

ENHANCING KURDISH SIGN LANGUAGE RECOGNITION THROUGH RANDOM FOREST CLASSIFIER AND NOISE REDUCTION VIA SINGULAR VALUE DECOMPOSITION (SVD)

Sara A. Ahmed^{**}, Bozhin Nabaz Mahmood^a, Diar Jamal Mahmood^a, Mohammed Mohammed Namq^a

^a Computer Science Department, Komar University of Science and Technology, Sulaymaniyah, Kurdistan Region, Iraq – sara.azad@komar.edu.iq

Received: 27 Feb., 2024 / Accepted: 21 Apr., 2024 / Published: 13 June., 2024.

<https://doi.org/10.25271/sjuoz.2024.12.2.1263>

ABSTRACT:

Deaf people around the world face difficulty communicating with others. Hence, they use their own language to communicate with each other. This paper introduces a new approach for Kurdish sign language recognition using the random forest classifier algorithm aiming to facilitate communication for deaf communities to communicate with others without relying on human interpreters. On the other side, for further enhancement of the images captured during recognition linear algebra techniques have been used such as singular value decomposition for image compression and Moore–Penrose inverse for blur removal. Kurdish language has 34 alphabets and (10 numeric numbers 10, . . . ,3 ,2 ,1). Additionally, three extra signs have been created and added to the dataset, such as space, backspace, and delete sentences for the purpose of real-time translation. A collection of 800 images has been gathered for each character, out of 800 images, only 80 per character were used due to their similar positions but varied alignment, totalling 3,520 images for the dataset (44 characters × 80 images each). Two simulation scenarios were carried out: one with optimal conditions - a white background and adequate lighting, and another with challenges such as complex backgrounds and varied lighting angles. Both achieved high match rates of 96% and 87%, respectively. Further, a classification report analyzed precision, recall, and F1 score metrics.

KEYWORDS: Sign Language, Kurdish Sign Language, Random Forest Algorithm, Real Time Recognition and Singular Value Decomposition.

1. INTRODUCTION

The most used code for communication by deaf people is sign language. Sign language is a rich and complex visual-gestural language and the primary communication standard for the deaf and hard hearing communities worldwide. Sign language combines gestures, facial expressions, body movements that give a meaning that exceeds spoken language (Valli, C., & Lucas, C., 2000). However, understanding these expressions can be challenging for those better versed in their nuanced language. The challenges or difficulties and the communication gap between the deaf or hard hearing and the hearing communities highlight the need for innovative solutions. The sign language recognition system is essential to bridge this gap by processing sign language and having an understandable output, typically text. Transforming these complex signs into a universally comprehended language with a sign language recognition system promises to promote inclusivity and connectivity between these communities.

Sign language recognition technology has evolved, driven by improved computer vision and machine learning techniques. It finds application in various fields, including self-driving cars, facial recognition, augmented reality, and healthcare (Cheok, M. J., et al., 2019).

Nevertheless, there is a growing awareness in this area of emerging concept called dynamic sign language. Rather than concentrating on understanding individual gestures, dynamic sign language gets to the heart of the flow and expressiveness of sign language communication. However, dynamic sign language recognition presents a unique set of challenges compared to static sign recognition. Its complexity arises from the fact that meaning depends on hand movements and, in a very complicated or detailed manner, is intertwined with facial expressions and body postures. Even a slight movement or pacing can change the entire

meaning of a sentence or word, amplifying the challenge (Huang, J., et al., 2018).

The rapid advancement of recognition technology for sign languages, driven by computer vision and machine learning techniques, has played a crucial role in addressing a wide range of issues. Undoubtedly, sign language serves as the primary mode of communication for deaf and hard of hearing individuals, highlighting the imperative role of technology in overcoming communication barriers. The recent emphasis on sign language recognition underscores the intricate nature of sign language communication, prompting the need for innovative solutions. Viewing this through the lens of machine learning techniques, particularly the Random Forest Classifier algorithm, offers opportunities to develop dynamic sign language recognition systems, enhancing accessibility for community members. The Random Forest Classifier algorithm, an extension of the bagging method, combines bagging and feature randomness to construct uncorrelated decision tree forests, demonstrating its versatility in various applications such as object detection and gesture recognition. Thus, the analysis delves into a diverse array of issues concerning the advancement of sign language recognition technology while also highlighting emerging challenges in uncovering the subtle nuances of dynamic sign language, all with the aim of fostering connectivity among linguistic communities (Parmar, A., Katariya, et al., 2019).

2. BACKGROUND

2.1 Challenges in Sign Language Recognition: A Global Perspective

The success and accuracy of sign language recognition systems, specifically Kurdish Sign Language, depend on various factors and challenges common to sign language recognition in general

* Corresponding author

This is an open access under a CC BY-NC-SA 4.0 license (<https://creativecommons.org/licenses/by-nc-sa/4.0/>)

The success and accuracy of Kurdish Sign Language Recognition (KSLR) systems lie in several factors and challenges that are common among sign language recognition on a general level. There are some potential issues that may arise as follow (Rastgoo, R., et al., 2021):

