Detecting Micro expressions for Early Identification of Mental Health Issues

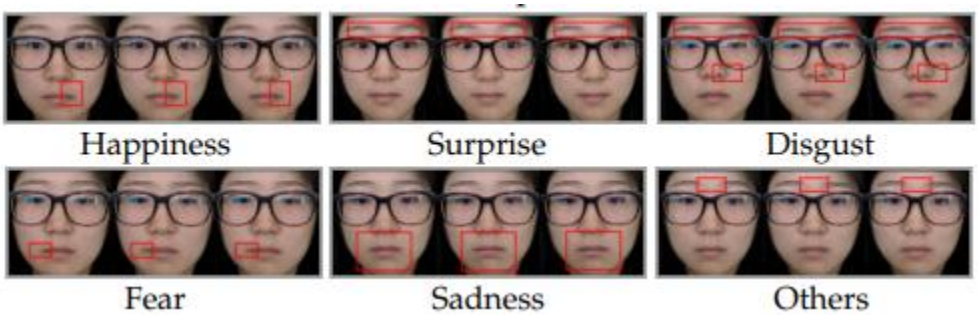
Abstract:

Now days people try hidden there expression from other.We can easily detect macro expression. But it is heard Micro expression. Specially under development country where people are not aware about there feeling. Longtime it give a big impact. And on this type of country do have enough psycratist to treatment. If detect this then we can treatment people or take them to psycratist for advance treatment. So we will to detect in micro expression from video where we can analysis video and get snape from those video we connect snape from there and analysis those image and detected this image. We will collect data in manuualy and also from online source as well.

Introduction:

Emotions are an inherent part of human life, manifesting either voluntarily or involuntarily through facial expressions during face-to-face communication[1]. As a typical form of nonverbal communication, facial expressions play an essential role in analyzing human emotion[2]. Broadly speaking, facial expressions can be classified into two categories: macro-expressions and micro-expressions (Figure 1). The primary difference between these two types lies in both their **duration** and **intensity**.

(a) **Macro-expressions**



(b) **Micro-expressions**

While macro-expressions can be analyzed from a single image, micro-expressions require analysis across an image sequence due to their low intensity. The subtle changes in micro-expressions, highlighted in the red boxes, are explained in the supplemental material. Face tracking is an important topic in machine vision, which is used to recognize faces, facial expressions, and even speech. Face detection is the first step in any facial processing system [3], [4]. Image-to-video translation can seem like an ill-posed problem because the output involves more unknowns to fill in than the input values. Despite various works in video generation [5-11], these approaches typically take multiple video frames as input and extrapolate future frames based on recurrent patterns, which prevents them from solving image-to-video translation where no temporal cues are provided in the input. Moreover, generating satisfactory video clips of facial expressions poses two challenges:

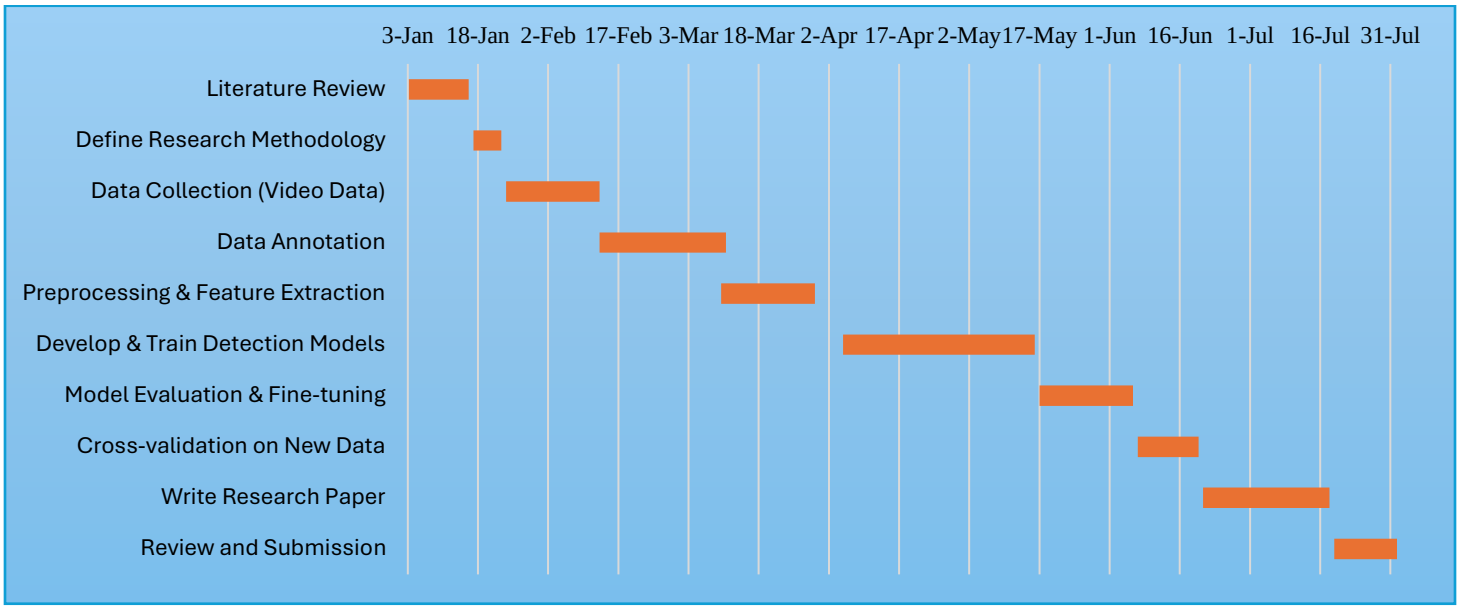
1. **Human sensitivity**: People are very familiar with facial expressions and can easily detect artifacts in spatial or temporal dimensions.
2. **Face identity preservation**: The generated video clips must preserve the identity of the face. This requires neural networks to learn the "imagination" capabilities rather than memorizing faces seen during training, enabling them to handle new faces during deployment.

