**B.Tech 2020-24 CSE- Project Phase 1**

**Proposal**

# I. Group No: D15

Project Title.: Deep Learning for Micro-expression Recognition

Team members :

|  |  |
| --- | --- |
| Roll No. | Name |
| AM.EN.U4CSE20316 | B SAI SARVAGNA |
| AM.EN.U4CSE20329 | GOUTHAM VIJAY |
| AM.EN.U4CSE20374 | VIGNESH A NAIR |

# II. Abstract

The goal of the research is to investigate the use of Deep Learning (DL) methods for Micro-expression Recognition (MER). Micro-expressions are unconscious, minute facial movements that indicate concealed emotions. Due to their spontaneous and quick nature, the challenge is in accurately identifying and recognising these fleeting emotions, which makes data collection and interpretation challenging. The practical uses of MER in a number of areas, such as emotion analysis, deception detection, and human-computer interaction, make this issue relevant. The project's goal is to use DL developments to improve MER performance and create reliable systems for use in the field. Limited accessible datasets, the difficulty of recording and annotating Micro-expressions, and the necessity for MER technology to be used in an ethically acceptable manner.

# III. Background Study

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title &year | Problem | Contributions | Limitations | Open problems/Future work |
| 1. CapsuleNet for Micro-Expression Recognition, **Published in:** [2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)](https://ieeexplore.ieee.org/xpl/conhome/8753205/proceeding) | The limited availability of publicly available facial micro expression datasets and the low intensity of facial movements in these datasets. This presents a significant challenge in training robust data models to recognize facial micro expressions. | CapsuleNet to micro-expression recognition, The work introduced a complete framework with a CapsuleNet for micro-expression recognition using only apex frames. Apex frames are peak frames within a specific dataset indicating maximum significance or intensity. | By using only spike frames, the method may lose some of the temporal context and all the information contained in the micro expression sequence. This can potentially result in less accurate detection, especially for longer and more complex micro expression patterns that evolve over time. | Focus on further optimizing the additional design. The primary goal would be to reduce the computational complexity while maintaining or even enhancing the model's performance on small micro-expression datasets. |
| 2.  Deep learning for micro-expression recognition: A survey in 2022  **Published in:** [IEEE Transactions on Affective Computing](https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=5165369) ( Volume: 13, [Issue: 4](https://ieeexplore.ieee.org/xpl/tocresult.jsp?isnumber=9964459&punumber=5165369), 01 Oct.-Dec. 2022) | The research tackles on the micro expression briefness, low intensity, and capacity for conscious suppression, micro expressions may be difficult to recognise and understand. Here, DL method is a "black box" and focuses on identifying patterns and learning new features by altering the weights of networks. | In order to capture temporal connections, the study offers a deep learning model for micro-expression identification using CNN and RNN. It also explores the advantages of multimodal fusion by using auditory or physiological data to improve accuracy of the recognition. | Large-scale, heterogeneous micro-expression datasets may be restricted. The accuracy and representativeness of the data have a significant impact on the model's performance. | Analysis of Action Units (AU) in Micro Expressions.  Action Units are specific facial muscle movements or configurations defined and coded in the Facial Action Coding System. The ME interpretation, however, suffers from ambiguity, It has been shown that the FACS effectively addresses the ambiguity problem. |
| 3.  A survey: facial micro-expression recognition Cite this article Takalkar, M., Xu, M., Wu, Q. *et al.* A survey: facial micro-expression recognition. *Multimed Tools Appl* **77**, 19301–19325 (2018). https://doi.org/10.1007/s11042-017-5317-2 | The problem identified is Lack of a standard micro-expression database paper aims to provide a survey of existing micro-expression recognition methods and their outcomes to help researchers understand the recent developments in this field and address the challenges in micro-expression analysis | The paper provides a survey of cutting-edge facial micro-expression recognition approaches, The paper compares handcrafted and deep learning approaches, highlighting the performance differences based on the datasets used.  The paper classifies and discusses various public micro-expression datasets, including CASME, CASME II, SMIC, USF-HD, Polikovsky, YorkDDT, and METT. | Micro-expressions can be easily mistaken for the standard macro-expressions, making it challenging to train a system based on existing macro facial expression databases without considering their dynamic information | More robust deep learning models, refining real-time efficiency by estimating pixel-level movement, expanding databases with diverse participants for better training, developing systems for less constrained situations, and conducting further research on deep learning features, like CNN, to improve recognition results on additional datasets. |
| 4. Zhang, S., Pan, X., Cui, Y., Zhao, X. and Liu, L., 2019. Learning affective video features for facial expression recognition via hybrid deep learning. IEEE Access, 7, pp.32297-32304. | The problem identified is the need for effective facial expression recognition (FER) in video sequences. While image-based methods work well for spatial information analysis from still images, they fail to capture the temporal variability in consecutive frames present in video sequences. | The paper introduces a novel approach for automatic facial expression recognition (FER) in video sequences using a hybrid deep learning model. The proposed approach combines the strengths of deep Convolutional Neural Networks (CNNs) for spatial and temporal feature learning and a Deep Belief Network (DBN) for deep fusion. | It provides a brief overview of the deep fusion network constructed with a Deep Belief Network (DBN). However, it lacks in-depth explanations regarding the mechanism of the deep fusion network, and information on its functioning. | The potential areas of future work the challenge of developing a real-time Facial Expression Recognition (FER) system based on their proposed method. As part of future work, they may focus on optimizing the model architecture, exploring hardware acceleration techniques, exploring deep compression techniques, such as model quantization, weight pruning, knowledge distillation, or network distillation. |
| 5.  Bodapati, J.D., Naik, D.S.B., Suvarna, B. *et al.* A Deep Learning Framework with Cross Pooled Soft Attention for Facial Expression Recognition. *J. Inst. Eng. India Ser. B* **103**, 1395–1405 (2022) | The problem statement is to develop a robust deep learning-driven framework for identifying facial expressions from given facial images. The goal is to recognize certain basic facial expressions such as anger, disgust, fear, happiness, sadness, and surprise. The proposed framework focuses on extracting discriminating features from localized facial regions, including eye-pair, mouth, and normalized face, using a pre-trained Convolutional Neural Network (CNN) with channel-wise convolutions. | The proposed attention-based deep learning model demonstrates efficient recognition of facial emotions on the JAFFE dataset. The inclusion of dropout at the input layer and batch normalization after every hidden layer in the proposed model helps in reducing model overfitting. In the experiments on the CK+ dataset, the model identifies potential noise present in the dataset. | Real-time processing presents a significant challenge as it requires the model to make predictions quickly and efficiently to be suitable for real-world applications. Incorporating features from multiple face regions, such as eye-pair, mouth, and normalized facial regions, poses a challenge in terms of feature fusion. Evaluating the model's performance on different datasets, such as JAFFE and CK+, introduces the challenge of generalization. Selecting appropriate evaluation metrics to assess the model's performance comprehensively is a challenge. | As a future extension, the proposed model's performance should be evaluated on facial images captured in unconstrained environments. Investigating the use of transfer learning with pre-trained models on large-scale datasets can accelerate the model's learning process and improve its performance in unconstrained environments. |

