Emotion detection using machine learning on EEG data

Bachelor's Project Thesis

Rik Vegter, s3147495, f.j.vegter.1@student.rug.nl,
Supervisor: Dr. Marieke van Vugt, Pallavi Kaushik

Abstract: Classiyfing emotions based on EEG data has grown in popularity in the last few decades. At the moment, literature trains classifiers on data sets created in a laboratory where emotions are artificially evoked, meaning they present a stimulus and expect a certain emotion. The EEG data during the emotion is then measured and analysed. In order to capture the EEG data of emotions that are naturally evoked, a special type of debate was analysed which evokes emotions for the debaters. The goal of this study was to create a classifier that is able to distinguish *happy* EEG data from *angry* EEG data when training the classifiers based on a data set where the emotions are evoked in a more daily life setting. Time-domain features were extracted from the preprocessed data set. A random forest classifier and KNN classifier were trained and returned accuracies of 92.5 % and 92.2 % respectively with parameter optimisation when tested on the training data. The confusion matrices of the classifiers both showed true positives and true negatives of above 90 %. K-fold cross validation showed accuracies of 65.1 % for the random forest classifier and 61.2 % for the KNN classifier.

1 Introduction

Emotion detection is a capability that humans are capable of by reading facial expressions (Ferretti & Papaleo, 2019). Humans detect emotions partly by reading facial expressions but also by reading body language (Barrett et al., 2011). However, it is not always the case that humans show their emotions by expressions that can be perceived with the eyes since facial expressions and body language can be consciously suppressed (Zheng et al., 2017). What humans cannot consciously suppress are their physiological response to a mental state that they are in (Shu et al., 2018). Examples of these physiological signals are heart rate, breathing rate, temperature, brain activity etc. Since humans cannot read the physiological signals off of others there has been a big interest in reading off these physiological signals with the help of measuring instruments.

A study of Ekman et al. (1983) used finger temperature, heart rate and skin conductance as physiological signals to distinguish the emotions of anger, fear, sadness, disgust and happiness. They were able to find significant differences between these physiological signals between the different emotions. Lisetti & Nasoz, (2011) provide a clear framework of the history of emotion detection with physiological signals.

Another popular physiological signal in order to classify emotions is electroencephalography (EEG). This is a technique that records electrical activity of the brain and is measured by placing multiple electrodes along the scalp. Every single electrode measures the voltage that is a result of the ionic current coming from the neurons.

This technique has found itself to be useful for identifying which parts of the brain are more active when people are in a certain mental state. When a person is listening to music that should evoke a feeling of happiness and joy, the left frontal area of the brain is more active and when a person is listening to negative valenced music the right frontal area of the brain seems to be more active (Schmidt & Trainor, 2001).

EEG data has also shown to be successful with classifying mental states with accuracies of 98.39% (Amin, 2017). In this study the classifier tried to predict based on a single trial whether the EEG data showed patterns of someone performing a cognitive complex task or someone performing a simple baseline task. Such classification of mental states is done by a particular workflow. Researchers first feed a stimulus to a participant where the researchers expect a certain mental state. The corresponding EEG data is collected and preprocessed. Then certain features of the signals are extracted and finally a classification model is trained with these features.

Most literature mentions three different types of features of EEG signals; time-domain features, frequency-domain features and time-frequency-domain features. Time-domain features are statistics about a signal that use a function of time (mean, amplitude etc.) where frequency-domain features use a function of frequency. Time-frequency domain features use statistics based on the frequency and the time. The type of feature extraction and the specific classification model that is used determines the performance of the model.

Edla et al., (2018) present a random forest classifier which was trained with time-domain features. This study tried to distinguish a concentration and meditation state of mind. The classifier got an accuracy of 75% using k-fold cross validation.

Moshfeghi et al., (2013) trained a support vector machine based on frequency-domain features. They used an emotional system where they had a positive and a negative valence. The accuracy of this binary classification model is 74% using k-fold cross validation.

Liu et al., (2016) trained a k-nearest neighbor classifier and a random forest classifier with features from the time-domain, frequency domain and the time-frequency domain. They got an accuracy of about 70% when using the DEAP data set (Koelstra et al., 2011). The accuracy was measured using k-fold cross validation.

The DEAP data set is used often in literature when classifying emotional states with EEG data. This data set handles emotions in a 2-dimensional model consisting of valence and arousal as can be seen in figure 1.1. This model is similar to the model of Schlosberg (1954) if the third dimension of Schlosberg called *level of activation* is ignored. The two dimensions are arousal and valence (originally called pleasantness and attention by Schlosberg). A high level of valence is associated with positive emotions like happy or joy. A low level of valence is associated with negative emotions like anger or frustration. The level of arousal determines how awake or excited someone is. If the level of arousal is low, a person is tired and when the level of arousal is high, someone is very alert and awake.

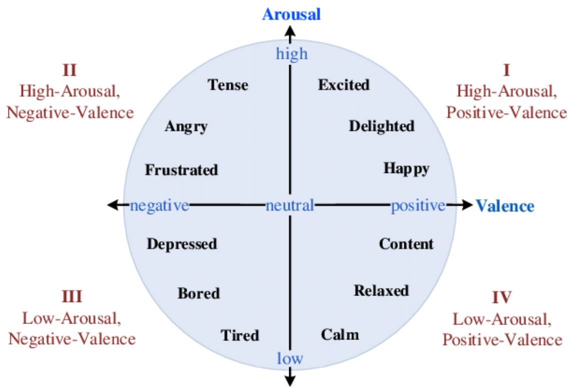

Figure 1.1: Two dimensional model of emotions. This model plots arousal versus valence. Inside this model are the associated emotions that corresponds to the levels of arousal and valence.