- **Limited Dataset:** The construction of a detailed dataset for Kurdish Sign Language could be an arduous task. Low amount of data could prevent the development of accurate and stable models.
- **Variability in Signs:** Regional variations, dialects or individual peculiarities are often reflected in the sign languages. The capture and representation of this variability in a recognition system can be challenging.
- **Complexity of Signs:** Some sign languages signs may be complicated and might involve slight hand moves or facial expressions, body gestures as well. However, capturing and interpreting these variations correctly is a technological challenge.
- **Real-world Conditions:** In real conditions, when it comes to distinct lighting, background noise or interference sign language recognition systems may have challenges in identifying signs precisely.
- **Integration with Technology:** Integrating sign language recognition in technology interfaces such as communication devices or educational websites demands smooth integration. Compatibility problems or the absence of common interfaces may be obstacles.
- **Limited Resources:** For sign languages such as Kurdish Sign Language, there may be limited resources for research and development in comparison to other widely studied sign languages.
- **Cultural Sensitivity:** Cultural considerations such as cultural differences in signing styles or gestures should be made. Interpretation must be culturally sensitive for the recognition systems to work accurately.
- **Dynamic Nature of Sign Languages:** Changes occur in sign languages with the emergence of new signs. Recognition systems need to be flexible or modifiable when there are changes and updates in the language.

2.2 Cultural and Linguistic Diversity in Sign Language Recognition Systems

It should be noted that the recognition systems of sign language should accommodate the diverse cultural and linguistic distribution of sign languages around the globe. Through this inclusivity people get to have equal access to and understanding of various communities. Inclusivity is achieved in the creation of technology that incorporates culturally sensitive design, constitutes distinct signing styles as well as embraces diverse communities facilitated by sign language users. In addressing these challenges, ethical considerations from an ethic of care perspective also user-centered design and adaptability mechanisms are crucial components to achieve recognition systems that address the technical robustness together with cultural and linguistic inclusion (Leigh, G., et al., 2015). To reach maximum accuracy and performance levels in recognizing Kurdish Sign Language, the design of the recognition system should be sensitized to both singularity in signing and the cultural richness embedded within the language. This comprehensive approach ensures equitable access and effective communication within the Kurdish Sign Language community (Salim, B. W., et al., 2023).

3. RELATED WORKS

Many research papers have adopted, implemented, or proposed new methods in Machine learning, Computer Vision, and Image Processing fields for the purpose of classifying,

analyzing, interpreting, recognizing, and tracking the required sign images or videos to be transferred into a language that ordinary people can understand and vice versa. Furthermore, the aforementioned fields can be combined or to be used alone to develop a sign language recognition system that can ease the communication between deaf or hard hearing community with ordinary people (Varshney, P. K., et al., 2023). Due to the most system's limitations in handling various hand shapes and conditions, pre-processing techniques are continuously being developed to identify the most precise methods. Researchers in both academia and industry have experimented with various algorithms using different techniques, and have occasionally combined different approaches to create hybrid methods that meet the specific requirements of Sign Language recognition. The authors of (Daniels, S., et al., 2021) have developed a highly accurate system for Indonesian sign language recognition using the YOLO method for both image and real-time video analysis. To create their dataset, the authors captured a set of images of the alphabet characters from A to Z, excluding J and R, relying on Bahasa Isyarat Indonesia (BISINDO) sign language. BISINDO alphabet signs typically involve static hand gestures, with the exception of J and R, which require dynamic gestures.

(Nimisha, K. P., et al., 2020), proposed a new method called the Vision Based Approach (VBA) has been introduced, which uses an image as input for recognizing different signs. Unlike the traditional method of image and signal processing, this approach involves giving the image to a pre-processor, extracting its feature vector, and then comparing it with a pre-existing dataset to determine the probabilities of various classes. 2 Based on the resulting probabilities, the sign is classified into its corresponding category. To extract features, several algorithms are used in VBA, such as YOLO, CNN, and PCA. The pre-trained model is the most recent algorithm and is considered the quickest since it can handle a large amount of data, which is critical for achieving high accuracy, the primary aim of this study.

Indeed, the lack of a unified Arabic sign language dialect represents a key obstacle for researchers in this field, as having no shared database impedes their ability to perform extensive studies and find the most appropriate solutions. To address this issue, the authors of a recent paper (Alzohairi, R., et al., 2018) created a small dataset of 20 words by capturing video inputs of hearing-impaired individuals, which were later analyzed as images. To enhance edge detection accuracy, the Canny algorithm was used by applying a predefined criterion to identify most edges while minimizing the error rate, thus improving edge localization when only one edge exists. The resulting recognition rate was 82.22%, which is considered high given the limited number of features used. Around 70 million people worldwide are deaf, and sign language is their primary language (United Nations, April 11, 2024). However, there is no advanced system that caters to the deaf community in the Arabic language. This creates a communication gap between hearing people and those who have deaf issues. To bridge this gap, a system that can translate sign language to words and vice versa is necessary. In (Youssif, A. A., et al., 2011), the authors designed an ArSL recognition system to recognize the Arabic alphabet in a sequence order. They used 30 gesture movement photos out of 900 as data and fed it individually into the classifier. However, they faced a dilemma with letters that have the same original gestures but with additional motions or different rotations. The translation of adequate sign language is also affected by the fact that many countries speak Arabic language but in different ways for writing. To test the system, the authors used different smartphones and processed the images using MATLAB. They used the Histogram of Oriented Gradients (HOG) descriptor as input to the soft-margin Support Vector Machine (SVM). The proposed ArSL system achieved an accuracy rate of 63.56%.

