Efficient Sign Language Recognition and Classification Using African Buffalo Optimization Using Support Vector Machine System

Karthikeyan M. P. [1†], Vu Cao Lam[2††] and Dac-Nhuong Le[2††],

[1]School of CS & IT, JAIN (Deemed-to-be University), Bengaluru, India
[2]Faculty of Information Technolgy, Haiphong University, Haiphong, Vietnam

Abstract
Communication with the deaf has always been crucial. Deaf and hard-of-hearing persons can now express their thoughts and opinions to teachers through sign language, which has become a universal language and a very effective tool. This helps to improve their education. This facilitates and simplifies the referral procedure between them and the teachers. There are various bodily movements used in sign language, including those of arms, legs, and face. Pure expressiveness, proximity, and shared interests are examples of nonverbal physical communication that is distinct from gestures that convey a particular message. The meanings of gestures vary depending on your social or cultural background and are quite unique. Sign language prediction recognition is a highly popular and Research is ongoing in this area, and the SVM has shown value. Research in a number of fields where SVMs struggle has encouraged the development of numerous applications, such as SVM for enormous data sets, SVM for multi-classification, and SVM for unbalanced data sets.Without a precise diagnosis of the signs, right control measures cannot be applied when they are needed. One of the methods that is frequently utilized for the identification and categorization of sign languages is image processing. African Buffalo Optimization using Support Vector Machine (ABO+SVM) classification technology is used in this work to help identify and categorize peoples' sign languages. Segmentation by K-means clustering is used to first identify the sign region, after which color and texture features are extracted. The accuracy, sensitivity, Precision, specificity, and F1-score of the proposed system African Buffalo Optimization using Support Vector Machine (ABOSVM) are validated against the existing classifiers SVM, CNN, and PSO+ANN.
Keywords:
African Buffalo Optimization, Support Vector Machine, Sign Language prediction, MNIST image dataset, image processing, confusion matrix.

1. Introduction

Deaf and dumb persons who are unable to talk or hear correctly have just one way of communication, which is nonverbal. It would be difficult for them to convey their message to ordinary people. Finding a skilled interpreter on a regular basis is difficult and expensive. We wanted to create a system that transforms sign language into text format using a vision-based technique that is also cost-effective. Sign language is made up of many face expressions, hand movements, hand transformations, and hand utilized for communication [1]. Each sign has a certain alphabet and meaning attached to it. ASL, BSL, and JSL are only a few of the languages that are spoken worldwide. The average person never makes an effort to learn sign language in order to converse with the deaf and dumb in Figure 1.

Figure 1. The letters A, S, L

This isolation can be lifted with the assistance of a computer. If a computer could be designed to convert sign language into written format. This article will tell you how we created a system, what the requirements were, and what sort of data we utilised for system training and testing. It will also tell you about past research done on sign language. It concludes with a conclusion. Children communicate and reason about spatial relationships from an early age. Previous research indicates a close relationship children's spatial language may be able to predict their cognition, including their spatial memory, between these two systems. This does not answer whether exposure to a language early on affects whether infants who have access to language from birth choose to speak a traditional spoken language as their first language or late life might alter link [2].

The majority of deaf children (85%) are born to hearing parents and may not have early exposure to a traditional signed or spoken language, even with hearing aids or cochlear implants, which may not allow appropriate access to surrounding speech.

Manuscript received June 5, 2024
Manuscript revised June 20, 2024
https://doi.org/10.22937/IJCSNS.2024.24.6.2

Particularly in non-Western countries, many deaf children who have hearing parents pick up sign language later in life.Usually after enrolling in a deaf school in Figure 2.

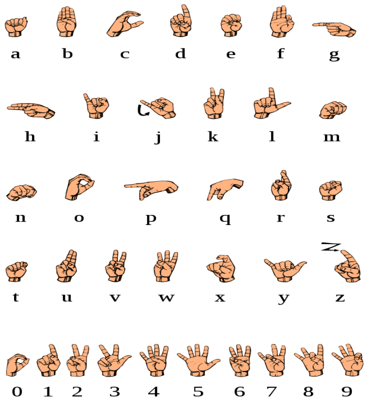

Figure 2. The letters A, S, L

El Ghoul *et al.* [3] obtain a result, they are termed late signers since they did not have first introduction to sign language after birth. On the other hand, Native signers are deaf children whose parents are deaf because they are exposed to sign language through their vocations from an early age. The visual technique understanding sign language has a definite advantage most user-friendly SLR systems are visual ones since they shorten the time it takes for the user to get used to the interface [4]. The developers are also having issues presents gesture information as a collection video devices are used to give the system with input.

