Expert Systems With Applications 228 (2023) 120453

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Double handed dynamic Turkish Sign Language recognition using Leap Motion with meta learning approach

Zekeriya Katılmış [a,*], Cihan Karakuzu [b]

[a] Department of Computer Engineering, Kütahya Health Sciences University, Kütahya 43020, Turkiye
[b] Department of Computer Engineering, Bilecik Seyh Edebali University, Bilecik 11210, Turkiye

A R T I C L E I N F O

Keywords:
Dynamic hand gesture recognition
Sign language
Feature fusion
Leap Motion
Meta-ELM
Constrained-ELMs

A B S T R A C T

Sign language is one of the most important communication tools for hearing impaired people. In this study, the recognition of two-handed dynamic words in Turkish Sign Language (TSL) was studied using the LMC device. As a result of the repetition of 26 dynamic words, which were determined by considering the similarities and differences between them, by 6 participants, two data sets were obtained by extracting two types of feature sets. By applying a three-stage strategy to these feature sets, word recognition performances are presented by considering many aspects. These stages are data regularization, feature selection and dimension reduction. By performing these stages, new datasets with less dimension were obtained. The recognition performance was tested with six different ELM networks and the results were compared. Five-fold cross-validation was used to test the validity of the proposed system and accuracy of the obtained results. According to the results obtained with a comprehensive analysis, it has been seen that the Meta-ELM classifier maintains its performance rate and gives the highest performance. At the same time it has been observed that the Meta-ELM classifier has a stable structure that offers less user intervention.

1. Introduction

Sign language is the most expressive way for deaf people to communicate and interact with others. It is the most popular in the deaf and hearing-impaired communities and is also the natural way of communicating with normal people. Those with hearing difficulties, often depend on sign language to participate in the real world. According to World Health Organization (WHO) statistics, the number of people with hearing difficulties is 432 million adults and 34 million children in the world and nearly 2,5 million in Turkey. The World Federation of the Deaf (WFD) states that there are approximately 300 active natural sign languages worldwide (WFD, 2018). Turkish Sign Language (TSL) is one of the sign languages with an unwritten grammar characterized by hand gestures and sometimes facial and body signs (Erten & Arıcı, 2022). TSL involves creating complex grammatical structures using dynamic word movements. Dynamic word movements are the most important building blocks during TSL sentence development and facilitate expressive communication.

Among the parts of the human body, the hand is the most effective and intuitive interaction tool in most human–computer interaction applications. Hand gesture recognition has been an active research area for the last 20 years, with various and different approaches being proposed. Hand gesture recognition is an active research area with numerous and diverse applications such as sign language recognition (Kim et al., 2013; Lu et al., 2016; Marin et al., 2016; Sohn et al., 2012; Katılmış & Karakuzu, 2021; Galván-Ruiz et al., 2023), interactive games (Rautaray & Agrawal, 2011; Lee & Hong, 2010), serious games (Avola et al., 2013; Placidi et al., 2013), emotional expression descriptors (Barrientos & Canny, 2002; Truong et al., 2016), robotics (Calinon & Billard, 2007; Goza et al., 2004), advanced computer interfaces (Ohn-Bar & Trivedi, 2014; Pierce & Pausch, 2002; Reale et al., 2011; Ren et al., 2013), educational systems, virtual reality systems and physical rehabilitation. It offers intuitive and easy-to-use interfaces for these and similar applications, provides support for the hearing impaired. Furthermore, it gives solutions for all environments that use contactless interfaces. This broad interest stems from the use of hands and fingers to communicate and interact with the physical world.

In general, approaches to hand gesture recognition can be divided into two main classes: 3D model/contact-based (Cheng et al., 2015) and vision-based (Li et al., 2013; Sharma & Singh 2023). The former uses basic elements of body parts to extract relevant 3D information, while the latter uses images or video sequences to extract main features. The

* Corresponding author.
E-mail addresses: zekeriya.katilmis@ksbu.edu.tr (Z. Katılmış), cihan.karakuzu@bilecik.edu.tr (C. Karakuzu).

https://doi.org/10.1016/j.eswa.2023.120453
Received 19 February 2023; Received in revised form 2 May 2023; Accepted 8 May 2023
Available online 11 May 2023

first group addressed detection for sign language recognition (SLR) using a contact-based system. It includes wearable sensor-based systems such as the smart glove (Oz & Leu, 2011; Huang et al., 2019), the Myo armband (Ding et al., 2018). In addition, a recognition and interaction pathway is being developed using sensors such as accelerometers (Kaya et al., 2018), magnetometers (Kim et al., 2016), gyroscopes (Mimouna et al., 2018) and electromyography (EMG) (Kosmidou et al., 2006). This approach has several advantages such as simplicity in pre-processing and computations, high accuracy in motion tracking and object detection. However, it has some limitations such as sensitive measurements and noise level in sensor signals (Nymoen et al., 2015). Furthermore, situations such as high cost, unnatural use, limitation of movement flexibility and calibration adjustment make the use of this technology difficult (Dipietro et al., 2008).

The advent of digital cameras and camera stereo gave rise to the second group, the vision-based SLR (Cui et al., 2019; Kumar et al., 2020; Kolivand et al., 2021; Poonia, 2023). The main advantage of vision-based systems is that the user does not need to use overly complex devices. In this category, systems using effective and 3D depth cameras such as Microsoft Kinect (Ganguly et al., 2020; Cardenas & Chavez, 2020), intel Realsense camera (Liao et al., 2018) and Leap Motion Controller (LMC) (Lee et al., 2020) have been developed. These provide detailed information about motion and even 3D motion data. Apart from these, there are also vision-based studies developed using a regular camera (Castro et al., 2023). These devices are increasingly used in the field of computer vision. The development of these sensors opens new opportunities for the field of hand gesture recognition. These sensor systems also have several challenges, such as sensitivity to lighting conditions, complex backgrounds and segmentation, occlusion, fixed interaction region in front of the sensor and noisy images (Almasre & Al-Nuaim, 2016; Jegham & Khalifa, 2017; Lejmi et al., 2017). In addition, they require extra computations in the preprocessing and image processing stages. Considering these devices and their limitations, the LMC sensor is accepted as the most natural, mobile, accurate and low-cost device dedicated to capturing hand gestures in an intuitive way. It is also of great interest due to several advantages such as less prone to environmental changes, low dimension of the features offered by skeletal data, and accurate knowledge of the positions of the hand joints. Hand gesture recognition is still a challenging computer vision topic. The main reason for LMC is that the hand is an object with a complex topology and has many possibilities to perform the same gesture. This device only focuses on tracking hand movements. Each point of the hand contains many features such as direction, length, distance, position, location, position and angle for finger, knuckle and joint points. It has the capacity to collect and extract skeletal features and movements from both hands simultaneously with greater precision than image sequences. LMC is also more robust and outperforms vision-based approaches against problems such as background subtraction and light variation. For this reason, LMC was preferred in our study. There are few LMC-based hand gesture recognition researches in the literature. These researches have analyzed words from sign languages such as ISL, ASL, SIBI and ArSLR.

Sign language is the ability to speak, mainly through the use of the hands, but in some cases the body, face and head. Sign language recognition (SLR) can be defined as a collaboration of multiple research areas that can include pattern matching, computer vision, natural language processing and linguistics (Al-Ahdal & Nooritawati, 2012; Cheok et al., 2019; Wadhawan & Kumar, 2021). Approaches dealing with 3D hand gesture recognition can be divided into two main categories: Static and dynamic hand gesture recognition. The former focuses only on posture, extracting hand silhouettes or hand regions of interest. The latter considers the progression of hand joint positions over time. Since it provides more information and is more suitable for interaction interfaces, dynamic hand gesture recognition using skeletal data is discussed in this study. In recent years, many approaches have been proposed to recognize dynamic hand gestures from hand skeleton sequences. Most TSL words are performed using both hands in deaf communication. The classification of dynamic sign words using one and two hands is a fundamental function for automatic sign language recognition applications. In particular, recognizing similar two-handed sign words is an important and useful research area in regards to accuracy. In this work, the recognition of two-handed dynamic words in TSL for towards the use of 3D skeleton-based features has been studied. It also contributed to the community with a new depth and skeleton-based dynamic hand gesture dataset. The original dataset obtained in this study can be downloaded from https://sites.google.com/view/zkdataset/.

In this study, two separate datasets were composed using two separate feature sets determined by us. It has been desired to investigate and compare the effects of the datasets obtained with the generated features on word recognition performance. The datasets of the feature sets were first reduced by implementing the data regularization and feature selection technique. In the second stage, their dimension was decreased by using dimension reduction techniques. Meta-ELM and Constrained Extreme Learning Machine networks and their specific learning methods were used to analyze the performance accuracy of the original dataset and other datasets formed. K-fold is a cross validation method to stabilize and to verify the performance of the recognition model. The recognition rate of the each classifier algorithm has been evaluated using the 3, 5, 7 and 9 fold cross-validation methods.

Since the results were close values, 5-fold cross validation results were used.

The overall block structure of the proposed hand gesture recognition approach contains 3 modules and 9 steps as presented in Fig. 1. The LMC module (blue colored blocks) contains pre-processing, hand tracking and frame detection and acquisition. The dataset module (green blocks) contains feature extraction, data regularization, feature selection and dimension reduction. The implementation module (yellow blocks) contains training, testing and recognition. The first step is the pre-processing step, which initializes the LMC service. The second step is the hand tracking, where the tracking layer matches the data by extracting the hand and fingers tracking data. In the third step, image frames are collected from the matched data using LMC API features. In the fourth step, the dataset is formed by extracting the specified features from the image frames captured from the LMC. In the fifth step, data with different lengths and durations are regularized to the same length and duration by applying time–frequency domain feature extraction (TFDFE) method. In the sixth step, effective and qualified features are

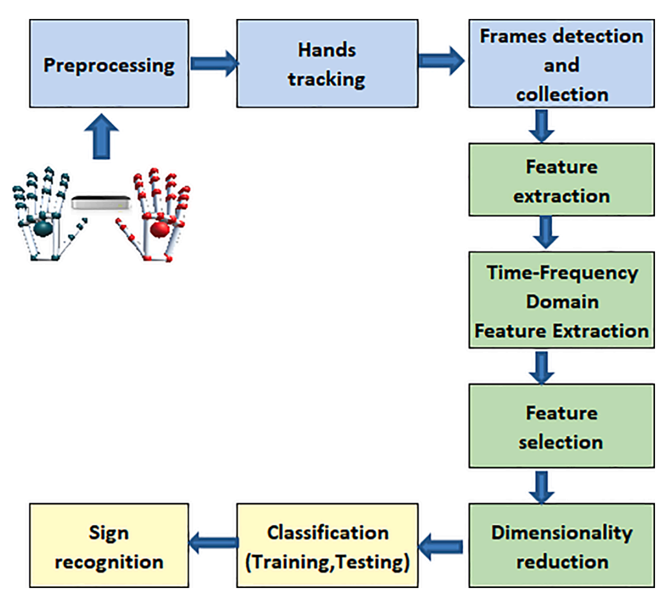

Fig. 1. Block structure of the proposed hand gesture recognition.

selected by applying the feature selection technique. In the seventh step, the selected features are decreased in dimension with the dimension reduction techniques. The eighth step is where these features are presented as a vector to the classifier and the classifier model is obtained by training and testing for word identification. The last step is sign recognition used to distinguish the hand gestures. The major contributions and novelty of this research briefly described above can be listed as follows:

a. Being the first application to use Meta-ELM and Constrained-ELMs network structures and their learning in dynamic word recognition
b. Using the time–frequency domain feature extraction (TFDFE) method in the data regularization process in TSL
c. Presenting a comprehensive datasets with 2 types of feature sets for sign language recognition in TSL
d. Using PCC feature selection, ZCA and SVD dimension reduction techniques in dynamic word recognition
e. Presenting recognition success analysis with 11 datasets formed by integrating 3-stage feature extraction, feature selection and dimension reduction approaches
f. Presenting the effect of making words in three different times as slow, medium and fast by the participants on recognition performance analysis.
g. Presenting recognition performance analysis with the best and an appropriate number of features by gradually reducing the features at the stage of feature selection and dimension reduction techniques.

