Real-time sign language framework based on wearable device: analysis of MSL, DataGlove, and gesture recognition

M. A. Ahmed[2] · B. B. Zaidan[1] · A. A. Zaidan[1] · A. H. Alamoodi[1] · O. S. Albahri[1] · Z. T. Al-Qaysi[2] ·
A. S. Albahri[3] · Mahmood M. Salih[2]

Abstract

Researchers have been inspired to use technology to enable people with hearing and speech impairment to communicate and engage with others around them. Sensory approach to recognition facilitates real-time and accurate recognition of signs. Thus, this study proposes a Malaysian Sign Language (MSL) recognition framework. The framework consists of three sub-modules for the recognition of static isolated signs based on data collected from a DataGlove. The first module focuses on the characteristics of signs, yielding sign recognition system requirements. The second module describes the different steps required to develop a wearable sign-capture device. The third module discusses the real-time SL recognition approach, which uses a template-matching algorithm to recognize acquired data. The final design of the DataGlove with 65 data channel fulfils the requirement identified from an analysis of MSL. The DataGlove is able to record data for all of the signs (both dynamic and static) of MSL due to the wide range of captured hand features. As a result, the recognition engine can accurately recognize complex signs.

Keywords Sign language · Gesture recognition · MSL · Data glove · Sensor · Malaysian sign language

1 Introduction

Individuals with hearing and speech impairment suffer discrimination and barriers that restrict their participation in various community activities. Due to the lack of proper communication, these individuals are deprived of their right to live, move, or even work independently. Language is a means of communication among humans. Different communities used different languages. Sign language (SL) is a means of communication among people who cannot hear or speak normally. SL is a visual-spatial language based on positional and visual components, such as finger and hand gestures, position, and orientation alongside arm and body movements. These elements serve to convey the meaning of an idea together. The phonological structure of SL generally has five elements, articulation point, configuration of the hands, type of motions, hand orientation, and facial expressions. Each gesture in SL is a combination of the five blocks mentioned. These five blocks are essential elements of SL and can be exploited for sign recognition by automated intelligent systems (Ramli 2012).

Sign language recognition system (SLRS) is intended to bridge the gap between people with hearing impairment and those around them by creating a common medium among them such as a translator. The translator interprets the signs performed into text or speech. Towards this end, two types of translators suggested in the academic literature, sensor-based sign language recognition, vision-based sign language recognition (VBSLR) (Ahmed et al. 2007; Rastgoo et al. 2020; Lee et al. 2021; Pezzuoli et al. 2021). VBSLR solution to hand-gesture recognition problems

✉ A. H. Alamoodi
Alamoodi@fskik.upsi.edu.my

B. B. Zaidan
bilalbahaa@fskik.upsi.edu.my

1 Department of Computing, Faculty of Arts, Computing and Creative Industry, Universiti Pendidikan Sultan Idris (UPSI), Perak, Malaysia

2 Department of Computer Science, Computer Science and Mathematics College, Tikrit University, Tikrit, Iraq

3 Informatics Institute for Postgraduate Studies (IIPS), Iraqi Commission for Computers and Informatics (ICCI), Baghdad, Iraq

depends on cameras and image processing techniques along with artificial intelligence (AI) methods. A single camera is a common solution adopted by many researchers to capture signs (Zhou et al. 2016; Sruthi and Lijiya 2019). In addition, using multiple cameras is another option adopted by researchers to localize the signer's body through saliency maps (Zamani and Kanan 2014; Tyagi and Bansal 2021) and skin color (Tang et al. 2018; Jebali et al. 2021). For a long time, the VBSLRS approach dominated in comparison with the sensor-based approach. However, VGR techniques can have occlusion problems, changes in illumination condition, changes in the distance between the signer and the camera, and high computation complexities (Kausar and Javed 2011). Thus, implementing real-time VBSLR is difficult because of high computation complexities.

An alternate approach to collecting gesture-related data is using instrumented gloves that are fitted with specific sensors such as flex, accelerometer, gyroscopes, and touch (Ahmed et al. 2021) (see Fig. 1). To date, wearable technologies (e.g., DataGlove) in real-world environments such as inadequate lighting and background complexity are considered a successful approach for SLRS. As opposed to vision-based systems (i.e., degree of bend, ordination, motion, and others), glove-based systems often have the advantage of being able to collect data directly, thereby eliminating the need to pre-process raw data (Vijayalakshmi and Aarthi 2016). Additionally, lightweight, low-cost, low-power embedded or wearable systems with minimal computing resources are another benefit of the VBSLR solution (Ahmed et al. 2007).

