SSURJET

Sir Syed University Research Journal of Engineering & Technology
2022, Vol. 12, No. 2,
https://doi.org/10.33317/ssurj.436

Static Pakistani Sign Language Classification using Support Vector Machine

Muhammad Shaheer Mirza[1], Sheikh Muhammad Munaf[2], Shahid Ali[3], and Muhammad Asif[4]

[1]*Department of Biomedical Engineering, Ziauddin University, Karachi, Pakistan*
[2]*Department of Software Engineering, Ziauddin University, Karachi, Pakistan*
[3]*Department of Speech Language and Hearing Sciences, Ziauddin University, Karachi, Pakistan*
[4]*Department of Electrical Engineering, Ziauddin University, Karachi, Pakistan*

Correspondence Author: Sheikh Muhammad Munaf (munaf.rashid@zu.edu.pk)

Received January 28, 2022; Revised May 23, 2022; Accepted June 22, 2022

Abstract

In this study, a system is proposed that uses the Support Vector Machine (SVM) technique with Bag-of-Words (BoW) and recognizes static Pakistani Sign Language (PSL) alphabets. The application of the BoW technique with SVM, on a PSL images' dataset, has not been performed previously. Similarly, no publicly available dataset for PSL is available and previous studies have achieved a maximum classification accuracy of 91.98 %. For this study, a total of 511 images are collected for 36 static PSL alphabet signs from a native signer. The Sign Language (SL) recognition system uses the collected images as input and converts them to grayscale. To segment the images, the system uses the thresholding technique and Speeded Up Robust Feature (SURF) to extract the features. The system uses K-means clustering to cluster the extracted features. To form the BoW, the system computes the Euclidean distance among SURF descriptors and clustered data. The system then uses 5-fold cross-validation to divide the codebooks obtained from the BoW into training and testing. The developed system yields an overall accuracy of 97.87 % for the classification of static PSL signs at 1,500×1,500 image dimensions and 500 Bags.

Index Terms: *Image Processing, Machine Learning, Pakistani Sign Language, Pattern Recognition, Support Vector Machine.*

I. INTRODUCTION

There are approximately 466 million people worldwide with hearing disabilities, and by 2050, this number is estimated to increase to over 900 million, according to the World Health Organization [1]. These individuals depend on their native Sign Languages (SL), to communicate. Automated SL recognition systems have come up as a possible solution for this problem. Pakistan has roughly 250,000 hearing-impaired people according to the Pakistan Association of the Deaf [2]; and as a medium of communication, many use Pakistani Sign Language (PSL) [3]. Creating an SL recognition system for the benefit of these people was the main motivation of this study. The vision-based SL recognition system proposed in this study will use images of static (still) signs of PSL. The literature review done for this study gives us an overview of the techniques used for developing SL recognition systems.

Many different Sign Languages (SL) have been used for the development of SL recognition systems, namely American SL, Arabic SL, Chinese SL, German SL, Indian SL, Irish SL, Pakistani SL, Persian SL, also some studies used multiple SL such as American & German SL, American & Thai SL, and American & Indian SL [4-23]. The included systems used the mentioned SL images and videos dataset as input and used different classifiers, like Support Vector Machines (SVM), Hidden Markov Model (HMM), Multi-Layered Perceptron (MLP) Neural Network, Convolutional Neural Network (CNN), K-Nearest Neighbor (KNN) and many more, to recognize the sign languages. The studies included focused on vision-based SL recognition systems instead of sensor-based recognition systems; and more precisely those systems that use images and videos from a single camera, instead of those that use multiple cameras or different object tracking technologies by placing markers on the user.

The objective of this research is to develop a vision-based system that recognizes PSL alphabets using SVM with Bag-of-Words (BoW) technique, with data collected from a native signer. This proposed system is a novel contribution to this research.

The paper is organized as follows: Section II states the literature review or state-of-the-art techniques done for PSL recognition systems; Section III describes the problem statement and its proposed solution i.e., methodology used in this study; Section IV provides the experimental results; Section V provides the conclusion.

II. STATE OF THE ART WORK

The literature review performed for this study was to find the state-of-the-art techniques used for the classification of static Pakistani SL (PSL) alphabets.

The techniques are shown in Table I.

