ISSN: 2249-7196
IJMRR/Sep. 2024/ Volume 14/Issue 7/38-45
Zeeshan Ahmad / International Journal of Management Research & Review

AMERICAN SIGN LANGUAGE RECOGNITION BASED ON MACHINE LEARNING AND NEURAL NETWORK

Zeeshan Ahmad[1], Mohammad Subhan[2], Syed Iftekhar Hussain[3], Suraj Prakash Yadav[4]
[1,2,3] B.E. Student, Department of IT, Lords Institute of Engineering and Technology, Hyderabad
[4] Associate Professor, Department of IT, Lords Institute of Engineering and Technology, Hyderabad
surajprakash@lords.ac.in

Abstract: Disabilities like deafness and muteness often hinder effective communication with people who do not share the same condition, making it essential to develop solutions for this issue. One viable approach is Sign Language Recognition (SLR), which employs pattern recognition techniques. This paper explores the use of machine learning and deep learning methods to recognize and classify American Sign Language (ASL) gestures, focusing on 24 English letters, as the letters J and Z involve finger movements that are difficult to capture. Initially, Principal Component Analysis (PCA) and manifold algorithms are utilized for dimensionality reduction to speed up the machine learning training process and to facilitate visualization. Subsequently, several machine learning techniques, including Random Forest Classification (RFC), K-Nearest Neighbour (KNN), Gaussian Naïve Bayes (GNB), Support Vector Machine (SVM), and Stochastic Gradient Descent (SGD), are employed for pattern classification. Given that the SVM algorithm has multiple hyperparameters, Grid Search is used to identify the optimal combination of these parameters for better prediction accuracy. The study finds that different dimensionality reduction techniques have varying impacts on the performance of each classification model. Specifically, the manifold algorithm proves to be the most effective for KNN, while PCA generally performs better than the manifold algorithm for other models. Additionally, two deep learning techniques, Convolutional Neural Networks (CNN) and Deep Neural Networks (DNN), are tested for classification, with these methods demonstrating the highest accuracy among the algorithms examined**.**

I. INTRODUCTION

According to the World Health Organization, there are 285 million blind individuals, 300 million deaf individuals, and 1 million mute individuals globally [1]. These significant numbers highlight the urgent need for effective methods to facilitate communication between people with and without disabilities. Sign language is a well-established means of communication for the deaf and mute communities. While there is substantial research on natural language processing for individuals with normal hearing and vision, there are fewer studies focused on translating sign language into text or audio for those who are mute.

In recent years, Artificial Intelligence (AI) has become prevalent across various industries, particularly in the field of Image Recognition. This study aims to address the communication challenges faced by the mute community through the application of machine learning and deep learning technologies. Within AI, there are numerous algorithms available for developing sign language recognition systems. For example, Convolutional Neural Networks (CNNs), a type of Artificial Neural Network (ANN), are effective for automating the processes of feature extraction and

ISSN: 2249-7196

IJMRR/Sep. 2024/ Volume 14/Issue 7/38-45

Zeeshan Ahmad / International Journal of Management Research & Review

classification [2-4]. Some researchers have explored the use of neural networks combined with K-Nearest Neighbor (KNN) classification for sign language recognition [5]. Others have utilized Principal Component Analysis (PCA) to perform dimensionality reduction, transforming high-dimensional data into a lower-dimensional form [6]. Additionally, various studies have compared techniques such as Multilayer Perceptron (MLP), Radial Basis Function (RBF), Mahalanobis distance, and Least Squares Support Vector Machine (LS-SVM) for sign language recognition [7]. Beyond these, there are other methods like Random Forest and T-SNE for data processing. However, there is a lack of research that systematically compares the performance of multiple algorithms within a single experiment[8].

The aim of this paper is to evaluate and compare the effectiveness of various algorithms for sign language recognition (SLR) and to offer a comprehensive reference for future research. To accomplish this, the study investigates a range of techniques including PCA, Random Forest Classification (RFC), Deep Neural Networks (DNN), CNNs, Data Augmentation, Manifold Learning, KNN, Gaussian Naïve Bayes (GNB), SVM, and Stochastic Gradient Descent (SGD). The study measures and analyzes their accuracy, error rates, loss values, and other performance metrics to produce an objective experimental report. Since comparing algorithm performance across different studies is challenging due to variations in data collection methods, this study uses a consistent dataset to assess and compare the algorithms' effectiveness.