# IV. Challenges

**Limited Data Availability:** The lack of publicly accessible large-scale datasets is one of the main problems with Deep Learning for Micro-expression Recognition. The spontaneous character of these emotions makes it difficult to gather and annotate the data, which leads to a lack of training data for DL models

**Temporal Modelling:** The term "temporal" highlights the importance of considering the time dimension in modeling micro-expressions accurately, as conventional image-based deep learning algorithms may not be able to fully capture the temporal dynamics exhibited in micro-expression sequences. For accurate MER, we need to consider temporal changes while taking geographical information into account.

**Overfitting and Generalisation:** DL models are vulnerable to overfitting with short datasets, where they memorise the few training instances and struggle to generalise successfully to new data. Building robust and dependable MER systems that function effectively in real-world circumstances requires the development of methods to reduce overfitting and enhance generalisation performance.

# V. Deliverables of Phase I

**Objectives**

* To perform a thorough assessment of research on deep learning techniques used to micro-expression identification and to entails locating pertinent papers, methodology, datasets, and difficulties associated with deep learning-based micro-expression analysis.
* To create a thorough survey questionnaire to get opinions from deep learning and micro-expression recognition practitioners, researchers, and experts.
* To conduct an analysis of the data gathered from the survey and literature study to find patterns, gaps, and certainly develop deep learning approaches.