2. Literature overview

There are many researches on sign language recognition (SLR) in the literature. Among these, numerous studies have been conducted with vision-based camera SLRs, of which Table 1 summarizes the most recent of these. Within these researches, LMC-based static alphabet (Chong & Lee, 2018; Yang et al., 2020; Lee et al., 2021; Enikeev & Mustafina, 2021) and dynamic sign recognition studies have been seen in the context of the subject. Especially dynamic word recognition is the scope of this study. Literature summary of dynamic sign language recognition studies is given in Table 2. In (Aliyu et al., 2017), Aliyu et al. proposed a dual LMC-based ArSLR recognition model to overcome the handicaps related to finger occlusions and missing data. For feature extraction, 17 geometric features were chosen from both sensors, while Bayesian Gaussian mixture model (GMM) and simple LDA approach was used for recognition. The data was acquired from one participant for 100 two-handed Arabic hand movements and a recognition performance of 94% was yielded. In the second research by the same authors (Deriche et al., 2019), 92% performance was achieved using 2 participants. In (Pramunanto et al., 2017), Pramunanto et al. analyzed 5 dynamic and 5 static words performed with one-handed hand gesture signs in SIBI and identified 19 features. In the research, the dataset was formed by performing 25 repeats for each sign. Using Naive Bayes method, 80% rate was achieved in the classification. In the research of Avola et al. (Avola et al., 2019), hand movements were represented by feature vector sets that change over time. Recurrent neural networks (RNNs) were trained with the angles formed by the finger bones of hand as features. The dataset contains 12 static and 12 dynamic ASL signs. The dataset contains 1200 hand movement sequences from 20 participants. Each gesture was performed by 15 males and 5 females. This proposed method achieved a performance of over 96%. In (Mittal et al., 2019), Mittal et al. proposed an LSTM model for sequential motion sequences and an SLR system that recognizes sequential connected motion sequences. The model was based on subdividing continuous signals into sub-units and modeling them with neural networks. The dataset including 36 dynamic features extracted was formed by 6 participants repeating each signal 15 times. The proposed model has been tested using 35 different isolated two-handed sign words and 942 sign sentences of Indian Sign Language (ISL). Average accuracies of 72% and 89% were yielded respectively. In Kumar et al. (Kumar et al., 2017a), a new combined multisensory structure for SLR was proposed using a combined hidden Markov model (CHMM). This framework was used to recognize dynamic isolated sign movements performed by hearing-impaired people. The best recognition rate was yielded with the CHMM, as high as 90%. For 25 dynamic one-handed sign gestures of ISL, 8 repeats were performed by 10 different participants to generate data. In this research, Kinect and LMC sensors were used in parallel to capture the motions. 66 features were identified for the dataset. In a similar study (Kumar et al., 2017b), Kumar et al. proposed a new multi-model structure for isolated SLR using sensor-based devices. LMC and Kinect sensors were used in the multi-model framework to capture palm positions and finger from two different views during movement. A feature-set is extracted from the data captured by both devices. Recognition was performed separately by sequential classifiers based on a hidden Markov model (HMM) and a bidirectional long short-term memory neural network (BLSTM-NN). The framework was tested on a dataset of 7500 ISL hand movements containing 50 different hand gestures. It has been reported that accuracies are healed if data from both devices are combined compared to single sensor based recognition. Performance rates were determined as 97% and 94% for HMM and BLSTM-NN, respectively. In (Hisham & Hamouda, 2017), Hisham et al. proposed a model for recognizing dynamic and static gestures in sign realization. The proposed system is based on two different sets of bone and palm

Table 1
Literature summary of vision-based camera SLR.

Reference.	Architecture	Sign Language	Dataset Properties	Dataset	Type of Gesture	Accuracy (%)
(Basnin et al., 2021)	Integrated CNN–LSTM	Bangla (BSL)	36 alphabets	13,400 samples	Static	88
(Venugopalan & Reghunadhan, 2021)	Hybrid GoogleNet-BiLSTM	Indian (ISL)	13 signs	932 samples	Dynamic	76
(Sharma & Singh 2021)	VGG CNN	Indian (ISL)	43 signs	2150 samples	Static	99
(Kothadiya et al., 2022)	LSTM, GRU	Indian (ISL)	16 signs	1100 samples	Dynamic	97
(Das et al., 2023)	Hybrid (VGG16 + RF)	Bangla (BSL)	10 digits	1075 samples	Static	97
	Hybrid (VGG16 + RF)		36 alphabets	1005 samples		91
(Castro et al., 2023)	Multi-stream 3D CNN	Brazilian (BSL) Indian (ISL) Korean (KSL)	56 signs 50 signs 77 signs	1200 samples 4287 samples 1540 samples	Dynamic	91 94 50
(Sharma & Singh 2023)	MobilenetV2	Indian (ISL)	18 words 10 alphabets	2150 samples	Static	99
(Poonia, 2023)	CNN + RNN PoseEST 3LayerLSTM LİST	Indian (ISL)	50 signs	4287 samples	Dynamic	65 66 75 92

Table 2
Literature summary of dynamic sign language recognition.

Reference	Sensor	Approach (Feature-set, number & language)			Hand	Performance (dataset, classifier & accuracy)		
(Aliyue et al., 2017)	Dual LMC	Hand sphere radius, hand Pitch, Roll, Yaw, fingertip positions, palm position, finger distance.	16	Arabic (ArSL)	Two-Hand	1 participant, 100 words, 10 repeats	LDA GMM	94
(Pramunanto et al., 2017)	LMC	Palm and finger angle, palm elevations, palm distance.	19	Indonesian (SIBI)	One-Hand	1 participant, 5 words, 25 repeats	NBC	80
(Kumar et al., 2017a)	LMC, Kinect	Fingertip direction, fingertip positions.	66	Indian (ISL)	One-Hand	10 participants, 25 words, 8 repeats	CHMM	90
(Kumar et al., 2017b)	LMC, Kinect	Fingertip direction, fingertip positions.	132	Indian (ISL)	One-Hand Two-Hand	10 participants, 50 words, 15 repeats	HMM BLSTM–NN	97 94
(Hisham & Hamouda, 2017)	LMC	Directional and positional features, each palm's position and phalanx's vectors, finger angles.	170	Arabic (ArSL)	One-Hand Two-Hand	3 participants, 16 words, 25 repeats	DTW SVM KNN ANN	97
		Positional and directional features, the phalanx's position vectors, finger angles.	140					96
(Deriche et al., 2019)	Dual LMC	Hand sphere radius, hand Pitch, Roll, Yaw, palm position, fingertip positions, finger distance.	16	Arabic (ArSL)	Two-Hand	2 participants, 100 words, 10 repeats	LDA GMM	92
(Avola et al., 2019)	LMC	Palm position, internal finger angles, fingertip positions.	23	American (ASL)	One-Hand Two-Hand	20 participants, 12 words, 2 repeats	RNN LSTM	96
(Mittal et al., 2019)	LMC	Palm position, fingertip positions.	36	Indian (ISL)	Two-Hand	6 participants, 35 words, 15 repeats	LSTM	89
(Bird et al., 2020)	LMC RGB	Finger tip and position distances, direction angles, hand, palm, arm, wrist and elbow positions.	90	British (BSL)	One-Hand Two-Hand	5 participants, 18 gestures, 2 repeats	DNN CNN	94
(Hisham & Hamouda, 2020)	LMC	Finger and palm position, direction, fingertips direction, finger angles, Pitch, Yaw, and Roll.	70	Arabic (ArSL)	One-Hand Two-Hand	28 participants, 40 words, 10 repeats	Ada-Boosting	93
(Katılmış & Karakuzu, 2021)	LMC	Hand, Palm, arm wrist and elbow positions, fingers extended, position distances, direction angles.	119	Turkish (TSL)	Two-Hand	4 participants, 50 words, 40 repeats	ELM ML-KELM	96 99
(Abdullahi & Chamnongthai, 2022a)	LMC	Finger joints, tips of hand, palm center, wrist center, hand velocity, orientation angles.	192	American (ASL)	Two-Hand	10 participants, 57 words, 10 repeats	FFV-Bi-LSTM	98
(Abdullahi & Chamnongthai, 2022b)	LMC	Hand shape, orientation, position and motion.	114	American (ASL)	Two-Hand	10 participants, 40 words, 10 repeats	BiLSTM	97
(Wu et al., 2022)	LMC	Position of the palm, fingers and joints, Pitch, Roll and Yaw of palm, distance between the fingertips and the palm, normal and direction of the palm.	190	Chinese (CSL)	One-Hand Two-Hand	5 participants, 20 words, 100 repeats	LSTM	99
(Galván-Ruiz et al., 2023)	LMC	Direction, position, orientation, length, width and speed.	276	Spanish (SSL)	Two-Hand	1 participants, 176 words, 30 repeats 15 participants, 50 words, 10 repeats	DTW	95 95
This work	LMC	Finger tip and palm position distances, bone and hand direction angles, and finger and hand position vectors.	124	Turkish (TSL)	Two-Hand	6 participants, 26 words, 15 repeats	Meta-ELM	93
		Finger joint position distance, direction angles and vectors.	124					92

feature-sets with common features between them. The study includes 20 dynamic and 38 static gestures of ArSL. It was reported that the KNN model achieved 99% and 98% performance rate for static and dynamic gestures respectively, while the DTW model achieved 97% and 96% performance rate respectively. Katılmış and Karakuzu proposed a two-handed dynamic TSL word recognition system. In their study (Katılmış & Karakuzu, 2021), a dataset consisting of 8000 instance with 40 repeats was formed by 4 different participants for 50 dynamic words. The number of features was gradually decreased using a feature selection technique. Then, PCA, LDA and hybrid PCA + LDA technique were applied and a total of 12 dimensionally reduced datasets were formed. The results were compared by testing with extreme learning (ELM) based classifiers. It was observed that the ML-KELM classifier maintained its performance rates and gave the best performance rates for the original and dimensionally reduced datasets. Bird et al. designed 3 models using a multimodal approach in SLR, RGB and LMC based and

RGB + LMC fused approaches. Recognition was performed with a large simultaneous dataset of 18 BSL gestures collected from multiple participants. Multimodality seems to be more successful for BSL classification. It also shows that weight transfer of the multimodal model is the most powerful approach for ASL classification (Bird et al., 2020). Abdullahi et al. presented an approach for recognizing highly correlated ASL words. The selected words belong to ASL words that are frequently used on a daily basis. The dataset was generated signer-independently using different types of participants. For complex gestures that are critical to recognize and do not require large abstraction for feature selection and encoding in traditional deep Bi-LSTM, fast fishing vector (FFV) was added. A large dynamic signal word recognition algorithm called FFV-Bi-LSTM was designed. This method was run on datasets containing a limited number of static and dynamic hand gestures, as well as on readily available published datasets, and was seen to give close results (Abdullahi & Chamnongthai, 2022a). In a second study by the same authors, a multi-stack deep BiLSTM recognition network was designed for two-handed dynamic similar and dissimilar ASL words. They presented innovative approaches in the process of tracking errors and uncertainties in each frame, controlling noise, selecting the most important features, and distinguishing between similar two-handed dynamic words in the acquisition of hand movement trajectories (Abdullahi & Chamnongthai, 2022b). Hisham et al. introduced an Arabic Sign Language (ArSL) recognition system using LMC and Latte Panda. The proposed model used two machine learning classifiers. DTW was applied to improve the accuracy of both algorithms and compared with AdaBoost. The research involves dynamic gestures that include more complex gestures variations than static gestures. The prototype was implemented on a single board (Latte Panda) to increase the reliability and mobility of the system and to ensure that it works efficiently (Hisham & Hamouda, 2021). Wu et al. proposed a dynamic Chinese Sign Language (CSL) word recognition method based on a modified LSTM model with attention mechanism. They developed a sign language human–computer interaction (HCI) synthesis system that validates the real-time performance and effectiveness of the proposed method (Wu et al., 2022). In (Galván-Ruiz et al., 2023), Galván-Ruiz et al. proposed an Spanish Sign Language (SSL) recognition system using 4 types of word groups and datasets. This study was tested using a DTW classifier with a total of 276 features and 176 words.

As can be seen from the studies given in Table 1 and briefly summarized above, this study proposes a different perspective and solution method. Briefly, the originalities of this study are as follows:

a. Meta-ELM based classifier is the first application in sign language recognition,
b. Selection of similar/close words in terms of construction in TSL,
c. Presentation of success analysis of datasets formed by integrating 3 different making rapidity,
d. Increasing the diversity of the dataset by making signs in 3 different rapidity (slow, medium and fast) by the participants.

3. The Leap Motion Controller (LMC)

The Leap Motion Controller (LMC) is an electronic device introduced by Leap Motion Company in 2013. The LMC contains many sensors (optical and infrared) that enable the detection of the pointer's hands and fingers. It can accurately track hand movement and has three infrared LEDs to capture the signer's hands in detail at 200 fps. The range of the device is from 25 to 600 mm (Mohandes et al., 2014). LMC has many advantages: Portability, cost-benefit value, plug and play via USB. The LMC service includes application programming interfaces (APIs) and SDKs that contain some algorithms and mathematical operations that interpret the captured 3D image and extract the positions of the detected objects (Weichert et al., 2013). This device is also capable of collecting movements from both hands simultaneously with an accuracy of more than 0.01 mm (LMC, 2023). Furthermore, LMC raw data is delivered as time series data, which offers important temporal information. While Kinect is used to interpret the movement of the whole body, LMC provides real-time tracking of finger and hand movements. LMC is widely used in research areas such as sign language recognition (SLR), computer interaction, medical diagnosis, robot control and rehabilitation. The interaction area and schematic view of the LMC device is given in Fig. 2.