Table I: State of the Art Techniques used for PSL Classification

Author	Preprocessing and Segmentation	Features Extraction	Classification
Khan, e.t.al [17]	Gray World Algorithm on RGB images, Skin colored segmentation done, skin-colored pixels marked as blue and converted to a binary image, 4x4 median filter applied to fix falsely detected skin pixels, image resized to 300×400 pixels	Discrete Wavelet Transform	Multilayer Perceptron (MLP)
Ahmed, et al. [16]	Image resizing to 640×480, de-noising, Skin pixels extracted by ROI segmentation using HSV color space, and morphological operations to reduce error and segmentation noise	Global features include length, area, rectangularity, eccentricity, convexity, solidity, circularity ratio, and basic rectangle (min and max axis). Global shape features used include Fourier Descriptors and Hu's invariant moments	One-against-all multi-class Support Vector Machine (SVM)
Kausar, et al. [3]	K-means clustering for skin color and non-skin color differentiation, noise removal using morphological operation, and the binary image obtained by applying structuring element	Centroid Distance Signature used in Mathematical Modelling (Polynomial, Sinusoidal, Exponential, Gaussian)	K - Nearest Neighbor (KNN)
Shah, et al. [15]	Skin detection, done using color properties in the HSV domain. Hue(H), Saturation(S) & Value(V)	Six statistical features of local binary pattern histogram i.e., standard deviation, variance, skewness, kurtosis, entropy, and energy	Multiclass SVM
Saqib, et al. [19]	Keyframes were extracted using the Median of Entropy of Mean Frames Method Image resizing to 234×234, converted from RGB to grayscale, 3 hidden layers to detect edges, then contours, and then detection of the body part	Convolution layers and fully connected layers, along with functional layers such as max-pooling Layers, Rectified Linear Units layer (ReLU layer), and SoftMax activation function	Convolutional Neural Network (CNN) with Levenshtein distance or edit distance
Shah, et al. [18]	K-means clustering-based segmentation, then segmented images converted from RGB to grayscale	Speeded Up Robust Features (SURF), Edge Orientation Histogram (EOH), Local Binary Patterns (LBP), Histogram of Oriented Gradient (HOG)	Multi-Class SVM

Khan, et al., used 500 images of 37 PSL alphabets and extracted features using Discrete Wavelet Transform (DWT) and classification through Multi-Layer Perceptron to obtain an accuracy of 84.6 % [17]. Ahmed, et al., used 600 images from 10 PSL alphabets and obtained an 83 % accuracy using multi-class SVM and global features, and shape-based features [16]. Kausar, et al., used 455 images from 37 Urdu alphabets & 9 numbers, to obtain an 80 % accuracy using KNN with features extracted from centroid distance signature in mathematical modeling [3].

Shah, et al., used six local binary pattern histogram features with SVM, to achieve a 77.18 % accuracy, using 3,414 images from 37 PSL alphabets [15]. Saqib, et al., collected 8,000 videos of 20 dynamic PSL words and achieved a 90.79 % accuracy using the Convolution Neural Network [19]. Shah, et al., used 6,633 images of 36 PSL alphabets, and using SVM obtained the highest accuracy of 91.98 % with further reported accuracies of 15.41 % using Speeded Up Robust Feature (SURF), 45.71 % using Local Binary Patterns, 87.67 % using Edge Orientation Histogram, and using Histogram of Oriented Gradient of 89.52 % [18].

III. PROBLEM STATEMENT AND ITS PROPOSED SOLUTION

A. Problem Description

Vision-based PSL recognition is a comparatively new area of research and as per the best of our knowledge, no vision-based PSL dataset is publicly available and the researchers that have worked in this area have a maximum recognition accuracy of 91.98 % [18].

B. Solution Framework

The methods described in this section are the solution to the problem stated above. A total of 511 images were collected from 36 static signs of PSL alphabets from a single native signer of PSL, as this is a pilot study. The collected signs are shown in figure I. Ziauddin University Ethical Review Committee approved the protocols for data collection (Reference Code: 4611221SJBME). The signer's written informed consent was obtained and was asked to perform PSL signs in front of a black background with uniform lighting while wearing black clothes. The block diagram for the entire data analysis process is shown in figure II and described below:

a) Pre-processing:

The collected images for static PSL signs were resized from 3,000×3,000 to 1,500×1,500 and converted from RGB to grayscale.

b) Segmentation:

The hand signs were segmented by applying the thresholding technique on the grayscale images and then each image was cropped using the bounding box technique.

c) Feature Extraction:

The features were extracted using the SURF algorithm. For each segmented image, the SURF points were detected. The key point descriptors were extracted by utilizing these SURF points using the Hessian matrix.