**Outcomes/Deliverables**

Here are some potential outcomes:

**Making a High-Performance Model:**

The study could result in the development of a reliable and precise deep learning model for micro-expression recognition. This model may be more accurate and efficient in identifying and classifying micro-expressions than current approaches.

**Fusion of Multiple Modalities to Improve Recognition**

Improved recognition models could result from investigating the integration of various modalities, such as audio or physiological data. Combining information from many sources may increase the model's resilience and improve its performance in real-world circumstances.

**Micro-Expression Analysis Ethical Conduct:**

The study could help in the creation of moral standards for microexpression analysis. For the technology to be used responsibly, privacy and consent concerns linked to video data consumption must be addressed.

**Applications in the Real World:**

The suggested study may open up a wide range of practical uses for the identification of micro-expressions. These applications may enhance security and lie detection technology, improve human-computer interaction, and provide insightful data for psychological studies.

**Making a Difference in the Field:**

The study's findings could advance knowledge and comprehension of deep learning-based micro-expression recognition among scientists.

Innovating the following areas:

a. Temporal modelling: To better capture the dynamics of micro-expressions, look at the efficacy of temporal modelling approaches like 3D CNNs and attention processes.

b. Multi-modal Fusion: Investigate the incorporation of other modalities, such as audio or physiological data, to improve the reliability and accuracy of the recognition model.

c. Transfer Learning: Determine if it is feasible to apply what has been learned from previously trained models to similar tasks, such as recognising facial expressions, in order to enhance the recognition of micro-expressions.