In the area of Kurdish sign language classifications and detection, several contributions have been published. Authors in (Hashim, A. D., & Alizadeh, F., 2018) introduced a real-time system that detects hand gestures and translates them immediately. Three algorithms were implemented for detecting KuSL, with a new algorithm called Grid-based. A recent study conducted by (Mahmood, M. R., et al., 2019) focuses on Kurdish sign language detection and utilizes an artificial neural network (ANN) classifier to recognize ten Kurdish words. Each word is represented by a single hand gesture, and dynamic word recognition is achieved through real-time video capture, followed by comparison of the captured image with two-line vectors of features stored in the dataset. In the pre-processing stage, ten different images have been used for each word, resulting in a dataset of 200 vectors.

However, this approach is not practical in real-life scenarios, as assigning all Kurdish language words to a hand gesture is difficult, and it is impossible for a deaf person to memorize them all. Moreover, new words are regularly added to the Kurdish dictionary, necessitating database updates and continuous training for mute users. Alternatively, using the Kurdish alphabet sign language along with numbers could be a more viable option as the total number of characters is 46. (Salim, B. W., & Zeebaree, S. R., 2023) utilize CNN, for distinguishing sign language, particularly Kurdish Sign Language (KSL). The study involves creating a KSL dataset to characterize the language using pre-trained algorithms for feature extraction and machine learning algorithms for classification. The proposed method was tested on KSL and American Sign Language (ASL) datasets, implementing algorithms such as VGG19 and RESNET101 for feature extraction and Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF) for classification. The CNN model designed for KSL showed promising results, with VGG19 outperforming RESNET101 and the proposed CNN in feature extraction. The Random Forest classifier achieved high accuracy rates of 95% for numbers and 97% for letters in both KSL and ASL datasets.

The lack of resources given to Kurdish sign language has led to a significant problem for researchers in the field, as there are currently few shared databases available for their use. To address this issue, the authors (Rawf, K. H., et al., 2022) have created a dataset for a real-time model recognition using Convolutional Neural Network (CNN) algorithm. The model has been trained and predicted on the KuSL2022 dataset using different activation functions for a number of epochs. The dataset consists of 71,400 images for the 34 Kurdish alphabets. The obtained results show that the proposed system's performance increased. On other hand, (Hama Rawf, et. al, 2024) present a study on the development of a model for real-time Kurdish sign recognition using a modified Convolutional Neural Network (CNN) technique. The primary focus is on recognizing the Kurdish alphabet, with training on the KuSL2023 dataset representing 34 sign languages and alphabets. The KuSL2023 dataset used in the research paper was created by merging and enhancing the ASL alphabet dataset and the ArSL2018 dataset, both of which are publicly available. The proposed technique demonstrates high accuracy in real-time classification and personal independence, achieving an average training accuracy of 99.05% for classification and prediction models.

4. DATASET

In this section, the researcher describes the source of the data, including its origin and collection methods, as well as pre-processing of the dataset. Furthermore, important characteristics of the dataset such as size, format, and variables are highlighted. Any relevant ethical considerations, such as data anonymization and consent procedures are also included. The dataset section plays a vital role in ensuring transparency and reproducibility of this research. To gather and validate our dataset, we engaged in discussions with trainers at "پهیمانگی هیوای نابیستان" an academic institute in Sulaymaniyah city which is dedicated to training the deaf community.

The aforementioned method has been applied on 34 characters and (10 numeric numbers 10, . . . ,3 ,2 ,1) for each letter 800 images has been taken with different positions, but subsequently only 80 images for each character were utilized for the labelling process due to their similarities in position but variations in alignment. This means that, with 44 characters multiplied by 80 images each, the total number of images in the dataset is equal to 3,520 images.

After the images were acquired, each image was prepared by adding the corresponding sign language label or gesture. This important labelling step is the spine of the supervised learning process and it grants the machine learning system an ability to effortlessly correlate photos with their classes. The later on, a variety of imagery cleaning techniques has been applied like resizing, cropping, and noise reduction. This precise mechanism guaranteed that all images were resized in a unified way and got the same aspect ratio, which, as a result, led to the generalization of the learning process. Hence, the images were rendered with a different appearance and the processing time was greatly reduced by using the 512×512 square images. The deliberately made choice is not only boosting the resolution but also improving the procedure by which the data can be processed so that the dataset is ready for a solid model training.

Following annotation and image enhancement, the dataset was partitioned into distinct subsets: datasets such as the training, validation and test sets. To get the best model outcome with the generalization property, the 4-fold cross-validation approach is utilized. This method presented the whole dataset as 4 groups that were exclusive of each other. For instance, one group was designated as testing data, in turn, the rest of the data became the training data, and the process was iterated through each fold. A subset of 80 images for each character was selected, comprising 75% of the total (60 images), which were used for training. An additional 20% (16 images) was allocated for validation purposes, while a mere 5% (4 images) was reserved for testing as described in (Xu, Y., & Goodacre, R., 2018). This means the total number of images used for training is 2640, with the remaining allocated for testing and validation. (refer to the Appendix-A for detail of Kurdish sign alphabetic gestures). A sentence comprises several elements, including spaces, characters, and the potential deletion of a sentence. Spaces between words are essential in a sentence, but occasionally errors occur, or the system may misinterpret the input signs, requiring the deletion of a character. At times, we may even want to delete an entire sentence with a single sign. To facilitate these actions, extra signs have been created and introduced for each operation. Therefore, these symbols differ from Kurdish Sign Language and are of our own invention due to the absence of external resources for reference or reliance for the aforementioned operations. The signs are shown in Figures 1-A to 1-C.

