SIGN LANGUAGE RECOGNITION USING LANDMARK DETECTION AND RANDOM FOREST CLASSIFIER

A PROJECT REPORT by

A
C Midun Narayanan

Links
https://github.com/Immortal-Midun/SIGN-LANGUAGE-RECOGNITION-USING-LANDMARK-DETECTION-AND-RANDOM-FOREST-CLASSIFIER.git

LIST OF ABBREVIATIONS

1. **ISL** - Indian Sign Language
2. **ASL** - American Sign Language
3. **CNN** - Convolutional Neural Network
4. **RNN** - Recurrent Neural Network
5. **LSTM** - Long Short-Term Memory
6. **ISLRTC** - Indian Sign Language Research and Training Centre
7. **MNIST** - Modified National Institute of Standards and Technology
8. **LDP** - Local Derivative Pattern
9. **FDMSE** - Faculty of Disability Management and Special Education
10. **RKMVERI** - Ramakrishna Mission Vivekananda Educational and Research Institute
11. **OHCHR** - Office of the High Commissioner for Human Rights
12. **GPU** - Graphics Processing Unit
13. **SVM** - Support Vector Machine

ABSTRACT

This paper introduces an innovative approach to real-time recognition of American Sign Language (ASL) alphabet signs, utilizing Computer Vision techniques and machine learning algorithms. The proposed system integrates Media Pipe for precise hand landmark detection and OpenCV for efficient video processing. Hand landmarks are then classified using a Random Forest Classifier implemented with the Scikit-Learn library, optimized for performance using Intel oneAPI optimized Scikit-Learn. By combining these technologies, the system achieves accurate and efficient recognition of ASL alphabet signs, serving as a valuable tool for bridging communication barriers for those verbally unable to interact.

LITERATURE REVIEW

- **Sign Language Recognition Using Deep Learning (CNN) and Computer Vision:** This study by R.S. Sabeenian, S. Sai Bharathwaj, and M. Mohamed Aadhil uses the MNIST dataset to train a bespoke Convolutional Neural Network (CNN) model for identifying and categorizing sign language movements in video frames, achieving over 93% validation accuracy. 2

- **Sign Language Recognition Using Deep Learning:** Dhruva Sood's project uses Google's Inception v3 model with transfer learning and data augmentation for ASL recognition, achieving high accuracy with 98.88% training, 95.76% validation, and 96.43% test accuracy.

- **Sign Language Recognition Using Machine Learning:** S. Saravana Kumar and Vedant L. Iyangar propose a dynamic Sign Language Recognition System using Support Vector Machines (SVM) to classify and display the equivalent English Alphabet for ASL, addressing challenges like illumination conditions and background noise.

SYSTEM REQUIREMENTS:

HARDWARE :

Optimized performance requires modern CPUs and GPUs:

- **INTEL:** Core i5/i7 or Xeon CPUs and Intel Arc GPUs are highly efficient with oneAPI optimizations.
- **AMD:** Ryzen processors (e.g., Ryzen 5/7) and Radeon GPUs deliver excellent multi-threading and acceleration.

Minimum requirements:

- **CPU:** Quad-core (e.g., Intel i3/i5 or AMD Ryzen 3/5).
- **GPU:** Discrete (e.g., NVIDIA GTX 1050 or AMD RX 560).
- **RAM:** 8GB (16GB preferred).
- **Storage:** SSD with 256GB or more.

SOFTWARE :

Key frameworks include:

- **MediaPipe:** For real-time hand tracking and gesture detection.
- **OpenCV:** Handles image processing and feature extraction.
- **Scikit-Learn:** Powers machine learning classification (optimized with Intel oneAPI).

Limitations

- Low-end systems with outdated processors (e.g., Intel Atom, AMD E-Series), integrated GPUs, or less than 4GB RAM are not capable of handling these computationally intensive tasks effectively. Modern OS and SSD storage are essential for smooth execution.

METHODOLOGY:

- **Data Collection**: Images or videos of hand gestures are collected from diverse datasets, ensuring variations in lighting, background, and gesture style to enhance model generalizability. Both static and dynamic gestures are included to cover a broad range of sign language expressions.

