National University of Computer & Emerging Sciences



**DEEP LEARNING FOR PERCEPTION (CS-4045)**

***“ Sign Language Detection Using LSTM ”***

**Submitted To**

**Ms. Sumaiyah Zahid**

**Members**

**Insha Samnani (20K-0247)**

**Ismail Ahmed Ansari (20K-0228)**

**Yusra Adam (20K-0207)**

1. **Objective**

The objective of this project is to develop a real-time hand gesture recognition system using deep learning techniques, specifically leveraging a combination of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. The system aims to accurately detect and classify hand gestures captured via depth maps or skeleton joint positions obtained from cameras or motion capture devices. This system holds potential applications in various domains such as education, entertainment, healthcare, and communication, particularly assisting individuals with hearing and speech impairments.

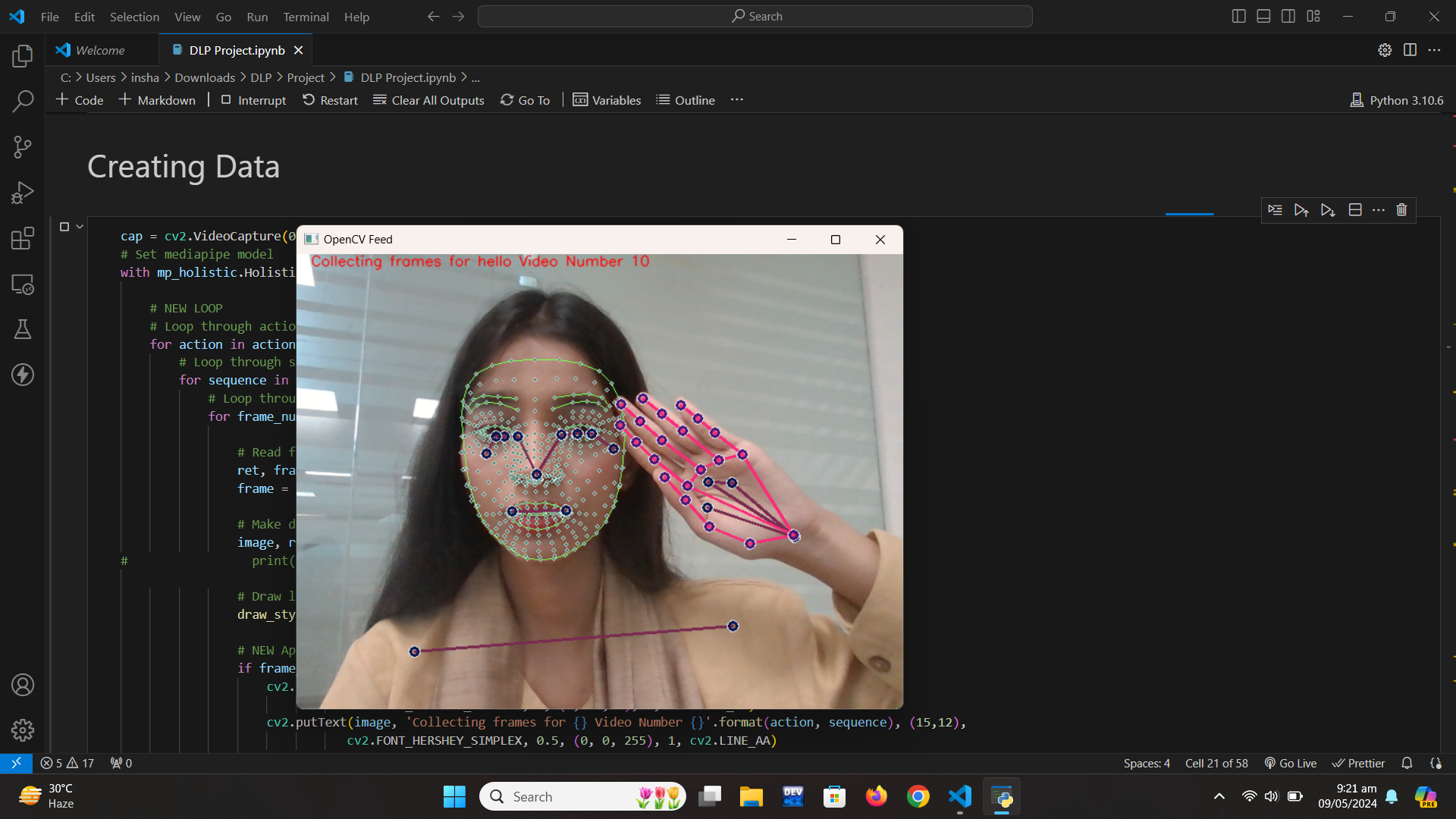
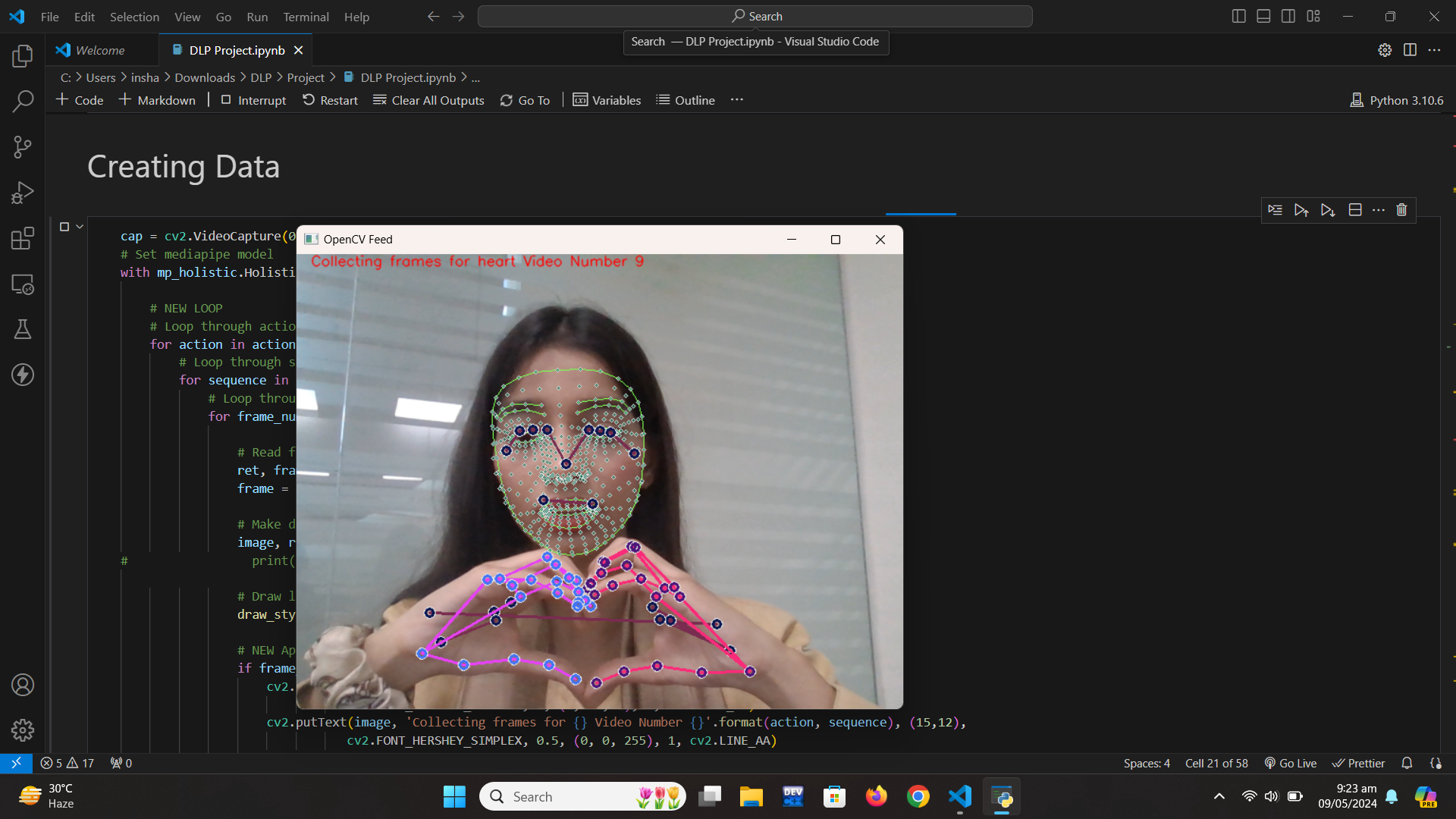
**B. Problem Statement**