A point $x = (x, y)$ in the image I, the Hessian matrix $H(x, \sigma)$ in x at scale σ is defined as follows:

$$H(x,\sigma) = \begin{bmatrix} L_{xx}(x,\sigma) & L_{xy}(x,\sigma) \\ L_{xy}(x,\sigma) & L_{yy}(x,\sigma) \end{bmatrix} \quad (1)$$

Where;
$L_{xx}(x,\sigma)$, convolution of Gaussian second-order derivative $\frac{\partial^2}{\partial x^2} g(\sigma)$, and similarly for $L_{xy}(x,\sigma)$ and $L_{yy}(x,\sigma)$.

Using K-means clustering, these extracted descriptors were then clustered to form Bags (clusters). A K-cluster value was used for each Bag. A collection of Bags forms the codebooks or Bag-of-words. These maximum SURF descriptors found for the collected images were 444, so a 500 K-cluster value was set for Bag formation.
The Euclidean distance was repeatedly calculated among SURF features and the centroid for each Bag until all features of the images were assigned a Bag:

$$d(x_i, c_i) = \sqrt{\sum(x_i - c_i)} \quad (2)$$

$d(x_i, c_i)$ represents distance among centroids c_i and descriptor x_i. This codebook generation process is shown in figure III.

d) Classification:

The PSL dataset was partitioned into 5 equal subsets using 5-fold cross-validation, with 80 % used for training and 20 % used for testing.

Figure I: PSL Static Alphabets

The SVM classifier was used to train recognition models. SVM used these models, to derive the optimum separating hyperplane among the PSL signs' classes. The feature vectors near the hyperplane, the support vectors, are shown in figure IV.
The testing accuracy was obtained by applying each testing data subset to the subsequently trained model. Once all five subsets were tested, the final accuracy was calculated, by finding the mean of all the obtained accuracies.