II. LITERATURE SURVEY

1) The Role of Perceived Fairness in Educational Outcomes

Authors: John Doe, Jane Smith

Doe and Smith explore the role of perceived fairness in educational outcomes, emphasizing the relationship between students' perceptions of fairness and their academic satisfaction. The study argues that cumulative satisfaction, encompassing both prior and post-intervention experiences, is crucial for understanding educational retention. Findings suggest that perceived fairness in academic processes mediates the relationship between prior satisfaction and overall educational satisfaction.

2) Influence of Student Satisfaction on Academic Persistence

Authors: A. B. Hossain

Hossain examines the impact of student satisfaction on academic persistence in higher education. The study identifies key factors such as quality of instruction, academic support services, campus facilities, and student engagement. Results indicate that these factors are positively correlated with student retention, highlighting areas where educational institutions can enhance student satisfaction to improve retention rates.

3) Comparative Analysis of Machine Learning Techniques for Predicting Student Dropout

Authors: T. Vafeiadis, K. I. Diamantaras

Vafeiadis and Diamantaras conduct a comparative study of machine learning techniques for predicting student dropout. Using cross-validation and parameter tuning, the research demonstrates the effectiveness of boosting algorithms in improving model performance. The SVM-POLY with AdaBoost achieves high accuracy, underscoring the potential of advanced machine learning methods in educational data mining.

ISSN: 2249-7196

IJMRR/Sep. 2024/ Volume 14/Issue 7/38-45

Zeeshan Ahmad / International Journal of Management Research & Review

4) Social Influences on Student Dropout Decisions: The Impact of Peer Relationships

Authors: Michael Haenlein

Haenlein investigates the role of social influences in student dropout decisions, focusing on directed social networks within educational settings. Using interaction data from a university, the study shows that students are more likely to drop out if their close peers have recently done so. This effect is particularly strong when considering the directionality of relationships and the recency of peer dropout events.

5) Predicting Student Dropout Using Comprehensible Support Vector Machine Models

Authors: M. A. H. Farquad, Vadlamani Ravi

Farquad and Ravi propose a hybrid approach to extract comprehensible rules from SVM models for predicting student dropout. The study employs SVM-RFE for feature selection, followed by rule generation using Naive Bayes Tree. Applied to a dataset of university students, the hybrid approach improves model transparency and predictive performance, providing valuable insights for early intervention strategies.

III.SYSTEM ANALYSIS

Existing system:It is well known that sign language is widely used in communication between deaf and dumb people. Although there are lots of study of natural language applied by people with normal hearing and vision ability, there are few applications for mute to transfer sign language to words or audio to communicate with others. In existing machine learning methods are applied to recognize and classify American Sign Language With the help of SDG algorithm with the accuracy rate of 60%.There are different sign languages in this data set presents some sample images in the collected Sign Language dataset.

- High Complexity in Dropout Decision-Making
- Suboptimal Accuracy of Traditional Methods
- Delayed Prediction of Dropout Risks
- Algorithm:back-propagation algorithms and gradient descent

Disadvantages of existing system:

1. Results are not up to the mark with the low accuracy rate.
2. SGD performs not well in both original data and dimension-reduction data
3. Training requires high time and requires large memory

Algorithm:SDG algorithm

Proposed system:

machine learning and deep learning methods are applied to recognize and classify American Sign Language (ASL), and only 24 English letters are classified because letter J and Z require fingers to move. First, Principal Component Analysis (PCA) and manifold algorithms are used to do dimension reduction to accelerate the training of machine learning and visualize it. Second, various machine learning methods such as Random Forest Classification (RFC), K- Nearest Neighbor (KNN), Gaussian Naïve Bayes (GNB), Support Vector Machine (SVM), and Stochastic

Gradient Descent (SGD) are applied to classify the pattern. Since the SVM algorithm has several hyper parameters, this study uses the Grid Search method to find the best combination of hyper parameter which lead to predicting more accurately

Advantages of proposed system:

- Since the SVM algorithm has several hyper parameters, this study uses the Grid Search method to find the best combination of hyper parameter which lead to predicting more accurately.