All of the classifiers mentioned were trained on data sets that were created in a laboratory setting. A participant wearing an EEG cap is presented a stimulus where the researchers expect a certain emotion. One could say that these emotions are artificially evoked. Furthermore, not every participant reaches the wanted intensity of the emotion that the researchers expect since some people are better at emotional regulation than others (Gessner, 2015). More ideal would be to record the EEG data of emotions that are evoked in a natural setting or a daily life setting since these emotions might result in EEG data that a person actually has when in a certain mental state of emotion.

Such a data set was collected by Van Vugt. In this study a monastic debate is researched. A monastic debate is a debate performed by Tibetan Buddhists. More experienced debaters improve 'the ability to handle high cognitive load situations as well as emotional changes' (Van Vugt et al., 2020). The setting of the debate is as follows: Two people participate in the debate. One of the two is the defender while the other is the challenger. The defender sits in a cross-legged position and the challenger is standing in front of the defender. The challenger makes statements. The subject of the debate can vary a lot. When the challenger makes a

Figure 1.2: Typical setting of a monastic debate. The challenger stands in front of the defender. Both the defender and the challenger are wearing an EEG cap.

statement, the defender has two options; accept the statement of the challenger or questioning the reasoning of the statement and requesting an explanation of the challenger. The defender is not allowed to accept a contradicting statement of the challenger. If the defender does accept two contradicting statements, the challenger creates awareness for that contradiction by screaming "tsa!". The challenger often makes intimidating verbal signs like clapping very loud after making a statement, or raising his voice in order to pressurize and the defender.

During such a debate the challenger and the defender undergo different emotions. For example when the challenger wants to pressurize the defender he may get angry. Or when the defender is in a tight spot during the debate he may get frustrated or angry since he cannot regulate his emotions anymore. During the debates, the two participants also laugh quite often. This means that we can distinguish two types of emotions during the debate, emotions with a high valence (happy) and emotions with a low valence (anger).

The question that this study tries to answer is whether we can distinguish happy EEG data from angry EEG data in a natural setting and whether we can build a classifier that is able to predict these two emotions on a single trial level.

2 Methods

In order to try to answer the research question the data set by Van Vugt was used. This data set was collected by creating a set up where a challenger and a defender debate with each other in a setting similar to figure 1.2. As mentioned in the introduction a monastic debate evokes naturally evoked emotions. The brain activity of both the challenger and the defender were collected. All debates were video recorded. Afterwards researchers annotated whether the debaters were happy or whether one of the debaters was angry. The annotations and the corresponding EEG data is combined and features of the data are extracted. Finally machine learning algorithms were trained on the data and the performance was analyzed.

2.1 Participants

The participants that participated during the debates were all Tibetan monks that all had some experience with monastic debates. Monks considered to be experienced had more than 15 years of experience where inexperienced monks had more than 3 years of experience. 10 experienced and 14 inexperienced monks participated in the debates.

2.2 Monastic debate

As mentioned in the introduction the debate consists of a challenger and a defender where the challenger makes statements. The defender can accept or question the statements of the challenger. A defender is not allowed to accept contradicting statements. During the debate some researchers were also in the room as can be seen in figure 1.2. These researchers were not allowed to interfere with the debate.

In total 50 debates were analyzed of which 24 debates were considered *easy* debates and 26 debates were considered *hard* debates. An easy debate handles a topic that is more accessible where a hard debate is handles a topic that is more difficult to reason about for the defender. An easy debate took approximately 10 minutes and a hard debate approximately 15 minutes. Every debater was paired up with another debater. Every pair performed 4 debates. 2 easy debates and 2 hard debates. For every type of debate both debaters once acted as a challenger and one time as a defender. Two debates were excluded from the data set due to unforeseen technical issues.

2.3 Video recording

All debates were video and audio recorded from the perspective that can be seen in figure 1.2. This means that the face of the challenger is not visible most of the time while the face of the defender is visible all the time.

2.4 EEG recording

The EEG system in this study uses the same setup as used in Jin et al., (2019). Both the debaters wore Biosemi EEG equipment with 32 electrodes instead of 128 electrodes as used in Jin et al., (2019). The two EEG caps were connected using a daisy chain methodology which means that the two systems are connected in a wiring scheme in a sequence. As mentioned in Jin et al. (2019) the sampling rate is 512 Hz which was down-sampled to 256 Hz after being passed to a band-filter with a range of 0.5 - 40 Hz.

2.5 Preprocessing the EEG data

Since we only want to capture the EEG data relevant to cognitive tasks in the brain we need to exclude irrelevant, unwanted EEG data, also called artifacts. Movements of the body generate artifacts. Think about eye blinking which is visible from spikes in the electrodes at the front side of the skull. It is also possible that the electrodes capture other electrical activity that is irrelevant for the person itself. For example a heater can also sent out a small amount of electrical activity that can be recorded by the EEG electrodes. In order to exclusively capture EEG data related to cognitive tasks, these artifacts should be removed in order to perform a meaningful analysis.