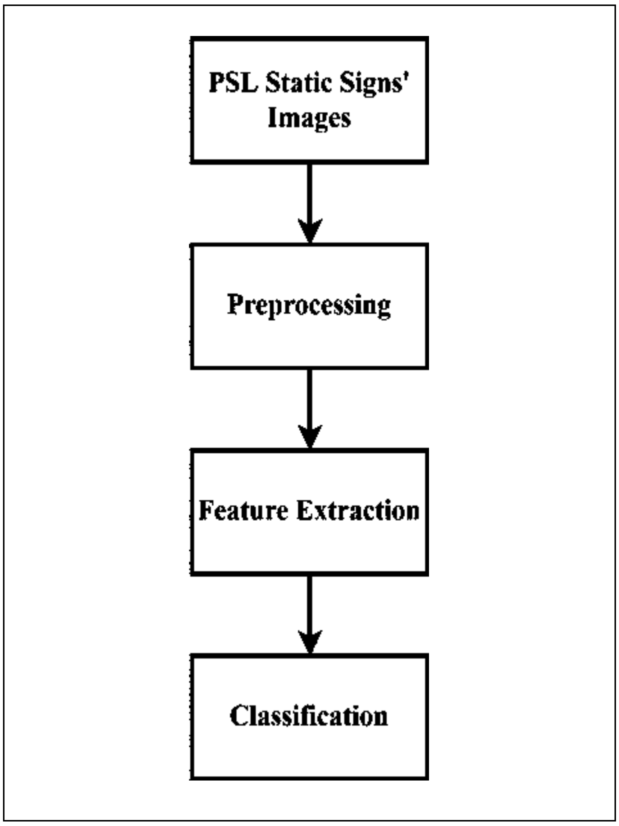

Figure II: PSL Recognition Block Diagram

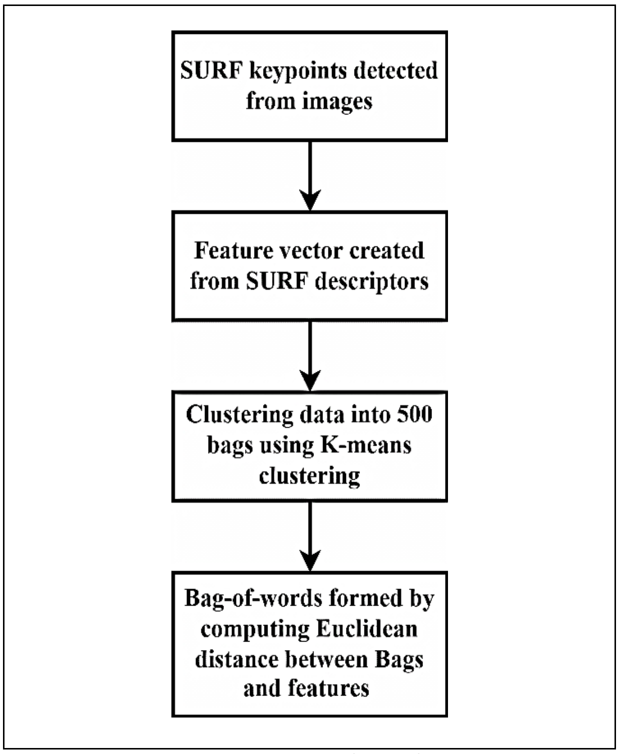

Figure III: Generation of Bag-of-Words

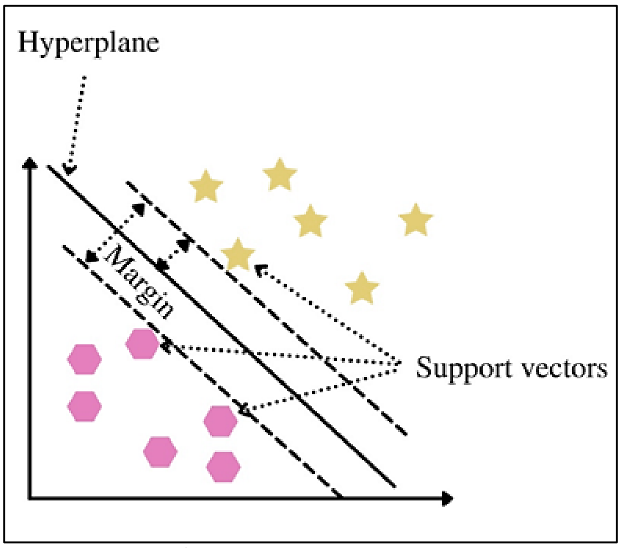

Figure IV: Margin Optimization

IV. RESULTS AND DISCUSSIONS

409 PSL images were used for training and 102 for testing from a total of 511 images. The consequent accuracies attained from these models or subsets were used to compute the training and testing accuracies shown in Table II. The overall accuracy was attained by calculating the mean of all the accuracies of each subset. At 500 Bags, the overall accuracy obtained for the recognition of static PSL alphabets was 97.87 %. Figure V displays the averaged testing confusion matrix, i.e., the true class labels vs the predicted labels of the used PSL static images.

Table II: Accuracies for Static PSL Signs

	Subset 1	Subset 2	Subset 3	Subset 4	Subset 5	Overall
Training	98.00	98.30	98.30	97.30	98.50	98.08
Testing	94.12	97.09	100.00	100.00	97.06	97.65
Overall	96.06	97.70	99.15	98.65	97.78	97.87

Static PSL Signs	True Classes																																			
Predicted Classes	Aa'in	Alif	Bari Ye	Bay	Chay	Choti Ye	Daal	Ddaal	Do Chashmi Hay	Fay	Gaaf	Gha'in	Hamza	Hay	Kaaf	Khay	Laam	Meem	Noon	Noon Ghunna	Pay	Qaaf	Ray	Say	Seen	Sheen	Suaad	Tay	Ttay	Tua'ay	Wow	Zaal	Zay	Zhay	Zua'ay	Zuaad
Aa'in	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Alif	0.0	2.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Bari Ye	0.0	0.0	2.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Bay	0.0	0.0	0.0	2.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Chay	0.0	0.0	0.0	0.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Choti Ye	0.0	0.0	0.0	0.0	0.0	3.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Daal	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Ddaal	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Do Chashmi Hay	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Fay	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.0	2.6	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Gaaf	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Gha'in	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Hamza	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Hay	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Kaaf	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Khay	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Laam	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Meem	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Noon	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Noon Ghunna	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.4	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0
Pay	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Qaaf	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Ray	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Say	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.4	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Seen	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Sheen	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Suaad	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Tay	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Ttay	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Tua'ay	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.2	0.0	0.0	0.0	0.0	0.0	0.0
Wow	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0	0.0	0.0	0.0	0.0
Zaal	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.8	0.0	0.0	0.0	0.0
Zay	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0
Zhay	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.8	0.0	0.0
Zua'ay	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0
Zuaad	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4

Figure V: Confusion Matrix of Static PSL Signs at 1,500×1,500 Image Dimensions and 500 Bags

For the collection of data, a single participant was selected to conduct this pilot study and to find out the effectiveness of using the BoW technique on the collected PSL data. The participant was asked to perform the signs naturally, and to add variations to the signs being performed. While using the thresholding technique during segmentation, the uniform background and clothing conditions facilitated the detection of the skin color in grayscale.