4. Data acquisition and feature extraction

Within the scope of this research, 26 TSL words were chosen considering the differences and similarities between them. These are composed of words that are similar to each other in terms of their structures under 8 different categories. The Turkish words used in our research are given in Table 3. Turkish words are listed in the first row of the table and their English corresponding in the second row. The words were made by the participants in different rapidity (slow, medium and fast) and data records were taken. Thus, it is aimed to increase the robustness of the recognition performance rate and reliability with the diversity of the dataset. In this research, first of all, data were acquired and feature extraction was performed for 26 dynamic two-handed words in TSL using 6 participants, two adult females and males, and a girl and a boy each. Special care was taken to ensure that the age range and hand geometry of the participants were different. In this context, for each of the hand, finger, joint and joint points, 124 discriminative features have been determined separately for 2 feature sets from features such as position, distance, direction and angle. A dataset raw consisting of a total of 2340 samples was formed by 6 participants containing 15 repeats with a single frame for each gesture. The dataset contains a total of 42,347 image frames taken at different times for the words. The arrayed image of the 2 similar-words used is shown in Fig. 3.

The large variation in the data is measured using a statistical analysis method called box plot. It uses 5 data points: median, first/third quartile, lowest/highest, first/third quartile, and highest/lowest for the level of dispersion, symmetry of the data and spread. A box plot for the hand movements of 2 similar-words for 6 participants is shown in Fig. 4. These movements were captured using LMC. Each color in the plot was used to distinguish the change obtained from one signer from the others. In addition, since a sequence of hand movements is made in the air, it can also be visualized using a 2D plot showing the variation in hand movements of similar-words when made by 6 participants. The representation in Fig. 5 was acquired by plotting the palm center points (Palm Position) over the entire gesture duration. The figure is an example 2D

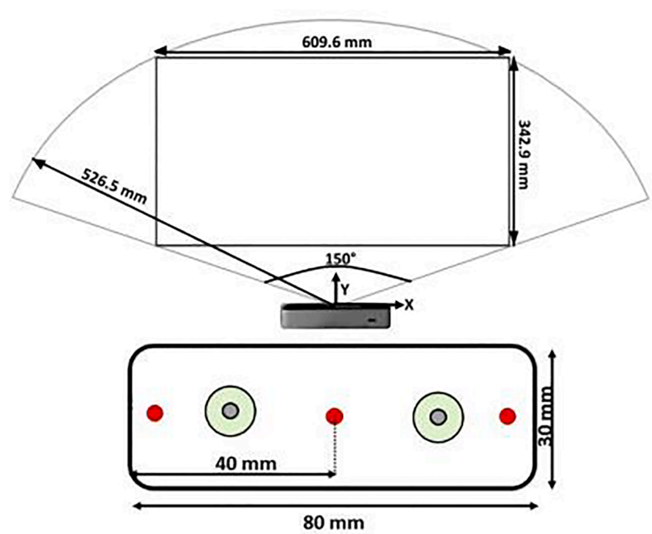

Fig. 2. LMC device: Interaction area (top) and schematic view (down) (Curiel-Razo et al., 2016).

Table 3
The dynamic words used in this study are grouped by their similarity to each other in regards to their construction.

	Words (Turkish, English)
1	Düzgün, ertelemek, karışmamak (Properly, to postpone, not to interfere)
2	Çatı, ev, komşu (Roof, house, neighbor)
3	Adım, bilgisayar, börek, sakin, terzi (Step, computer, patty, quiet, tailor)
4	Büyük, geniş, kocaman (Big, wide, huge)
5	Ağır, kabul (Heavy, acceptance)
6	Bozmak, bozuk, işaret dili (Disrupt, corrupt, sign language)
7	Yaymak, yetmek (Spread, suffice)
8	Hafif, ihmal, kaldırmak, kantar, okumak (Light, neglect, remove, scale, read)

plot for the hand movements of 2 similar-words for 6 participants. In the drawing, the same color line was used for 6 participants in order to distinguish the change (Palm Position) during the construction of the two words. In Figs. 3, 4 and 5; it is aimed to visualize the data obtained for two words (heavy/heavy, accept/accept) that are similar in structure and to reveal the diversity and relationship for the data. An illustration of hand and fingers skeleton points obtained from the LMC is shown in Fig. 6.

The implementations have been performed using a computer (Intel core i7 CPU at 2.2 GHz and 6 GB DDR4 RAM) with Matlab program. A raw data acquisition was implemented in via C# programming language using LMC SDK version 4.1 to obtain the dataset. Hand movements were collected using an LMC connected to the laptop through USB port. In the process of performing the hand movements of the participants, daylight lighting and natural environments were used without changing the lighting conditions. The conversion of the raw file to Matlab file format, normalization, feature extraction, data regularization, feature selection, dimension reduction, classification and recognition stages were performed using Matlab.

Two group of feature extraction sets (Feature Set 1, Feature Set 2) were prepared for 26 dynamic words selected from TSL. The features were chosen by considering the structure of each sign and every point needed, and divided into two equal groups. Each group consists of 3 subcategories: distance, angle and vector info. In order to compare the success impact of the feature sets, an equal number of features were chosen.

In the first feature set (Feature Set 1), the distance data among the fingertip (TipPosition) and palm (PalmPosition) position points is the first referenced. A total of 30 features were defined among these 2 position points. **Distance_set1={f1,…f30}**.

In addition to these, since finger sizes and dimensions differ, finger bone direction (Bone.Direction) and finger-hand direction (Direction) angle data was needed. Furthermore, angle data about the longitudinal

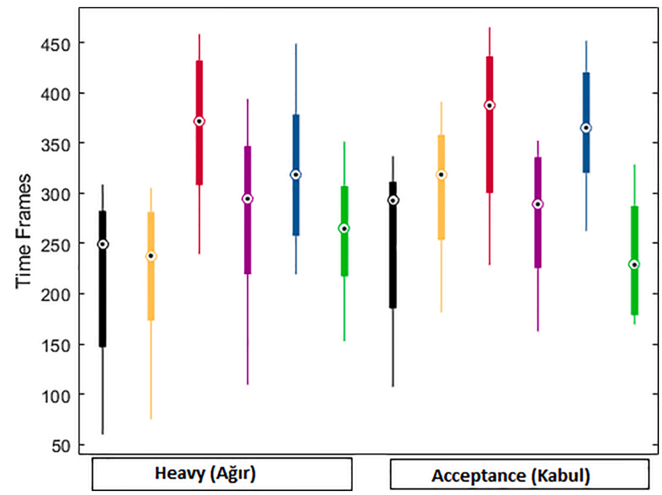

Fig. 4. Box-plots for the hand movements of two similar-words.

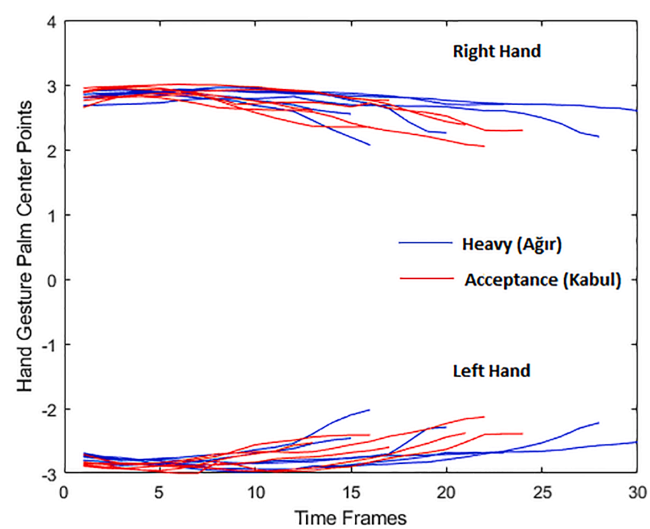

Fig. 5. 2D plot of for the hand movements of two similar-words.

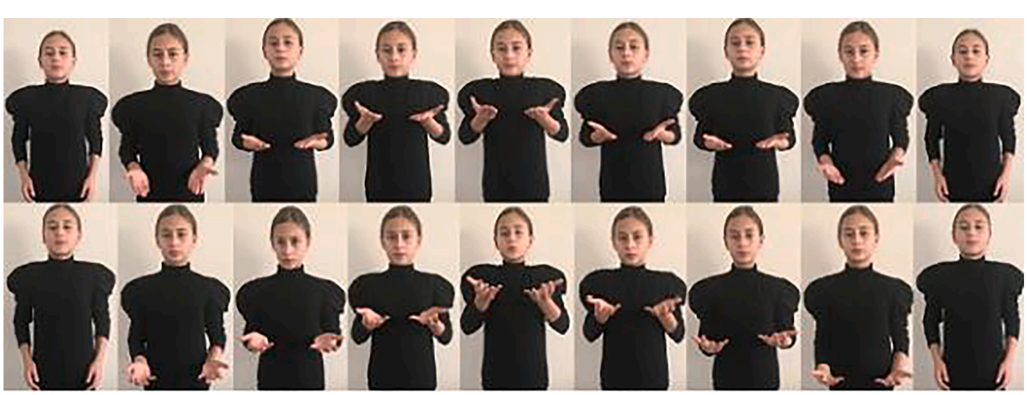

Fig. 3. Arrayed image of the 2 similar-words: “Accepted (Kabul)” (top row) and “Heavy (Ağır)” (bottom row).

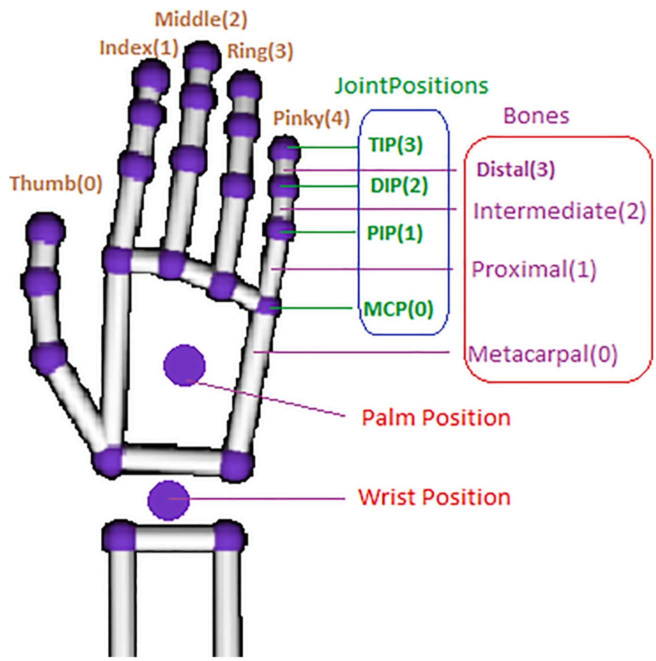

Fig. 6. Illustration of hand and fingers skeleton points obtained from the LMC (LMC, 2023).

horizontal and vertical axis (pitch, yaw, roll) directions of the hand has been also utilized. A total of 40 features were defined for the angle data in each direction. **Angle_set1= {f31,……,f70}**.

In addition, some feature vectors were utilized to enhance discrimination. These are fingertip (TipPosition) and palm (PalmPosition) position points of hands, and the vector data of hands and its horizontal axis (PalmNormal) direction. Moreover, the arm-elbow position point vector data was also utilized. A total of 54 features were utilized for each vector data.

Additionally, some feature vectors are utilized to enhance discrimination. These are the fingertip (TipPosition) and palm (PalmPosition) position points of the hand and the vector data of the horizontal axis (PalmNormal) directions of the hands. Furthermore, arm-elbow position point vector data is also utilized. A total of 54 features were utilized for hands. **Vector_set1={f71,…..,f124}**.

In the second feature set (Feature Set 2), firstly, the distance data among the fingertip (TipPosition) and joint (JointPosition) position points is referenced. A total of 24 features were defined for these position points. **Distance_set2={f1,…f24}**.

In addition to these, since finger sizes and dimensions differ, finger direction (Direction) and finger-hand horizontal and vertical axis direction (Direction) angle data was needed. Angle data about the direction of the hand (pitch, yaw, roll) has been also utilized. A total of 34 features were defined for the angle data among each direction. **Angle_set2= {f25,……,f58}**.