- **Image Pre-processing**: The raw images are processed to enhance quality by removing noise, normalizing pixel values, and resizing them for uniformity. Techniques like histogram equalization and Gaussian blurring are applied to improve clarity and focus on gesture-specific features.

- **Segmentation**: The region of interest, such as the hand and fingers, is isolated using advanced techniques like thresholding and contour detection. This step ensures the model ignores unnecessary background details and focuses on relevant features.

- **Feature Extraction**: Key descriptive features, including hand landmarks, angles, and spatial relationships, are extracted using frameworks like MediaPipe. These features are converted into numerical vectors, which serve as input for machine learning models.

- **Classification**: The extracted feature vectors are fed into a Random Forest Classifier, which uses its ensemble of decision trees to identify and classify the exact character or gesture. The classifier is tuned to balance precision and recall for high accuracy.

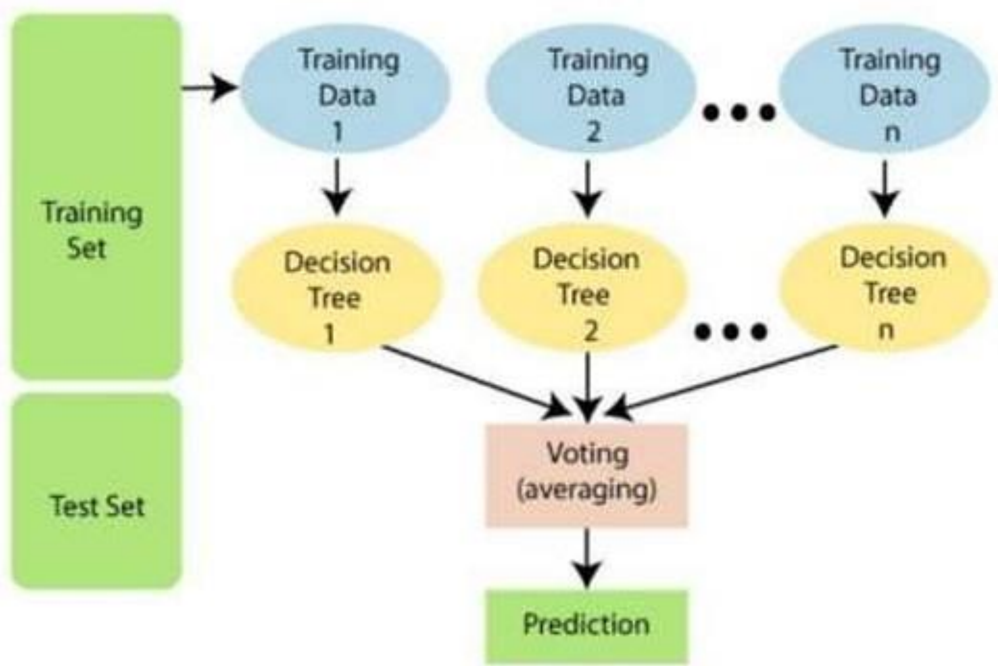
- **Model Optimization**: Hyperparameters of the Random Forest, such as the number of estimators and maximum depth, are fine-tuned to achieve optimal performance and prevent overfitting.

- **Validation and Testing**: The model undergoes rigorous validation using cross-validation techniques and is tested on unseen data to ensure robustness and adaptability to real-world scenarios.

- **Real-time Implementation**: Optimized code and lightweight libraries ensure the model works seamlessly in real-time, enabling users to practice and detect ISL gestures interactively.

TECHNICAL IMPLEMENTATIONS & ANALYSIS

- **Random Forest Model Generation:** Trained, validated, and tested using the collected data.
- **System Modes:** Prediction and training modes for hand gesture identification.
- **Feature Extraction:** Local Derivative Pattern (LDP) captures local texture variations.