Extraction of key frames**:** A video camera can record images between 15 and 120 frames per second. The average spelling rate of sign language users per minute is constrained by the finite hand movement speed, the requirement to lock it for a small amount of time, and the length of the transitional process between the signalsan adult may safely sign 300 letters per minute. The computational expense of the signature and camera speed overhead results in a significant lengthening of the recognition time. Consequently, it's essential to examine the most informative frames. Figure 3 explains about the flow of proposed system sign language detection using image processing technique.

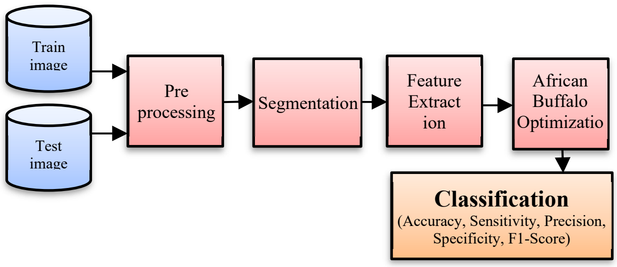

Figure 3. The flow of proposed system

Preprocessing**:** Key image frames must be filtered at the preprocessing step in order to collect relevant data, reduce distracting noise, a static background, and undesired items.The suggested system's image preprocessing entails image resizing, grayscale conversion, and edge recognition for effectively extracting its features. To adhere to a standard cropping ratio, the test and training photos are first downsized to $a \times b$ pixels.

Segmentation**:** The basis for object recognition and computer vision is image segmentation. Generally speaking, image preprocessing should remove image noise. And following the primary operation of image segmentation, there is some explicitly assigned work (such as region extraction and picture marking) to complete for the purpose of acquiring a better visual effect.

Feature extraction**:** The algorithms that identify fingers, palms, faces, and their configuration, orientation, etc. are then fed the processed photos. The process transforms the image into a data array that represents the input sign.

Classification**:** One of the more common issues in image processing is classifying images. Use the attributes of the supplied image to predict the categories is the aim of image classification. The k closest neighbor, adaptive boost, ANN, and SVM are some of the methods that can be used to solve this problem.

This paper is divided into five sections: Session 2, which describes the current or existing system using CNN, ANN and other classification methods; Session 3 deals with African Buffalo Optimization using Support Vector Machine (ABOSVM), which describes the methodology of the proposed system; Session 4, which presents the results of the proposed system; and Session 5, which concludes the proposed system.

2. Literature Review

This section talks about the problems with the existing methods that led to the creation of a sign language prediction that performed poorly during the categorization stages. According to the literature review, there have been various recent developments and deficiencies are identified.

It was advised to use a suitable ASL recognition system. Then, after analyzing images of various regional architecture, each sign was translated into language in accordance with its significance, and this process was repeated for each symbol. This method detects 10 numbers and 24 ASL characters. Singh *et al.* [5] took an experiment contrasted CNN-based and contour support vector (SVM)-based methods. Three of the most often used standard information sites are SLD, ASL, and ASL-FS, and they have all been put to the test in a variety of rotation and backdrop measurement circumstances. Using a concrete-based process, the concert was purchased and painted as a convex shell depending on length and convex angle. SVM was used to further identify the touch this. The proposed comparison has an accuracy rate of 69 percent and a rate of 98.31%. The CNN algorithm was applied to the MNIST and CIFAR-10 databases. Rastgoo *et al.* [6] plan to evaluate the algorithm's capacity for symbol recognition and identification, several image databases were employed. In order to recognize signal quantities with quick assembly and minimal crowding, deep CNN has been used. Implement batch normalization, pause for merging, and reduce consecutive overlap to restore the layers.

Camgoz *et al.* [7] choose to use CNN models to identify signs in automated sign language translation. Between hearing-impaired people and people with adequate hearing, a sign language identification and translation system can aid. People who are deaf or hard of hearing can become contributing members of society by being allowed to participate in social, economic, and political activities. With 265 million speakers, Bangla is the sixth most spoken language in the world, although there is little research and employment available in Bengali sign language automatic detection. To make a contribution to this field, we developed a deep learning-based BSL identification system that can correctly classify BSL letters and digits from sign pictures. Due to the unavailability of extensive open-source datasets for sign language recognition, many earlier experiments used a little amount of data and did terribly. Mittal *et al.* [8] study, examining normally developing youngsters with instant access to a spoken language would not have allowed.