Additionally, some feature vectors were utilized to enhance discrimination. These are the finger and arm (Direction) direction, arm-wrist position and arm center point vector data. Furthermore, the vector data of the base points of the hands was also utilized. A total of 66 features were utilized for hands. **Vector_set2={f59,…..,f124}**.

In this research, the same number of features was defined for each feature set. The measures, units of measurement, names and numbers of the features of the feature sets are given in Table 4 and Table 5, respectively.

5. Time-Frequency domain feature extraction (TFDFE)

Time-frequency domain feature extraction (TFDFE) method was used in the data regularization phase on the obtained dataset. There are many criteria used for this feature extraction. Among these, 5 criteria has been used, namely the root of mean squares (RMS), standard deviation (STD), mean (AVG), fast Fourier transform (FFT) and discrete wavelet transform (DWT) (Risqiwati et al., 2013; Risqiwati et al., 2020; Çalışkan, 2019). In the first stage, these criteria were applied separately for each feature of each signal sample and a value was taken for each of them. In addition, the highest 2 peak values were taken for the frequency domain criteria (FFT, DWT). A new data set was formed by combining the 7 values obtained in total for each sample feature separately. The values obtained for each feature of each signal sample are as much as the number of criteria. In the second stage, this process is performed for 124 features data belonging to each feature set. As a result of this process, a (2340 × 868) sized dataset is obtained. Through this process, the size of the dataset with different lengths is regularized and at the same time reduced. At the end of this process, the dataset is reduced by approximately 2,5 times.

Table 4
Feature information extracted from hands and fingers (Feature Set 1).

Measure	Unit	Features	Number of features
Distance	mm	Fingers.(TipPositionToTipPosition)	20
		Fingers.TipPositionToHands.PalmPosition	10
Angle	radian	Fingers.(Bone.DirectionToBone.Direction)	18
		Fingers.DirectionToHands.Direction	10
		Hands.PalmNormal(Pitch,Yaw,Roll)	6
		Hands.PalmPosition(Pitch,Yaw,Roll)	6
Vector	float	Fingers.TipPosition(X,Y,Z)	30
		Hands.PalmPositions(X,Y,Z)	6
		Hands.Direction(X,Y,Z)	6
		Hands.PalmNormal(X,Y,Z)	6
		Arm.ElbowPosition(X,Y,Z)	6

Table 5
Feature information extracted from hands and fingers (Feature Set 2).

Measure	Unit	Features	Number of features
Distance	mm	Fingers.(TipPositionToJointPosition)	24
Angle	radian	Fingers.(DirectionToDirection)	8
		Fingers.DirectionToHands.PalmNormal	10
		Fingers.DirectionToHands.PalmPosition	10
		Hands.Direction(Pitch,Yaw,Roll)	6
Vector	float	Fingers.Direction(X,Y,Z)	30
		Hands.Basis(X,Y,Z).Basis(X,Y,Z)	18
		Arm.WristPosition(X,Y,Z)	6
		Arm.Direction(X,Y,Z)	6
		Arm.Center(X,Y,Z)	6

6. Feature selection

Feature selection is defined as the selection of the best, efficient and qualified features from an original dataset. It offers advantages such as resource optimization, result accuracy, execution time, and effective use in datasets with large size and high number of features. In this study, PCC technique was used.

Pearson Correlation Coefficient (PCC) is a technique used to statistically measure strength and direction of relationship between features. It is based on the covariance method. It is determined by dividing the covariance of two features by their standard deviation. With this technique, the correlation coefficient is found for each feature in the dataset and the linear dependency relationship between the features is revealed (Sun et al., 2013; Sun et al., 2023; Sadiq et al., 2013, Sadiq et al., 2021; Daru et al., 2013; Daru et al., 2021; Sugianela & Ahmad, 2020). The correlation coefficient takes a value in interval of [−1, 1]. The values obtained for the features are ranked after taking their absolute value and the ones with the highest value are selected. P represents the Pearson correlation coefficient between the features and is defined as in equation (1). The P coefficient (effect ratio) values of the 868 features belonging to the first feature set are presented in Fig. 7.

$$P = \frac{cov(x,y)}{\sigma_x \sigma_y} \tag{1}$$

7. Dimension reduction techniques

Dimension reduction technique is the process of transforming a high-dimensional dataset with a large number of features into a low-dimensional dataset. In this study, SVD and ZCA techniques were used for dimension reduction.

7.1. Singular value decomposition (SVD)

Singular value decomposition (SVD) is a technique used to extract dominant features and reduce dimension for large datasets. Matrices are decomposed using this technique to create a semantic vector space. SVD is also used both classification and dimension reduction (Li & Park, 2009; Song et al., 2013). Its dimension reduction technique is similar to PCA but more effective. In this technique, a matrix is expressed as the product of three different matrices (U, S, V). Moreover, these three matrices have certain constraints associated with them. Singular value decomposition for a matrix A is defined as in Eq. (2). Through SVD decomposition, the orthogonal matrix U and V, and the diagonal matrix S of singular values are obtained. After finding the eigenvalues, eigenvectors are formed using singular values by sorting. In this technique, a new dataset is obtained by multiplication the largest singular values (S) matrix and the orthogonal (U) matrix as in Eq. (3).

$$[U, S, V] = svd(X) \tag{2}$$

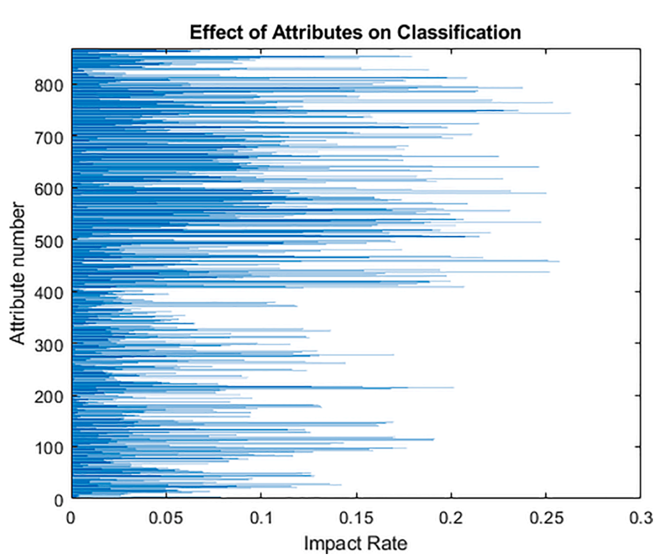

Fig. 7. P coefficients of features belonging to Feature Set 1.

$$\check{X} = U \times S \tag{3}$$

7.2. Zero-phase component analysis (ZCA)

Zero-phase component analysis (ZCA) is another dimension reduction technique utilized in many areas such as statistics and machine learning. ZCA is an effective technique to reduce the correlation between different features making the features more independent. It was proposed to ensure that all features have the same covariance (Lin et al., 2019). In the ZCA technique, for an n- instance dataset $\boldsymbol{X} = \{x_1, x_2, \ldots, x_n\}$ of dimension d, the matrix $\overline{X}$ is obtained by subtracting the mean vector. Through SVD decomposition, the orthogonal based eigenvector **U** and the diagonal eigenvalues matrix **S** are obtained as shown in Eq. (4). **ε** is fixed to a small number to avoid zeros in the diagonal matrix. Here, the eigenvectors with the highest value in the feature vector are sorted to obtain a new size-changed $\check{X}$ dataset. The zero-phase component analysis for matrix $\check{X}$ is defined as in Eq. (5).

$$[U, S] = svd\left(\overline{X}\overline{X}^T\right) \tag{4}$$

$$\check{X} = U\left(\frac{1}{\sqrt{S+\varepsilon}}\right)U^T\overline{X} \tag{5}$$

8. Meta-extreme learning Machines (Meta-ELM)

Over the last few decades, many neural network architectures have been developed. Feed-forward neural networks are among the most widely studied. The model has been extensively analyzed and applied by researchers due to its simplicity and relatively high learning and responsiveness. The learning capacity of single-layer feed-forward networks (SLFNs) is higher than that of multilayer feed-forward neural networks, as proven in (Tamura & Tateishi, 1997). Gradient-based learning, optimization-based learning, and least mean square (LMS)-based learning are three approaches used to train SLFNs.

Recently, Huang et al (Huang et al., 2006) proposed ELM, a new extremely fast learning model for SLFNs. ELM has some advantages such as no need to tune the hidden layer parameters, fast learning and good generalization performance. In ELM, it is sometimes not possible to achieve the desired level of learning performance due to the random selection of input connection weights and cell thresholds. Moreover, for large-scale problems, the high computational cost was a problem due to the repetitive use of the entire dataset. Various studies have been carried out to overcome these drawbacks and various ELM network architectures have been developed. One of them is a network architecture that combines several ELMs together, called Meta-ELM, which stands out for its generalization capability (Liao & Feng, 2014). Recent studies have shown that a group combination of ELMs achieves better generalization performance than the original ELM. Meta-ELM is a learning algorithm that combines a group of traditional ELMs. It can be seen as an ensemble of standard ELMs. Meta-ELM can also be called a hierarchical learning model. The Meta-ELM architecture is achieved by a basic configuration shown in Fig. 8, which has several basic ELMs and a meta structure that connects them (Liao & Feng, 2014; Zou et al., 2018). ELM structures are widely used in classification, regression, semi-supervised, supervised and unsupervised learning areas (Huang et al., 2014; Karakuzu, 2019; Huang et al., 2011; Bartlett, 1996).

9. Constrained - extreme learning Machines (CELMs)

Constrained-ELM is a single hidden layer feed-forward neural network based on ELM. CELMs randomly select the parameters of the hidden neurons based on the sample distribution. The main contribution of C-ELM is that it provides better discriminative feature matching by incorporating the sample distribution before the construction of the hidden layer (Zhu et al., 2015; Zhu et al., 2014). Experimental results

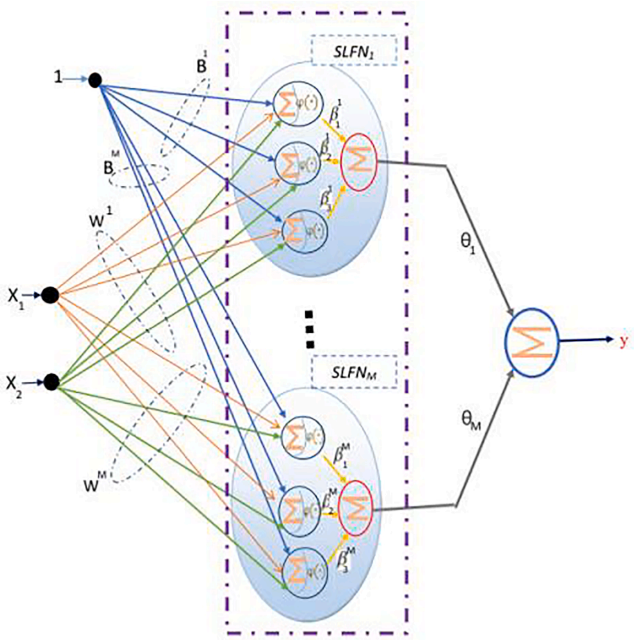

Fig. 8. Meta-ELM architecture with M SLFNs with 3 cells each (Karakuzu & Bakırcı, 2019).

show that CELMs have better generalization ability than traditional ELM and some other methods. Moreover, CELMs retain the similar fast learning characteristics of ELM. Sub-models of CELMs with different constraints have also been defined to balance the high discriminative feature learning and fast training speed of ELM. A model called DELM (Constrained Difference ELM) is presented which uses a random subset of the difference vectors of the between-class examples (Zhu et al., 2015; Zhu et al., 2014). Other CELM models has been proposed by using the sum vectors of in-class samples SCELM (Constrained Sum ELM), sample vectors of all classes SELM (Constrained Sample ELM), the sum vectors of randomly selected sample vectors RSELM (Rand Sample ELM), and the mixed vectors containing difference and sum vectors MELM Constrained Mixed ELM) (DELM + CSELM) (Zhu et al., 2015; Zhu et al., 2014).

