation movement I tended to decide it was more likely to be unchanging due to a lack of attention overall”.

When asked “How often did seeing the orientation/position of the phone or amount of movement of the phone affect choosing an 'Attention'?” Both analysts answered that they consistently looked at the amount of the phone movement, sporadically looked at just the specific orientation of the phone, and usually looked at both combined.

Emotions

Contrary to the analysts’ mostly consistent reasons for labelling attentions, the reasoning for labelling emotions highly depended on the individual. When asked their reason behind labelling the more intensive emotions, excitement and frustration, Analyst 1 mentioned that they conferred high pressure and rapid movements in orientation to those emotions, while Analyst 2 thought it highly depended on the context. For instance, Analyst 2 assumed that if the movements of the phone were slowing down then it was likely excitement. However, if there were rapid movements in orientations and pressure, it indicated excitement too.

When labelling the more sedate emotions, relaxed and bored, Analyst 1 associated low pressure intensities and lesser changes in orientations and pressure locations as relaxed or bored. Analyst 2 reported that they were more likely to label relaxed after labelling excitement and more likely to label boredom at the end of a segment.

When making decisions in labelling emotions, Analyst 1 found the following categories most to least helpful: both (orientation + pressure), pressure, and then orientation. Analyst 2 found the both category to be the most helpful, then orientation, followed by pressure.

Analyst 1 thought orientation was very ambiguous. They reported that, “I felt like orientation did not tell me much. It could just be that the person is used to moving their phone around more or less. In addition, I was sometimes confused how the orientation would be linked to someone's emotion, especially since the phone was constantly shifting orientations.“

Analyst 2 reported that the category Both was helpful as it gave them opportunities to cross check between modes. However, when inferring emotions, they looked at orientation first in order to decide on two possible labels, and then looked at pressure when deciding on the final choice.

Machine Learning Model Results

All of the models’ accuracies were compared across all categories. The Random Forest and Remotion Model performed better with higher accuracy and F1-scores compared to the SVM and LR model in every task in both emotions and attention.

First, consider simple accuracy, the percentage of correct predictions out of all predictions. When it comes to the pressure and both combined categories, the Remotion Model produced higher accuracies than Random Forest. However, Random Forest consistently produced better accuracy results in the Orientation category compared to the Remotion Model.

F1-Scores provide an alternative way to calculate accuracy by using Precision and Recall. The F1-Score can be calculated as follows

$$F1-Score = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (1)$$

assuming $\Pr ecision = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}$ (2) and $\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}$ (3)

Precision and Recall are made up of the following elements:

TP - True Positive

TN - True Negative

FP - False Positive

FN - False Negative

Attention

The below tables show F1-scores for all of the models in each input category (pressure, orientation, both) averaged across all three tasks.

Pressure	LR	Random Forest	SVM	Remotion Model
Task A	81.19%	**90.47%**	82.39%	**90.49%**
Task B	66.53%	**71.76%**	66.75%	**79.70%**
Task C	84.61%	**90.73%**	85.82%	**91.69%**

Table(1) - F1-Score Comparison between three tasks among different models with only pressures as input and attention as output

Orientations	LR	Random Forest	SVM	Remotion Model
Task A	59.29%	**88.18%**	50.22%	**79.66%**
Task B	47.21%	**72.86%**	45.86%	**56.05%**
Task C	72.22%	**87.82%**	72.23%	**80.36%**

Table(2) - F1-Score Comparison between three tasks among different models with only orientation as input and attention as output

Both	LR	Random Forest	SVM	Remotion Model
Task A	81.79%	**92.35%**	83.23%	**92.42%**
Task B	67.21%	**78.31%**	68.06%	**83.25%**
Task C	84.96%	**93.21%**	86.13%	**93.43%**

Table(3) - F1-Score Comparison between three tasks among different models with both orientation and pressure as input and attention as output

In conclusion, when predicting attention, the Remotion Plus Model performed the best out of all the models in the pressure and both categories, and Random Forest performed the best in the orientation category.