The preprocessing of the raw EEG data was done with Fieldtrip (Oostenveldt et al., 2011) which is software developed in Matlab that allows EEG data processing. In order to remove the artifacts, an independent component analysis (ICA) was performed (Radüntz et al., 2015). This analysis recognizes fixed patterns that are in EEG data. For example an eye blink has a very characteristic EEG signal and can be identified by an ICA. Other fixed patterns are muscle movement, heartbeats, sweating and clapping. After identifying these fixed patterns, they can be removed from the signal which leaves us with only the relevant brain activity.

2.6 Annotating the emotions

Four researchers were given the task to watch the videos of the debates and annotate whether they saw the defender or challenger being in a mental state of *happiness* or *anger*. The annotations were done with Boris which is software that allows to create annotations with states. When the researchers saw one of the debaters entering a state of *happiness* or *anger* this was annotated. When, according to the researchers, that debater entered a *neutral* state of emotion again, the end of that emotion was annotated. This allowed to create a time interval of which a debater is in a certain mental state. Due to the subjective nature of annotating emotions, every video was annotated by three researchers.

When one of the researchers annotated a video, these annotation were saved in a *.csv* file. This means that per video, there are three *.csv* files which have to be combined in order to reach a final annotation file per video. The three annotation files per video were combined as follows: When two time intervals of the same emotion of the same debater overlaps for two or more researchers, this overlapping time interval is annotated as a final annotation. Imagine one researcher annotated the defender as being in a *happy* state at time interval 6s - 10s. Now consider another researcher annotating the defender being in a *happy* state at time interval 8s - 11s. When the annotation files are combined the final annotation will be 8s - 10s because two researchers annotated the defender being in a *happy* state at this time interval.

2.7 Synchronize the EEG data with the annotations

Since the EEG data did not always start at the same time as the video recording, a problem arises when synchronizing the timing of the EEG data and the annotations. This problem was solved with multiple measures. In some videos a technique was used where the challenger blinks three times. Blinking creates very characteristic spikes at electrodes above the forehead. If this is done three times, this can be easily red off the EEG data and is easily visible from the video recordings since these blinks were done in front of the camera. This means that it is possible to synchronize the EEG data with the video recordings and the annotations. Unfortunately this accurate technique was done for only 30 videos. For the other 20 videos the delay of the EEG data recording was manually timed.

2.8 Feature extraction

In order to prepare the data so it can be put in the model to train, we first need to extract features from the EEG signals. Features from the time-domain were extracted from the EEG signals since these features have shown to be able to classify basic emotions such as happy and anger(Chai, Woo, Rizon, & Tan, 2010; Takahashi, 2004). More specifically, the mean, standard deviation and the power were extracted. For the sake of completeness all three statistic calculations are shown in equation 2.1, 2.2 and 2.3.

Mean:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{2.1}$$

Standard deviation:

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{2.2}$$

Power:

$$P = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \tag{2.3}$$

The features are measured over the entire time interval of a final annotation. This means that if a final annotation was 3 seconds long, the mean, SD and power for all 32 channels were extracted over the entire 3 seconds resulting into one data point with 96 values (32 means, 32 SDs and 32 powers).

It is a possibility to exclude the non significant channels but these channels might be relevant in combination with other channels in order to provide us information about whether a person is angry or happy. Because of this reason no channels were excluded for the classification.

2.9 Classification

A random forest classifier and a k-nearest neighbors were used in order to classify the data points. 5-fold cross validation was used to test the performance of the classifiers. The classifiers were implemented with functions from the python library *sklearn*. The following two functions were used to create the classifiers in python3: *RandomForestClassifier() and KNeighborsClassifier()*

2.9.1 Random forest classifier

A random forest classifier is an ensemble method for classification. This classifier creates an N amount of randomly initialized decision trees. All of these decision trees are trained. All outputs of the decision trees are calculated and the mean prediction of all output of the trees will be the output of the classification. Due to the bootstrapping nature of this algorithm the random forest classifier reduces overfitting which single decision trees do not.

2.9.2 K-nearest neighbor classifier

K-nearest neighbors (KNN) is a simple but powerful classifier that relies on a parameter K. If a data point needs to be classified, the K closest data points are determined. The class with a majority of labels of these K-nearest neighbors determine the class of the to be classified data point.

3 Results

The data set showed an imbalanced number of data points. After the data was collected there were 1953 *happy* data points and 439 *angry* data points. Since the durations of the annotations of the angry data points were longer than the annotations of the happy data points, all data points were multiplied by the duration of that annotation so that the data became more balanced. If for example a data point's duration was 5 seconds, that data point was repeated 5 times. This resulted in 10842 *happy* data points and 13415 *angry* data points.

3.1 Statistics

In order to explore the data an exploratory analysis was done between *happy* and *angry* EEG data. This analysis focused on finding differences between the channels and to get a global overview of the data. Appendix B shows the results of dependent t-tests for all channels for the mean, standard deviation and power for *happy* and *angry* EEG data. For the power and standard deviation only 2 channels were not significantly different from each other. Surprisingly this was the same channel for both statistics, namely channel FP1 and T8. For the mean, 5 channels on the anterior ventral side of the skull were found to be significantly different from another (these are denoted by a * in appendix B and are highlighted red in figure 3.1). More specifically, channels AF3, F7, F3, FC5 and CP1 are significantly different between *happy* and *angry* EEG data when looking at the mean. In appendix A a table is shown which show the mean per channel per class.

