Generating a personalized diet and fitness routine using emotion, activity, and food tracking

D.S. Jiffry   
Software Engineering Department  
Sri Lanka Institute of Information TechnologyKandy, Sri Lanka.  
dsjiffry@gmail.comPremarathna K.S.N.  
Software Engineering Department *Sri Lanka Istiture of Information*  
TechnologyKandy Road, Malabe, Sri Lanka  
sathiranimhana@gmail.comYahathugoda S.R  
Software Engineering DepartmentSri Lanka Institute of Information  
TechnologyKandy Rd, Malabe, Sri Lanka  
radishanyahathugoda@gmail.com

*Abstract*—The evolution of technology has reduced the need for physical activities, because of this we see people feel the need to allocate time in their day for physical activities. To assist in this requirement, in this paper we propose an android app, which will monitor a user’s lifestyle and emotional wellbeing and based on that data, generate workouts and diet plans. Emotional wellbeing is monitored via the front camera, by performing analysis on the user’s facial expressions. Lifestyle is tracked by performing activity tracking on the user throughout the day. Furthermore, the app would identify common unhealthy practices and attempt to assist the user in avoiding them.

Keywords—fitness app, emotion analysis, activity tracking, deep learning, diet plan, workout plan,

# Introduction

Technology has made modern lifeless physical, the introduction of motorized transport and the internet means the physical activities a person needs to perform daily have reduced.

Studies show that lack of physical activity can even lead to a premature exit from employment [1]. Obese and overweight individuals are more likely to exit their paid employment via disability pension, the study recommends long-term interventions to encourage physical activities in the population [1]. Lack of physical activity can also be a primary cause for chronic diseases [2], The study says that the body can mal-adapt to insufficient physical activity and can result in a decrease in the total and the quality of years of life [2].

Due to these kinds of health concerns, we see a trend of people purposely allocating time in their schedule for physical fitness. We aim to assist in this regard by offering the user an app that can dynamically adjust to how they perform their physical activities and how their emotional state varies. The app is able to learn the user’s current lifestyle via sensors and also obtain their emotional state via the camera. The user will be given a diet and workout plan and their commitment to it will be monitored. Based on this data along with the emotional analysis and lifestyle tracking, our future recommendations will change. Furthermore, if the app identifies any habits that would adversely affect the user’s health, then it would notify the user so that they may act. We hope that by using our app the user would be able to avoid the adverse health issues that could arise due to a lack of physical activity.

# Methodology

## Obtaining the Dataset

The app consists of three main sections, identifying the user’s emotion, activities, and the food consumed. In order to do this, we needed three separate datasets, since we would be building three separate Models to identify each of these.

In order to obtain the user’s activities, we decided to require them to wear a smartwatch. While it would be possible to track the activities from a smartphone, the accuracy would not be to our satisfaction. Looking at research done with smartphone activity tracking showed that their accuracy would max out at about 72% [3]. Whereas studies that had used both a smartphone and a smart-watch for activity tracking would boast a 19% better accuracy when using smart-watch sensors [4]. Therefore, we decided to use extraction from the WISDM dataset [5], which contains accelerometer readings along the x-axis, y-axis, and z-axis taken from a smartwatch. The data distribution is shown in Fig 1.

The dataset consisted of readings from 51 different subjects each of them had performed the 5 activities shown and sensor readings had been recorded. Readings were taken at a rate of 20Hz as each participant did each activity for 4 minutes apiece. This gives us many readings which would be beneficial when training the model. Furthermore, we will be considering only the smart-watch accelerometer readings. Since that sensor is what we decided to use, considering that power consumption of the sensor must also be given attention. Running a gyroscope would consume more power than an accelerometer, hence we decided to only use the accelerometer to predict the user’s activity.

To capture the user’s emotion, the author uses a mobile phone’s camera. A mobile phone’s camera is the most suitable way to track the emotion of users. The research was done in emotion extraction, the accuracy of their project 91% [9]. The research used a picture to extract the emotion [9]. The research used 8 emotions, Angry, Contempt, Disgust, Fear, Happy, Sad, Surprise, and Natural [9]. For this research, the author used 7 emotions. The research regarding the Development of Deep Learning-based Facial Expression Recognition System shows, 7 emotions used for research more accurate than 8 emotions [10]. 7 emotions as follows, Anger, Disgust, Fear, Happiness, Sadness, Surprise, and Chart, bar chart