Emotions

The F1-Scores for all of the models in each input category (pressure, orientation, both) averaged across all three tasks are shown below

Pressure	LR	Random Forest	SVM	Remotion Model
Task A	72.95%	**85.75%**	73.50%	**85.99%**
Task B	66.53%	**71.76%**	66.75%	**79.70%**
Task C	72.77%	**78.87%**	73.56%	**82.99%**

Table(4) - F1-Score Comparison between three tasks among different models with

only pressure as input and emotions as output

Orientations	LR	Random Forest	SVM	Remotion Model
Task A	49.92%	**84.06%**	33.92%	**63.51%**
Task B	47.21%	**72.86%**	45.86%	**56.05%**
Task C	36.33%	**69.67%**	34.88%	**58.30%**

Table(5) - F1-Score Comparison between three tasks among different models with only orientation as input and emotions as output

Both	LR	Random Forest	SVM	Remotion Model
Task A	74.25%	**88.46%**	75.96%	**88.58%**
Task B	67.21%	**78.31%**	68.06%	**83.25%**
Task C	73.59%	**83.54%**	74.26%	**85.31%**

Table(6) - F1-Score Comparison between three tasks among different models with both orientation and pressure as input and emotions as output

The F1-score results for predicting emotions ended up being just like the F1-scores for predicting attention. The Remotion Plus model scored highest for the pressure and both and Random Forest scored highest for the orientation category.

Model vs Analyst

The table below compares the model predictions and analyst predictions. As the Random Forest model scored highest in the Orientation category for both emotions and attention, the analyst results in orientation were compared with the results from Random Forest. Furthermore, since the Remotion Model scored highest in the pressure and both categories, the Remotion Model predictions were used to compare with the predictions from the analysts.

Attention	M-Pressure	M-Orientation	M-Both	A-Pressure	A-Orientation	A-Both
Shopping	88.7%	84.2%	86.2%	24.4%	23.1%	17.7%
Maze	83.3%	72.5%	80.2%	32.3%	29.4%	26.2%
DSS	85.5%	68.4%	83.2%	32.1%	28.6%	34.8%
Average	85.8%	75.1%	83.2%	29.6%	27.0%	26.2%

Table(7) - The machine learning model's average accuracy at predicting attention (left) compared with the average accuracy of the Analysts (right) for all tasks; M- indicates a machine learning model and A- indicates analysts.

Emotions	M-Pressure	M-Orientation	M-Both	A-Pressure	A-Orientation	A-Both
Shopping	88.7%	84.2%	86.2%	24.4%	23.1%	17.7%
Maze	83.3%	72.5%	80.2%	32.3%	29.4%	26.2%
DSS	85.5%	68.4%	83.2%	32.1%	28.6%	34.8%
Average	85.8%	75.1%	83.2%	29.6%	27.0%	26.2%

Table(8) - The machine learning model's average accuracy at predicting emotions (left) compared with the average accuracy of the Analysts (right) for all tasks; M- indicates a machine learning model and A- indicates analysts.

These tables show that the machine learning models consistently performed better than analysts by more than 40%. In some cases, the models even outperformed the analysts by more than 60%.

Chapter 4: Discussion and Conclusions

The results of this research have shown that there is a correlation between smartphone orientation and rear finger pressures, and user emotions and attention. Furthermore, this research shows that computer models can predict emotions and attention levels of smartphone users as they use apps in real-time. This section presents the results of each procedure.

Labelling Task

The Remotion Plus platform is capable of sending data remotely from a smartphone to a remote computer program. As mentioned earlier, this data includes microphone audio, screen capture video, device orientation, and rear finger pressures. However, transmitting audio and video can be very costly. Even worse, sharing audio or video data can be intrusive privacy-wise.

With this in mind, for the Labelling Task, audio and video data are not shared with analysts. Analysts solely depended on orientation data and pressure data from the actors in their decision making process. The analysts had lower accuracies on average compared to the model due to humans generally not associating hand orientation or finger pressures with emotions or attention level.

It was observed that there was an overall trend where the analysts' predictions generally increased in accuracy over time for each input category. This is interesting, especially considering that the analysts were never given any feedback on how they were doing. It was also observed that analysts' accuracy decreased each time they started working with a new type of data. This is only to be expected when switching from viewing just orientations to just pressures as each input type is very different from each other, but the same decrease also happened when switching from just pressures to viewing both orientations and pressures together.