Li *et al.* [9] states that because it enables us to express our feelings and thoughts, communication is a crucial part of our existence. People who are dumb or deaf struggle because they are unable to speak in their native languages. Communication relies heavily on language, whether it is nonverbal (using facial expressions and sign language) or verbal (using words to speak, read, and write). Therefore, those who are deaf or dumb have no alternative except to communicate through nonverbal Sign language. But others who are not familiar with sign language find it difficult. Wadhawan *et al.* [10] plan to develop a few automated movements that we use frequently. The project uses a learning algorithm and a collection of datasets that include images of each letter of the integer's images is performed using a convolutional neural network. In order to access the camera and collect data, also used an open CV.

Bazarevsky *et al.* [11] achieved this by comparing the signers (8-year-old youngsters) with that of their classmates who sign natively.The results showed that young children and those who start signing later frequently memorize the spatial relationship, much like their counterparts who sign more naturally. However, in contrast to their peers who signed in their mother tongue, late signers used a certain type of spatial language but not children. However, it was discovered that neither the frequency nor type of spatial language use, nor exposure to late sign language, had an impact on the accuracy of spatial memory. Although late language exposure appears to have an impact on the way spatial language is used, it does not seem to be a reliable indicator of future spatial memory.

Krebs *et al.* [12] studies lone computer scientists continue to conduct the most of the research in this area. This article compiles extensive research on SL translation from various areas. To demonstrate the importance of automatic SL translation, we first provide a high-level overview of SL linguistics and machine translation (MT). The state of the art in the field is then thoroughly reviewed in the literature. De Coster *et al.* [13] list crucial open research questions.We hear that research on spoken language MT has resulted in substantial advancements. The methods that are now being used, however, either discourage language learning or aren't designed for SLs. We look at issues with SL data representation, dataset gathering, and evaluation of SL translation models. We encourage future linguistic studies on SLs as well as multidisciplinary study. In order to develop useful SL translation models, use case research, data collection, and evaluation must also include both hearing and deaf end users of SL translation apps.

JyothiRatnam *et al.* [14] discusses how a novel structure for a deep learning neural network was created to recognize static hand gestures in sign language. The suggested structure combines a standard, non-intelligent feature extraction technique with convolutional neural networks (CNN). Three distinct feature extraction streams are used to route the hand gesture image after preprocessing background removal in order to extract meaningful features and create three widely employed methods for categorizing hand gestures that each extract their own unique characteristics independently. The suggested structure gains increased resistance to hand gesture ambiguities like rotation and ambiguity as well as extraordinarily high classification accuracy by combining these effective approaches. When compared to previous methodologies, another unique feature of the recommended structure is its coverage of multiple picture databases by Murtagh *et al.* [15]. The transfer learning method was used to show how the suggested structure then used to analyze images in Massey, ASL Alphabet, and ASL databases, respectively.80 percent in each case. The results show that the recommended structure's mean accuracy photographs is 75.92%, for ASL, it is 78.8%, and for ASL Alphabet, it is 79.80%.

Angelova *et al.* [16] claims that for people who are hard hearing or deaf, complete form of public communication. Since they are unable to speak or hear, they cannot exchange information by speech or sound signals. Instead, individuals are required to use visual signals convey emotions in day-to-day interactions. They do by using a variety of hand gestures and other body language. There are two groups of sign language concepts Numbers in the form of digits and alphabetical characters. In this paper, we suggested the datasets "Ishara-Lipi" and "Ishara-Bochon"are the first extensive, versatile, open-access datasets BSL and feature isolated numerals and alphabets, respectively. Nair *et al.* [17] Additionally, we developed a backdrop removal strategy to take unnecessary details out of sign photos. For digit recognition, the system achieves accuracy, precision, recall, and f1-score values of 87.33%, 87.89%, 87.33%, and 87.37%.