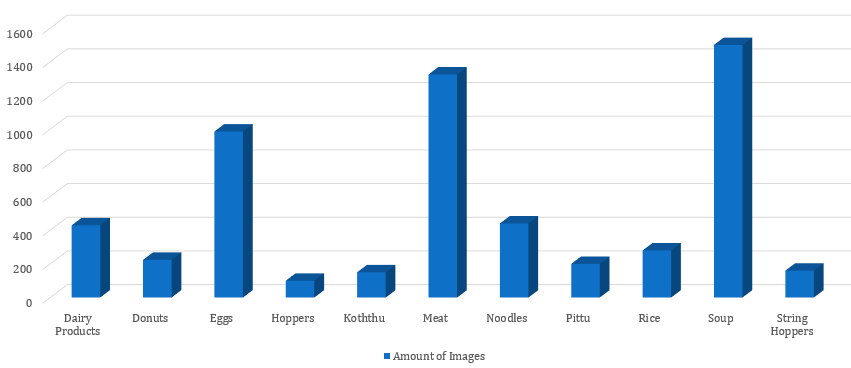
Description automatically generatedNatural [10]. The emotion class distribution is shown in Fig 6.

The dataset includes 7 classes of emotions. Which are as follows, Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral [10]. The current data set has enough data to improve accuracy for extract users' emotions. The dataset includes pictures from 7 emotion classes. The smartphone camera takes input as a picture from the user’s face. The model analyses the input using the dataset’s pictures.

So in the process of recommending the food diet, we are getting the data like Age, Height, and Weight. So, from those data, we calculate the Body Mass Index (BMI) of the user and categories them into three main body types as underweight, healthy, and overweight. So, according to that category, the system recommending the food diet. Due to the freedom of using any food according to the user's wish we give a function to the user to upload a picture taken by the camera of his mobile phone of the foods which he eats. So, through that, we calculate the calorie amount of the food in that food and continue it for further dynamic change of the app.

In sense of that, we took many images of 11 food categories which many used in Sri Lanka as dairy products, donuts, eggs, hoppers, Koththu, Meat, Noodles, Pittu, Rice, Soup, String Hoppers.

So, as previously mentioned the data set consists of 11 food categories. So it helps to predict the food through an image and get the calorie count of the food for further process of the app



## Building the Model

The creation of deep learning models is a resource-intensive task, therefore we decided to make use of Google Colab. Which is a service by Google which offers free GPU access for the creation of the model.

To build a model for the emotion tracking section, the dataset’ data put into 7 labels, such as Anger, Disgust, Fear, Sad, Surprise, and Neutral. Architecture use for this model Convolutional Neural Network. Building the layers for the model, the author went for the sequential Keras model. The first layers of the neural network study a small number of convolutional filters apart from that deep in neural network study more convolutional filters comparing to first layers. Batch normalization technique will use for increasing the rate of training in the neural network. Max Pooling was put to decrease the spatial dimensions in output. In the network, the input shape takes as 48x48. The first convolutional layer has 5x5 filters and learns 64 filters, the second convolutional layer has 5x5 filters and learns 128 filters, the third convolutional layer has 3x3 filters and learns 256 filters. Dropout use for network generalization and dropout use for not overfitting the training data. Add the Softmax classifier to the network to get the output from the layer.

The current model shows 61% of accuracy. The loss and confusion matrix is given below.

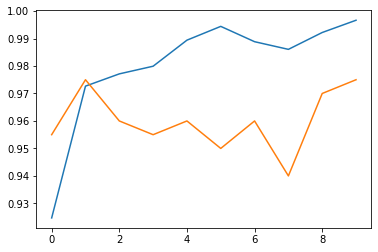
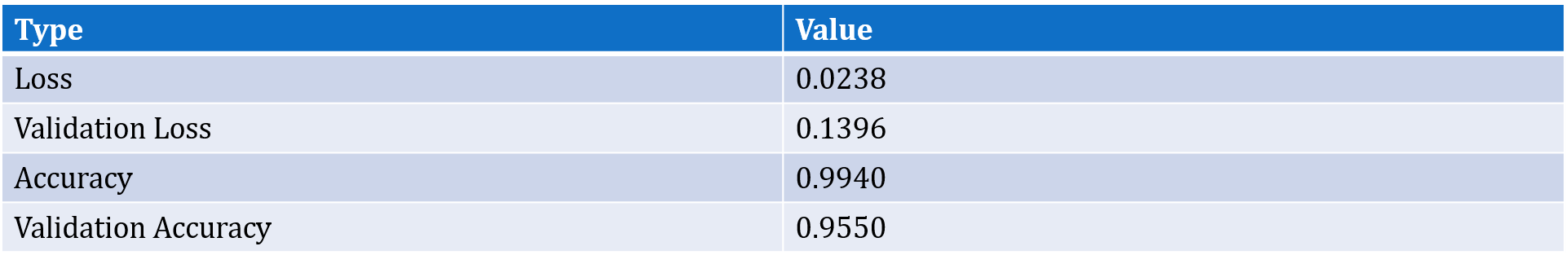
A picture containing box and whisker chart

