IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 3, April 2024

Review Paper on SignSense : An AI Framework for Sign Language Recognition

Prof. V. M. Dilpak[1], Rewa S. Joshi[2], Harshada K . Sonje[3]

Professor, Department of Artificial Intelligence and Machine Learning[1]

Students, Department of Artificial Intelligence and Machine Learning[2,3]

All India Shri Shivaji Memorial Society Polytechnic, Pune, Maharashtra, India

Abstract: *In this project, we propose an ensemble learning-based system for Sign Language Recognition (SLR) integrated with an Explainable AI (XAI) component called SignExplainer. Our goal is to enhance transparency and trust in SLR systems by providing interpretable predictions. The ensemble learning architecture is designed to recognize sign gestures from images, and the SignExplainer module generates statistical values to evaluate prediction correctness. Performance evaluation on benchmark datasets like ASL and BSL demonstrates the effectiveness of our approach in interpreting predictions from various machine learning and deep learning models. Future work aims to extend this methodology to real-time applications and other Sign Languages, advancing accessibility and inclusivity for the hearing-impaired community*

Keywords: Deep learning, computer vision, explainable AI, SignExplainer, classification, sign language, technological development

I. INTRODUCTION

Sign Language recognition is an innovative framework aimed at advancing Sign Language Recognition (SLR) through Ensemble Deep Learning models. This research focuses on enhancing SLR accuracy, resilience, and interpretability by leveraging the InceptionResNetv2 architecture within an ensemble learning approach.

InceptionResNetv2 combines ideas from Inception and ResNet to achieve state-of-the-art performance in image classification. It features inception modules for capturing features at different scales and residual connections to enable effective training of deep networks. The model also uses factorized convolutions to reduce computational complexity while maintaining representational capacity.

By integrating InceptionResNetv2, this framework excels in capturing intricate details crucial for accurate sign language interpretation. It is designed to scale seamlessly for expanding sign vocabularies, diverse users, and dynamic environments, ensuring high performance. This approach aims to enhance accessibility and inclusivity for the deaf and hard-of-hearing communities using cutting-edge deep learning techniques.

II. MATERIALS & METHODS

The proposed methodology employs Explainable AI using DeepExplainer for sign language recognition, focusing on model interpretability with SHAP (Shapley Additive exPlanations). SHAP offers comprehensive insights into model predictions by interpreting single feature effects on target variables. The approach consists of three key stages: Ensemble Learning, Prediction of Learning, and Sign Explainer.

Ensemble Learning:

The methodology incorporates ensemble learning with attention-based models, particularly utilizing ResNet50 to mitigate the vanishing gradient problem in deep convolutional networks. Attention modules are designed to enhance feature extraction, combining feature extraction (F(x)) and attention (A(x)) modules. The resulting global feature embedding (G(x)) is then fed into a fully connected DCNN network for classification.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, April 2024

Classification and Prediction:

The output from the fully connected layer is processed using a Deep Forward Neural Network (DFNN) implemented with ReLU activation. Class-wise performance metrics including accuracy, precision, recall, and F1-score are computed using NumPy and Scikit-learn.

Sign Explainer:

The Sign Explainer module employs SHAP as an interpretability method for deep learning models. It focuses on agnostic interpretability techniques to explain gesture-based signs, allowing for comprehensive model insights and interpretation of model predictions.

By leveraging these techniques, the proposed methodology achieves explainable AI in sign language recognition, facilitating a deeper understanding of model predictions and improving model transparency.

III. MACHINE LEARNING METHODS

1. Ensemble Learning: The methodology utilizes ensemble learning techniques, specifically employing an ensemble of deep learning models such as ResNet50. Ensemble learning combines predictions from multiple models to improve overall accuracy and robustness.
2. Attention Mechanisms: Attention-based models are integrated within the ensemble learning framework. These mechanisms enhance feature extraction and focus on relevant parts of input data, contributing to improved performance.
3. Deep Learning Models: Deep learning architectures like ResNet50 and Deep Forward Neural Networks (DFNN) are utilized for feature extraction, classification, and prediction tasks. These models are capable of learning complex patterns and representations from raw input data.
4. Multi-layer Perceptron (MLP):MLP is employed as part of the classification and prediction stage. MLPs are feedforward neural networks with multiple layers of nodes (neurons) and are commonly used for classification tasks.
5. SHAP (Shapley Additive exPlanations): SHAP is utilized for model interpretability. It is an advanced technique that provides explanations for individual model predictions, allowing insights into feature importance and contribution to model outcomes.
6. Cross-validation: The methodology likely incorporates k-fold cross-validation to evaluate model performance and generalize well on unseen data.

In summary, the proposed methodology leverages a combination of ensemble learning, attention mechanisms, deep learning architectures, and interpretability techniques like SHAP, along with standard machine learning tools and practices, to achieve accurate and interpretable sign language recognition.

IV. RESULTSANDREVIEWS

We analysed four different ASD datasets to build prediction models for different stages of people. In order to do this, we applied various FS methods to those ASD datasets and classified them utilizing eight different simple but effective ML classifiers and also determined how the FS methods affect the classification performance. Furthermore, we also employed four different FSTs to compute the importance of the features which are more responsible for ASD prediction. Inspecting the experimental findings, the best performing classifiers model predicted ASD with AB (99.25%), AB (97.95%), LDA (97.12%), LDA (99.03%) accuracy; AB, LR (99.99%), GNB

V. CONCLUSION

Explainable AI (XAI) is crucial for ensuring trust in deep learning models, especially in computer vision and NLP. This study proposes an ensemble learning-based Sign Language Recognition system with SignExplainer for interpretable predictions. Performance evaluation on ASL and BSL datasets shows promising results. SignExplainer effectively interprets predictions from various machine learning and deep learning models, enhancing transparency in sign language recognition.

Future work aims to extend this methodology to real-time applications and other Sign Languages.

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 3, April 2024

REFERENCES

[1]. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems (NeurIPS) (pp. 4765-4774).

[2]. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1135-1144).

[3]. Camgoz, N. C., Hadfield, S., & Bowden, R. (2018). Neural Sign Language Translation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops (pp. 1989-198906).

[4]. Assaleh, K., Al-Nafjan, A., & Al-Mualla, M. (2018). Sign Language Recognition Using Machine Learning Techniques: A Review. Journal of King Saud University - Computer and Information Sciences, 30(3), 395-405.