10. Experimental results and discussion

In this study, 2 separate datasets consisting of 2340 samples and belonging to 2 different feature sets were prepared by 6 participants to evaluate the performance rates. For each of the 2 original datasets, a 3-stage process of data regularization, feature selection and dimension reduction was performed. These stages were applied separately for each original dataset or feature set. First of all, the data regularization stage is realized, by using the TFDFE method on data with different lengths due to the fact that the words were made by the participants at different rapidity. At this stage, although the number of features (inputs) increases, dimension-reduction is also performed. In the second stage, the **PCC** feature selection technique was implemented. In this stage, four different number of feature selections (250, 150, 100, 50) were made considering to determine the most effective and appropriate feature numbers. Hence, four datasets were formed using the feature selection technique. In the third stage, **ZCA** and **SVD** dimension reduction techniques were applied to the feature-selected datasets. In this process, it was desired to form the best dimensional subsets by performing 3-gradual (150, 100, 50) dimension reduction to the new datasets. Using each dimension reduction technique, 3 datasets belonging to the gradually composed feature sub-sets were formed. As a result, a total of 11 datasets were used in the classification process for the purposes of investigation. Table 6 presents the formation of 11 datasets for the 3 stages. The performance of the datasets using ELM-based Meta-ELM and CELMs classifiers was analyzed. In Meta-ELM and CELMs classifiers, neuron activation is sigmoid function and 100 neurons were used in each ELM structure. For the Meta-ELM, the number of ensembles (M) was taken as 150. These values were determined by trial-and-error analogy.

First of all, the recognition results of 5 sub-datasets obtained by applying PCC to the TFDFE applied datasets were analyzed. The recognition results and execution times of Meta-ELM and CELMs classifiers for these datasets are given in Table 7 and Table 8, respectively, for the datasets we named Feature Set 1 and Feature Set 2. The testing accuracy graphs are given in Fig. 9 and Fig. 10 respectively. In the tables, the performance analysis for each sub-dataset was labeled as “application”. The first application is for the TFDFE dataset and the other 4 applications are for the sub-datasets obtained by applying PCC. These datasets were subjected to classification algorithms. Considering the results obtained, the best performance rate for the first 2 applications was obtained with the Meta-ELM classifier. As can be seen from the tables, Meta-ELM maintained its performance rate despite the decrease in the number of features in the sub-datasets. The other 3 applications (with 150, 100, 50 features) were not used in the dimension reduction stage due to their low performance rates. Although the performance rates for the other classifiers were close in the first application, the performance rates gradually decreased in the other 4 applications (PCC applied). SELM is the classifier with the least training and testing time. Considering the execution times of the algorithms, a decrease was observed in the execution time as the number of features decreased. Except for Meta-ELM, CELMs based algorithms have been seen to have lower performance rates within the scope of PCC-based generated applications.

The second subgroup of datasets is the datasets formed using ZCA dimension reduction technique on the dataset having 250-features selected by applying PCC. The performance rates of these datasets were also analyzed. This examination consists of 3 hybrid (PCC(250) + ZCA) applications. The recognition results and execution times of Meta-ELM and CELMs classifiers for these datasets are given in Table 9 and Table 10, respectively, for each feature set. The testing accuracy graphs are given in Fig. 11 and Fig. 12, respectively. When the results of this

Table 6
The formation stages of 11 dataset.

Original dataset	First stage	Second stage	Third stage
	Feature extraction	Feature selection	Dimension reduction
(124 feature)	Time-Frequency Domain Feature Extraction (TFDFE) (868 feature)	PCC (250 feature)	ZCA (150 feature)
			ZCA (100 feature)
			ZCA (50 feature)
			SVD (150 feature)
			SVD (100 feature)
			SVD (50 feature)
		PCC (150 feature)	
		PCC (100 feature)	
		PCC (50 feature)	

Table 7
Recognition results and execution times of PCC datasets (Feature Set 1).

1.PHASE (DATA FEATURE SET 1)																				
Classifier	1. Implementation				2. Implementation				3. Implementation				4. Implementation				5. Implementation			
	TFDFE (868 features)				PCC (250 features)				PCC (150 features)				PCC (100 features)				PCC (50 features)			
	Training.		Testing.		Training.		Testing.		Training.		Testing.		Training.		Testing.		Training.		Testing.	
	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)
Meta-ELM	95,63	7,0803	92,94	0,3608	95,77	6,5200	92,89	0,1687	94,82	6,4025	90,15	0,1327	92,97	6,2529	88,63	0,1237	90,98	6,0462	87,14	0,1125
DELM	93,91	0,0239	90,43	0,0046	92,83	0,0173	89,57	0,0030	91,99	0,0094	88,42	0,0019	89,76	0,0079	86,45	0,0017	88,39	0,0075	85,51	0,0017
CSELM	94,68	0,0459	92,05	0,0044	94,35	0,0369	91,54	0,0030	92,94	0,0307	90,00	0,0019	91,05	0,0311	87,14	0,0019	88,71	0,0323	85,81	0,0019
MELM	94,26	0,0348	91,24	0,0047	93,22	0,0268	90,64	0,0030	92,30	0,0198	88,85	0,0020	90,41	0,0173	86,92	0,0018	88,87	0,0165	85,56	0,0019
RSELM	95,00	0,0372	91,71	0,0048	93,75	0,0213	90,81	0,0029	92,91	0,0140	89,70	0,0019	90,98	0,0119	87,82	0,0018	88,49	0,0104	85,17	0,0017
SELM	94,28	0,0205	92,01	0,0046	94,21	0,0140	91,50	0,0030	92,85	0,0092	89,91	0,0020	90,87	0,0078	87,78	0,0018	88,48	0,0073	85,60	0,0017

Table 8
Recognition results and execution times of PCC datasets (Feature Set 2).

1.PHASE (DATA FEATURE SET 2)																				
Classifier	1. Implementation				2. Implementation				3. Implementation				4. Implementation				5. Implementation			
	TFDFE (868 features)				PCC (250 features)				PCC (150 features)				PCC (100 features)				PCC (50 features)			
	Training.		Testing.		Training.		Testing.		Training.		Testing.		Training.		Testing.		Training.		Testing.	
	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)
Meta-ELM	95,76	7,1983	92,03	0,3346	95,07	6,2856	91,76	0,1727	93,23	6,2169	87,95	0,1392	91,96	6,1820	87,95	0,1246	87,62	6,0687	81,67	0,1171
DELM	93,44	0,0237	90,34	0,0043	92,24	0,0184	89,40	0,0029	90,57	0,0101	86,24	0,0021	89,52	0,0082	85,38	0,0018	84,24	0,0075	80,47	0,0017
CSELM	94,85	0,0465	91,41	0,0048	93,31	0,0363	89,83	0,0031	91,23	0,0322	87,18	0,0021	89,28	0,0302	85,26	0,0019	82,63	0,0325	78,50	0,0018
MELM	93,99	0,0355	90,43	0,0046	92,45	0,0277	88,80	0,0031	90,77	0,0189	87,14	0,0022	89,33	0,0180	85,90	0,0020	83,92	0,0168	80,68	0,0018
RSELM	94,46	0,0392	91,32	0,0050	93,30	0,0250	90,21	0,0032	91,50	0,0148	87,61	0,0023	90,59	0,0123	86,54	0,0019	84,89	0,0110	81,54	0,0018
SELM	94,81	0,0204	91,41	0,0044	93,15	0,0139	89,83	0,0030	91,27	0,0105	87,78	0,0023	89,11	0,0080	85,09	0,0018	82,75	0,0078	78,80	0,0018

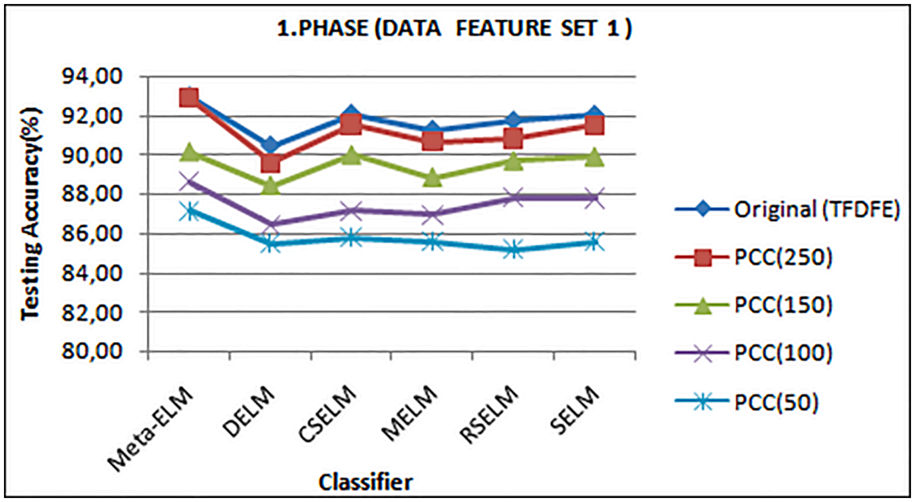

Fig. 9. Testing accuracy graphics of PCC datasets (Feature Set 1).

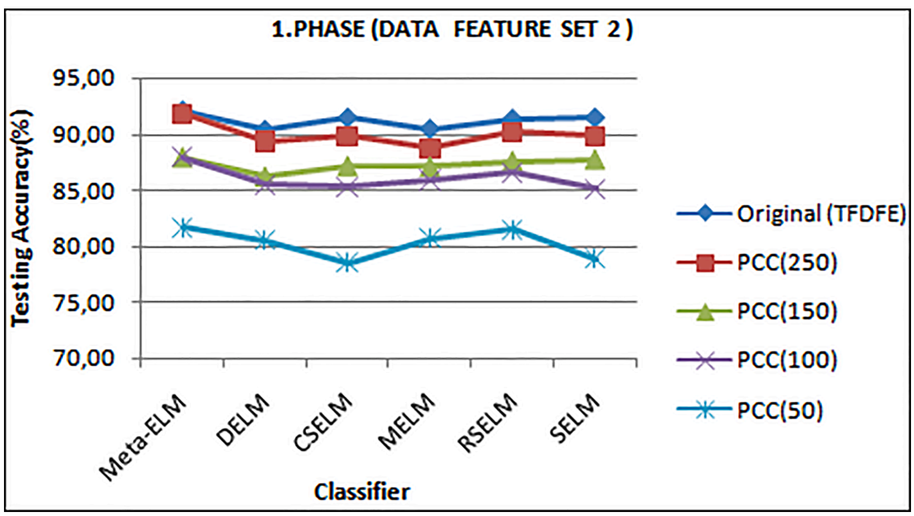

Fig. 10. Testing accuracy graphics of PCC datasets (Feature Set 2).

gradual 5 implementation are analyzed from the tables, it is seen that the best performance rate is obtained with the Meta-ELM classifier. In fact, we can say that Meta-ELM maintains its success rate even though the number of features decreases. The classifier with the least time in terms of training and testing times is again SELM. Considering the execution times of the algorithms, a gradual decrease was observed in the execution time as the number of features decreased. Except for *meta*-ELM, performance rate of CELM-based algorithms decreased in sub-datasets created by applying hybrid PCC + ZCA.

The third subgroups of datasets are the datasets formed by applying SVD dimension reduction technique to the dataset having 250-features selected by applying PCC. The performance rates of these datasets were also analyzed. This analysis also consists of 3 hybrid (PCC(250) + SVD) applications. The recognition results and execution times of Meta-ELM and CELMs classifiers for these datasets are given in Table 11 and Table 12, respectively, for each feature set. The testing accuracy graphs are given in Fig. 13 and Fig. 14, respectively. For each of the 5 cascaded applications in this analysis, the best performance rate was obtained with the Meta-ELM classifier. Again, it is seen that this classifier keeps its success accuracy despite the decrease in the number of features for these datasets. The fastest working classifier in terms of training and testing times is again SELM. Looking at the execution times of the algorithms, it has been observed that the execution time also decreased gradually as the number of features gradually decreased. Except for Meta-ELM, it has been observed that the performance rates of CELM-based algorithms have decreased within the scope of hybrid PCC + SVD-based generated applications.

Considering the above-mentioned applications it seems that among the classifiers, Meta-ELM outperforms the others in terms of performance robustness. On the other hand, the Meta-ELM classifier seems to have a more stable performance rate. However, it has the largest elapsed time for both training and testing. The reason for this difference in computation and execution time is the difference in the architecture as well as the computational style used in this architecture. The main reason for this is the training for each group during Meta-ELM training and the process of determining the output connection weights between the base ELMs and the outputs by *meta*-learning with the whole dataset after the base ELMs have been trained. Therefore, relatively large dimensional matrix operations take a long time. However, it is clear that the computation structure mentioned briefly brings a more robust and stable generalization ability compared to other structures. On the other hand, it has been observed that CELMs classifiers cannot maintain their performance rates and show a decrease. Besides, due to the different structure of the algorithms, there have been differences in execution time. On the basis of all applications, the fastest operating classifier is SELM. In addition, when all applications are analyzed, it is seen that the

Table 9
Recognition results rates and execution times of PCC(250) + ZCA datasets (Feature Set 1).