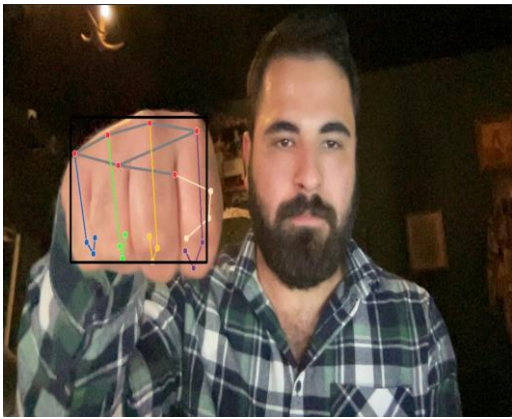


Figure 1-A: Delete Sentence Sign

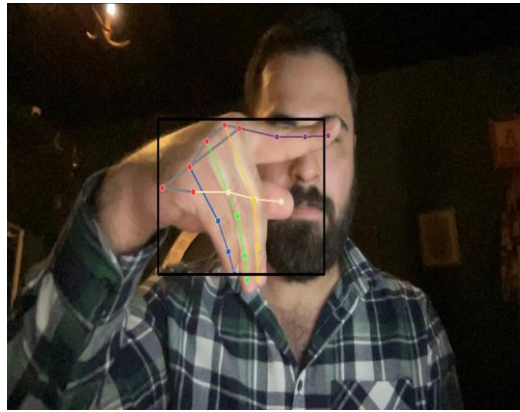


Figure 1-B: Delete Character Sign



Figure 1-C: Space Character Sign

5. METHODOLOGY

5.1 Implementation of RFC algorithm for Kurdish sign language

In this paper, Random Forest algorithm, which is a widely used machine learning algorithm developed by Leo Breiman and Adele Cutler, has been employed. Decision trees are the main structure that supports the RFC algorithm. This decision support technique forms a tree-like structure consisting of three components: decision nodes, leaf nodes, and the root node. It divides the training dataset into branches, which further segregate

into other branches. This sequence continues until a leaf node is attained, which cannot be segregated further.

RFC classification utilizes ensemble strategy to reach the result. The training data is used to train a set of decision trees. This dataset comprises observations and features, which are extracted randomly with the process of dividing nodes into parts. Each tree's terminal node is the final output resulted from that particular decision tree. The output is selected by the majority voting method. In the case, the chosen output by a majority of the decision trees is the final output of the rain forest system problems (Liu, Y., et al., 2012). The architecture of RFC classification can be seen in Figure 2.

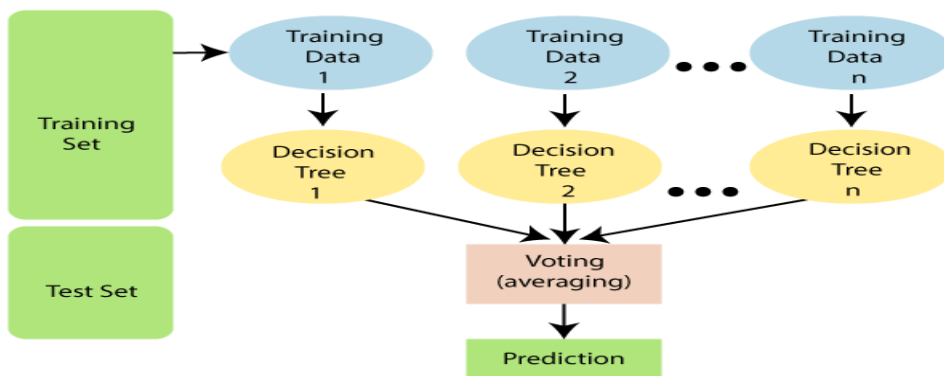


Figure 2: Random Forest Classification Architecture

Moreover, MediaPipe library which is an open-source library developed by Google for building applications with perception capabilities, particularly in the domain of computer vision has been employed (Lugaresi, C., et al., 2019). To make predictions based on the tracked hand's features and training the given data, our proposed method has employed MediaPipe and Random Forest algorithm. The algorithm's strength lies in its ability to handle complex datasets and mitigate overfitting, making it a valuable tool for various predictive tasks in machine learning.

One of the most essential features of RFC ability to handle datasets is containing both continuous variables, as seen in regression, and categorical variables.

As shown in Figure 3, to initiate the process of translation from sign to text, the model begins by loading the given sign will be taken and transformed into an array labelled with the Kurdish alphabet/number images. Subsequently, the program opens the webcam, utilizing the MediaPipe library with the algorithm to extract holistic key points from the hand gesture and detect it. Following this, predictions are made based on the provided and trained dataset, then results are displayed that satisfy the

similarity between the trained data and the real time given data. Finally, upon user command (pressing 'Q' key), the system concludes. Figure 5 offers a visual representation, elucidating each step of this transformative procedure for Kurdish Sign Language recognition.