Table (7) and Table (8) showed the analysts' predictions by task for each input category. For Task A, the analysts scored an average of 24.51% across both emotions and attentions, and for Task B and Task C analysts scored an average of 31.29% and 36.45% respectively. This shows an upward trend in accuracy. The analysts performed the labeling unsupervised, which means that analysts did not know if their answers were correct or not at the time. However, the prediction results still became better with time. It is difficult to determine whether the results of Task A were harder to analyze than Task C or if the analysts naturally evolved in predicting emotion and attention state over time with each input type.

The average analyst accuracy for emotions and attentions across all tasks and input types were 28% and 35% respectively.

As a reminder, this labeling task took approximately 15 hours to complete. At the beginning of the task, three analysts were hired. However, one analyst dropped out due to the tediousness of the task. It is unknown if this tediousness also affected the two analysts who did complete the entire fifteen-hour task.

Analyst Decision Making Process

The analysts were asked "When provided with both types of data, does seeing orientation or pressure affect choosing an `Emotion' more than the other?" Analyst

1 mentioned that it was more pressure than an orientation that affected their decision-making process, where Analyst 2 said it was about half and half.

They were also asked regarding Attentions “When viewing both types of data at the same time, which was more helpful to come to a decision: seeing orientation or pressure?” Analyst 1 answered that pressure was more helpful than orientation while Analyst 2 answered that it was about half and half.

Finally, analysts were asked whether attention or emotion was easier to label for each of the three tasks. Analyst 1 answered that neither was easier to label than the other for all three tasks. Analyst 2 answered that for task A, attention was easier to label. For task B, it was emotions that were easier to label. Finally, in task C, both emotions and attention were equally easy to label for task C.

Compare Contrast Among Models

In the first steps of designing the Remotion Plus model, this research compared the Remotion Plus Model with Linear Regression as a sanity check. The Remotion Plus Model consistently produced better results. The Remotion Plus Model was also compared to other existing models such as Logistic Regression (a linear classifier), Random Forest (Ensemble Learning), and SVM (radial classifier).

When a computer model derives answers, it only considers what it was programmed to do so. As the labeling task only considers rotation and pressure data, it was decided that the computer model should also only be provided rotation and pressure data to provide a one-to-one comparison.

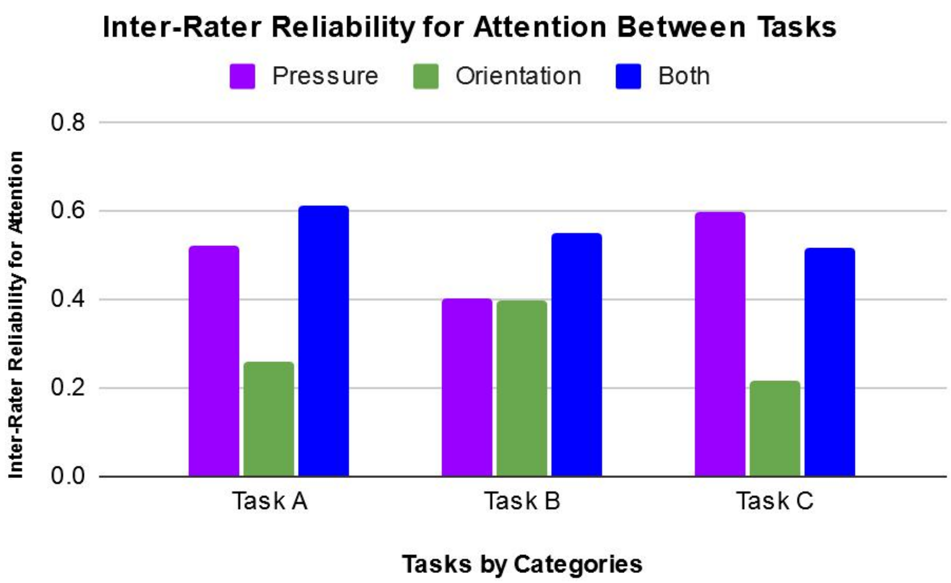
It was observed that Random Forest produced higher accuracies than all of the other tested models in the orientation category for all tasks. Contrariwise, the Remotion Plus model scored higher in accuracy than all of the other tested models when using only pressure as input and when using both orientation and pressure together.