Description automatically generatedChart, line chart

Description automatically generated

When building the model for the activity tracking section, initially the dataset was loaded and then split into training and validation subsets. The splitting was done via the user id column, this column values range from 1600 to 1650, so we took the readings with user id less than 1641 for training and the others for validation. When constructing the layers of our model, we went with a sequential Keras model for simplicity, then a Long Short-Term Memory (LSTM) layer was used. This layer was selected since it is suitable for making predictions based on time series data and helps overcome the vanishing gradient issue that can occur in traditional Recurrent Neural Networks (RNN). The next layer was a dropout layer, intended to reduce overfitting to the training data. A dense fully connected layer was used to interpret features of the LSTM layer and then an output layer would give us the predictions.

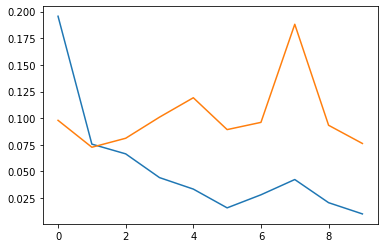
The model was then trained, and we observed an accuracy of about 88%. The loss graph and confusion matrix are given in Fig 3 and Fig 4.

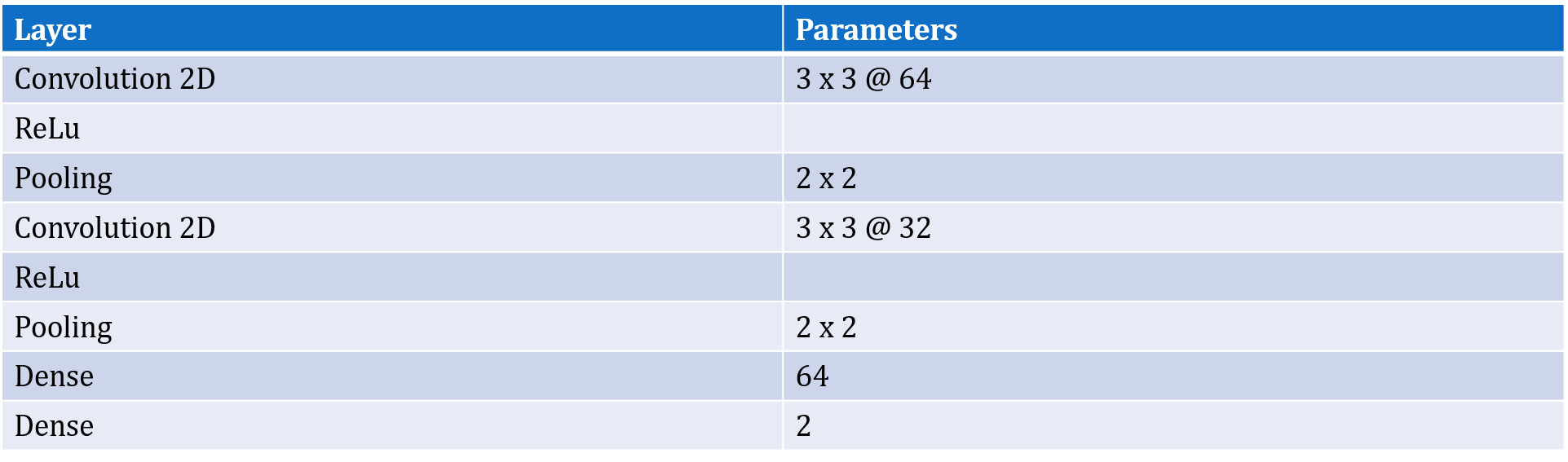


When considering the Food Recommendation function, we divide it into 3 main parts. As

1. Identify the food from a picture taken by the user's camera on his mobile phone and predict the average calorie amount of the food in that picture.
2. Analyze and identify the user’s body type and recommend a food diet according to that body type
3. Analyze and identify the remaining calorie amount in the past 4 days of the user and changing the next day food diet to a low-calorie input food diet if the user is getting increase the remaining calorie amount

So, for the first part, we use CNN because there is a prediction part. (Convolutional Neural Network).so already we have trained the data model of the first part.