Additionally, certain exceptions have been introduced to avoid the duplication of characters or, more precisely, the entire alphabet. This involves implementing a technique at the initial stage of our method to store the predicted characters. Following this, the variable was employed to establish a condition: if the currently predicted character matches the previous one, duplicated characters are not appended to the sentence. However, a challenge arises concerning the Kurdish language, as some words contain repeated characters (e.g., کۆتایی، دوور). Hence, we established additional conditions. If the current predicted character is either 'و' or 'ئ', the program allows for duplication within the word. This condition holds true when these characters are attempted to be added consecutively without any intervening character; the same condition applies to space characters.

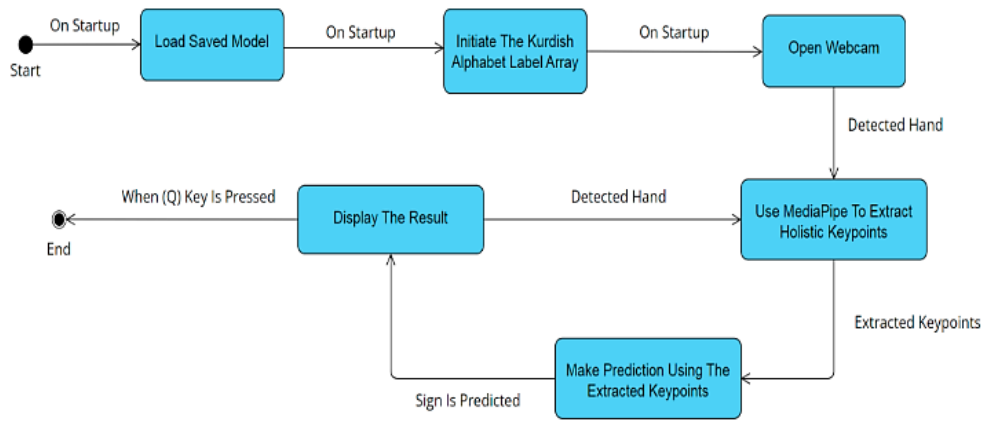


Figure 3: State Diagram of the System

Additionally, a noteworthy feature involves the incorporation of a timer for character prediction. During our initial testing phase, we omitted the introduction of a delay in the prediction process, resulting in chaotic outcomes as the system instantaneously processed input. To address this issue, a specialized technique was developed to capture the time value when the application first runs. Subsequently, upon predicting a character, the current time of the prediction is recorded. In the later stages of implementation, this time value is employed in a condition: if the last predicted time exceeds the desired delay time, the system proceeds to predict the next character (or alphabet). If not, it must wait until the time surpasses the last predicted time. This approach proves beneficial in preventing potential errors that could arise if the system identifies the gesture too quickly. Considering that a swift change in hand gesture might lead to misidentification, this method ensures a more accurate and reliable prediction process.

RFC, it doesn't involve optimization techniques like gradient descent, learning rates, or batch sizes as in deep learning models. Instead, it relies on the ensemble of decision trees to collectively make predictions based on the input features. To achieve high accuracy, several parameter adjustments have been considered. Firstly, we set the number of trees to 120 to capture the complexity of the dataset. Additionally, the maximum depth value is adjusted to 35, allowing trees to grow to their maximum

depth. To prevent overfitting, splitting nodes and leaf nodes is set to 12. Considering large image size, fewer features for splitting nodes is chosen, hence, maximum features setting it to " $\log_2 x$ ". Finally, both "gini" and "entropy" criteria is experimented for splitting nodes to determine which performed better.

5.2 Blur Restoration and Size Optimization

Given that our method focuses on real-time sign recognition, there is a heightened risk of image degradation compared to other systems. To tackle this challenge, we have employed linear algebra techniques to enhance performance. By utilizing Singular Value Decomposition (SVD) (Abdi, H. 2007), in linear algebra SVD is a factorization of a real or complex matrix into a rotation, followed by a rescaling followed by another rotation. Each pixel's saturation level in the original image is treated as numerical entries in a matrix denoted by $A_{m \times n}$. Furthermore, computing the SVD of this matrix, it breaks down matrix A into a special form as:

$$A = U_{m \times m} \Sigma_{m \times n} V_{n \times n}^T \tag{1}$$

This calculation enables us to retain the crucial components of the image while discarding fewer essential elements, thereby preserving image quality. The nonzero singular values σ_i are arranged in a specific order based on their magnitude. Hence,

considering only k th sufficiently large singular values and their corresponding singular vectors, the image matrix can be reconstructed as follow:

$$\hat{A} = U_{k \times k} \Sigma_{k \times k} V_{k \times k}^T \quad (2)$$

This decomposition breaks down matrix \hat{A} into its essential components, resulting in a modified version with smaller components. This reduction in size means less memory needed to store the information. To incorporate this approach into the proposed method involves treating the saturation levels of each pixel as entries in a matrix. The SVD of this matrix is computed, and singular values are removed following the procedure outlined in (2), starting from the smallest. Subsequently, the modified matrices are converted back into images. This process effectively reduces the required storage without compromising the image quality.