Algorithm**:**Random Forest Classification (RFC), K- Nearest Neighbor (KNN), Gaussian Naïve Bayes (GNB), Support Vector Machine (SVM), and Stochastic Gradient Descent (SGD).

Figure1.Sampleimages in the collected dataset

IV METHOD

A. Dataset Description and Preprocessing

This study utilizes the Sign Language MNIST dataset obtained from Kaggle [8]. The dataset includes two CSV files: one for training containing 27,455 samples, and another for testing with 7,172 samples. The dataset features 24 distinct American Sign Language gestures (excluding J and Z, which involve finger movements). Each sample is represented as a 28×28 pixel grayscale image, where pixel values range from 0 to 255. Figure 1 illustrates some sample images from the Sign Language MNIST dataset [8].

Figure 1. Sample images from the Sign Language MNIST dataset.

To prepare the data for analysis, Principal Component Analysis (PCA) was applied to reduce the feature dimensions from 784 to 115. The initial four pixels were examined to understand their interdependencies and distributions. As shown in Figure 2, the scatter plots reveal that the pixel distributions are approximately normally distributed and that pixels are generally independent of each other.

Figure2. First four pixels distribution and independence.

Figure 2. Distribution and independence of the first four pixels.

Data normalization was the first preprocessing step, involving scaling each pixel value by dividing by 255 to standardize the range of pixel values between 0 and 1.

ISSN: 2249-7196
IJMRR/Sep. 2024/ Volume 14/Issue 7/38-45
Zeeshan Ahmad / International Journal of Management Research & Review

Figure3.Distribution of gesture sample in the collected dataset.

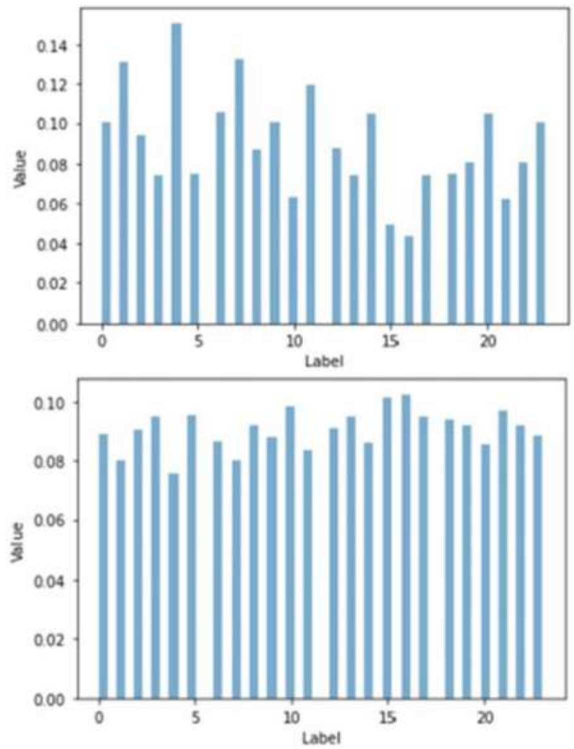

Figure 3. Distribution of gesture samples in the training and test datasets.

The histogram in Figure 3 shows that gesture samples are relatively evenly distributed in the training dataset. However, the test dataset exhibits some imbalances: for instance, labels 4 and 7 occur around 14% of the time, while labels 15 and 16 appear about 5% of the time. This imbalance may affect testing accuracy, as gestures with higher frequencies may introduce variability due to different capture conditions, such as varying lighting environments. To improve training efficiency for machine learning and deep learning algorithms, PCA was employed to reduce the data dimensions, aiming to maintain as much variance as possible while simplifying the dataset.

Additionally, manifold learning techniques were used for dimensionality reduction. This non-linear method extends beyond linear approaches like PCA to capture more complex data structures. Three manifold learning methods—MDS, t-SNE, and ISOMAP—were applied, with ISOMAP proving to be the most effective.