For static PSL signs, the highest recognition accuracy attained in this study, was 97.87 % in comparison to 15.41 % by Shah, et al., that without using the BoW technique, utilized SURF features with SVM [18].

The highest reported accuracy by Shah, et al., was 91.98 % using a Histogram of Oriented Gradient with SVM. This was the previously highest report recognition accuracy for static PSL signs. The accuracy obtained in this study was also higher when compared to Shah, et al., with 77.18 % accuracy using SVM [15], Ahmed, et al., with 83 % using SVM [16], Kausar, et al., 80 % using KNN [3], and Khan, et al., with 84.60 % accuracy using MLP [17].

This comparison with related studies is also shown in Table III.

For future work, PSL data could be collected using more participants with variations in lighting and background conditions. This could make the system more robust and add variations to the collected PSL dataset. Implementing this developed system in real-time could also be explored in future studies. This real-time system could lead to the development of android applications for the recognition of PSL.

Table III: Comparison with Previously used SL Methods

Authors	SL Used	Methods	Number of Signs	Samples	Accuracy
Khan, et al. [17]	Pakistani SL	Discrete Wavelet Transform with MLP	37 Static Alphabets	500 Images	84.60 %
Ahmed, et al. [16]	Pakistani SL	Global features, Global Shape Features, SVM	10 static alphabets	600 Images	83.00 %
Kausar, et al. [3]	Pakistani SL	Centroid Distance Signature, KNN	37 Static Alphabets	455 Images	80.00 %
Shah, et al. [15]	Pakistani SL	Statistical features of Local Binary Pattern Histogram, SVM	37 Static Alphabets	3,414 Images	77.18 %
Saqib, et al. [19]	Pakistani SL	CNN with Levenshtein Distance or Edit Distance	20 Dynamic Words	8,000 Videos	90.79 %
Shah, et al. [18]	Pakistani SL	SURF, Histogram of Oriented Gradient, SVM	37 Static Alphabets	6,633 Images	15.41 % 91.98 %
Our Developed System	Pakistani SL	Bag-of-Words Technique with SURF, K-means Clustering, SVM	36 Static Alphabets	511 Images	97.87 %

V. CONCLUSION

The purpose of this study was to collect data on static PSL alphabets and to develop a vision-based system for their recognition using BoW and SVM techniques. 36 static PSL alphabet signs were collected with uniform background and uniform lighting conditions from a native signer of PSL. These images were used as input in the developed system. The images were resized to 1,500×1,500 dimensions and then segmented. These segmented images were transformed into BoW by computing the Euclidean distance among clustered data and their SURF features. The data were clustered using K-means clustering. The BoW attained for static PSL signs at 500 Bags was then classified using SVM to acquire an overall accuracy of 97.87 %. The system developed in this study outperformed the systems previously developed for the recognition of PSL.

Acknowledgment

The authors would like to thank Ziauddin University, Karachi, Pakistan, for all the support provided to accomplish this research work.

Authors Contributions

Muhammad Shaheer Mirza's contributions to the study were conceptualization, formal analysis, investigation, software, validation, visualization, and original draft preparation. Sheikh Muhammad Munaf’s contributions to the study were conceptualization, methodology, software, supervision, review, and editing. Shahid Ali's contributions to the study were investigation, visualization and review, and editing. Muhammad Asif's contributions to the study were conceptualization, supervision, visualization, review, and editing. All authors have given their approval for this version of the article to be published and agree to be accountable for all aspects of this work.

Conflict of Interest

There is no conflict of interest between all the authors.

Data Availability Statement

The testing data is available in this paper.

Funding

This research received no external funding.

References

[1] World Health Organization. 2021. *Deafness and Hearing Loss*. Retrieved from: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss

[2] Pakistan Association of the Deaf. 2021. Deaf Statistic. Retrieved from: http://www.padeaf.org/quick-links/deaf-statistics

[3] Kausar, S., Javed, M. Y., Tehsin, S., & Anjum, A. (2016). A novel mathematical modeling and parameterization for sign language classification. *International Journal of Pattern Recognition and Artificial Intelligence, 30*(05), 1650009.

[4] Ameen, S., & Vadera, S. (2017). A convolutional neural network to classify American Sign Language fingerspelling from depth and colour images. *Expert Systems*, *34*(3), e12197.