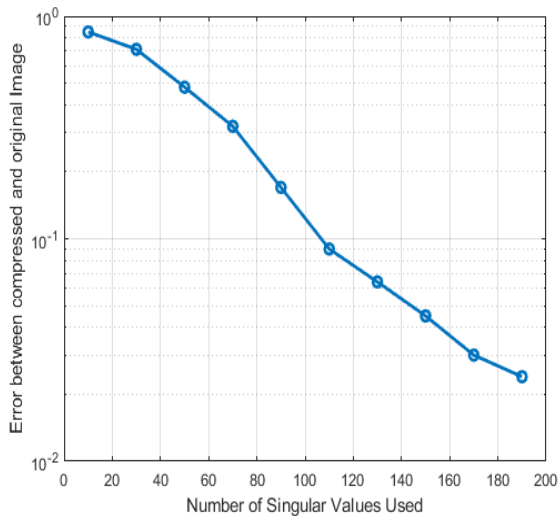


Figure 4: Compression error as a function of number of singular values considered

As depicted in Figure 4, the number of singular values and their corresponding singular vectors have a proportional relationship with the quality of image compression. Therefore, an adequate number of singular values must be considered during the comparison process.

Additionally, Moore-Penrose inverse based SVD are implemented for resizing images (Barata, J. C. A., et al., 2012). This step helps in eliminating noise or blur caused by linear motion, ensuring clearer and more precise recognition. The Moore-Penrose inverse matrix, often denoted as A^{\dagger} , a powerful tool in linear algebra used to solve systems of equations, particularly when dealing with matrices that may not have an exact inverse. In the context of image restoration, the Moore-Penrose inverse can be used to recover the original image $f(x, y)$ from the convolved image $g(x, y) = f(x, y) * h(x, y)$, where $*$ denotes convolution and $h(x, y)$ represents the blurring function. By approximating the inverse of the blurring matrix using its Moore-Penrose inverse H^{\dagger} , the original image F can be estimated as:

$$F = G H^{\dagger} \quad (3)$$

This will result in an improvement in the signal-to-noise ratio (SNR) of the image and increase the overall system performance when the images are input into the suggested model.

6. RESULTS AND DISSCUSSION

In conclusion, in this paper a new dataset has been generated for Kurdish Sign Language with the addition of three new special characters for the purpose of real time recognition. Random forest classifier algorithm and SVD have been used by considering the unique characteristics of the KuSL. The Random Forest classifier trains the data through a process of bootstrapped sampling, where multiple decision trees are constructed from randomly selected subsets of the dataset. During training, each tree "votes" on the predicted class, and the final prediction is determined by aggregating the votes from all trees.

Furthermore, the paper addressed the challenge of image degradation in real-time sign recognition by employing linear algebra techniques, specifically SVD implementation of pseudoinverse for the image compression and blur filtering. To evaluate the system performance, 6 characters have been selected for the testing under two different scenarios; Scenario 1 (ideal) involves a white background with appropriate lighting, while Scenario 2 (not ideal) entails a complex background with spotlighting from different angles (refer to Appendix-B for detail). As shown in Table 1, a perfect match of recognition is obtained for scenario 1 and slightly lower performance for Scenario 2.

Table 1: The real-time precision of the characters

No.	Character	Accuracy (Scenari-1) %	Accuracy (Scenario-2) %
1	ژ	97	90
2	ڤ	100	93
3	ب	92	85
4	ه	91	82
5	د	96	83
6	Delete Sign	100	92
Average		96.1	87.5

Table 2 presents a classification report for the proposed model utilized in the recognition of Kurdish sign language. These findings were derived from an experiment conducted using cross-validation. As indicated in the table, misclassification occurred infrequently for certain selected letters within the dataset under

investigation. These results highlight the minimized misclassification rate, underscoring the efficacy of the proposed model.

Table 2: Classification Report

Letters	Precision	Recall	F1-score
ڤا	0.99	0.96	0.97
ڤا	1.00	0.99	0.99
ڤا	0.93	0.88	0.90
ڤا	1.00	1.00	1.00
ڤا	0.94	1.00	0.97
ڤا	0.92	0.79	0.85
ڤا	0.99	1.00	0.99
ڤا	0.94	0.98	0.96
ڤا	1.00	1.00	1.00
ڤا	0.88	0.90	0.89

For dynamic gesture recognition, the signs must be stored in video format, with N frames per file. Consequently, the real-time system needs upgrades to evaluate each sign over a period of K seconds to distinguish between static and dynamic gestures. This adjustment inevitably slows down system response compared to the current model since it requires evaluating each sign before accepting the next sign letter. On the other hand, for pre-processing techniques, we now employ polynomial SVD to

extract polynomial singular values from video files represented as 3D images. This approach enhances the feature extraction process, enabling more nuanced analysis of dynamic gestures. On the other side, a comparative analysis of the proposed approach and current state-of-the-art methodologies concerning accuracy within the domain of Kurdish Sign Language is presented in Table 3.