Othman *et al.* [18] describes this More than 5% of the world's population, according to the World Health Organization (WHO), is deaf and faces significant communication difficulties with hearing people. They struggle to explain themselves when there is no interpretation their signals. Sign recognition systems have a number of challenges, including low accuracy, complex motions, high levels of noise, and the ability to generalize or be constrained by such restrictions. As a result, numerous academics offered diverse answers to these problems.It could be challenging to cover all the signals because each language has its own unique set. Seligman [19] explains about two objectives of the current study are to provide a dataset of 20 Arabic words and to suggest a deep learning (DL) architecture that mixes recurrent and convolutional neural networks. The proposed architecture obtained 88% accuracy on the given dataset. On the UCF-101 dataset, it also reported top-1 and top-5 accuracies of 83.4% and 88.8%, respectively.

Chen *et al.* [20] indicates that hand gestures are a versatile method of human-to-human communication. People with speech impairments use them frequently for communicative purposes since they are a natural form of interaction. In actuality, this category includes about 1% of Indians. For segmentation, backdrop removal and skin tone are both used. In order to map the signs with proper names, histograms were constructed using the photographs' SURF features. Support vector machines SVM and CNN are used for categorization [22-25].

3. System Design

The proposed system for sign language prediction for deaf people who can hear a voice is discussed in this session. As a result, when compared to current systems like SVM, CNN, and ANN, the suggested method, African Buffalo Optimization using Support Vector Machine (ABOSVM), produces better results.

3.1 Image Acquisition

We needed to gather a dataset of photos or videos of Norwegian sign language in order to create our NSL (Norwegian sign language) recognition model. We conducted a thorough search but were unable to locate any publicly available NSL datasets, so we had to start from scratch to build our own. Choosing which sign language category to establish was the initial step in generating our NSL dataset.

3.2 Preprocessing

The preprocessing block normalizes the frames and ensures input homogeneity by processing the different frame sizes in videos. Input signal homogeneity is necessary for the model to handle inter-database inconsistencies brought on by the use of multiple and digitalization processes. More specifically, the model can accept input in a variety of circumstances that strike a balance between reducing input and maintaining the signal characteristics required to complete the sign scoring assignment. Similar to epoch-based scoring methods, 30-s window video segmentation produces input patterns of a specified size that are fed into the next convolutional neural network processing block. The amplitude of each of these input patterns is then normalized using a Gaussian standard process. Although it is not necessary, the input signal can be filtered as a first step in the processing process. This section focuses on eliminating non-essential but dataset-specific signals to increase the generalizability of the final model.

3.3 Segmentation

While edge-based segmentation divides an image based on sudden changes in intensity along the edges. The key examples of approaches in this area are thresholding, region increasing, region splitting, and region merging. This research merits widespread attention because it can serve as a vital source of reference for those segmentation methods. Universal applicability of evaluation algorithms, their simplification and dependability, and if referent pictures or manual involvement are required. In general, two fundamental techniques are used to evaluate image segmentation objectively.

3.4 Feature Extraction

Although the use of the hands and postural estimate are crucial to sign language, DSL presents a number of difficulties because to the constant movement. Identifying the hands and figuring out their size, shape, and motion are some of these difficulties. The critical points are extracted for three *X*, *Y*, and *Z* dimensions of both hands after it calculates postures for each frame. Using the posture estimation technique, the hand location in relation to the body was predicted and tracked.

3.5 Hybrid African Buffalo Optimization with Support Vector Machine

A closer examination of the algorithm ABO finds that which illustrates buffalos' behavior has three components. The initial m.k stands for the buffalo's recall of its previous location. The buffalo has an intrinsic memory that allows recall its previous whereabouts. This is essential in helping it avoid areas that gave poor outcomes in its hunt for answers. Each buffalo has a memory that has a list of alternatives to the current local maximum location. The third portion, $lp_2\left(bpmax_k - w_k\right)$ denotes the intelligence of the buffalos. The second part, $lp_1\left(bgmax_k - w_k\right)$ refers to the buffalo's sociable and information-sharing behavior and is concerned with the compassionate or cooperative aspect of the species. Therefore, ABO uses the buffaloes' intelligence, memory, and compassion in the democratic equation.

African Buffalo Optimization algorithm essentially simulates the three main traits of the African Buffalo listed above. The 'waaa' sound is used to symbolize the buffalos'maaa' sound $\left(k = 1,\ 2,\ 3\ ...n\right)$.