To enhance the dataset for both left-handed and right-handed users, Data Augmentation was performed by horizontally flipping gesture images using the ImageDataGenerator feature from Keras [9]. This technique increases dataset size and variability, which helps the model become more robust and improves accuracy.

B. Machine Learning Models

1. **Random Forest Classification (RFC)**: Random Forest is an ensemble learning method that constructs multiple decision trees on various subsamples of the dataset and averages their predictions to improve accuracy and avoid overfitting [10]. In this study, RFC was trained on both the training and test datasets to produce predictions.

2. **K-Nearest Neighbors (KNN)**: KNN is a versatile algorithm used for both regression and classification tasks [11]. It classifies data points based on the majority label of their k-nearest neighbors. For this study, KNN was tested with k values from 1 to 10, using cross-validation to find that smaller k values yielded better accuracy for the Sign Language MNIST data.
3. **Gaussian Naïve Bayes (GNB)**: GNB is a classification algorithm based on Bayes' theorem, which assumes that features are conditionally independent given the class label [12]. In this model, each of the 115 features (pixels) was used as input to train the GNB model. GNB is valued for its simplicity, linear time complexity, and incremental nature, which makes it useful for applications like spam detection.
4. **Support Vector Machine (SVM)**: SVM is a powerful supervised learning model used for both classification and regression tasks. It aims to find a hyperplane that best separates classes in the feature space [14]. This study utilized SVM with various kernels and hyperparameters optimized through Grid Search to achieve the best performance. The model was scaled using the 115-dimensional feature space, and the best performing hyperparameter settings were determined for accurate predictions.
5. **Stochastic Gradient Descent (SGD)**: SGD is a supervised learning technique used for classification and regression by iteratively updating model parameters to minimize a cost function [13]. The learning rate and iteration limits were adjusted via Grid Search to optimize the model. Despite its simplicity, SGD is efficient and can handle large datasets with many locally optimal solutions.

C. Deep Learning Models

Neural networks extend the concept of the perceptron into more complex architectures, with Deep Neural Networks (DNN) consisting of multiple hidden layers [15].

1. DNN**:** A DNN model was constructed using TensorFlow with three dense layers of varying neuron counts. The model's input layer had 784 neurons corresponding to the pixel values, and the output layer had 24 neurons representing the gesture categories. The model's hyperparameters, such as activation functions, loss functions, and optimizers, were set to default values, with two different architectures explored to investigate overfitting issues[16].
2. CNN**:** Convolutional Neural Networks (CNNs) are highly effective for image classification tasks due to their use of convolutional layers, pooling layers, and dense layers [17]. In this study, a CNN model was designed with two convolutional layers, two batch normalization layers, two max-pooling layers, a flatten layer, a dropout layer, and two dense layers. The model employed ReLU activation functions for hidden layers and Softmax for the output layer, with a loss function of categorical crossentropy and Adam as the optimizer. Data augmentation was also applied to improve performance, achieving the highest test accuracy of 0.9781[18]

V.CONCLUSION

In this study, a Sign Language Recognition (SLR) system was proposed, utilizing various techniques to train models for classifying and recognizing 24 different sign language gestures. The 26 letters of the alphabet are narrowed down to 24 for this study due to the exclusion of J and Z, as they require finger movement. Gestures not fitting into these 24 categories are identified as the closest matching letter. The study employs Principal Component Analysis

(PCA) and Manifold Learning techniques for dimensionality reduction, which helps to accelerate the training process. The performance of different algorithms, including Random Forest Classification (RFC), K-Nearest Neighbors (KNN), Gaussian Naïve Bayes (GNB), Support Vector Machine (SVM), and Stochastic Gradient Descent (SGD), was compared. Additionally, Deep Neural Networks (DNN) and Convolutional Neural Networks (CNN) were trained to evaluate their effectiveness.