Hand gesture recognition, also known as sign language recognition, presents several challenges due to the complexity and variability of gestures, different lighting conditions, occlusions, and variations in hand shapes and positions. Traditional approaches to hand gesture recognition often rely on handcrafted features and require extensive preprocessing, making them less adaptable to diverse environments. Hence, there is a need for a robust, real-time hand gesture recognition system that can automatically extract relevant features from raw sensor data and accurately classify hand gestures.

**C. Methodology**

1. **Data Collection:**

* The system collects hand gesture data through video sequences captured in real-time using a camera.
* Thirty videos are recorded for each action (e.g., 'quiet', 'hello', 'heart'), with each video containing 30 frames.
* Depth maps or skeleton joint positions are obtained using the MediaPipe library, and key point values are extracted from these data.

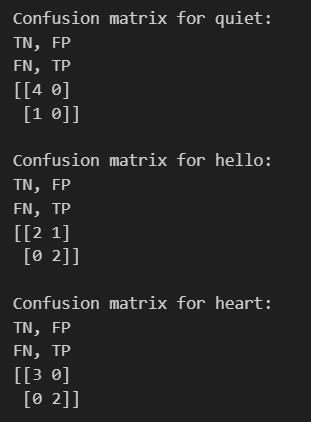
1. **Data Preprocessing:**
   * Key point values extracted from the hand gesture data are preprocessed to generate feature representations suitable for input into the LSTM model.
   * The data is organized into sequences, where each sequence represents a series of key point values corresponding to a single hand gesture video.
2. **Model Training:**
   * An LSTM-based deep learning model is constructed using the TensorFlow and Keras libraries.
   * The model architecture consists of multiple LSTM layers followed by fully connected dense layers.
   * The model is trained on preprocessed hand gesture data to learn the temporal dependencies and patterns in the sequences.