So, part 2 is doing using the FFNN and part 3 is doing using RNN.



## Building the App

When it came to building our app, we decided to have a Django server as our backend. The server would contain each of our exported models and when we send the inputs required via POST requests, the server would make the prediction and respond with the result. This would help us conserve battery life on the user’s phone since the task of making predictions would require the phone to do more work and be constantly awake. Fig 5 shows a diagram of how our app would work.

For the app to function we need the user’s age, height, weight, and gender, so upon the first launch of the app, they would be asked to fill in these details. The app consists of four main sections.

1. Emotion Tracking
2. Lifestyle Tracking
3. Diet Planning

In the lifestyle tracking section, the watch would supply accelerometer readings which the phone would then send to the server and obtain an activity prediction. These activities would then be stored on the user’s phone on a local database. For privacy reasons it was decided that the server would not store any kind of log of the user’s activity. The Lifestyle tracking section would run in the background for one week initially. During this time, it would attempt to identify what the user’s current lifestyle pattern is. Using a combination of the activity history and the other smartphone sensor readings the app would identify the user’s

* Wake-up and sleep time
* Home and work location
* Hours spent at the workplace
* Exercise Time and type
* How they get to work

After the one-week analysis period is complete the app would start analyzing this data to see where the user can improve their lives so that they may be healthier.

* By analyzing the wake-up and sleep time we can find out how many hours of sleep the user gets. Ideally, an adult should get at least 7 hours of sleep [6], so by looking at the user’s wake up time, our app would notify the user about 7 hours in advance that if they go to sleep shortly they will be able to get the required hours of sleep, while still waking up at their usual time.
* Knowing the work hours and location is beneficial for the workout and diet sections of the app since we can avoid disturbing the user when they are at work.
* Sitting continuously for long periods of time has been shown to have adverse effects on a person’s health [7]. If the app detects the user sitting continuously for one hour, it would send a notification suggesting that they stand up and move about for a bit.
* Meditation is shown to help with managing stress and anxiety [8]. The app would consider suggesting the user try out meditating if it detects they are at home and are not engaging in any taxing activities.
* To track the user’s progress the app would be requesting a weekly update on the user’s weight, which would be used to plot a Body Mass Index (BMI) history. This would be considered when updating the user’s diet and/or workout plan.

In this way, the lifestyle tracking section would assist the user in maintaining their health.

When considering the recommending food diet, at the beginning we get some data from the user like name, gender, height, and weight as an input. Using them we calculate the Body Mass Index (BMI) of the user. After that using the BMI value we categorized the user as underweight, healthy, and overweight. So, according to the category system recommending a diet plan in each and every week. In that diet plan, each and every day has a specific calorie intake amount. so that amount differs from day by day according to the body category. But we do not know whether the user properly getting the food that we recommend. So we give another feature to the user “if he taking the food according to his wish”. So, the user can take a picture of food from the camera of his mobile phone, which he eats, and upload it to the system. So using that image we identify the food and identify the average calorie amount of the food in the picture. Then we take that value as a “Calorie Intake”. Then at the end of the day, the system calculates the total calorie intake amount of the day.

Then the amount of calories burn which has been taken from the recommending workout function is moved forward to this function. And then create a new formula,

*Remaining Calories = Calories Intake – Celeries Burn*

So using this formula, the system calculates the remaining calorie amount of the user at the end of the day.

Afterward, the system collects and stores the remaining calorie amount of the past 4 days and analyzes the pattern in between that 4 days. If the system detects that the remaining calorie is getting higher day by day, the system changes the next day food diet and give a diet that has a low-calorie intake to balance the calorie intake of user and break the increasing of remaining calorie amount.

So, it will be a great help to the user to maintain a good healthy life by considering the factor ‘food diet’

For Emotion tracking sector includes two main functions,

1. Facial expression analysis function
2. Mobile game function

The facial expression analysis function runs on the backend server and android game implemented in the app. To play the game, the game function checks whether the user completed the Daily task or not. After validating that parameter, the android game function allows access to the game. When the start of the game difficulty is set to beginner level. The gameplay is limited to 25 minutes because the daily average time spent on mobile games 2015 and 2016 is 25 minutes and 30 minutes [11].