Inter-Rater Reliability

To determine if there was any consistency between the two analysts' responses, we performed Inter-Rater Reliability comparisons.

Attention

When predicting attention levels, the analysts had the most consistent answers for Task A (Shopping) and Task B (Maze) when viewing both pressure and orientation data together, scoring 61% and 55% respectively. Their highest score for Task C (Dual Stick Shooter) was 60% when viewing just pressure data.

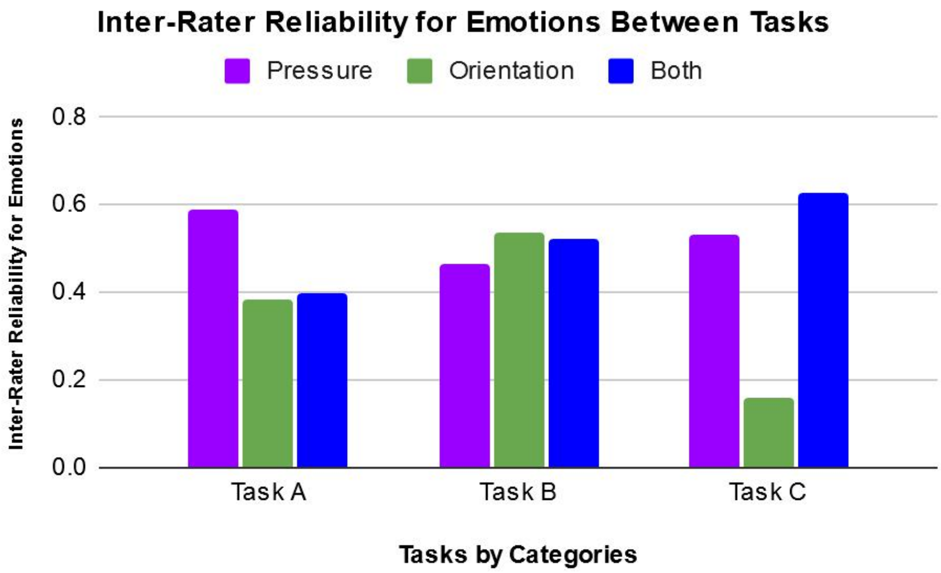


Figure(14) - Inter-Rater Reliability for Attention

Emotions

When predicting emotions, the analysts scored the highest inter-rater reliability for Task A (Shopping) with pressure data at 59%. They then scored the best with 54% for Orientation in Task B (Maze) and 63% with both data types together in Task C (Dual Stick

Shooter).



Figure(15) - Inter-Rater Reliability for Emotions

Limitations

This research has a few limitations. Firstly, the age groups of the actors involved in Pre-Data Collection were limited to just college undergraduate and graduate students. Secondly, the study's tasks were limited to just three different specific activities, whereas in real life the types of apps used would be very diverse. Third, the required amount of attention for each task depended on the individual, which could result in a bias. Furthermore, the ground truth could be incorrect if actors were reluctant to mention their true emotions or attention levels during the pre-data collection process. For instance, if the actor was feeling bored during a session and felt embarrassed about reporting it honestly, they may have given another emotion instead. It is also unclear if other factors such as cultural backgrounds [1], age [2], or race could influence emotion and attention state. Some emotions might also last longer than the other emotions [3], and this can differ from individual to individual. Lastly, the study's sample size is relatively small, so it may be difficult to generalize these results to other situations.

Conclusion

This research shows that emotions and attention levels impact the way smartphone users use different apps. To better analyze this data, the Remotion Plus platform was created. Remotion Plus has three main parts: a computer program, an Android phone app, and a physical pressure pad that goes on the back of the phone. The android app can stream data such as microphone audio, screen capture video, smartphone orientation, and finger pressure intensities. The smartphone app streams this data to the remote computer program remotely using the UDP protocol.

