

# Research Proposal: Machine Learning for Emotion Unveiling in a multidimensional World

Alan Teesdale

Supervisors: Qi Chen, Bing Xue

November 18, 2024

## Part 1 Problem Statement

The problem of Facial Expression Recognition (FER) is a classic task for computer vision, where the aim is to take an image of a person and produce and be able to learn to recognise the expression on their face. These can be broadly categorised into two buckets. The discrete/categorical models, where emotions are divided into discrete categories (happy, sad, surprised, tired, ...), and dimensional models, where emotions are represented on a continuous spectrum. One such model is the circumplex model<sup>[1]</sup> where emotions are represent by two values Valence, how positive the emotion is, and Arousal, how intense the emotion is.

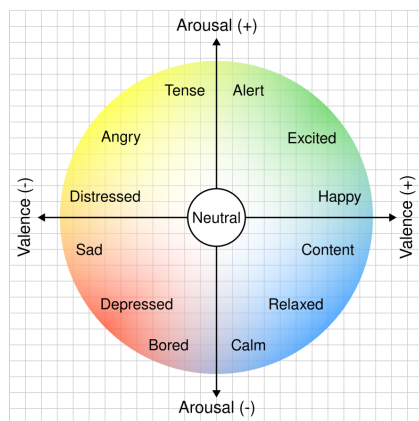


Figure 1: A figure depicting the circumplex model. Image taken from wikimedia<sup>[2]</sup>

For this project we aim to use a Genetic Programming approach to construct a human interpretable model to take a to take an image of a face and convert and extract the valance and arousal corresponding to the facial expression.

## Part 2 Motivation

Facial expressions is an incredibly powerful methods of human subconscious communication, allowing machine systems to sense and change behaviour based on the emotions of the user could allow

for new<sup>1</sup> types of machine human interaction and potentially smooth out others that already exist. Because of this an other reasons, FER is a classic task in computer vision that has been studied for many decades<sup>[3]</sup>.

Since the field of FER is so mature, many models have been developed for this task (although the majority of the work focuses on categorical models, which by it's nature can struggle to capture mixed emotions), including similar GP models<sup>[4]</sup>. However, a historical constraint of many of these models is the available training data, as historically the photos are often taken in laboratory conditions, well lit, posed, or are of smaller datasets<sup>[5]</sup>. Both of these effects mean that these models tend to do much worse “in the wild” where the data is “messier” (variable lighting, faces at different angles, e.t.c.). To address these concerns *AffectNet*<sup>[5]</sup> was created, a database with more than 1 million faces in these “messier” conditions. In this project we will use the *AffectNet* dataset to train a new FER model that will hopefully be able handle the variable conditions needed to FER “in the wild”.

The preferred model in literature appears to be CNN based methods<sup>[6]</sup>, as they often provide very good results in practice. However, the explainability of these models is very poor, especially with deep CNN's, so often these models get treated like black boxes. GP methods have been used<sup>[7]</sup> for similar tasks, with comparable results to other models, including CNN based methods.

While there has been use of GP for similar tasks, the use of GP based models on the *AffectNet* dataset is seemingly unexplored, and GP based models that predict dimensional models. In this project we aim to extend the use GP models for FER to use the *AffectNet* database and to predict the values for arousal and valence. We hope to achieve the following:

1. Use genetic programming to create a reasonably simple and explainable FER model
2. Maintain roughly comparable performance to current CNN models
3. Find a minimal but effective function set, so that the training time is reduced.

---

<sup>1</sup>and potentially dystopian

## Part 3 How?

To achieve this we will use Strongly typed GP to enable us to create a multi-layered GP model where we have layers corresponding to different tasks. Figure 2 depicts what a model might look like. The first layer represents the inputs to the system, where the image is given by the user and the other constants are evolved throughout the running of the program. The second layer takes these inputs and converts them into regions of interest, the third layer takes these regions and transforms them into a numerical feature (an example might take a Region R run it through a transformation to pull out the edges, then finally get the std of that to convert it to a number). The fourth layer takes the numbers produced by the previous layer and operates on those number (scaling, division, addition, subtraction...) till only 2 values remain. These two values are combined in the 5th layer into a pair and then outputted as the final result.

