REAL-TIME SIGN GESTURE RECOGNITION USING AN ADAPTABLE CONVOLUTIONAL NEURAL NETWORK MODEL

Introduction:

Technological advancement in communication is a necessity in today's society since those with a disability need to have equal media and service access. Nevertheless, the members of the deaf and hard-of-hearing population remain incapacitated by the communication challenges that plague the oral language-dominated society. Communication is yet limited to sign language but there is little provision of sign language in technologies. Our approach within the scope of this project, "LSL SignBot," is to fill this gap by employing an alternating CNN model based on MobileNetV2 that allows sign gesture recognition in real-time. It serves the need of the deaf and hard-of-hearing community and incorporates the concepts of deep learning and computer vision to tailor this system for the aforementioned category of people.

Background and Rationale:

Such a need for equal media was evidenced in the advancement of communication technologies. Deaf and hard-of-hearing people are deprived of spoken language imposed on a society to an extent of facing problems with communication. Communication, especially with sign language, which is their main means of interaction, is another area that is insufficiently represented and often left beyond the scope of digitalization. As a result, the present study targets this challenge through the proposed project called "LSL SignBot" where a flexible Convolutional Neural Network (CNN) model based on MobileNetV2 is employed to identify real-time sign gestures. Not only does this system fulfill the communication requirements of the deaf and hard-of-hearing population, but it also uses deep learning and computer vision to provide additional accessibility to the modern technological world.

Research Questions:

What are the difficulties that one has to face while recognizing different regional sign languages?

In this case, what strategies can be adopted to enhance real-time sign gesture recognition when operated on mobile devices?

Which preprocessing techniques should be applied to increase the recognition rate of data of the sign language?

Methodology:

Theoretical Resources: Based on the concepts of deep learning, computer vision, and sign language recognition, provided in the literature.

Research Approach: Implement a CNN that has been modified to work in real-time for sign gesture recognition.

Research Methods:

- **Data Gathering and Cleaning:** Gather a large variety of sign gestures where a broad population familiar with LSL has been captured. Slightly pre-process the data to get rid of any errors or inconsistencies that might have been found.
- **Image Preprocessing:** Pre-process the input image by resizing the image, normalizing the image, converting the image to grayscale, reducing the amount of noise, and finally enhancing the contrast of the image. Apply Region of Interest (ROI) cropping to extract the hand or the body parts that are involved in the production of sign gestures.
- **Region Extraction:** To detect hands, it is advisable to apply background subtraction algorithms, skin color detection, and motion detection and hand detection models to help in pinpointing the right areas in the images and videos.

- **Feature Extraction:** These include Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP), CNN features, landmark points, and for the dynamic signals, temporal features are extracted.
- **Model Configuration and Training:** Propose a tailored MobileNetV2 architecture CNN. Fine-tune the model with the prepared input data augmenting the training data to diversify it to enhance the result.
- **Real-Time Processing:** Implement OpenCV for the real-time video processing, detection of hands, and tracking within the application of the mobile. Assure that the system may run with acceptable performance on the aspects related to computational load, memory, power consumption, and battery life of a smartphone.
- **User Interaction Design:** Introduce easy-to-understand prompts and feedback to guide the users and accompany the proposed graphic design of the interface for LSL.
- **Performance Evaluation:** As for the assessment of the model, the most common ones have been given, including accuracy, precision, and recall. Carry out comprehensive functional tests to determine how the fix performs in various scenarios.
- **Ethical Considerations:** Secure the user’s data as well as make sure that user consent and responsible use are integrated into the project. Perform a usability test on the proficient LSL users and be able to have feedback from users to modify the system.
- **Impact Assessment:** Assess the effectiveness of the system to enhance the quality of life of the deaf and hard-of-hearing customers by facilitating efficient communication and alleviating social exclusion.

Plan of Work and Time Schedule:

1. **Initial Phase (Month 1-2):**
    - Literature review
    - Data collection
    - Compounding of the data set
2. **Development Phase (Month 3-5):**
    - Data pre-processing
    - Model setting
    - Training
3. **Testing Phase (Month 6-7):**
    - Further assessment of the model
    - Performance checking
    - Performance testing combined with real-time information processing
4. **Implementation Phase (Month 8-9):**
    - Designing of the interfaces that users are going to interact with
    - Testing interfaces
    - Developing a method for obtaining feedback from users
5. **Final Phase (Month 10):**
    - Feedback from the actual users
    - Final assessment of system and component performance
    - Documentation

References:

- Alashhab, S., Gallego, A. J., & Lozano, M. Á. (2019). Hand gesture detection with convolutional neural networks. In Distributed Computing and Artificial Intelligence 15th International Conference 15 (pp. 45-52). Springer International Publishing.
- Ren, Y., Lu, J., Beletchi, A., Huang, Y., Karmanov, I., Fontijne, D., ... & Xu, H. (2021, March). Hand gesture recognition using 802.11 ad mmWave sensor in the mobile device. In 2021 IEEE Wireless Communications and Networking Conference Workshops (WCNCW) (pp. 1-6). IEEE.
- Bao, P., Maqueda, A. I., del-Blanco, C. R., & García, N. (2017). Tiny hand gesture recognition without localization via a deep convolutional network. IEEE Transactions on Consumer Electronics, 63(3), 251-257.
- Zhao, H., Ma, Y., Wang, S., Watson, A., & Zhou, G. (2018). MobiGesture: Mobility-aware hand gesture recognition for healthcare. Smart Health, 9, 129-143.
- Liu, P., Li, X., Cui, H., Li, S., & Yuan, Y. (2019). Hand gesture recognition based on single-shot multibox detector deep learning. Mobile Information Systems, 2019, 1-7.
- Mahesh, M., Jayaprakash, A., & Geetha, M. (2017, September). Sign language translator for mobile platforms. In 2017 International Conference on Advances in Computing Communications and Informatics (ICACCI) (pp. 1176-1181). IEEE.