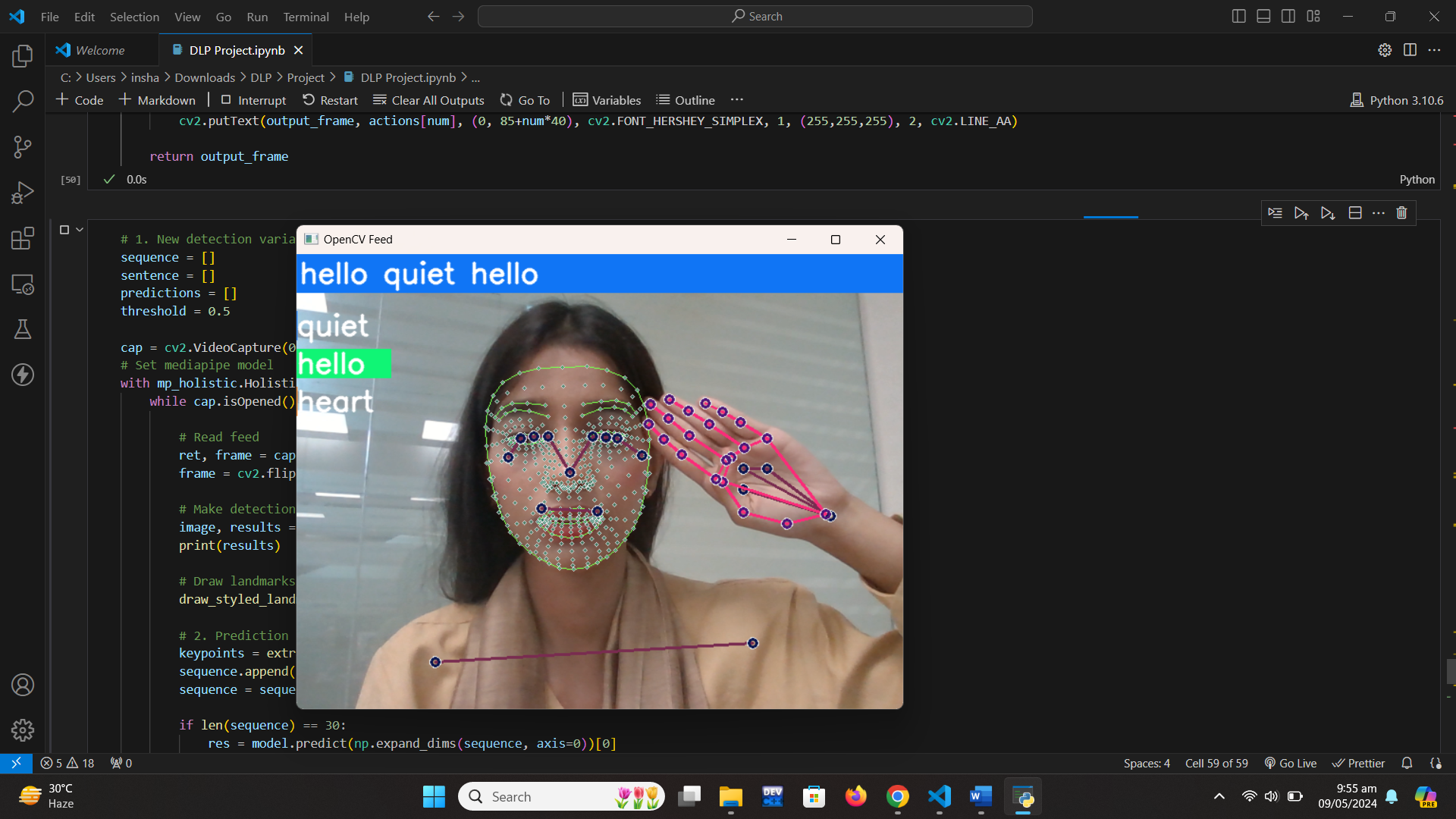
1. **Model Evaluation:**
   * The trained model is evaluated using a separate test set to assess its performance in hand gesture recognition.
   * Metrics such as accuracy, loss, and confusion matrices are used to evaluate the model's performance and identify any potential areas for improvement.
2. **Real-Time Gesture Recognition:**
   * The trained model is deployed to perform real-time hand gesture recognition.
   * Using the MediaPipe library, hand gesture data is continuously captured from the camera feed, and key point values are extracted.
   * The LSTM model predicts the hand gesture based on the sequential key point data, and the predicted gestures are visualized in real-time.
   * A confidence threshold is applied to filter out low-confidence predictions, and a sentence formation logic is implemented to construct meaningful sequences of detected gestures.
3. **Visualization and Interaction:**
   * The system provides real-time visualization of detected hand gestures, enhancing user interaction and feedback.
   * Predicted gestures are displayed on the screen, along with confidence scores and constructed sentences, enabling seamless communication or interaction with the system.

**D. Results**

**Confusion Matrix:**



**Inference:**



A screenshot of a computer

Description automatically generated

**E. Reference**

[1] Dabwan, B. a. (n.d.). *A CNN-LSTM Model for Arabic Sign Language*. Retrieved from https://www.researchgate.net/publication/372639670\_A\_CNN-

[2] LSTM\_Model\_for\_Arabic\_Sign\_Language\_Recognition

Pranav Sheth, S. a. (n.d.). Sign Language Recognition Application using LSTM and GRU (RNN).