2.PHASE (DATA FEATURE SET 1)																				
Classifier	1. Implementation				2. Implementation				3. Implementation				4. Implementation				5. Implementation			
	TFDFE (868 features)				PCC (250 features)				PCC(250) + ZCA (150 features)				PCC(250) + ZCA (100 features)				PCC(250) + ZCA (50 features)			
	Training.		Testing.		Training.		Testing.		Training.		Testing.		Training.		Testing.		Training.		Testing.	
	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)
Meta-ELM	95,63	7,0803	92,94	0,3608	95,77	6,5200	92,89	0,1687	95,89	6,3813	92,86	0,1737	95,80	6,6047	93,03	0,1672	94,16	6,4763	92,75	0,1504
DELM	93,91	0,0239	90,43	0,0046	92,83	0,0173	89,57	0,0030	92,81	0,0168	89,74	0,0030	92,78	0,0118	90,09	0,0024	92,14	0,0101	89,40	0,0023
CSELM	94,68	0,0459	92,05	0,0044	94,35	0,0369	91,54	0,0030	92,53	0,0378	89,70	0,0031	92,45	0,0345	90,00	0,0023	92,18	0,0340	89,66	0,0022
MELM	94,26	0,0348	91,24	0,0047	93,22	0,0268	90,64	0,0030	92,40	0,0279	89,70	0,0032	92,65	0,0208	89,96	0,0024	92,28	0,0189	88,97	0,0022
RSELM	95,00	0,0372	91,71	0,0048	93,75	0,0213	90,81	0,0029	94,20	0,0217	90,94	0,0030	94,22	0,0171	90,34	0,0024	92,30	0,0147	89,49	0,0021
SELM	94,28	0,0205	92,01	0,0046	94,21	0,0140	91,50	0,0030	92,20	0,0140	90,21	0,0032	92,62	0,0107	89,57	0,0024	91,92	0,0092	89,19	0,0022

Table 10
Recognition results rates and execution times of PCC(250) + ZCA datasets (Feature Set 2).

2.PHASE (DATA FEATURE SET 2)																				
Classifier	1. Implementation				2. Implementation				3. Implementation				4. Implementation				5. Implementation			
	TFDFE (868 features)				PCC (250 features)				PCC(250) + ZCA (150 features)				PCC(250) + ZCA (100 features)				PCC(250) + ZCA (50 features)			
	Training.		Testing.		Training.		Testing.		Training.		Testing.		Training.		Testing.		Training.		Testing.	
	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)
Meta-ELM	95,76	7,1983	92,03	0,3346	95,07	6,2856	91,76	0,1727	95,01	6,4137	90,98	0,1555	95,42	6,4989	92,15	0,1622	94,49	6,5299	90,93	0,1652
DELM	93,44	0,0237	90,34	0,0043	92,24	0,0184	89,40	0,0029	92,12	0,0159	88,85	0,0028	91,91	0,0116	88,38	0,0024	91,04	0,0099	88,46	0,0021
CSELM	94,85	0,0465	91,41	0,0048	93,31	0,0363	89,83	0,0031	92,32	0,0338	89,06	0,0027	91,92	0,0376	89,44	0,0024	92,22	0,0331	89,32	0,0024
MELM	93,99	0,0355	90,43	0,0046	92,45	0,0277	88,80	0,0031	92,83	0,0254	88,80	0,0028	92,66	0,0213	88,89	0,0023	92,04	0,0199	88,68	0,0022
RSELM	94,46	0,0392	91,32	0,0050	93,30	0,0250	90,21	0,0032	93,89	0,0195	89,74	0,0028	93,59	0,0174	89,57	0,0024	91,00	0,0166	88,12	0,0023
SELM	94,81	0,0204	91,41	0,0044	93,15	0,0139	89,83	0,0030	92,03	0,0134	88,93	0,0030	92,13	0,0107	88,72	0,0024	91,85	0,0109	88,76	0,0024

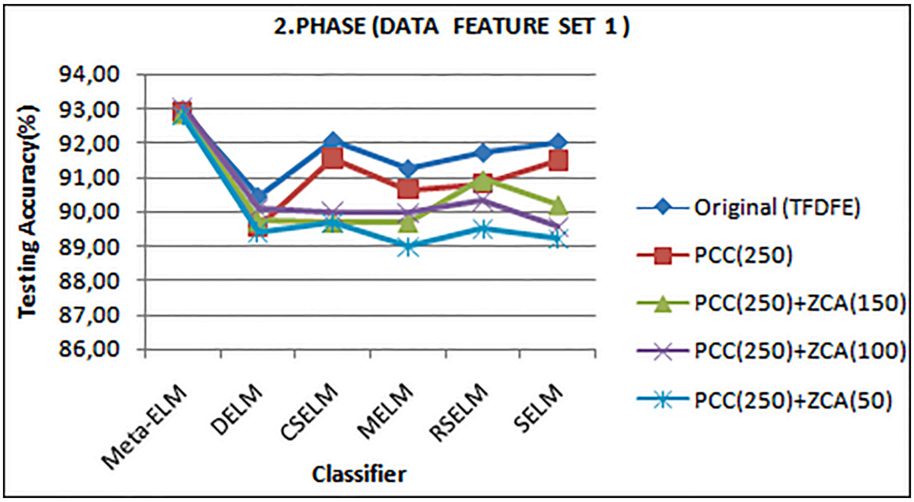

Fig. 11. Testing accuracy graphics of PCC(250) + ZCA datasets (Feature Set 1).

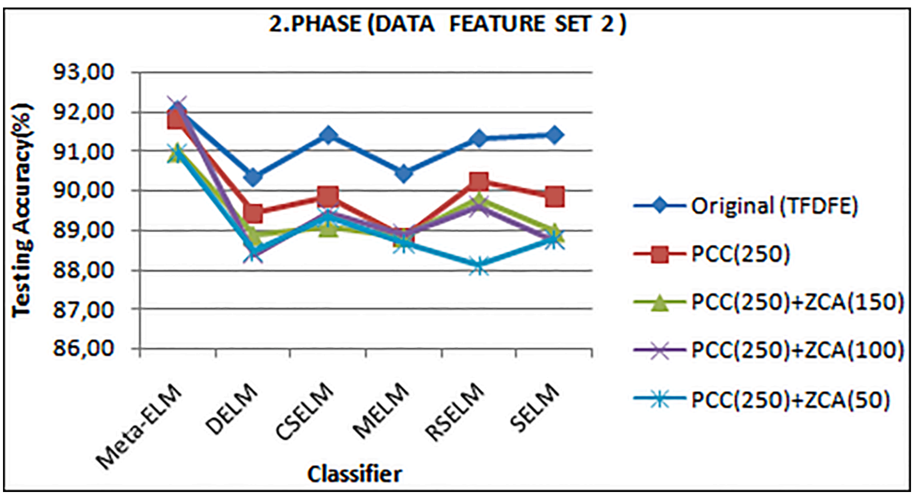

Fig. 12. Testing accuracy graphics of PCC(250) + ZCA datasets (Feature Set 2).

first feature set (Feature Set 1) is more advantageous than the other in terms of performance rate. When we compare the datasets obtained by applying the PCC + ZCA and PCC + SVD hybrids in terms of classification success, number of features and execution times, hybrid feature selection and dimension reduction, it is seen that the classifier algorithms are successful. For this reason, it can be said that a balanced feature distribution has been made for the two feature sets. Moreover, it can be said that the success rates of CELMs classifiers decrease in dimensionally reduced (hybrid applications) datasets. We can say that the RSELM network provides the highest performance rate among CELMs classifiers.

11. Conclusions and future work

In this research, an efficient dynamic gesture acquisition approach has been proposed by using sequential data captured by LMC, which is preferred over similar sensors due to its superiority in many ways, in order to approach the goal of people communicating more easily through sign language. To test the performance of the proposed approach, two original datasets/feature sets with two different feature sets were formed for 26 dynamic words with different lengths and durations according to their chosen structures. The word recognition performance of the sub-datasets formed from these feature sets by data regularization, feature selection and dimension reduction techniques were handled in many ways. Datasets which are obtained performing feature selection, dimension reduction, and their hybrid application offer advantages in terms of computation, performance, cost, storage and time. In this study, a high performance rate was obtained from the datasets formed with the developed strategy. This proposed approach is robust as it efficiently evaluates the sign inputs for similar and differently structured words using the presented strategy in this study. Recognition was performed using the Meta-ELM approach, where the best accuracy of approximately 93% and 92% was recorded for the 2 feature sets, respectively. Considering that the stable generalization ability of this network will contribute to improvement in the field of sign language recognition, it is recommended to be used in studies in this field. The specified feature set datasets and the datasets obtained from these sets with the described methods can be easily adapted to other sign languages.

Since the LMC device and its API software are new and open to technologically development, it is thought that the limitations and problems occurred especially in perception of letters performed with two-handed can be solved and better results can be obtained with the development of the device in the future. Our future research will focus on developing real-time SLR and an effective sign language human–computer interaction system.

Table 11
Recognition results and execution times of PCC(250) + SVD datasets (Feature Set 1).

3.PHASE (DATA FEATURE SET 1)																				
Classifier	1. Implementation				2. Implementation				3. Implementation				4. Implementation				5. Implementation			
	TFDFE (868 features)				PCC (250 features)				PCC(250) + SVD (150 features)				PCC(250) + SVD (100 features)				PCC(250) + SVD (50 features)			
	Training.		Testing.		Training.		Testing.		Training.		Testing.		Training.		Testing.		Training.		Testing.	
	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)
Meta-ELM	95,63	7,0803	92,94	0,3608	95,77	6,5200	92,89	0,1687	96,10	6,2254	93,08	0,1472	95,91	6,1681	93,03	0,1374	94,48	6,1082	92,70	0,1213
DELM	93,91	0,0239	90,43	0,0046	92,83	0,0173	89,57	0,0030	92,70	0,0156	90,90	0,0027	92,91	0,0093	90,38	0,0020	92,02	0,0073	88,38	0,0017
CSELM	94,68	0,0459	92,05	0,0044	94,35	0,0369	91,54	0,0030	94,19	0,0350	91,58	0,0026	93,97	0,0313	91,50	0,0020	92,22	0,0328	89,19	0,0019
MELM	94,26	0,0348	91,24	0,0047	93,22	0,0268	90,64	0,0030	93,08	0,0264	90,26	0,0027	93,26	0,0218	90,98	0,0021	92,66	0,0176	89,96	0,0018
RSELM	95,00	0,0372	91,71	0,0048	93,75	0,0213	90,81	0,0029	94,01	0,0183	90,73	0,0027	93,88	0,0126	91,11	0,0020	92,60	0,0103	89,53	0,0017
SELM	94,28	0,0205	92,01	0,0046	94,21	0,0140	91,50	0,0030	94,39	0,0126	91,54	0,0027	94,23	0,0089	91,75	0,0020	91,98	0,0072	89,32	0,0017

Table 12
Recognition results and execution times of PCC(250) + SVD datasets (Feature Set 2).

3.PHASE (DATA FEATURE SET 2)																				
Classifier	1. Implementation				2. Implementation				3. Implementation				4. Implementation				5. Implementation			
	TFDFE (868 features)				PCC (250 features)				PCC(250) + SVD (150 features)				PCC(250) + SVD (100 features)				PCC(250) + SVD (50 features)			
	Training.		Testing.		Training.		Testing.		Training.		Testing.		Training.		Testing.		Training.		Testing.	
	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)	Results (%)	Runtime (s)
Meta-ELM	95,76	7,1983	92,03	0,3346	95,07	6,2856	91,76	0,1727	95,14	6,2191	91,62	0,1582	95,30	6,1190	91,67	0,1435	94,16	6,0347	91,32	0,1250
DELM	93,44	0,0237	90,34	0,0043	92,24	0,0184	89,40	0,0029	92,18	0,0152	88,72	0,0026	92,47	0,0102	88,80	0,0020	91,21	0,0074	88,16	0,0017
CSELM	94,85	0,0465	91,41	0,0048	93,31	0,0363	89,83	0,0031	93,28	0,0368	89,62	0,0026	93,30	0,0317	90,30	0,0019	91,23	0,0303	88,33	0,0018
MELM	93,99	0,0355	90,43	0,0046	92,45	0,0277	88,80	0,0031	92,26	0,0256	88,38	0,0027	92,87	0,0179	89,10	0,0019	91,93	0,0166	88,97	0,0017
RSELM	94,46	0,0392	91,32	0,0050	93,30	0,0250	90,21	0,0032	93,26	0,0185	89,91	0,0026	93,38	0,0128	89,53	0,0019	91,83	0,0103	89,36	0,0017
SELM	94,81	0,0204	91,41	0,0044	93,15	0,0139	89,83	0,0030	93,04	0,0133	89,66	0,0029	93,56	0,0091	90,00	0,0020	91,10	0,0080	88,21	0,0018

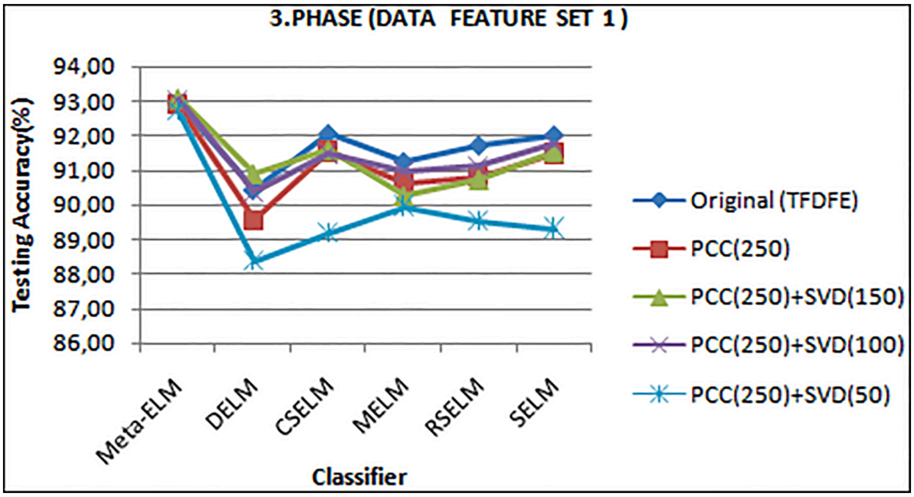

Fig. 13. Testing accuracy graphics of PCC(250) + SVD datasets (Feature Set 1).