When the game started system take a picture of the user by using the Phone's camera within a given time. After that, taken pictures send to the facial expression analysis function that runs on the backend server. The facial expression function includes a trained neural network. Function analyzes the picture. Analyzing procedures give an accurate result. The result is the user's emotion. Result send back to game function. According to the result, the mobile game's difficulty is chang to beginner, hard, challenging, or epic; not only that, the result is used for display emoji during gameplay. Emoji choose according to the user's emotions. Purpose of display emoji to assist the current player mood [12]. This process runs until the end of the gameplay. After ending gameplay, the system sends data to the Daily task generating component that the user completed the daily game level

In the facial expression analysis function, include 7 emotions. Neutral, Happiness, Disgust, Anger, Fear, Sadness, and Surprise. In this feature, using the above facial expressions helps to get an accurate result of this component. Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision" research use Happy, Sad, Surprise, Fear, Disgust, and Angry for CNN [13]. However, in this component, the author uses Neutrality emotion to improve component' performance.

In this feature, the user's face picture is used for emotion extraction. This picture was taken within a given time while the user plays the game. Previous research used the user's video frames for emotion extraction [13]. In previous research, the emotion extraction process is needing more resources and take a long time to train neural network [13].

The mobile game function has game modes, first, is difficulty balance activated, second is difficulty balanced deactivated. Previous research shows that by using two methods [15], users can play games without disturbing player experience. The game level starts on the difficulty balance deactivated method. This method sets beginner difficulty [15]. After a given time, difficulty balance activates [15]. The second method concurrently engages until to end of gameplay.

Gain interaction with the game; this function includes displaying the emoji method. According to previous research [16], if the game assists the player, the player willingly more interacts with the game. It helps a stable player mindset [16]. Most computer games assist the player by giving hints [16]. In this component, the mobile game displays emoji according to emotion given by facial expression function.

Fig 7 shows a diagram of the emotion analysis diagram.

Diagram

Description automatically generated

# Results and Discussion

When considering emotion tracking, we are able to track the emotions of the user while he playing the game. When the play user playing the game, the system detects his emotion and sends it to the game then the game automatically changes the difficulty according to the emotion. Along with that to change the current mind of the user, the system pops up an emoji that shows the current emotional status of the user.

When considering the food diet recommendation, we are able to detect and analyze the body type of the user and recommend a food diet weekly according to his detected body type. If the user taking food according to the user’s wish, the user can take a picture from the camera in his mobile phone of the food and predict the average calorie amount of the food in the picture. And calculate the total calorie intake at the end of each day. Then detect and analyze the pattern of calorie remaining amount in past 4 days and change the next day's food diet if the user is increasing his remaining calorie amount to break his pattern of increasing remaining calorie amount.

When considering the tracking of user’s activities, we are able to track and detect all the activities of the user from wake-up to sleep and then give and recommend tips and methods if the user maintains a bad routine. By detecting the user's daily routing make and adjust the whole app user-friendly to the user. Like, the app will not send any notifications when the user in the office or during his working time

# Conclusion

Health and fitness are a major concern in modern society, with technological advancements making life easier. So People use apps to track their daily fitness and there are no existing apps that can be used to make decision making learn considering the user’s lifestyle and expressions as the main input factors. Also, there is not an app that combines daily tasks, routines, nutrition patterns, and user expression monitoring. For a particular person to examine all these tasks mentioned, they need to use multiple apps. Thus, then all these functionalities cannot be obtained.

Due to the static nature of such apps, people tend to move away from such fitness apps and ultimately impact a healthy life.

So, at the end, we develop and introduce an app that takes the user’s habits and expressions as inputs and leads to a better decision-making platform in combining daily tasks routines, nutrition patterns, and emotion prediction. Further, considering the inputs, the system adapts to match the user’s different lifestyles, dynamically changing food diet and analyze and detect user’s moods. So, in this way, the user will be motivated to keep using the app and thereby leading to a healthy life

## Figures and Tables

* Fig 1. Distribution of smart-watch sensor data
* Fig 2. Layers in the activity tracking model
* Fig 3. Activity tracking model loss graph
* Fig 4. Activity tracking model confusion matrix
* Fig 5. The architecture of the app
* Fig 6. Class distribution of 7 emotions
* Fig 7. System diagram of emotion analysis function
* Fig 8. Emotion tracking model loss graph
* Fig 9. Distribution of food category data
* Fig 10. Accuracy of the model
* Fig 11. Accuracy of the model
* Fig 12. Accuracy of the model
* Fig 13. CNN architecture of the model

##### Acknowledgment

We wish to show our appreciation to Mr. Prasan Yapa (Lecturer at SLIIT) for acting as the supervisor of this project and for his valuable insights.