African Buffalo Optimization (ABO) Algorithm

1. **Objective function**: $f\left(x\right)$: $x = \left(x_1, x_2, ..., x_n\right)^T$

2. **Initialization:** Add buffaloes at random to the solution space's nodes;

3. Update the buffalo's fitness values using Eq (1).

$$m_{k+1} = m_k + lp_1\left(bgmax_k - w_k\right) + lp_2\left(bpmax_k - w_k\right) \quad (1)$$

where *bgmax* is the herd's best fitness and $bpmax_k$ is the individual buffalo's best, and w_k and m_k represent the k^{th} buffalo's exploration and exploitation moves, respectively $\left(k = 1, 2, ..., N\right)$.

4. Update the location of buffalo *k* in relation to $bgmax_k$ and $bpmax_k$ using Eq(2)

$$W_{k+1} = \frac{\left(w_k + m_k\right)}{\pm 0.5} \quad (2)$$

5. **If** (*bgmax* is updating) **then** Go Step 6 **else** Goback Step 2.

6. **If** (the stoping criteria is not met) **then** Goback Step 3.

7. Output the best solution.

The Eq (1) determines the democratic movement of the buffaloes mathematically.

This mammal's harmony with other herd members is its second characteristic. The “waa” sound is an alarm call that the herd uses to signal exploration of new territory or, occasionally, assistance for a member in need. The herd members should continue moving forward because the current area is unsafe or unfavorable based on this sound. Sometimes a buffalo will make this sound when it is in danger and in need of assistance from other animals. However, the “maa” sound is a beautiful sound that lets the herd know that the current environment is lovely and secure. This sound assures coworkers that they are welcome to stay and make use of the location. The third and last element represents the parliamentary characteristics of the herd. If there are conflicting viewpoints within the herd, the majority decides the next course of action with the aid of elections. After the optimization the features are feed to SVM classifier to predict the letter with more accuracy.

3.6 Support Vector Machine

A classification tool, Support Vector Machine is a kernel-based supervised learning method. The margin between the training data and class boundary is maximized using SVM training algorithm. The training data, known as support vectors, that are closest to the decision border as depicted are the only data sources on which the decision function is based. It works well in high-dimensional spaces where there are more dimensions than training data. Using kernel functions, SVM converts data from the input space into a high-dimensional feature space. Hyper-plane separation of nonlinear data is also possible in high-dimensional space. Reduced computational complexity. Support vector machines are designed to build a hyper plane between data sets to identify the class to which each set of data belongs.

The classifier receives the feature vector as input. Training and testing feature vectors are separated for the database photographs. The training set serves as the basis for applying the classifier to the testing set. A comparison of the predicted labels and actual values is used to determine the classifier's effectiveness. One of the most well-known techniques for classifying patterns and images is SVM.

SVM Algorithm

Require: Sample x to classifier;

Testing set $T = \{(x_1, y_1),\ (x_2, y_2), ..., (x_n, y_n)\}$;

Number of nearest neighbour k.

Ensure: Decision $Y_p \in \{-1, 1\}$

1. Find the K sample (x_i, y_i) with minimal values

$$K(x_i, y_i) - 2 \times K(x_i, y_i) \quad (3)$$

2. Test an *SVM* model on the K selected samples
3. Classify x using this model, get the result Y_p
4. Return Y_p

By eliminating the sign function, it is also possible to obtain the result as a real number rather than in binary form the SVM algorithm pseudocode for test image. Using the formulas $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$, where x_i is a d-dimensional feature space in Rd and y i is the class label in the range of (-1,+1), where i=1..n [1], the objective is to divide a series of training images into two unique classes. The best separating hyper planes are produced using SVM using a kernel function (K). All photographs with feature vectors on one side of the hyperplane are categorized as class -1 photographs, while all other photographs are categorized as class +1 photographs.

This pattern provides a starting point for it is clear that sign language uses similar language processing prediction mechanisms to spoken language interpretation when considering how language and the visual environment may be processed through separate channels. The theoretical foundations of language prediction, which up to now have been mostly fashioned by studies in the spoken domain, have also been greatly influenced by our study. Our work is a first, crucial step in describing the extent to which prediction may be a crucial strategy for directing attention during language processing, regardless of modality, even though more research is required.