Table 3: Comparative Evaluation of the Proposed Model Against State-of-the-Art Sign Language Classifier

References	Sign Language (SL)	Approach	Accuracy
Daniels, S., et al., 2021.	Indonesian SL	YOLO	77%
Taskiran, M., et al., 2018.	American SL	CNN	98.5%
Tang, A., et al., 2015.	American SL	DBN	98.1%
Li, S. Z., et al., 2015.	American SL	SAE PCA-Net	99%
Alzohairi, R., et al. 2018.	Arabic SL	SVM using HOG descriptors	63%
Aly, S., et al. (2016).	Arabic SL	PCA Net SVM	99%
Mirza, S. F., & Al-Talabani, A. K., 2021.	Kurdish SL	RNN	97%
Salim, B. W., & Zeebaree, S. R., 2023	Kurdish SL	CNN	96%
Hama Rawf, et al., 2024.	Kurdish SL	CNN	99%
Proposed model	Kurdish SL	RFC	96%

Instead of concentrating mainly on accuracy, several studies look into alternative metrics including complexity which consider training time and throughput besides the accuracy gestures. With this example, (Daniels et al. 2021) declared that their model needed one second to train. However, (Tang et al. 2015) discovered that the training times were substantially longer as their will deep belief network required 485.1 minutes and their CNN used 790.4 minutes to accomplish the training. Consequently, their deep learning resulted in over 16 hours of training. Similarly, a recent study by (Hama Rawf et al. 2024) showed that the procedure had the duration of 1 hour 40 minutes and 13.49 seconds. Resulting in throughput of around 10.70 samples per second. In this paper RFC has been implemented which exchanges the CNNs architectural complexity and training

difficulty for greater stability. The CNNs that have a more complicated architecture and heavy-duty training procedure, require more extensive computing resources for this procedure. During model training, the accuracy and loss for validation data may fluctuate in different scenarios. Ideally, as the number of training epochs increases, the loss should decrease, and accuracy should improve. In our suggested approach, both validation and training losses consistently improved over 40 epochs, indicating convergence. The validation loss decreased during training, and the curves of all metrics became nearly identical after 28 epochs. Figures 5-A and 5-B demonstrate that the proposed method predict a perfect match, ensuring the results neither overfit nor underfit.

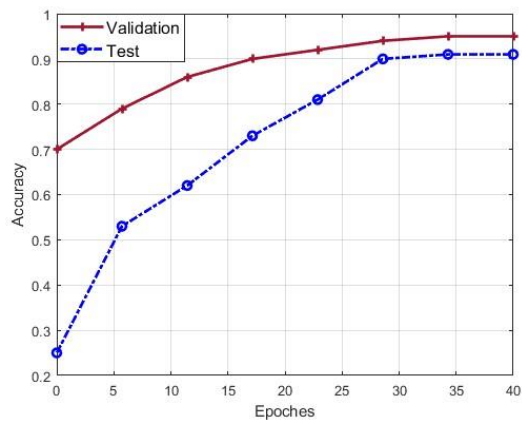


Figure 5-A: Accuracy for the proposed model

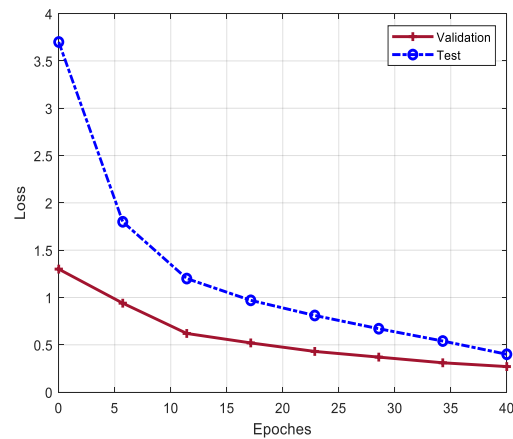


Figure 5-B: Loss evaluation for the proposed model

CONCLUSION

This study introduced real-time Kurdish sign language recognition using the random forest classifier algorithm. A dataset was generated, comprising multiple images captured for training purposes for each sign. Previously, KuSL lacked specific characters for interpretation as other gestures; thus, three special characters (delete character, delete sentence, and space) were incorporated. Furthermore, linear algebra techniques such as singular value decomposition and Moore-Penrose were employed for real-time processing to facilitate image comparison and blur removal. Simulation scenarios were conducted under two conditions: (i) an ideal scenario with a white background and proper lighting and (ii) a challenging scenario featuring complex backgrounds and lighting from various angles affecting hand gestures. In both cases, a high matching performance of 96% and 87%, respectively, was achieved. Additionally, a classification report analysis was performed using precision, recall, and F1 score metrics.

REFERENCES

Abdi, H. (2007). Singular value decomposition (SVD) and generalized singular value decomposition. *Encyclopedia of measurement and statistics*, 907, 912.

Alzohairi, R., Alghonaim, R., Alshehri, W., & Aloqeely, S. (2018). Image based Arabic sign language recognition system. *International Journal of Advanced Computer Science and Applications*, 9(3).

Aly, S., Osman, B., Aly, W., & Saber, M. (2016). Arabic sign language fingerspelling recognition from depth and intensity images. In 2016 12th International Computer Engineering Conference (ICENCO) (pp. 99-104). IEEE.

Barata, J. C. A., & Hussein, M. S. (2012). The Moore–Penrose pseudoinverse: A tutorial review of the theory. *Brazilian Journal of Physics*, 42, 146-165.

Cheok, M. J., Omar, Z., & Jaward, M. H. (2019). A review of hand gesture and sign language recognition techniques. *International Journal of Machine Learning and Cybernetics*, 10, 131-153.

Daniels, S., Suciati, N., & Fathichah, C. (2021, February). Indonesian sign language recognition using yolo method. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1077, No. 1, p. 012029). IOP Publishing.