Several challenges related to the development of sensory DataGlove SLR approaches are reported in the academic literature. Performing a particular sign multiple times by the same user results in different values (Oz and Leu 2007) because hand size differences between users (Dipietro et al. 2008), position of sensors, material and usage of the gloves

Fig. 1 Glove-based system to recognize sign language gestures

affect the accuracy of the technology (Borghetti et al. 2013; Kau et al. 2015). In addition, actual size and length of the finger required changing the position of sensors or pre-calibration process whenever the user changed (Oz and Leu 2007; Khan et al. 2009), which reflects a tradeoff between the number of sensors, and the accuracy of recognition and data processing complexity (Kanwal et al. 2014; Ibarguren et al. 2010). Two cost aspects are related to the cost challenge: the first aspect is related to the cost of commercial gloves priced between $1000 and $20,000. The users relate the second issue to glove usage. Due to the high cost, the new technology might help rich people rather than poor people (Kau et al. 2015; Bajpai et al. 2015). Therefore, low-cost development of the sensory DataGlove SLR system is necessary (Abualola et al. 2016; Ahmed et al. 2010; Vijay et al. 2012). System accuracy is measured with small-scale data in terms of number and complexity (Fu and Ho 2008). Real-time recognition is another issue in recognition accuracy (Ibarguren et al. 2009). In practice, the real-time application is more preferable in developing the final product, but accuracy remains a barrier in developing such an approach (Bajpai et al. 2015; Gupta et al. 2015). The similarities between the signs were signed with a closed fist, for instance, the letters M, N, S, and T in the ASL alphabet (Bui and Nguyen 2007). Another example is the gesture for V and U characters, which produce error (Abualola et al. 2016).

Different SLs have different rules (Pradhan et al. 2008; Kadam et al. 2012; Maarif et al. 2018). Perhaps SL analysis is one of the fundamental challenges when a large-scale project aims to develop a proposed SLR (Dipietro et al. 2008; Zhang et al. 2011; Tanyawiwat and Thiemjarus 2012). In addition, understanding the gesture angles, orientations, and other features alongside the sensors that can acquire the relevant data per sign is a must to develop a robust SLR system (Das et al. 2015). Extra features acquired from extra sensors can help to recognize two different signs with a similar shape (e.g., letters U and V) (Aguiar et al. 2016).

Software, an essential component of every system, plays an important role in data processing in addition to the possibility of improving system outputs. Software development for SLR systems is related to the methods used in the classification process to recognize gestures. One of the common direct methods to perform static posture recognition is prototype matching (also known as statistical template matching), which operates on the basis of statistics to determine the closest match of acquired information values with pre-defined training samples called templates (Bhatnagar et al. 2015). In fact, this method is characterized by the lack of a need for complex training processes or wide calibration, thus increasing its speed. From a pattern recognition standpoint, the artificial neural network is the
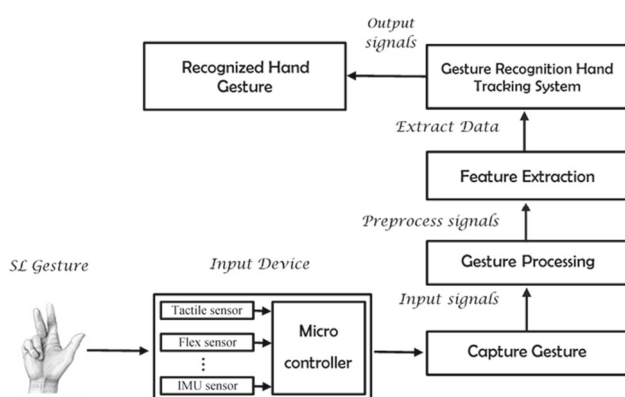
most popular method used in ML in the recognition field (Iwasako et al. 2014; Abdulateef et al. 2020). Therefore, this technique can be trained to distinguish both static and dynamic gestures, as well as posture classification, based on the data obtained from the data glove (Pradhan et al. 2008; Iwasako et al. 2014; Mehdi and Khan 2002; Lei and Dashun 2015; Adnan et al. 2012). Long-term fuzzy logic has been used in many fields that need human decision-making; one of these areas is recognition of SL (Kau et al. 2015; Tanyawiwat and Thiemjarus 2012; Das et al. 2015; Swee et al. 2007a; Qiu et al. 2020; Jianbin et al. 2020). Another useful algorithm falls within the scope of ML employed to provide accurate and less complex classification through dimensionality reduction with improved clustering, which is known as linear discriminant analysis (Abhishek et al. 2016; Kong and Ranganath 2014). Hidden Markov model (HMM) is a popular technique that has shown its potential in numerous applications such as computer vision, speech recognition, molecular biology, and SLR (Swee et al. 2007b; Oz and Leu 2011; Gałka et al. 2016; Anupreethi and Vijayakumar 2012; Kosmidou and Hadjileontiadis 2009). Besides the HMM, the KNN is also used to classify hand gestures (Ani et al. 2014), and the KNN classifier with support vector machine (SVM) has been applied in posture classification (Zhang et al. 2011; Ahmad 2016). The KNN is applied in the recognition of ASL signs (Tubaiz et al. 2015).