Despite these challenges, tackling image-to-video translation, especially for facial expression generation, remains feasible for several reasons:

- **Similar emotional expressions**: Different people express emotions in similar ways. For instance, people tend to open their mouths when excited or surprised.
- **Unimodal expressions**: Facial expressions generally follow a gradual change from a neutral state to an extreme expression. For example, happiness increases monotonically until reaching its peak expression.
- **Focus on the face**: In profile photos, the human face draws most of the viewer's attention, making the quality of the background less significant. These factors help reduce variability in video frames, making image-to-video translation more plausible.

In this study, we aim to develop a robust framework for facial expression generation through image-to-video translation, with a particular focus on micro-expressions. Our approach leverages advanced machine learning techniques to address the dual challenges of human sensitivity to facial expressions and the preservation of facial identity. We will explore methodologies to enhance realism, maintain expression continuity, and optimize the generation of subtle changes in micro-expressions. The ultimate goal is to provide a practical tool for applications in mental health analysis, emotion detection, and other fields requiring detailed understanding of human expressions.

Gantt Chart:



References

[1] X. Ben et al., "Video-based Facial Micro-Expression Analysis: A Survey of Datasets, Features and Algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–1, 2021, doi: 10.1109/tpami.2021.3067464.
[2] A. Freitas-Magalhaes, "The Psychology of Emotions: The Allure of the Human Face," University Fernando Pessoa Press, Oporto, 2007.
[3] N. Esau, E. Wetzel, L. Kleinjohann, and B. Kleinjohann, "Real-time facial expression recognition using a fuzzy emotion model," in *Proc. Fuzzy Systems Conf.*, 2007, pp. 1-6.
[4] I. Cohen, N. Sebe, A. Garg, M. S. Lew, and T. S. Huang, "Facial expression recognition from video sequences: temporal and static modeling," *Comp*.
[5] Junhyuk Oh, Xiaoxiao Guo, Honglak Lee, Richard L. Lewis, and Satinder Singh, "Action-conditional video prediction using deep networks in Atari games," *Advances in Neural Information Processing Systems*, pp. 2863–2871, 2015.
[6] Michael Mathieu, Camille Couprie, and Yann LeCun, "Deep multi-scale video prediction beyond mean square error," *arXiv preprint arXiv:1511.05440*, 2015.

[7] William Lotter, Gabriel Kreiman, and David Cox, “Deep predictive coding networks for video prediction and unsupervised learning,” *arXiv preprint arXiv:1605.08104*, 2016.
[8] Ruben Villegas, Jimei Yang, Yuliang Zou, Sungryull Sohn, Xunyu Lin, and Honglak Lee, “Learning to generate long-term future via hierarchical prediction,” *arXiv preprint arXiv:1704.05831,* 2017.
[9] Sergey Tulyakov, Ming-Yu Liu, Xiaodong Yang, and Jan Kautz, “Mocogan: Decomposing motion and content for video generation,” *arXiv preprint arXiv:1707.04993*, 2017.
[10] Ziwei Liu, Raymond Yeh, Xiaoou Tang, Yiming Liu, and Aseem Agarwala, “Video frame synthesis using deep voxel flow,” in *International Conference on Computer Vision (ICCV)*, vol. 2, 2017.
[11] Carl Vondrick and Antonio Torralba, “Generating the future with adversarial transformers,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.