4. Result and Discussion

In this session explains the outcomes of the suggested system African Buffalo Optimization using Support Vector Machine (ABO+SVM). Sign language prediction image based and MNIST image when validating the newly built classification technique, dataset are considered. The training dataset includes 4000 photos each of the letter and numbers. The proposed study employs MATLAB 2013a to assess the effectiveness of classification. The SVM, CNN, PSO+ANN, and newly constructed ABO with SVM are contrasted in terms of sensitivity and accuracy metrics.

Accuracy: The accuracy formula defines accuracy as the deviation of the error rate from 100%. We must first determine the error rate in order to determine accuracy. Additionally, by dividing the observed value the actual value, the error rate is determined as a percentage.

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \quad (4)$$

The results shows that the accuracy of the proposed system, measured by ABO+SVM, is 91.23%, 1.88 times greater than the accuracy of the present system, measured by 89.35%.

Table 1. The comparison of accuracy values

Algorithm	SVM	CNN	PSO+ANN	ABO+SVM
Accuracy	88.35%	87.23%	89.35%	**91.23%**

Sensitivity: Sensitivity is defined as the ratio of the total number of positives to the total number of valid positive predictions.

$$Sensitivity = \frac{TP}{TP + TN} \quad (5)$$

Table 2. The comparison of sensitivity values

Algorithm	SVM	CNN	PSO+ANN	ABO+SVM
Sensitivity	86.37%	88.23%	89.06%	**92.83%**

Table 2's findings reveal that the proposed system's sensitivity, as determined by ABO+SVM, is 92.83%, which is 3.77 times more than the current system's sensitivity, as determined by 89.06%.

Specificity: The proportion of total negatives to positives is correctly foreseen negative outcomes (TN) is used to calculate specificity (N).

$$Specificity = \frac{TN}{TP + TN} \quad (6)$$

Findings from Table 3 show that the Specificity of the proposed system, as determined by ABO+SVM, is 93.65%, which is 6.60 times greater than the Specificity of the current system, as determined by 87.05%.

Table 3. The comparison of specificity values

Algorithm	SVM	CNN	PSO+ANN	ABO+SVM
Specificity	86.89%	89.23%	87.05%	**93.65%**

Precision: By dividing the total positive predictions by the total positive predictions that were accurate (TP + FP), the precision is determined.

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

Table 4's findings demonstrate that the proposed system's precision, as determined by ABO+SVM, is 92.87%, which is 3.33 times higher than the present system's precision, as determined by 89.54%.

Table 4. The comparison of precision values

Algorithm	SVM	CNN	PSO+ANN	ABO+SVM
Precision	89.88%	85.55%	89.54%	**92.87%**

F1 score: The precision of a model is rated using the F1 score, a machine learning evaluation metric. It considers the recall and precision ratings of a model. The F1 score statistic calculates the proportion of times a model accurately forecast the complete dataset.

$$F1 - Score = 2 \times \frac{percision \times recall}{percision + recall} \quad (8)$$

The results of Table 5 show that the F1 score of the proposed system, as calculated by ABO+SVM, is 94.28, which is 5.27 times higher than the F1 score of the current system, as calculated by 89.01.

Table 5. The comparison of F1 score values

Algorithm	SVM	CNN	PSO+ANN	ABO+SVM
F1 score	82.28%	86.89%	89.01%	**94.28%**

Performance of the suggested system is significantly better than that of the existing system, as shown by the Tables (1, 2, 3, 4 and 5) and Figure 4. As a result, the suggested system performs better when employing the ABOSVM approach to forecast sign language. This indicates that the offered strategies performed better in the study of the deaf.