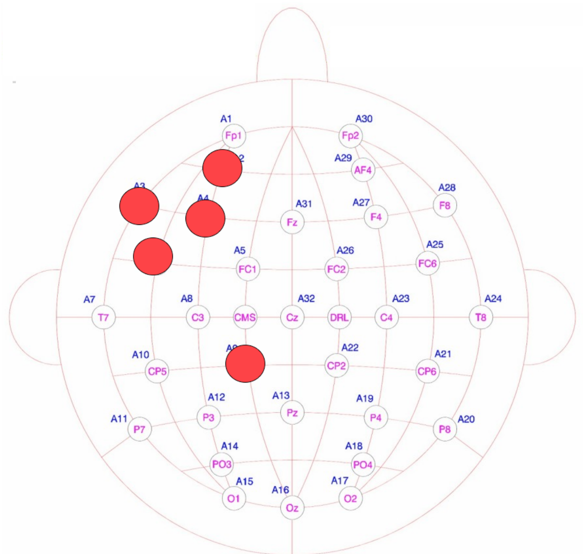

Figure 3.1: Overview of the electrodes placed on the scalp. The electrodes in red are the electrodes that turned out to be significantly different for *happy* and *angry* EEG data when using the mean as a statistic for a dependent t-test. Appendix C provides the same figure but bigger.

3.2 Classifiers

After exploring the data I tried to build a classifier that is able to distinguish *happy* and *angry* emotions on a single trial. As previously mentioned, a random forest classifier and k-nearest neighbor algorithm were used. In the next sections I will go over the performance of these classifiers.

3.3 Random forest classifier

Multiple random forest classifiers were used with different numbers of trees. Table 3.1 shows the accuracy, precision and recall of three random forest classifiers with 8, 10 and 12 trees.

Trees	Accuracy	Recall	Precision
8	92.3 %	94.8 %	81.4 %
10	92.5 %	95.4 %	81.4 %
12	92.4 %	95.2 %	81.3 %

Table 3.1: Accuracy, precision and recall of a random forest classifier with 8, 10 and 12 trees when tested on the training data

As can be seen from table 3.1 the the random forest classifier performs slightly better when using 10 trees in comparison with 8 and 12 trees. The accuracy is 92.5 % when tested on the training data. For all forests with more than 12 trees or with less than 8 trees, the accuracy, recall and precision decreases. Table 3.2 shows the confusion matrix when testing the random forest classifier on the training data.

	Predict: happy	Predict: angry
True: happy	0.952	0.0484
True: angry	0.0867	0.913

Table 3.2: Normalized confusion matrix of the Random Forest Classifier with 10 trees when tested on the training data

As can be seen from the normalized confusion matrix in table 3.2, 95.2 % of the happy data points and 91.3 % of the angry data points are correctly classified from the training data. In order to test the classifier on data it has not seen before a 5-fold cross validation was used. This test returned 5 accuracies with a mean of 65.1 % and a standard deviation of 7.2.

3.4 K-nearest neighbors

This KNN classifier uses the Euclidean distance as a distance function. Multiple values for K were tried. The ones with the best performances were K = 3, 5, 7, 9. The accuracy, precision and recall are in table 3.3 when tested on the training data.

K	Accuracy	Recall	Precision
3	92.2 %	95.5 %	80.7 %
5	90.8 %	88.3 %	81.2 %
7	89.2 %	81.2 %	80.8 %
9	86.2 %	79.2 %	74.1 %

Table 3.3: Accuracy, precision and recall of a k-nearest neighbor algorithm with K = 3, 5, 7 and 9 when tested on the training data

As can be seen from table 3.3 the KNN classifier works best for K = 3 where the accuracy is 92.2 % when tested on the training data. The confusion matrix of the KNN classifier for K = 3 when tested on the training data is shown in table 3.4.

	Predict: happy	Predict: angry
True: happy	0.955	0.0453
True: angry	0.0907	0.909

Table 3.4: Normalized confusion matrix of K-nearest neighbors for K = 3 when tested on the training data

The normalized confusion matrix in table 3.4 shows that 95.5 % of the happy data points and 90.9 % of the angry data points were classified correctly. A 5-fold cross validation was also used for the KNN with K = 3. This test returned 5 accuracies with a mean of 61.2 % and a standard deviation of 5.2 mV.

4 Discussion

This study attempted to train classifiers that are able to classify *happy* EEG data and *angry* EEG data. This classifier was trained on data that was created in a setting that evokes emotions similar as in daily life instead of emotions that are artificially evoked in a laboratory based on a stimulus and an expected emotion.

After the raw EEG data got preprocessed, features from the time-domain were extracted, namely the mean, power and standard deviation. Then a random forest classifier and a k-nearest neighbors classifier were trained with parameter optimisation.

The random forest classifier with 10 trees got an accuracy of 92.5% when tested on the training data where the confusion matrix showed true positives and true negatives of above 91% which suggest that this classifier recognizes patterns in the data and reaches accuracies similar to studies using the DEAP data set (Alhagry et al., 2017; Lui et al., 2016). When 5-fold cross validation is used an accuracy of 65.1% is used which is similar to other studies using a random forest classifier with k-fold cross validation (Nascimben et al., 2019).