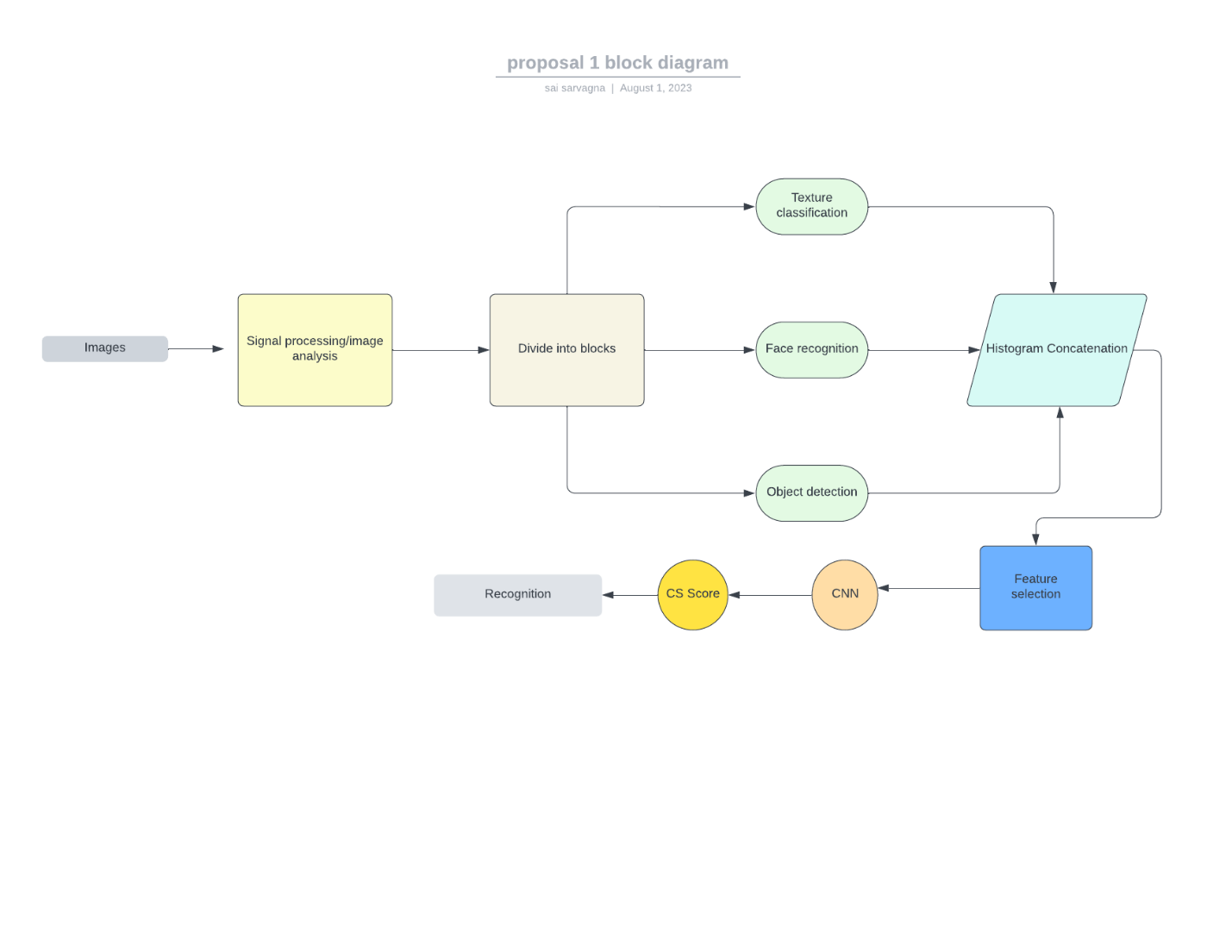
# VI. Assumptions/Declarations

**Data Access:** Adequate and varied dataset of micro-expressions will be accessible for the deep learning model's training and evaluation. The dataset needs to include a variety of people, feelings, and environmental factors.

# VII. Tools to be used

|  |  |
| --- | --- |
| **Software** | **Specifications** |
| **Python language**  **Jupyter notebook software**  **Libraries**  TensorFlow  Keras  OpenCV  NumPy | **Python:** With well-known tools like TensorFlow and Keras for creating and refining deep learning models, Python is a commonly utilised programming language in the area of deep learning.  **TensorFlow:** TensorFlowis an open-source deep learning framework that Google created that offers a high-level API for creating and training neural networks.  A high-level neural network API called **Keras** may be used with TensorFlow.  **OpenCV** is a potent library for computer vision applications, such as analysing images and videos. It may be used to enhance datasets and preprocess picture data.  **NumPy:** A essential library for Python's numerical processing, NumPy is widely used to manage numerical data, including picture data.  **Jupyter Notebook:** Jupyter Notebook is an interactive computing environment that makes it easier to explore data, construct models, and visualise the results. |

# VIII. High Level Design



**Figure 1: Block diagram**

**Input data:**

A dataset including video clips or picture sequences of microexpressions recorded from diverse people used as system's input.

**Data preparation**

The input data is preprocessed to deal with noise, conduct normalisation, and enlarge the dataset to enhance model generalisation. The photos may need to be resized, cropped, and normalised during preprocessing.

**Deep learning:**

The deep learning model for recognising microexpressions is built in this module, which is the project's main focus. In order to extract spatial and temporal information from the input data, the model may make use of CNNs, RNNs, or their variations. Optimising hyperparameters and training models

**Model assessment:**

On a different test dataset, the trained model is tested for performance. The model's efficiency is evaluated using measures which includes accuracy, precision, recall, and F1-score.

**Phase I Results:**

A trained deep learning model that can identify microexpressions from input video clips or picture sequences, In the phase 1 we will be doing the expression detection and then move on to the micro expression detection in further phase for precised and accurate results.

**References**

Zhang, S., Pan, X., Cui, Y., Zhao, X. and Liu, L., 2019. Learning affective video features for facial expression recognition via hybrid deep learning. IEEE Access, 7, pp.32297-32304.

# CapsuleNet for Micro-Expression Recognition,

**Published in:** [2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)](https://ieeexplore.ieee.org/xpl/conhome/8753205/proceeding)

Bodapati, J.D., Naik, D.S.B., Suvarna, B. *et al.* A Deep Learning Framework with Cross Pooled Soft Attention for Facial Expression Recognition. *J. Inst. Eng. India Ser. B* **103**, 1395–1405 (2022)

Li, Y., Wei, J., Liu, Y., Kauttonen, J. and Zhao, G., 2022. Deep learning for micro-expression recognition: A survey. IEEE Transactions on Affective Computing.

A survey: facial micro-expression recognition

### Cite this article

Takalkar, M., Xu, M., Wu, Q. *et al.* A survey: facial micro-expression recognition. *Multimed Tools Appl* **77**, 19301–19325 (2018). https://doi.org/10.1007/s11042-017-5317-2