Data was collected from nine actors with the Remotion Plus Smartphone app. The actors were asked to perform three tasks. After completing these tasks, the actors used the Remotion Plus computer program to annotate their sessions with color coded emotions and attentions. These annotations were used as the ground truth for all future phases.

From here, a pair of analysts were shown the recorded data and asked to reannotate the ground truth labels. The analysts were required to reannotate the entire dataset three separate times, starting with being able to see only the orientation data, then just pressure, then with orientation and pressure data together. This was done to understand how accurately humans can predict emotions and attention given this limited dataset.

Analysts found that when predicting emotions and attention, seeing both the orientation and pressure data together was helpful as they had more context. It was also determined that the analysts were worse at predicting emotions than they were at predicting attentions.

To prove that a machine learning model can predict emotion and attention state with a higher accuracy than humans, the Remotion Plus Model was created. This model was also compared with existing machine learning algorithms like Logistic Regression, Random Forest, and SVM. Remotion Model consistently produced the highest results in the Pressure and Both input categories across all task and output types while Random Forest had the highest accuracy predictions with just Orientation as input.

This research found that prediction accuracy depended on the individual, not on the specific data being analyzed. However, when using machine learning to predict emotion and attention accuracies, some tasks were easier to predict by the model than others.

When predicting attention, the Remotion Plus Model reached accuracy percentages above 80 and 90 percent, while humans remained in the 20s and 30s. The Remotion Plus Model scored approximately 60% more accurately than the average human being. However, humans did become better with time as they became better at pattern recognition.

This work shows that it is possible to correlate user emotional and attention state when using apps to specific device and user physical states. The Remotion Plus Platform and Model will be valuable tools for researchers and developers who are testing their newly developed apps.

References

[1] Russell JA. Culture and the categorization of emotions. Psychological bulletin. 1991 Nov;110(3):426.

[2] Phillips LH, MacLean RD, Allen R. Age and the understanding of emotions: Neuropsychological and sociocognitive perspectives. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences. 2002 Nov 1;57(6):P526-30.

[3] Philippe Verduyn and Saskia Lavrijsen. Which emotions last longest and why: The role of event. Motiv Emot,39:119–127, 2015.

Appendix A

A few selected plots for emotions and attention levels are shown below with F1-Score and Accuracy comparisons for different models.

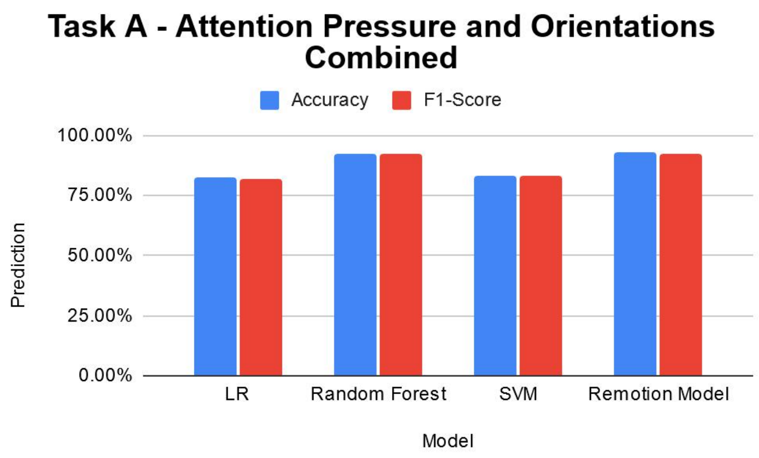


Figure (1) - Task A Attention- Pressure and Orientations Combined

Remotion Model Performed better than all the models.