The KNN classifier (K = 3) got an accuracy of 92.2% when tested on the training data. The confusion matrix showed true positives and true negatives of above 90% which also suggests that the KNN classifier recognizes patterns in the data and is able to perform similar to classifiers trained on the DEAP data set. 5-fold cross validation returned an accuracy of 61.2% which is also similar to Nascimben et al., (2019).

Both classifiers got accuracies above 90% when tested on the training data but the cross fold validation shows that the accuracy decreases by almost 30% when tested on data the classifier has not yet seen before which makes the classifiers less generic.

The exploratory analysis of independent t-tests for the mean found some interesting results. Namely that certain channels were significantly different from each other. This analysis found that channels FC1, FC5, F3, C3 and P7 are significantly different for *happy* EEG data and *angry* EEG data. From figure 3.1 it is visible that four of these electrodes are in the frontal lobe. One electrode is in the parietal lobe. From appendix A it is visible that for all of these electrodes the mean of the brain activity is higher for *happy* EEG data than for *angry* EEG data. This finding suggests that these areas are more active when a person is in a *happy* state than in an *angry* state which is partly in accordance with Machado et al., (2017) which also suggests that the prefrontal cortex is associated with being in a mental state of positive valence. However, the electrode in the parietal lobe is more difficult to place in the context of the literature since research suggests that the parietal lobe is more active when being in a mental state of negative valence (Luo et al., 2016). As mentioned in the introduction, Schmidt et al., 2001 found that negatively valenced emotions were visible in the right frontal lobe. Although the data suggests that the mean of the electrical activity is higher in the right frontal lobe (this is true for electrodes FC2, FC6, F4, F8, AF4, FP2) when in an *angry* state than the mean when someone is in a *happy* state, no significant differences were found in the electrodes placed on this part of the scalp when looking at the mean.

4.1 Problems and improvements

Some aspects of the methods might need improvement. As previously mentioned in the methods section, three researchers were given the task to watch all of the videos of the debates and annotate the challenger or defender was *happy* or *angry*. Annotating emotions is a subjective matter, therefore two researchers had to agree in order for an annotation to become a definitive annotation. This method reduces the aspect of subjectivity but does not minimize it. An improvement would be to provide the researchers examples of debaters being *happy* and *angry* in order to make the annotations of the different researchers more consistent. This could also provide more data which is important for machine learning.

Every final annotation was seen as a single data point. If there was a final annotation of 3 seconds, this was processed as a single data point. However after balancing the data set, all data points were repeated relative to the duration of that annotation since the annotations of the *angry* data points were longer and there was a lack of *angry* data points. Even though this resulted in a more balanced data set, a potential improvement might be to process the EEG data by creating data points with a time interval of 1 second. So for example if a final annotation has a time interval of 4 seconds, the data could be split into four different data points and the features of the EEG data could be extracted per second instead of over the entire time interval.

Furthermore, as cited in Popescu, (2019) 'people from within a culture are better able at recognizing emotions from that culture than people from outside it (Gendron et al., 2014)'. If other debaters who are culturally related to the Tibetan monks (or who are actually Tibetan monks) annotated the emotions of the videos, the annotations might have been more accurate.

4.2 Future research

In order to build upon this research and this data set, multiple features and classifiers should be tested. Since literature presents many promising classifiers with a support vector machine and time-frequency domain features (Iscan et al., 2011) or a neural network instead of machine learning (Lahane & Sangaiah, 2015) this is definitely something worth looking into.

Another very promising addition to the project could be to use channel selection. In this project all channels were used but a genetic algorithm could provide the optimal channels which could increase the accuracy of the classifier (Wen & Zhang, 2017).

5 References

Alhagry, S., Fahmy, A. A., El-Khoribi, R. A. (2017). Emotion recognition based on EEG using LSTM recurrent neural network. Emotion, 8(10), 355-358.

Amin, H. U., Mumtaz, W., Subhani, A. R., Saad, M. N. M., Malik, A. S. (2017). Classification of EEG signals based on pattern recognition approach. Frontiers in Computational Neuroscience, 11(November). https://doi.org/10.3389/fncom.2017.00103

Barrett, L. F., Mesquita, B., Gendron, M. (2011). Context in emotion perception. Current Directions in Psychological Science, 20(5), 286-290.

Chai, T. Y., Woo, S. S., Rizon, M., Tan, C. S. (2010). Classification of human emotions from EEG signals using statistical features and neural network. In International (Vol. 1, No. 3, pp. 1-6). Penerbit UTHM.

Edla, D. R., Mangalorekar, K., Dhavalikar, G., Dodia, S. (2018). Classification of EEG data for human mental state analysis using Random Forest Classifier. Procedia Computer Science, 132(Iccids), 1523–1532. https://doi.org/10.1016/j.procs.2018.05.116

Ekman, P., Levenson, R. W., Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. science, 221(4616), 1208-1210.

Ferretti, V., Papaleo, F. (2019). Understanding others: emotion recognition in humans and other animals. Genes, Brain and Behavior, 18(1), e12544.

Gendron, M., Roberson, D., van der Vyver, J. M., Barrett, L. F. (2014). Perceptions of emotion from facial expressions are not culturally uni- versal: Evidence from a remote culture. Emo- tion, 14(2), 251-262.

Gessner, S. N. (2015). The Effect of Emotion Stimulus Intensity on the selection and Implementation of Distraction and Reappraisal as Emotion Regulation Strategies (Doctoral dissertation).