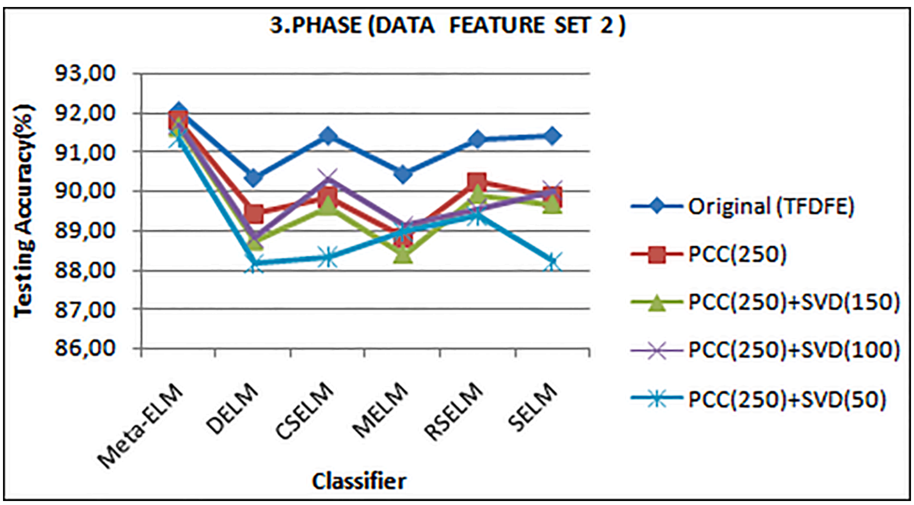

Fig. 14. Testing accuracy graphics of PCC(250) + SVD datasets (Feature Set 2).

CRediT authorship contribution statement

Zekeriya Katılmış: Conceptualization, Methodology, Software, Data curation, Formal analysis, Writing – original draft, Investigation, Visualization. **Cihan Karakuzu:** Conceptualization, Methodology, Supervision, Software, Writing – review & editing, Validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The link of the data used in the study is shared in the paper.

References

Abdullahi, S. B., &Chamnongthai, K. (2022a). American Sign Language Words Recognition Using Spatio-Temporal Prosodic and Angle Features: A Sequential Learning Approach. volume 10, page 15911-15923. https://doi.org/10.1109/access.2022.3148132.

Abdullahi, S. B., &Chamnongthai, K. (2022b). ASL Words Recognition of Skeletal Videos Using Processed Video Driven Multi-Stacked Deep LSTM. Sensors, 22, 1406. page 1-28. https://doi.org/10.3390/s22041406.

Al-Ahdal, M. E., & Nooritawati, M. T. (2012). Review in sign language recognition systems. In *2012 IEEE Symposium on Computers & Informatics (ISCI)* (pp. 52–57). IEEE.

Aliyu, S., Mohandes, M., & Deriche, M. (2017). Dual LMCs Fusion for Recognition of Isolated Arabic Sign Language Words. In *14th International Multi-Conference on Systems, Signals & Devices (SSD)*. https://doi.org/10.1109/SSD.2017.8167010

Almasre, M. A., & Al-Nuaim, H. (2016). A real-time letter recognition model for Arabic sign language using kinect and Leap Motion controller v2. *International Journal of Advanced Engineering, Management and Science, 2*(5), 514–523, 239469.

Avola, D., Bernardi, M., & Cinque, L. (2019). Exploiting recurrent neural networks and LMC for the recognition of sign language and semaphoric hand gestures. *IEEE Transactions on multimedia, 21*(1). https://doi.org/10.1109/TMM.2018.2856094

Avola, D., Spezialetti, M., & Placidi, G. (2013). Design of an efficient framework for fast prototyping of customized human–computer interfaces and virtual environments for rehabilitation. *Computer Methods and Programs in Biomedicine, 110*(3), 490–502.

Barrientos, F. A., & Canny, J. F. (2002). Cursive: Controlling expressive avatar gesture using pen gesture. In *Proceedings of the 4th international Conference on Collaborative Virtual Environments* (pp. 113–119).

Bartlett, P. (1996). The sample complexity of pattern classification with neural networks: The size of the weights is more important than the size of the network. *IEEE Transactions on Information Theory, 44*(2), 525–536.

Basnin, N., Nahar, L., & Hossain, M. S. (2021). An integrated CNN-LSTM model for Bangla lexical sign language recognition. In *Proceedings of international conference on trends in computational and cognitive engineering* (pp. 695–707). Springer.

Bird, J. J., Ekart, A., & Faria, D. R. (2020). British Sign Language Recognition via Late Fusion of Computer Vision and Leap Motion with Transfer Learning to American Sign Language. *MDPI, Sensors, 20*, 5151. https://doi.org/10.3390/s20185151

Calinon, S., & Billard, A. (2007). Incremental learning of gestures by imitation in a humanoid robot. In *Proceedings of the ACM/IEEE international conference on Human-robot interaction* (pp. 255–262).
Çalışkan, A. (2019). Improvement Of The Hybrid Feature Extraction Method For EMG Signals. *Omer Halisdemir University Journal of Engineering Sciences, 8*(2), 652–664.
Cardenas, E. J. E., & Chavez, G. C. (2020). Multimodal hand gesture recognition combining temporal and pose information based on CNN descriptors and histogram of cumulative magnitudes. *Journal of Visual Communication and Image Representation, 71*, Article 102772. https://doi.org/10.1016/j.jvcir.2020.102772
Castro, G. Z., Guerra, R. R., & Guimarães, F. G. (2023). Automatic translation of sign language with multi-stream 3D CNN and generation of artificial depth maps. *Expert Systems with Applications, 215*, Article 119394.
Cheng, H., Yang, L., & Liu, Z. (2015). Survey on 3D hand gesture recognition. *IEEE Transactions on Circuits and Systems for Video Technology, 26*(9), 1659–1673.
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.
Chong, T. W., & Lee, B. G. (2018). American sign language recognition using leap motion controller with machine learning approach. *Sensors, 18*(10), 3554. https://doi.org/10.3390/s18103554
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.
Curiel-Razo, Y. I., Icasio-Hernández, O., Sepúlveda-Cervantes, G., Hurtado-Ramos, J. B., & González-Barbosa, J. J. (2016). Leap Motion Controller three dimensional verification and polynomial correction. Measurement, 93, 258-264.https://doi.org/10.1016/j.measurement.2016.07.017.
Daru, A. F., Hanif, M. B., & Widodo, E. (2021). Improving neural network performance with feature selection using Pearson correlation method for diabetes disease detection. *JUITA: JurnalInformatika, 9*(1), 123–130. https://doi.org/10.30595/juita.v9i1.9941
Das, S., Imtiaz, M. S., Neom, N. H., Siddique, N., & Wang, H. (2023). A hybrid approach for Bangla sign language recognition using deep transfer learning model with random forest classifier. *Expert Systems with Applications, 213*, Article 118914.
Deriche, M., Aliyu, S. O., &Mohandes, M. (2019). An intelligent Arabic Sign language recognition system using a pair of LMCs with GMM based classification. IEEE Sensors Journal, 19(18), 8067–8078. https://doi.org/10.1109/JSEN.736110.1109/JSEN.2019.2917525.
Ding, J., Lin, R. Z., & Lin, Z. Y. (2018). Service robot system with integration of wearable Myo armband for specialized hand gesture human–computer interfaces for people with disabilities with mobility problems. *Computers & Electrical Engineering, 69*, 815–827. https://doi.org/10.1016/j.compeleceng.2018.02.041
Dipietro, L., Sabatini, A. M., & Dario, P. (2008). *A survey of glove-based systems and their applications., 38*(4), 461–482. https://doi.org/10.1109/tsmcc.2008.923862. IEEE transactions on systems, man, and cybernetics, part c (applications and reviews.
Enikeev, D., & Mustafina, S. (2021). Russian Fingerspelling Recognition Using Leap Motion Controller. In *2021 3rd International Conference on Control Systems, Mathematical Modeling* (pp. 604–606). IEEE.
Erten, H., & Arıcı, N. (2022). Historical Headway of Sign Language and Turkish Sign Language. *AfyonKocatepe University Journal of Social Sciences, 24*(1), 1–14.
Galván-Ruiz, J., Travieso-González, C. M., Pinan-Roescher, A., & Alonso-Hernández, J. B. (2023). Robust Identification System for Spanish Sign Language Based on Three-Dimensional Frame Information. *Sensors, 23*(1), 481.
Ganguly, B., Vishwakarma, P., & Biswas, S. (2020). Kinect Sensor Based Single Person Hand Gesture Recognition for Man–Machine Interaction. In Computational Advancement in Communication Circuits and Systems: Proceedings of ICCACCS 2018 (pp. 139-144), Springer Singapore. https://doi.org/10.1007/978-981-13-8687-9_13.
Goza, S. M., Ambrose, R. O., Diftler, M. A., & Spain, I. M. (2004). Telepresence control of the NASA/DARPA robonaut on a mobility platform. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 623–629).
Hisham, B., & Hamouda, A. (2017). Arabic static and dynamic gestures recognition using leap motion. *Journal of Computer Science, 13*(8), 337–354. https://doi.org/10.3844/jcssp.2017.337.354
Hisham, B., &Hamouda, A. (2021). Arabic sign language recognition using Ada-Boosting based on a LMC. International Journal of Information Technology 13(3):1221–1234, Springer. https://doi.org/10.1007/s41870-020-00518-5.
Huang, G. B., Zhou, H., Ding, X., & Zhang, R. (2011). Extreme learning machine for regression and multiclass classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 42*(2), 513–529.
Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2006). Extreme learning machine: Theory and applications. *Neurocomputing, 70*(1–3), 489–501.
Huang, G., Song, S., Gupta, J. N., & Wu, C. (2014). Semi-supervised and unsupervised extreme learning machines. *IEEE transactions on cybernetics, 44*(12), 2405–2417.
Huang, X. A., Wang, Q., Zang, S., Wan, J., Yang, G., Huang, Y., & Ren, X. (2019). Tracing the motion of finger joints for gesture recognition via sewing RGO-coated fibers onto a textile glove. *IEEE sensors journal, 19*(20), 9504–9511. https://doi.org/10.1109/JSEN.2019.2924797
Jegham, I., & Khalifa, A. B. (2017). Pedestrian detection in poor weather conditions using moving camera. In *2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)* (pp. 358–362). IEEE. https://doi.org/10.1109/AICCSA.2017.35.
Karakuzu, C. (2019). Performance Comparison of a Neural Network and a Fuzzy Network Trained by ELM for Dynamic System Identification Problems. In Proceeding Book of 2nd International Congress on Engineering and Architecture (ENAR), 22-24 April Marmaris, Turkey, (Vol. 22, No. 24, pp. 146-147).
Karakuzu, C., & Bakırcı, A. (2019). Dynamic System Identification Based on an Ensemble of ELMs. In Proceeding Book of 2nd International Congress on Engineering and Architecture (ENAR)22-24 April Marmaris, Turkey, pp. 1455-1463.
Katılmış, Z., & Karakuzu, C. (2021). ELMBased Two-Handed Dynamic Turkish Sign Language (TSL) Word Recognition. *Expert Systems With Applications (ESWA), 182*, 1–12. https://doi.org/10.1016/j.eswa.2021.115213
Kaya, F., Tuncer, A. F., & Yildiz, Ş. K. (2018). Detection of the Turkish sign language alphabet with strain sensor based data glove. In *2018 26th Signal Processing and Communications Applications Conference (SIU)* (pp. 1–4). IEEE.
Kim, K. W., Lee, M. S., Soon, B. R., Ryu, M. H., & Kim, J. N. (2016). Recognition of sign language with an inertial sensor-based data glove. *Technology and Health Care, 24* (s1), S223–S230. https://doi.org/10.3233/THC-151078
Kim, T., Shakhnarovich, G., &Livescu, K. (2013). Fingerspelling recognition with semi-Markov conditional random fields. In Proceedings of the IEEE international conference on computer vision (pp. 1521-1528).
Kolivand, H., Joudaki, S., Sunar, M. S., & Tully, D. (2021). A new framework for sign language alphabet hand posture recognition using geometrical features through artificial neural network (part 1). *Neural Computing and Applications, 33*, 4945–4963.
Kosmidou, V. E., Hadjileontiadis, L. J., &Panas, S. M. (2006). Evaluation of surface EMG features for the recognition of American Sign Language gestures, Proceedings of the 28th IEEE EMBS Annual International Conference New York City, USA. https://doi.org/10.1109/IEMBS.2006.259428.
Kothadiya, D., Bhatt, C., Sapariya, K., Patel, K., Gil-González, A. B., & Corchado, J. M. (2022). Deepsign: Sign language detection and recognition using deep learning. *Electronics, 11*(11), 1780.
Kumar, E. K., Kishore, P. V. V., Kumar, M. T. K., & Kumar, D. A. (2020). 3D sign language recognition with joint distance and angular coded color topographical descriptor on a 2–stream CNN. *Neurocomputing, 372*, 40–54.
Kumar, P., Gauba, H., Roy, P. P., & Dogra, D. P. (2017a). A multimodal framework for sensor based sign language recognition. *Neurocomputing*, 1–18. https://doi.org/10.1016/j.neucom.2016.08.132
Kumar, P., Gauba, H., Roy, P. P., & Dogra, D. P. (2017b). Coupled HMM-based multisensor data fusion for sign language recognition. *Pattern Recognition Letters, 86*, 1–8. https://doi.org/10.1016/j.patrec.2016.12.004
Lee, A. R., Cho, Y., Jin, S., & Kim, N. (2020). Enhancement of surgical hand gesture recognition using a capsule network for a contactless interface in the operating room. *Computer Methods and Programs in Biomedicine, 190*, Article 105385. https://doi.org/10.1016/j.cmpb.2020.105385
Lee, C. K., Ng, K. K., Chen, C. H., Lau, H. C., Chung, S. Y., & Tsoi, T. (2021). American sign language recognition and training method with recurrent neural network. *Expert Systems with Applications, 167*, Article 114403. https://doi.org/10.1016/j.eswa.2020.114403
Lee, D. H., & Hong, K. S. (2010). Game interface using hand gesture recognition. In *5th International Conference on Computer Sciences and Convergence Information Technology* (pp. 1092–1097). IEEE.
Lejmi, W., Mahjoub, M. A., &Khalifa, A. B. (2017,). Event detection in video sequences: Challenges and perspectives. In 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD) (pp. 682-690), IEEE. https://doi.org/10.1109/FSKD.2017.8393354.
Li, C. H., & Park, S. C. (2009). An efficient document classification model using an improved back propagation neural network and singular value decomposition. *Expert Systems with Applications, 36*(2), 3208–3215. https://doi.org/10.1016/j.eswa.2008.01.014
Li, X., Hu, W., Shen, C., Zhang, Z., Dick, A., & Hengel, A. V. D. (2013). A survey of appearance models in visual object tracking. *ACM transactions on Intelligent Systems and Technology (TIST), 4*(4), 1–48.
Liao, B., Li, J., Ju, Z., & Ouyang, G. (2018). Hand gesture recognition with generalized hough transform and DC-CNN using Realsense. In *2018 Eighth International Conference on Information Science and Technology (ICIST)* (pp. 84–90). IEEE. https://doi.org/10.1109/ICIST.2018.8426125.
Liao, S., & Feng, C. (2014). Meta-ELM: ELM with ELM hidden nodes. *Neurocomputing, 128*, 81–87.
Lin, J., Zhang, L., Li, J., & Zhou, W. (2019). A feature domain space transfer method for improving identification of maize haploid seed based on near-infrared spectroscopy. In *2019 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS)* (pp. 238–241). IEEE. https://doi.org/10.1109/HPBDIS.2019.8735481.
LMC. (2023). Leap Motion Controller C# SDK Documentation. Retrieved from https://developer-archive.leapmotion.com/documentation/csharp/index.html. Accessed January 11, 2023.
Lu, W., Tong, Z., & Chu, J. (2016). Dynamic hand gesture recognition with leap motion controller. *IEEE Signal Processing Letters, 23*(9), 1188–1192.
Marin, G., Dominio, F., & Zanuttigh, P. (2016). Hand gesture recognition with jointly calibrated leap motion and depth sensor. *Multimedia Tools and Applications, 75*(22), 14991–15015.
Mimouna, A., Khalifa, A. B., & Amara, N. E. B. (2018). Human action recognition using triaxial accelerometer data: Selective approach. In *In 2018 15th International Multi-Conference on Systems, Signals & Devices (SSD)* (pp. 491–496). IEEE. https://doi.org/10.1109/SSD.2018.8570429.
A. Mittal P. Kumar P.P. Roy R. Balasubramanian B.B. Chaudhuri A modified LSTM model for continuous sign language recognition using leap motion IEEE Sensors Journal 19 16 2019 7056 7063 10.1109/JSEN.736110.1109/JSEN.2019.2909837.
Mohandes, M., Aliyu, S., & Deriche, M. (2014). Arabic sign language recognition using the leap motion controller. In *2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE)* (pp. 960–965). IEEE.
Nymoen, K., Haugen, M. R., & Jensenius, A. R. (2015). Mumyo–evaluating and exploring the myo armband for musical interaction: Proceedings of the International