Hama Rawf, K. M., Abdulrahman, A. O., & Mohammed, A. A. (2024). Improved Recognition of Kurdish Sign Language Using Modified CNN. *Computers*, 13(2), 37.

Hashim, A. D., & Alizadeh, F. (2018). Kurdish sign language recognition system. *UKH Journal of Science and Engineering*, 2(1), 1-6.

Huang, J., Zhou, W., Zhang, Q., Li, H., & Li, W. (2018, April). Video-based sign language recognition without temporal segmentation. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 32, No. 1).

Kamal, Z., & Hassani, H. (2020). Towards Kurdish text to sign translation. In *Proceedings of the LREC2020 9th Workshop on the Representation and Processing of Sign Languages: Sign Language Resources in the Service of the Language Community, Technological Challenges and Application Perspectives* (pp. 117-122).

KP Nimisha and Agnes Jacob. "A brief review of the recent trends in sign language recognition". In: 2020 International Conference on Communication and Signal Processing (ICCSP). IEEE, 2020, pp. 186–190.

Leigh, G., & Crowe, K. (2015). Responding to cultural and linguistic diversity among deaf and hard-of-hearing learners. *Educating deaf learners: Creating a global evidence base*, 69-92.

Liu, Y., Wang, Y., & Zhang, J. (2012). New machine learning algorithm: Random forest. In *Information Computing and Applications: Third International Conference, ICICA 2012, Chengde, China, September 14-16, 2012. Proceedings 3* (pp. 246-252). Springer Berlin Heidelberg.

Li, S. Z., Yu, B., Wu, W., Su, S. Z., & Ji, R. R. (2015). Feature learning based on SAE–PCA network for human gesture recognition in RGBD images. *Neurocomputing*, 151, 565-573.

Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., ... & Grundmann, M. (2019). Mediapipe: A framework for building perception pipelines. *arXiv preprint arXiv:1906.08172*.

Mahmood, M. R., & Abdulazeez, A. M. (2019, April). Different model for hand gesture recognition with a novel line feature extraction. In *2019 International Conference on Advanced Science and Engineering (ICOASE)* (pp. 52-57). IEEE.

Mayyadah Ramiz Mahmood, Adnan Mohsin Abdulazeez, and Zeynep Orman. "Dynamic hand gesture recognition system for kurdish sign language using two lines of features". In: 2018 International Conference on

- Advanced Science and Engineering (ICOASE). IEEE. 2018, pp. 42– 47.
- Mirza, S. F., & Al-Talabani, A. K. (2021). Efficient kinect sensor-based kurdish sign language recognition using echo system network. *ARO-The Scientific Journal of Koya University*, 9(2), 1-9.
- Nimisha, K. P., & Jacob, A. (2020, July). A brief review of the recent trends in sign language recognition. In 2020 International Conference on Communication and Signal Processing (ICCSP) (pp. 186-190). IEEE.
- Parmar, A., Katariya, R., & Patel, V. (2019). A review on random forest: An ensemble classifier. In *International conference on intelligent data communication technologies and internet of things (ICICI) 2018* (pp. 758-763). Springer International Publishing.
- Rastgoo, R., Kiani, K., & Escalera, S. (2021). Sign language recognition: A deep survey. *Expert Systems with Applications*, 164, 113794.
- Rawf, K. H., Abdulrahman, A., & Mohammed, A. (2022). Effective Kurdish Sign Language Detection and Classification Using Convolutional Neural Networks.
- Salim, B. W. ., & Zeebaree, S. R. M. . (2023). Kurdish Sign Language Recognition Based on Transfer Learning. *International Journal of Intelligent Systems and Applications in Engineering*, 11(6s), 232–245.
- Tang, A., Lu, K., Wang, Y., Huang, J., & Li, H. (2015). A real-time hand posture recognition system using deep neural networks. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 6(2), 1-23.
- Taskiran, M., Killioglu, M., & Kahraman, N. (2018). A real-time system for recognition of American sign language by using deep learning. In 2018 41st international conference on telecommunications and signal processing (TSP) (pp. 1-5). IEEE.
- United Nations. (April 11, 2024). International Day of Sign Languages. Retrieved from <https://www.un.org/en/observances/sign-languages-day>
- Valli, C., & Lucas, C. (2000). *Linguistics of American sign language: An introduction*. Gallaudet University Press.
- Varshney, P. K., Kumar, G., Kumar, S., Thakur, B., Saini, P., & Mahajan, V. (2023). Real Time Sign Language Recognition.
- Xu, Y., & Goodacre, R. (2018). On splitting training and validation set: a comparative study of cross-validation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning. *Journal of analysis and testing*, 2(3), 249-262.
- Youssif, A. A., Aboutabl, A. E., & Ali, H. H. (2011). Arabic sign language (arsl) recognition system using hmm. *International Journal of Advanced Computer Science and Applications*, 2(11).

APPENDIX-A

Table of Kurdish language alphabet signs

Label	Image	Label	Image	Label	Image
ئا		ب		گ	
ـ		ز		ل	
ـ		ژ		ئ	
ـ		س		م	
ـ		ش		ن	
ـ		ع		و	

چ		غ		ز	
ح		فا		ه	
خ		قا		ه	
د		ق		ی	
ر		ک		ئ	

APPENDIX-B

Real life difficult conditions for Kurdish sign language