Iscan, Z., Dokur, Z., Demiralp, T. (2011). Classification of electroencephalogram signals with combined time and frequency features. Expert Systems with Applications, 38(8), 10499–10505. https://doi.org/10.1016/j.eswa.2011.02.110

Jin, C. Y., Borst, J. P., van Vugt, M. K. (2019). Predicting task-general mind-wandering with EEG. Cognitive, Affective, Behavioral Neuroscience, 19(4), 1059-1073.

Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., ... Patras, I. (2011). Deap: A database for emotion analysis; using physiological signals. IEEE transactions on affective computing, 3(1), 18-31.

Lahane, P., Sangaiah, A. K. (2015). An approach to eeg based emotion recognition and classification using kernel density estimation. Procedia Computer Science, 48(C), 574–581. https://doi.org/10.10

Lisetti, C. L., Nasoz, F. (2004). Using noninvasive wearable computers to recognize human emotions from physiological signals. EURASIP Journal on Advances in Signal Processing, 2004(11), 929414.

Liu, J., Meng, H., Nandi, A., Li, M. (2016). Emotion detection from EEG recordings. 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery, ICNC-FSKD 2016, 1722–1727. https://doi.org/10.1109/FSKD.2016.7603437

Liu, W., Zheng, W. L., Lu, B. L. (2016, October). Emotion recognition using multimodal deep learning. In International conference on neural information processing (pp. 521-529). Springer, Cham.

Luo, Y., Kong, F., Qi, S., You, X., Huang, X. (2016). Resting-state functional connectivity of the default mode network associated with happiness. Social cognitive and affective neuroscience, 11(3), 516-524.

Machado, L., Cantilino, A. (2017). A systematic review of the neural correlates of positive emotions. Brazilian Journal of Psychiatry, 39(2), 172-179.

Moshfeghi, M., Bartaula, J. P., Bedasso, A. T. (2013). Emotion Recognition from EEG Signals using Machine Learning.

Nascimben, M., Ramsøy, T. Z., Bruni, L. E. (2019, October). User-independent classification of emotions in a mixed arousal-valence model. In 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE) (pp. 445-449). IEEE.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Computational intelligence and neuroscience, 2011.

Popescu, M. (2019). Distinguishing anger in the brain using machine learning on EEG data.

Radüntz, T., Scouten, J., Hochmuth, O., Meffert, B. (2015). EEG artifact elimination by extraction of ICA-component features using image processing algorithms. Journal of neuroscience methods, 243, 84-93.

Schlosberg, H. (1954). Three dimensions of emotion. Psychological review, 61(2), 81.

Schmidt, L. A., Trainor, L. J. (2001). Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions. Cognition Emotion, 15(4), 487-500.

Shu, L., Xie, J., Yang, M., Li, Z., Li, Z., Liao, D., Xu, X., Yang, X. (2018). A Review of Emotion Recognition Using Physiological Signals. Sensors (Basel, Switzerland), 18(7), 2074. https://doi.org/10.3390/s18072074

Takahashi, K. (2004, September). Remarks on SVM-based emotion recognition from multi-modal bio-potential signals. In RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759) (pp. 95-100). IEEE.

van Vugt, M. K., Pollock, J., Johnson, B., Gyatso, K., Norbu, N., Lodroe, T., ... Lobsang, J. (2020). Inter-brain synchronization in the practice of Tibetan monastic debate. Mindfulness, 1-15.

Wen, T., Zhang, Z. (2017). Effective and extensible feature extraction method using genetic algorithm-based frequency-domain feature search for epileptic EEG multiclassification. Medicine (United States), 96(19), 1–17. https://doi.org/10.1097/MD.0000000000006879

Zheng, W. L., Zhu, J. Y., Lu, B. L. (2017). Identifying stable patterns over time for emotion recognition from EEG. IEEE Transactions on Affective Computing.

Appendix A

Channel	happy mean	angry mean
Fp1	-0.0545	0.395
AF3	0.252	-0.152
F7	0.135	0.0329
F3	0.0251	-0.266
FC1	0.0489	-0.280
FC5	0.0344	-0.452
T7	0.0393	0.113
C3	0.0301	-0.278
CP1	0.00277	-0.188
CP5	0.0105	-0.103
P7	0.00620	0.639
P3	-0.0289	-0.104
Pz	0.0162	-0.142
PO3	0.00882	0.00808
PO1	0.00531	0.215
OZ	-0.0111	0.228
O2	-0.0912	0.176
PO4	-0.0313	0.0252
P4	-0.0342	0.0102
P8	-0.0987	0.112
CP6	0.0238	0.0646
CP2	0.0151	-0.191
C4	0.00859	-0.0535
T8	-0.0757	0.193
FC6	-0.0706	0.150
FC2	-0.0768	-0.210
F4	-0.0170	-0.0122
F8	-0.0149	0.211
AF4	-0.120	0.232
FP2	-0.00382	0.144
Fz	0.0325	-0.342
Cz	0.00895	-0.179

Table A.1: Mean of the mean of all data points for *happy* and *angry* EEG data per channel