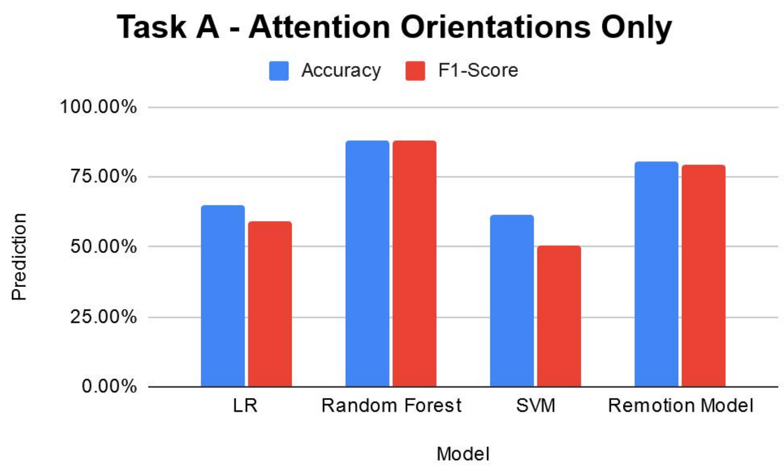


Figure (2) - Task A Attention- Orientations

Random Forest Performed better than all the models.

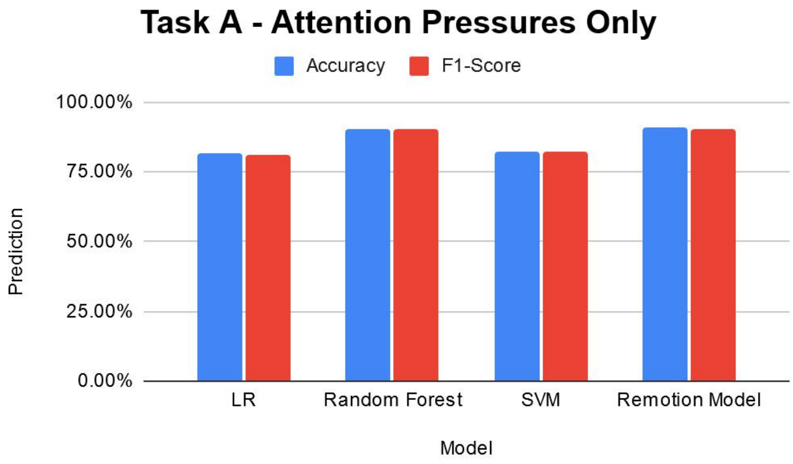


Figure (3) - Task A Attention- Pressures

Once again, Remotion Model out performed all the other models.

Analyst Results

Figure (4) - Analyst 1 a) Emotion Accuracy percentage for three tasks b) Attention

Accuracy for three tasks.

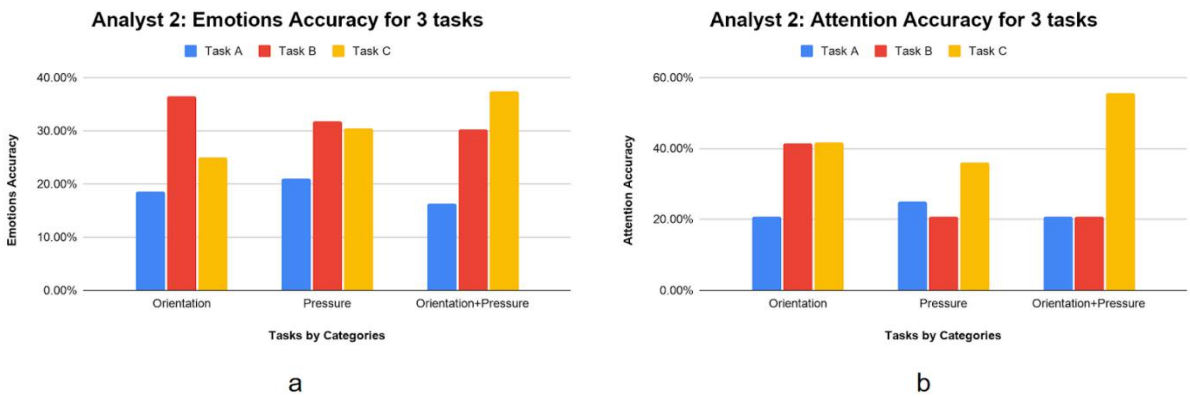


Figure (5) - Analyst 2 a) Emotion Accuracy percentage for three tasks b) Attention Accuracy for three tasks.