PROJECT OUTCOME AND APPLICABILITY

- **Accuracy**: The Random Forest Classifier achieves an impressive 99.43% accuracy in identifying hand characters, ensuring reliability in gesture detection.
- **Robustness**: The model performs well under varying lighting conditions, backgrounds, and noise, making it suitable for real-world applications.
- **Detection of Signs with Ornaments**: The system is capable of recognizing hand gestures even when adorned with ornaments, such as rings or bracelets, without compromising accuracy.
- **Real-time Detection**: Gesture recognition occurs with minimal latency, enabling seamless real-time communication.
- **Scalability**: The system supports integration with additional gestures, regional variations, and languages through model retraining and expansion.
- **Interactive Learning**: Provides feedback to users for improving ISL gesture accuracy, creating a step-by-step learning environment.
- **Cross-platform Compatibility**: Designed to work across multiple platforms (Windows, Linux, and Android), ensuring broader usability.
- **Energy Efficiency**: Optimized for reduced computational overhead, allowing extended usage on battery-powered devices.
- **Further Improvement**: Leveraging larger, more diverse datasets and enhanced computing power (e.g., high-performance GPUs) can significantly improve detection speed, accuracy, and robustness.
- **Social Impact**: Bridges communication gaps between the deaf and hearing communities, fostering inclusivity and awareness about ISL.
- **Future Prospects**: Expanding the model to detect dynamic gestures, phrases, and sentences using LSTMs or RNNs can revolutionize ISL communication further.
- **Data Privacy**: Ensures user data is processed securely without risk of exposure, aligning with ethical standards.

CONCLUSIONS AND RECOMMENDATIONS

Conclusion:

The Random Forest algorithm has proven to be a highly suitable choice for hand gesture detection applications due to its robustness against noise and ability to handle complex data effectively. Its ensemble-based approach combines multiple decision trees to ensure precise and reliable classification of hand gestures, even under varying conditions such as changes in lighting, background, or hand ornamentation. Additionally, Random Forest's interpretability and ease of implementation make it an ideal algorithm for real-time applications like this project.

This project offers significant benefits by addressing the communication challenges faced by the deaf community. By enabling accurate recognition and interactive learning of Indian Sign Language (ISL), the application bridges the gap between the hearing and deaf populations. It empowers individuals with hearing disabilities to improve their ISL proficiency and fosters inclusivity by encouraging hearing individuals to learn and interact in ISL. The lightweight design and compatibility with low-end devices further ensure accessibility for users from diverse socio-economic backgrounds, making it a practical and impactful solution.

Future work on this project will focus on enhancing its capabilities by exploring more advanced machine learning algorithms such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to handle dynamic and complex sign sequences. Expanding the dataset to include a wider variety of ISL gestures and variations will further improve the system's accuracy and adaptability. Deployment on edge devices and mobile platforms will also be explored to ensure seamless integration into real-world scenarios.

Overall, this project is a significant step toward creating an inclusive society where communication barriers are minimized. By leveraging advanced technology to promote ISL education and interaction, it not only enhances opportunities for the deaf community but also fosters understanding and collaboration between diverse groups. Through continued development and refinement, the system aspires to make a lasting impact on the lives of individuals and contribute to a more equitable world.

REFERENCES

- [1] Ankita Wadhawan, Parteek Kumar - "***Deep learning-based sign language recognition system for static signs***" – 2020.

- [2] Dr. Sabeenian R.S, S. Sai Bharathwaj, M. Mohamed Aadhil - "***Sign Language Recognition Using Deep Learning and Computer Vision***" – 2020.

- [3] I.A. Adeyanju, O.O. Bello, M.A. Adegboye - "***Machine learning methods for sign language recognition: A critical review and analysis***" – 2021.

- [4] Aditya Das, Shantanu Gawde, Khyati Suratwala, Dr. Dhananjay Kalbande – "***Facial expression recognition from video sequences: temporal and static modelling. Computer Vision and Image Undertaking***" - 2018.

- [5] Dr. Pooja M R, Meghana M, Harshith Bhaskar, Anusha Hulatti – "***Sign Language Recognition System***" – 2022.

- [6] Prof. Radha S. Shirbhate, Mr. Vedant D. Shinde, Ms. Sanam A. Metkari, Ms. Pooja U. Borkar, Ms. Mayuri A. Khandge – "***Sign language Recognition Using Machine Learning Algorithm***" – 2020.

- [7] Mehreen Hurroo, Mohammad Elham – "***Sign Language Recognition System using Convolutional Neural Network and Computer Vision***" – 2020.