Appendix B

Mean T-test results:

t-test channel: Fp1
(statistic=-0.592, p-value=0.554)
t-test channel: AF3
(statistic=1.378, p-value=0.171)
t-test channel: F7
(statistic=0.439, p-value=0.661)
t-test channel: F3
(statistic=2.391, p-value <0.05) *
t-test channel: FC1
(statistic=2.645, p-value <0.05) *
t-test channel: FC5
(statistic=2.087, p-value <0.05) *
t-test channel: T7
(statistic=-0.322, p-value=0.747)
t-test channel: C3
(statistic=2.113, p-value <0.05) *
t-test channel: CP1
(statistic=1.1470, p-value=0.255)
t-test channel: CP5
(statistic=0.641, p-value=0.523)
t-test channel: P7
(statistic=-2.050, p-value <0.05) *
t-test channel: P3
(statistic=0.984, p-value=0.328)
t-test channel: Pz
(statistic=1.011, p-value=0.314)
t-test channel: PO3
(statistic=0.004, p-value=0.996)
t-test channel: O1
(statistic=-1.103, p-value=0.273)
t-test channel: OZ
(statistic=-1.019, p-value=0.310)
t-test channel: O2
(statistic=-1.325, p-value=0.189)
t-test channel: PO4
(statistic=-0.478, p-value=0.633)
t-test channel: P4
(statistic=-0.359, p-value=0.720)
t-test channel: P8
(statistic=-0.965, p-value=0.336)
t-test channel: CP6
(statistic=-0.220, p-value=0.826)
t-test channel: CP2
(statistic=1.534, p-value=0.129)
t-test channel: C4
(statistic=0.406, p-value=0.685)

t-test channel: T8
(statistic=-1.366, p-value=0.175)
t-test channel: FC6
(statistic=-1.251, p-value=0.214)
t-test channel: FC2
(statistic=0.901, p-value=0.369)
t-test channel: F4
(statistic=-0.008, p-value=0.993)
t-test channel: F8
(statistic=-0.836, p-value=0.405)
t-test channel: AF4
(statistic=-1.529, p-value=0.129)
t-test channel: FP2
(statistic=-0.236, p-value=0.813)
t-test channel: FZ
(statistic=1.876, p-value=0.064)
t-test channel: CZ
(statistic=1.190, p-value=0.237)

Table B.1: Dependent t-test between the means for happy and angry EEG data per channel, DF = 24255

Power T-test results:
t-test channel: Fp1
(statistic=-0.741, p-value=0.459)
t-test channel: AF3
(statistic=-2.771, p-value <0.05) *
t-test channel: F7
(statistic=-3.857, p-value <0.05) *
t-test channel: F3
(statistic=-3.387, p-value <0.05) *
t-test channel: FC1
(statistic=-4.199, p-value <0.05)*
t-test channel: FC5
(statistic=-5.956, p-value <0.05) *
t-test channel: T7
(statistic=-2.431, p-value <0.05) *
t-test channel: C3
(statistic=-4.858, p-value <0.05) *
t-test channel: CP1
(statistic=-3.554, p-value <0.05) *
t-test channel: CP5
(statistic=-4.582, p-value <0.05)
t-test channel: P7
(statistic=-2.705, p-value <0.05) *
t-test channel: P3
(statistic=-3.776, p-value <0.05) *
t-test channel: Pz
(statistic=-4.339, p-value <0.05) *
t-test channel: PO3
(statistic=-4.230, p-value <0.05) *
t-test channel: O1
(statistic=-4.951, p-value <0.05 *
t-test channel: Oz
(statistic=-3.209, p-value <0.05) *
t-test channel: O2
(statistic=-5.047, p-value <0.05) *
t-test channel: PO4
(statistic=-4.007, p-value <0.05) *
t-test channel: P4
(statistic=-4.682, p-value <0.05) *
t-test channel: P8
(statistic=-3.693, p-value <0.05) *
t-test channel: CP6
(statistic=-3.946, p-value <0.05) *
t-test channel: CP2
(statistic=-3.336, p-value <0.05) *
t-test channel: C4
(statistic=-4.322, p-value <0.05) *
t-test channel: T8
(statistic=0.0217, p-value=0.982)
t-test channel: FC6
(statistic=-3.796, p-value <0.05) *
t-test channel: FC2
(statistic=-2.951, p-value <0.05) *
t-test channel: F4
(statistic=-3.734, p-value <0.05) *
t-test channel: F8
(statistic=-4.699, p-value <0.05) *
t-test channel: AF4
(statistic=-3.747, p-value <0.05) *
t-test channel: Fp2
(statistic=-2.214, p-value <0.05) *
t-test channel: Fz
(statistic=-3.743, p-value <0.05) *
t-test channel: Cz
(statistic=-4.739, p-value <0.05) *

Table B.2: Dependent t-test between the powers for happy and angry EEG data per channel, DF = 24255