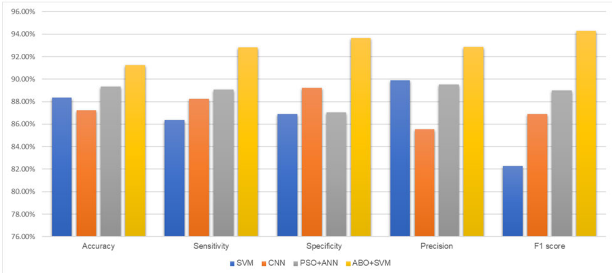

Figure 4. The comparison of algorithms

5. Conclusion

Using MNIST image dataset as the training image, we have described how to utilize African Buffalo Optimization, Support Vector Machine (ABO+SVM), and how to detect hand movements in front of a plain background and in good lighting in this study.The qualities of sign language and problems with translation have been researched. Even though finger spelling is a sluggish and ineffective form of communication, it offers a more basic linguistic. Research emphasizes that, rather than being an inclusive SL translator, the major goal of the SL prediction system is to match an active interpreter with a corresponding deaf person's needs population.African Buffalo Optimization with SVM method, which combined the classification outcomes of the other current ML techniques, produced the best classification performance when utilizing machine learning to predict sign language. The ABOSVM technique had a classification accuracy rating of 91.23%. In comparison to past work on sign language recognition, the new system used real-time classification and outperformed the previous system in terms of person-independent high classification accuracy. It was also tested by a single user from the training dataset.More symbols and symbols in multiple languages can be added to enhance the functionality, which aids deaf students in understanding. A user-friendly smartphone application could be created in the future.

References

[1] Zhou, H., Zhou, W., Zhou, Y., & Li, H. (2020, April). Spatial-temporal multi-cue network for continuous sign language recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 34, No. 07, pp. 13009-13016).

[2] Mittal, A., Kumar, P., Roy, P. P., Balasubramanian, R., &Chaudhuri, B. B. (2019). A modified LSTM model for continuous sign language recognition using leap motion. IEEE Sensors Journal, 19(16), 7056-7063.
[3] El Ghoul, O., & Othman, A. (2022, March). Virtual reality for educating Sign Language using signing avatar: The future of creative learning for deaf students. In *2022 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1269-1274). IEEE.
[4] Muzata, K. K. (2020). Interrogating parental participation in the education and general development of their deaf children in Zambia. *Afr. Disability Rts. YB*, *8*, 71.
[5] Singh, A., Singh, S. K., & Mittal, A. (2022). A review on dataset acquisition techniques in gesture recognition from Indian sign language. *Advances in Data Computing, Communication and Security: Proceedings of I3CS2021*, 305-313.
[6] Rastgoo, R., Kiani, K., &Escalera, S. (2021). Sign language recognition: A deep survey. *Expert Systems with Applications*, *164*, 113794.
[7] Camgoz, N. C., Koller, O., Hadfield, S., & Bowden, R. (2020). Sign language transformers: Joint end-to-end sign language recognition and translation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10023-10033).
[8] Mittal, A., Kumar, P., Roy, P. P., Balasubramanian, R., &Chaudhuri, B. B. (2019). A modified LSTM model for continuous sign language recognition using leap motion. *IEEE Sensors Journal*, *19*(16), 7056-7063.
[9] Li, D., Rodriguez, C., Yu, X., & Li, H. (2020). Word-level deep sign language recognition from video: A new large-scale dataset and methods comparison. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision* (pp. 1459-1469).
[10] Wadhawan, A., & Kumar, P. (2020). Deep learning-based sign language recognition system for static signs. *Neural computing and applications*, *32*, 7957-7968.
[11] Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., &Grundmann, M. (2020). Blazepose: On-device real-time body pose tracking. *arXiv preprint arXiv:2006.10204*.
[12] Krebs, J., Roehm, D., Wilbur, R. B., &Malaia, E. A. (2021). Age of sign language acquisition has lifelong effect on syntactic preferences in sign language users. *International journal of behavioral development*, *45*(5), 397-408.
[13] De Coster, M., Shterionov, D., Van Herreweghe, M., &Dambre, J. (2023). Machine translation from signed to spoken languages: State of the art and challenges. *Universal Access in the Information Society*, 1-27.
[14] JyothiRatnam, D., Soman, K. P., Bijimol, T. K., Priya, M. G., &Premjith, B. (2021). Hybrid machine translation system for the translation of simple english prepositions and periphrastic causative constructions from english to hindi. *Applications in Ubiquitous Computing*, 247-263.
[15] Murtagh, I., Nogales, V. U., &Blat, J. (2022, September). Sign language machine translation and the sign language lexicon: A linguistically informed approach. In *Proceedings of the 15th biennial conference of the Association for Machine Translation in the Americas (Volume 1: Research Track)* (pp. 240-251).
[16] Angelova, G., Avramidis, E., &Möller, S. (2022, May). Using neural machine translation methods for sign language translation. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop* (pp. 273-284).
[17] Nair, J., Nithya, R., &VinodJincy, M. K. (2020). Design of a morphological generator for an English to Indian languages in a declension rule-based machine translation system. In *Advances in Electrical and Computer Technologies: Select Proceedings of ICAECT 2019* (pp. 247-258). Springer Singapore.
[18] Othman, A., &Jemni, M. (2019). Designing high accuracy statistical machine translation for sign language using parallel corpus: case study English and American Sign Language. *Journal of Information Technology Research (JITR)*, *12*(2), 134-158.
[19] Seligman, M. (2019). The evolving treatment of semantics in machine translation. *Adv. Empir. Transl. Stud. Dev. Transl. Resour. Technol.*, 53.
[20] Chen, K., Wang, R., Utiyama, M., &Sumita, E. (2020, July). Content word aware neural machine translation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 358-364).
[21] Le, D. N., Nguyen, G. N., Bhateja, V., & Satapathy, S. C. (2017). Optimizing feature selection in video-based recognition using Max–Min Ant System for the online video contextual advertisement user-oriented system. *Journal of computational science*, *21*, 361-370.
[22] AL-kubaisy, W. J., Yousif, M., Al-Khateeb, B., Mahmood, M., & Le, D. N. (2021). The red colobuses monkey: a new nature-inspired metaheuristic optimization algorithm. *Int J Comput Intell Syst*, *14*(1), 1108-1118.
[23] Le, D. N. (2017). A New Ant Algorithm for Optimal Service Selection with End-to-End QoS Constraints. *Journal of Internet Technology*, *18*(5), 1017-1030.
[24] Dey, A., Biswas, S., & Le, D. N. (2023). Recognition of Human Interactions in Still Images using AdaptiveDRNet with Multi-level Attention. *International Journal of Advanced Computer Science and Applications*, *14*(10).
[25] Dey, A., Biswas, S., & Le, D. N (2024). Workout action recognition in video streams using an attention driven residual DC-GRU network. *Computers, Materials & Continua*, *79(2)*, 3067-3087.