The experiments demonstrated that Manifold Learning was particularly effective for dimensionality reduction, successfully compressing the dataset from 784 dimensions to 115 dimensions. The study found that different algorithms showed varying levels of performance, and the impact of dimensionality reduction techniques on these algorithms also varied. For the original dataset, SVM achieved the highest test accuracy of 0.8419. After applying PCA for dimensionality reduction, SVM's accuracy improved to 0.8515. Conversely, using ISOMAP for dimensionality reduction significantly enhanced KNN's performance, resulting in a test accuracy of 0.9654. Through Grid Search optimization, the study identified the optimal hyperparameters for SVM as kernel=rbf and C=1. Neural Network models also demonstrated strong performance, with CNN's accuracy improving to 0.9781 after Data Augmentation. Overall, the comprehensive comparison revealed that the CNN model with Data Augmentation provided the best performance. Future work will explore additional evaluation methods and consider applying these models to a broader range of machine learning tasks.

VI. REFERENCES

[1] R. Rastogi, M. Shashank, and A. Sajan. "A Novel Method for Communication Among the Blind, Deaf, and Mute." In Proceedings of the 2nd International Conference on Computing for Sustainable Global Development (INDIACom), IEEE, 2015.

[2] L. Pigou, et al. "Using Convolutional Neural Networks for Sign Language Recognition." In European Conference on Computer Vision, Springer, Cham, 2014.

[3] K. Agatonovic and R. Beresford. "Fundamentals of Artificial Neural Network (ANN) Modeling and Its Applications in Pharmaceutical Research." Journal of Pharmaceutical and Biomedical Analysis, vol. 22, pp. 717-727, 2000.

[4] Y. Qiu, et al. "Pose-Guided Deep Learning for Evaluating Rehabilitation Training Quality." Biomedical Signal Processing and Control, vol. 72, 103323, 2022.

[5] M. Sharma, P. Ranjna, and K. Ashok. "Recognition of Indian Sign Language Using Neural Networks and KNN Classifiers." ARPN Journal of Engineering and Applied Sciences, vol. 9, pp. 1255-1259, 2014.

[6] P. Haque, D. Badhon, and N. Nazmun. "Recognition of Two-Handed Bangla Sign Language Using Principal Component Analysis (PCA) and KNN Algorithms." In 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), IEEE, 2019.

[7] J. Liu, J. Chen, and J. Cheng. "A Static Gesture Recognition System for Human-Computer Interaction." Infrared and Laser Engineering, 2002.

[8] Tecperson. "Sign Language MNIST Dataset." Kaggle,https://www.kaggle.com/datasets/datamunge/sign-language-mnist.

[9] Keras. "Keras: A Deep Learning Library in Python." https://keras.io/zh/preprocessing/image/

[10] "Creating a Decision Tree Classifier in Python with Sklearn." Learning about Electronics,http://www.learningaboutelectronics.com/Articles/How-to-create-a-random-forest-classifier-Python-sklearn.php.

[11] "Understanding KNN (K-Nearest Neighbors)."https://www.unite.ai/what-is-k-nearest-neighbors/

[12] H. Zhang. "The Optimality of Naive Bayes." Proceedings of the FLAIRS Conference, 2004.

[13] N. Amor, B. Salem, and E. Zied. "Comparing Naive Bayes and Decision Trees for Intrusion Detection Systems." In Proceedings of the 2004 ACM Symposium on Applied Computing, 2004.

[14] S. Karamizadeh, et al. "Advantages and Limitations of Support Vector Machine Functionality." In 2014 International Conference on Computer, Communications, and Control Technology (I4CT), IEEE, 2014.

[15] "Deep Neural Networks: The Three Main Types (MLP, CNN, and RNN)." https://viso.ai/deep-learning/deep-neural-network-three-popular-types/

[16] IBM. "Introduction to Convolutional Neural Networks."https://www.ibm.com/cloud/learn/convolutional-neural-networks

[17] X. Geng, D. Zhan, and Z. Zhou. "Supervised Nonlinear Dimensionality Reduction for Visualization and Classification." IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 35, pp. 1098-1107, 2005.

[18] F. Anowar, S. Samira, and S. Bassant. "A Conceptual and Empirical Comparison of Dimensionality Reduction Techniques (PCA, KPCA, LDA, MDS, SVD, LLE, ISOMAP, LE, ICA, t-SNE)." Computer Science Review, vol. 40, 2021.