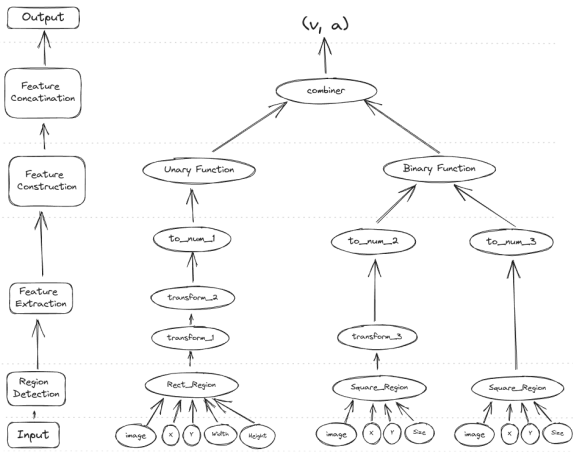


Figure 2: An example of what a tree generated by this program could look like. Image and general architecture inspired by<sup>[8]</sup>

## Part 4 Milestones

The Plan of this project will hopefully be something like the following:

- Research related work.
- Implement simple models.
- Implement more complex and powerful GP models.
- Increase the complexity of the function set.
- Decide on most effective model (taking into account accuracy and complexity of produced models).
- See if the complexity of the models/function set can be reduced without impacting the performance too much.
- Compare to existing models.

## Part 5 Resources

The resources needed for this project should be pretty light, training the GP models could be computationally expensive, especially if lots of runs with different function sets to determine are needed to good function sets.

## References

- [1] J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, pp. 1161–1178, 6 Dec. 1980. DOI: [10.1037/h0077714](https://doi.org/10.1037/h0077714). [Online]. Available: <https://doi.org/10.1037/h0077714>.
- [2] W. Commons, *File:circumplex model of emotion.svg — wikimedia commons, the free media repository*, [Online; accessed 11-November-2024], 2024. [Online]. Available: [https://commons.wikimedia.org/w/index.php?title=File:Circumplex\\_model\\_of\\_emotion.svg&oldid=861227692](https://commons.wikimedia.org/w/index.php?title=File:Circumplex_model_of_emotion.svg&oldid=861227692).
- [3] X. Zhao, J. Zhu, B. Luo, and Y. Gao, "Survey on facial expression recognition: History, applications, and challenges," *IEEE MultiMedia*, vol. 28, no. 4, pp. 38–44, 2021. DOI: [10.1109/MMUL.2021.3107862](https://doi.org/10.1109/MMUL.2021.3107862).
- [4] H. Ghazouani, "A genetic programming-based feature selection and fusion for facial expression recognition," *Applied Soft Computing*, vol. 103, p. 107173, 2021, ISSN: 1568-4946. DOI: <https://doi.org/10.1016/j.asoc.2021.107173>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S156849462100096X>.
- [5] A. Mollahosseini, B. Hasani, and M. H. Mahoor, "Affectnet: A database for facial expression, valence, and arousal computing in the wild," *IEEE Transactions on Affective Computing*, vol. 10, no. 1, pp. 18–31, Jan. 2019, ISSN: 2371-9850. DOI: [10.1109/taffc.2017.2740923](https://doi.org/10.1109/taffc.2017.2740923). [Online]. Available: <http://dx.doi.org/10.1109/TAFFC.2017.2740923>.
- [6] M. Sajjad, F. U. M. Ullah, M. Ullah, et al., "A comprehensive survey on deep facial expression recognition: Challenges, applications, and future guidelines," *Alexandria Engineering Journal*, vol. 68, pp. 817–840, 2023, ISSN: 1110-0168. DOI: <https://doi.org/10.1016/j.aej.2023.01.017>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1110016823000327>.
- [7] B. Evans, H. Al-Sahaf, B. Xue, and M. Zhang, "Evolutionary deep learning: A genetic programming approach to image classification," in *2018 IEEE Congress on Evolutionary Computation (CEC)*, 2018, pp. 1–6. DOI: [10.1109/CEC.2018.8477933](https://doi.org/10.1109/CEC.2018.8477933).
- [8] M. Z. Ying Bi Bing Xue, *Genetic Programming for Image Classification*. Springer, 2021.