M.P.Karthikeyan, is presently working as Assistant Professor, School of Computer Science and Information Technology, JAIN (Deemed-to-be University), Bangalore, India. His area of Specialization Machine Learning and Deep Learning. He obtained his Bachelors in Computer Science, from Sree Saraswathi Thyayagaraja College, Pollachi. He obtained his MCA, from Sree Saraswathi Thyayagaraja College, Pollachi. He Pursuing his PhD in Computer Science from Sri SRNM College, Sattur in the area of Query Optimization in Deep Reinforcement Learning. He Completed Industry Integrated Course Data Science and Machine Learning from Nunnari Labs, Coimbatore. He has published 6 Indian Patents. He has acted as resource person for DBT Sponsored workshop on "Big data Analytics using Machine Learning Techniques". He has authored and coauthored over 30 research papers in international journal (SCI/SCIE/ESCI/Scopus) and conferences including IEEE, Springer and Elsevier Conference proceeding. He is a co-author of 5 books and 15 books Chapters. He has presented various papers at international conferences. He has Completed more than 70 Online Course in Coursera, LinkedIn, etc. He is an IEEE Member and Member of Computer Society of India (CSI). He is a Board Studies Member for Various Autonomous Institutions and Question Paper Setter for Various Colleges. He acted as a Resource Person for Several Institutions. I am also Volunteer of Bangladesh Robotics Foundation.

https://orcid.org/0000-0002-2346-0283

Vu Cao Lam has graduate on computer science from the Faculty of Information Technology at Haiphong University, Vietnam. His core research is in soft computing and image processing.

Dac-Nhuong Le has an M.Sc. and Ph.D. in computer science from Vietnam National University, Vietnam in 2009 and 2015, respectively. He is an Associate Professor of Computer Science, Dean of the Faculty of Information Technology at Haiphong University, Vietnam. He has a total academic teaching experience of 20+ years with many publications in reputed international conferences, journals, and online book chapter contributions (Indexed By: SCI, SCIE, SSCI, Scopus, ACM, DBLP). His area of research includes Soft computing, Network communication, security and vulnerability, network performance analysis and simulation, cloud computing, IoT, and Image pro- cessing in biomedical. His core work is in network security, soft computing and IoT, and image processing in biomedical. Recently, he has been on the technique program committee, the technique reviews, and the track chair for international conferences under Springer-ASIC/LNAI Series. Presently, he is serving on the editorial board of international journals and he authored/edited 25+ computer science books by Springer, Wiley, CRC Press, Bentham, IET, IGI Global.

https://orcid.org/0000-0003-2601-2803