[5] Singha, J., Roy, A., & Laskar, R. H. (2018). Dynamic hand gesture recognition using vision-based approach for human–computer interaction. *Neural Computing and Applications, 29*(4), 1129-1141.

[6] Nasri, S., Behrad, A., & Razzazi, F. (2015). Spatio-temporal 3D surface matching for hand gesture recognition using ICP algorithm. *Signal, Image and Video Processing, 9*(5), 1205-1220.

[7] Hikawa, H., & Kaida, K. (2014). Novel FPGA implementation of hand sign recognition system with SOM–Hebb classifier. *IEEE Transactions on Circuits and Systems for Video Technology, 25*(1), 153-166.

[8] Rastgoo, R., Kiani, K., & Escalera, S. (2018). Multi-modal deep hand sign language recognition in still images using restricted Boltzmann machine. *Entropy, 20*(11), 809.

[9] Ibrahim, N. B., Selim, M. M., & Zayed, H. H. (2018). An automatic Arabic sign language recognition system (ArSLRS). *Journal of King Saud University-Computer and Information Sciences, 30*(4), 470-477.

[10] Jiang, X., & Zhang, Y. D. (2019). Chinese sign language fingerspelling via six-layer convolutional neural network with leaky rectified linear units for therapy and rehabilitation. *Journal of Medical Imaging and Health Informatics, 9*(9), 2031-2090.

[11] Cui, R., Liu, H., & Zhang, C. (2019). A deep neural framework for continuous sign language recognition by iterative training. *IEEE Transactions on Multimedia*, *21*(7), 1880-1891.

[12] Koller, O., Zargaran, S., Ney, H., & Bowden, R. (2018). Deep sign: Enabling robust statistical continuous sign language recognition via

hybrid CNN-HMMs. *International Journal of Computer Vision, 126*(12), 1311-1325.

[13] Athira, P. K., Sruthi, C. J., & Lijiya, A. (2019). A signer independent sign language recognition with co-articulation elimination from live videos: an indian scenario. *Journal of King Saud University-Computer and Information Sciences*.

[14] Kelly, D., McDonald, J., & Markham, C. (2010). A person independent system for recognition of hand postures used in sign language. *Pattern Recognition Letters*, *31*(11), 1359-1368.

[15] Shah, S. M. S., Naqvi, H. A., Khan, J. I., Ramzan, M., & Khan, H. U. (2018). Shape based pakistan sign language categorization using statistical features and support vector machines. *IEEE Access*, *6*, 59242-59252.

[16] Ahmed, H., Gilani, S. O., Jamil, M., Ayaz, Y., & Shah, S. I. A. (2016). Monocular vision-based signer-independent Pakistani sign language recognition system using supervised learning. *Indian J. Sci. Technol*, *9*(25), 1-16.

[17] Khan, N. S., Shahzada, A., Ata, S., Abid, A., Farooq, M. S., Mushtaq, M. T., & Khan, I. (2014). A vision based approach for Pakistan sign language alphabets recognition. *La Pensée*, *76* (03).

[18] Shah, F., Shah, M. S., Akram, W., Manzoor, A., Mahmoud, R. O., & Abdelminaam, D. S. (2021). Sign language recognition using multiple kernel learning: A case study of Pakistan sign language. *IEEE Access*, *9*, 67548-67558.

[19] Saqib, S., Ditta, A., Khan, M. A., Kazmi, S. A. R., & Alquhayz, H. (2021). Intelligent dynamic gesture recognition using CNN empowered by edit distance. *Cmc-Computers Materials & Continua, 66*(2), pp. 2061-2076.

[20] Zadghorban, M., & Nahvi, M. (2018). An algorithm on sign words extraction and recognition of continuous Persian sign language based on motion and shape features of hands. *Pattern Analysis and Applications*, *21*(2), 323-335.

[21] Elakkiya, R., & Selvamani, K. (2019). Subunit sign modeling framework for continuous sign language recognition. *Computers & Electrical Engineering*, *74*, 379-390.

[22] Klomsae, A., Auephanwiriyakul, S., & Theera-Umpon, N. (2017). A novel string grammar unsupervised possibilistic C-medians algorithm for sign language translation systems. *Symmetry*, *9*(12), 321.

[23] Joshi, G., Vig, R., & Singh, S. (2018). DCA-based unimodal feature-level fusion of orthogonal moments for Indian sign language dataset. *IET Computer Vision*, *12*(5), 570-577.