Conference on New Interfaces For Musical Expression, 2015, pp. 1–4, https://doi.org/10.5555/2993778.2993834.

Ohn-Bar, E., & Trivedi, M. M. (2014). Hand gesture recognition in real time for automotive interfaces: A multimodal vision-based approach and evaluations. *IEEE Transactions on Intelligent Transportation Systems, 15*(6), 2368–2377.

Oz, C., & Leu, M. C. (2011). American Sign Language word recognition with a sensory glove using artificial neural networks. *Engineering Applications of Artificial Intelligence, 24*(7), 1204–1213. https://doi.org/10.1016/j.engappai.2011.06.015

Pierce, J. S., & Pausch, R. (2002). Comparing voodoo dolls and HOMER: exploring the importance of feedback in virtual environments. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 105-112).

Placidi, G., Avola, D., Iacoviello, D., & Cinque, L. (2013). Overall design and implementation of the virtual glove. *Computers in biology and medicine, 43*(11), 1927–1940.

Poonia, R. C. (2023). LiST: A Lightweight Framework for Continuous Indian Sign Language Translation. *Information, 14*(2), 79.

Pramunanto, E., Sumpeno, S. & Legowo, R. S. (2017). Classification of Hand Gesture in Indonesian Sign Language System using Naive Bayes. International Seminar on Sensor, Instrumentation, Measurement and Metrology (ISSIMM), Surabaya, Indonesia. https://doi.org/10.1109/ISSIMM.2017.8124288.

Rautaray, S. S., & Agrawal, A. (2011). Interaction with virtual game through hand gesture recognition. In *2011 International Conference on Multimedia, Signal Processing and Communication Technologies* (pp. 244–247). IEEE.

Reale, M. J., Canavan, S., Yin, L., Hu, K., & Hung, T. (2011). A multi-gesture interaction system using a 3-D iris disk model for gaze estimation and an active appearance model for 3-D hand pointing. *IEEE Transactions on multimedia, 13*(3), 474–486.

Ren, Z., Yuan, J., Meng, J., & Zhang, Z. (2013). Robust part-based hand gesture recognition using kinect sensor. *IEEE transactions on multimedia, 15*(5), 1110–1120.

Risqiwati, D., Wibawa, A. D., Pane, E. S., Islamiyah, W. R., Tyas, A. E., & Purnomo, M. H. (2020). Feature selection for EEG-based fatigue analysis using pearson correlation. In 2020 International Seminar on Intelligent Technology and Its Applications (ISITIA) (164-169), IEEE. https://doi.org/10.1109/ISITIA49792.2020.9163760.

Sadiq, A., Yahya, N., & Tang, T. B. (2021). Diagnosis of Alzheimer's Disease Using Pearson's Correlation and Relief Feature Selection Approach. In *2021 International Conference on Decision Aid Sciences and Application (DASA)* (pp. 578–582). IEEE. https://doi.org/10.1109/DASA53625.2021.9682409.

Sharma, S., & Singh, S. (2021). Vision-based hand gesture recognition using deep learning for the interpretation of sign language. *Expert Systems with Applications, 182*, Article 115657.

Sharma, S., & Singh, S. (2023). ISL Recognition System using Integrated Mobile-Net and Transfer Learning Method. *Expert Systems with Applications, 119772*.

Sohn, M. K., Lee, S. H., Kim, D. J., Kim, B., & Kim, H. (2012). A comparison of 3D hand gesture recognition using dynamic time warping. In Proceedings of the 27th Conference on Image and Vision Computing New Zealand (pp. 418-422).

Song, M., Yang, H., Siadat, S. H., & Pechenizkiy, M. (2013). A comparative study of dimensionality reduction techniques to enhance trace clustering performances. *Expert Systems with Applications, 40*(9), 3722–3737. https://doi.org/10.1016/j.eswa.2012.12.078

Sugianela, Y., & Ahmad, T. (2020). Pearson correlation attribute evaluation-based feature selection for intrusion detection system. In *2020 International Conference on Smart Technology and Applications (ICoSTA)* (pp. 1–5). IEEE. https://doi.org/10.1109/ICoSTA48221.2020.1570613717.

Sun, S., Hu, M., Wang, S., & Zhang, C. (2023). How to capture tourists' search behavior in tourism forecasts? A two-stage feature selection approach. *Expert Systems with Applications, 213*, Article 118895. https://doi.org/10.1016/j.eswa.2022.118895

Tamura, S., & Tateishi, M. (1997). Capabilities of a four-layered feedforward neural network: Four layers versus three. *IEEE Transactions on Neural Networks, 8*(2), 251–255.

Truong, A., Boujut, H., & Zaharia, T. (2016). Laban descriptors for gesture recognition and emotional analysis. *The visual computer, 328*(1), 83–98.

Venugopalan, A., & Reghunadhan, R. (2021). Applying deep neural networks for the automatic recognition of sign language words: A communication aid to deaf agriculturists. *Expert Systems with Applications, 185*, Article 115601.

Wadhawan, A., & Kumar, P. (2021). Sign language recognition systems: A decade systematic literature review. *Archives of Computational Methods in Engineering, 28*, 785–813.

Weichert, F., Bachmann, D., Rudak, B., & Fisseler, D. (2013). Analysis of the accuracy and robustness of the leap motion controller. *Sensors, 13*(5), 6380–6393. https://doi.org/10.3390/s130506380

With Sign Language, Everyone is Included. (2018). World Fed. Deaf (WFD), Helsinki, Finland.

Wu, B., Lu, Z., & Yang, C., (2022). A Modified LSTM Model for Chinese Sign Language Recognition Using Leap Motion. In 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 1612-1617).

Yang, L., Chen, J. A., & Zhu, W. (2020). Dynamic hand gesture recognition based on a LMC and two-layer bidirectional recurrent neural network (BRNN). Sensors, 20(7), 2106. https://doi.org/doi:10.3390/s20072106.

Zhu, W., Miao, J., & Qing, L. (2014). Constrained extreme learning machine: A novel highly discriminative random feedforward neural network. In *2014 International Joint Conference on Neural Networks (IJCNN)* (pp. 800–807). IEEE.

Zhu, W., Miao, J., & Qing, L. (2015). Constrained extreme learning machines: A study on classification cases. arXiv preprint arXiv:1501.06115.

Zou, W., Yao, F., Zhang, B., & Guan, Z. (2018). Improved Meta-ELM with error feedback incremental ELM as hidden nodes. *Neural Computing and Applications, 30*, 3363–3370.