We thank Dr. Evone Jayaweera (Senior Registrar in Clinical Nutrition at Teaching Hospital Peradeniya, Sri Lanka.) for her knowledge and expertise in the nutrition section of our project.

##### References

1. S. J. Robroek, K. G. Reeuwijk, F. C. Hillier, C. L. Bambra, R. M. V. Rijn, and A. Burdorf, “The contribution of overweight, obesity, and lack of physical activity to exit from paid employment: a meta-analysis,” Scandinavian Journal of Work, Environment & Health, vol. 39, no. 3, pp. 233–240, 2013.
2. F. W. Booth, C. K. Roberts, and M. J. Laye, “Lack of Exercise Is a Major Cause of Chronic Diseases,” Comprehensive Physiology, 2012.
3. X. Sheng, J. Tang, J. Wang, T. Li, G. Xue, and D. Yang, “LIPS: Lifestyle Learning via Mobile Phone Sensing,” 2016 IEEE Global Communications Conference (GLOBECOM), 2016.
4. G. M. Weiss, J. L. Timko, C. M. Gallagher, K. Yoneda, and A. J. Schreiber, “Smartwatch-based activity recognition: A machine learning approach,” 2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), 2016.
5. G. M. Weiss, K. Yoneda, and T. Hayajneh, “Smartphone and Smartwatch-Based Biometrics Using Activities of Daily Living,” IEEE Access, vol. 7, pp. 133190–133202, 2019.
6. H. H. Publishing, “Repaying your sleep debt,” Harvard Health. [Online]. Available: https://www.health.harvard.edu/fhg/updates/Repaying-your-sleep-debt.shtml. [Accessed: 05-Mar-2021].
7. H. H. Publishing, “Why you should move - even just a little - throughout the day,” Harvard Health. [Online]. Available: https://www.health.harvard.edu/heart-health/why-you-should-move-even-just-a-little-throughout-the-day. [Accessed: 05-Mar-2021].
8. H. H. Publishing, “How meditation helps with depression,” Harvard Health. [Online]. Available: https://www.health.harvard.edu/mind-and-mood/how-meditation-helps-with-depression. [Accessed: 05-Mar-2021].
9. Liliana, D Y (2019). Emotion recognition from facial expression using deep convolutional neural network. Journal of Physics: Conference Series, 1193(), 012004–. doi:10.1088/1742-6596/1193/1/012004
10. Jung, Heechul; Lee, Sihaeng; Park, Sunjeong; Kim, Byungju; Kim, Junmo; Lee, Injae; Ahn, Chunghyun (2015). [IEEE 2015 21st Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV) - Mokpo, South Korea (2015.1.28-2015.1.30)] 2015 21st Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV) - Development of deep learning-based facial expression recognition system. , (), 1–4. doi:10.1109/FCV.2015.7103729
11. C. Gough, "Time spent with mobile games worldwide 2016 | Statista", Statista, 2020. [Online]. Available: <https://www.statista.com/statistics/667679/time-spent-mobile-games-world/>.
12. *Eprints.lancs.ac.uk*, 2020. [Online]. Available: https://eprints.lancs.ac.uk/id/eprint/12587/1/Gilleade\_Affective\_Gaming\_DIGRA\_2005.pdf.
13. Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision," ISCAS 2010 - 2010 IEEE Int. Symp. Circuits Syst. Nano-Bio Circuit Fabr. Syst., pp. 253–256, 2010
14. X. Pan et al., "Video-Based Facial Expression Recognition using Deep Temporal–Spatial Networks", IETE Technical Review, pp. 1-8, 2019. Available: 10.1080/02564602.2019.1645620
15. J. Moniaga, A. Chowanda, A. Prima, Oscar and M. Tri Rizqi, "Facial Expression Recognition as Dynamic Game Balancing System", 2020. .
16. *Eprints.lancs.ac.uk*, 2020. [Online]. Available: https://eprints.lancs.ac.uk/id/eprint/12587/1/Gilleade\_Affective\_Gaming\_DIGRA\_2005.pdf. [Accessed: 18- Jul- 2020]