Standard Deviation T-test results:
t-test channel: Fp1
(statistic=-0.770, p-value=0.441)
t-test channel: AF3
(statistic=-2.765, p-value <0.05 *
t-test channel: F7
(statistic=-3.861, p-value <0.05 *

t-test channel: F3
(statistic=-3.373, p-value <0.05 *
t-test channel: FC1
(statistic=-4.191, p-value <0.05 *
t-test channel: FC5
(statistic=-5.907, p-value <0.05 *
t-test channel: T7
(statistic=-2.429, p-value <0.05 *
t-test channel: C3
(statistic=-4.844, p-value <0.05 *
t-test channel: CP1
(statistic=-3.538, p-value <0.05 *
t-test channel: CP5
(statistic=-4.557, p-value <0.05 *
t-test channel: P7
(statistic=-2.692, p-value <0.05 *
t-test channel: P3
(statistic=-3.779, p-value <0.05 *
t-test channel: Pz
(statistic=-4.339, p-value <0.05 *
t-test channel: PO3
(statistic=-4.233, p-value <0.05 *
t-test channel: 01
(statistic=-4.930, p-value <0.05 *
t-test channel: Oz
(statistic=-3.196, p-value <0.05 *
t-test channel: O2
(statistic=-5.034, p-value <0.05 *
t-test channel: PO4
(statistic=-4.029, p-value <0.05 * *
t-test channel: P4
(statistic=-4.690, p-value <0.05 * *
t-test channel: P8
(statistic=-3.686, p-value <0.05 * *
t-test channel: CP6
(statistic=-3.939, p-value <0.05 *
t-test channel: CP2
(statistic=-3.325, p-value <0.05 *
t-test channel: C4
(statistic=-4.311, p-value <0.05 *
t-test channel: T8
(statistic=0.0520, p-value=0.9585)
t-test channel: FC6
(statistic=-3.768, p-value <0.05 *
t-test channel: FC2
(statistic=-2.946, p-value <0.05 *
t-test channel: F4
(statistic=-3.722, p-value <0.05 *
t-test channel: F8
(statistic=-4.684, p-value <0.05 *
t-test channel: AF4
(statistic=-3.738, p-value <0.05 *
t-test channel: Fp2
(statistic=-2.225, p-value <0.05 *
t-test channel: Fz
(statistic=-3.730, p-value <0.05 *
t-test channel: Cz
(statistic=-4.725, p-value <0.05 *

Table B.3: Dependent t-test between the standard deviations for happy and angry EEG data per channel, DF = 24255

Appendix C

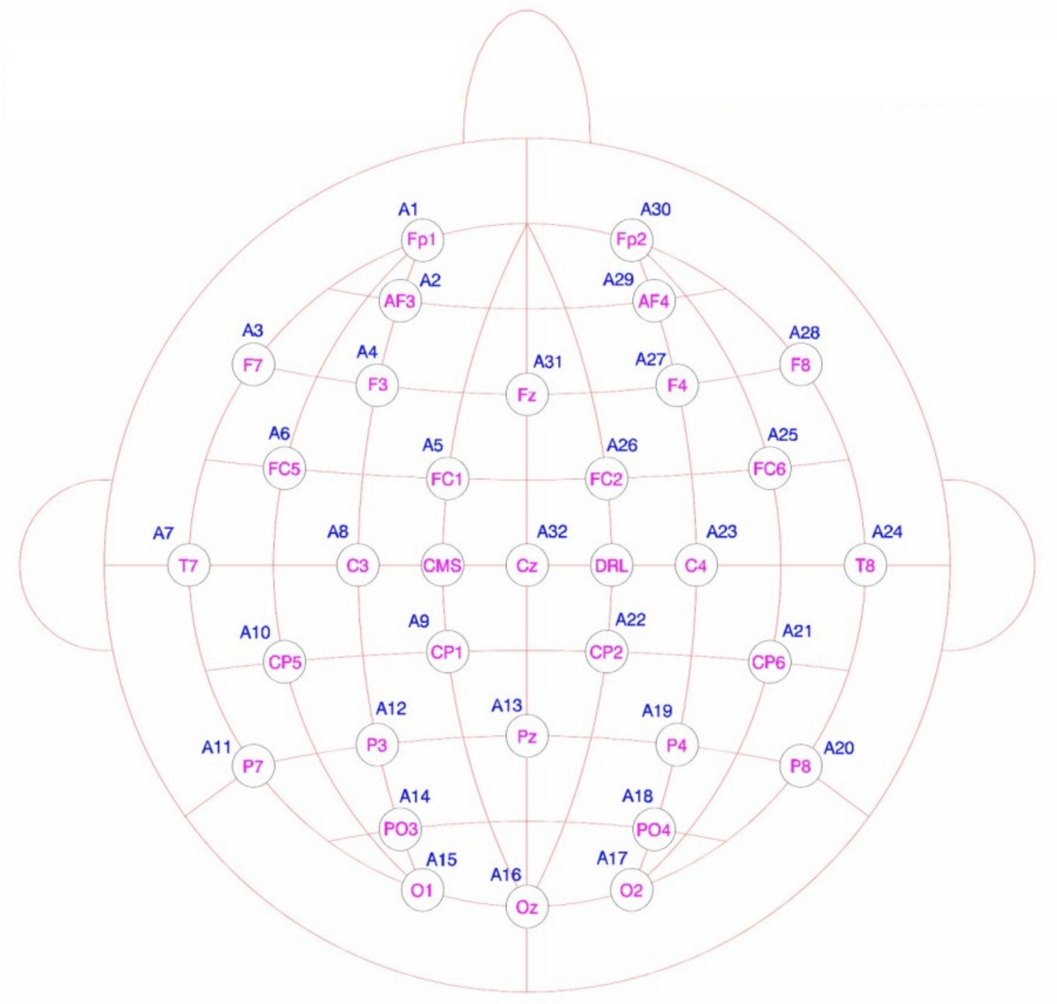

Figure .1: Overview of the locations of the electrodes as used in the experiment.