With the scope of MSL, current work is insufficient to solve this humanitarian problem. In particular, few articles have been published to address the problem of MSL gesture recognition, where only three studies in the literature discussed the development of an MSL recognition system (one development article and two frameworks) (Swee et al. 2007a, b; Shukor et al. 2015). Several issues were reported in these articles. In particular, these articles developed in a small-scale test, no published dataset was associated with these articles, the glove tested was limited to tilt sensor to measure finger angle, no SL analysis was provided, the glove was limited to SLR while the accuracy was high due to the number of signs used during the experiment (as few as nine signs). Therefore, the development of MSL lacks certain elements.

The major contributions of this study are to develop a framework for MSL recognition based on the DataGlove. Other contributions of this study are the following:

1. A novel DataGlove for capturing most of the hand attributes is designed. This DataGlove guarantees the required attribute to recognize different SL postures with no limitation in the complexity of signs.

2. The presented work is considered as the first real-time MSL recognition system modeled to recognize a wide range of static gestures.

3. System hardware components include a mechanism to distinguish different signs with similar hand attributes but different hand locations.

2 Related work

Looking at the academic literature from the view of critical analysis, we find that none of the presented articles has developed an approach of DataGlove SLR that is capable of handling a wide range of gestures, for example, the gender of the noun. However, certain developments have covered several aspects related to hand attributes. (Vijayalakshmi and Aarthi 2016) applied a sensor-based approach for English alphabet gesture recognition by developing a module to convert gesture to text using statistical template matching. This approach converts the output into text and finally to voice using HMM, which has been used to construct the speech corresponding to the text. However, the glove does not include a mechanism for reading thumb movements, which has an active role in shaping many of the signs. Another factor is the small size of the training data (eight characters only). Nevertheless, the recognition accuracy ranged from 80 to 89%. A wireless DataGlove was designed and developed (Tanyawiwat and Thiemjarus 2012) for American Sign Language. A combination of five contact sensors and five flex sensors in addition to 3D accelerometer was used in this glove for fingerspelling gesture recognition. The gesture recognition engine performed statistical template matching. Gestures were collected from six deaf subjects and one healthy subject. However, low recognition accuracy (76.1%) was recorded with the 21 features from the new sensor glove. This result is due to the large number of misclassifications of letters such as D, H, X, Z, and SP. The work presented by Elmahgiubi et al. (2015) produced a sensory glove that captures the signs of American Sign Language (ASL) and converts these postures into text using template matching teaching. Three types of sensors were used for gesture acquisition: five flex sensors along with five force sensors and 6 degrees of freedom MPU6050. However, the system is designed to recognize 26 ASL alphabet only. Furthermore, the system was able to interpret 20 out of the 26 letters. This condition means that the system misclassified six letters: E, M, N, S, T, and Z. The electronic portable hand glove was developed by Arif et al. (2016). The glove consisted of five flex sensors and one accelerometer to capture hand form. A contact sensor was also involved in the process of distinguishing some signs. The system was used to translate gestures of ASL with through statistical template matching to produce a useful solution for those with hearing impairment. The system

outputs are in the form of a voice through a speaker and text format on liquid crystal display (LCD). However, the system has been tested only on the alphabetical data of the ASL. In addition, it has a large system design and heavy equipment weight that makes it inconvenient to use. Additionally, the loss of usability privilege as a portable device in public areas is considered as a failure. The system LabVIEW was used by Sharma et al. (2015) to interpret the 26 letters of the ASL. The data of gestures are obtained by using a haptic glove and converting signs into text and voice. The system also could be trained for learning the SL. The glove consists of nine flex sensors, four contact sensors, and one accelerometer. However, this work focused on the technical aspects of the device design rather than discussing the recognition accuracy rate.

With respect to MSL, a selected set of articles suggested three articles that have employed the glove-based approach to solving the problem of Malaysian SL recognition. Two of these articles belong to the framework category, while the third falls into the development class. (Swee et al. 2007b) and (Swee et al. 2007a) described hardware design for a sensory glove system to recognize the 25 commonly used gestures of MSL as well as the setting and configuration of the system. A system consists of a set of flex sensors, and accelerometers have been proposed to measure the motion of the elbow, hand wrist, and fingers. However, these two works only presented a framework for designing a glove to distinguish the SL; they did not address the problem of similar signs in the language. (Shukor et al. 2015) developed a translation system for MSL with the capacity to interpret MSL gestures into text. The configuration of the right-hand sensory glove comprises 10 tilt sensors to capture the flexion of fingers, a single three-axis accelerometer to identify the orientation of the hand, a microcontroller to process the data and convert gestures into text and finally, a Bluetooth module to transmit the recognized data to a smartphone. However, the data on Malaysian signs that were used in experiments are as few as three letters (A, B, and C), three numbers (1–3), and three isolated words (i.e., "saya," "Apa," and "makan"). Although the experiment samples were very few, a contrast exists in the accuracy reported during system testing, with an average accuracy of 95% for characters, 93% for numbers, and 78.3% for isolated words. In addition, this study did not provide a solution for similar gesture problem concerning MSL (e.g., U, R, and V) characters and other issues. Furthermore, a flex sensor was replaced, which was proved to have high ability to determine the amount of finger bending with high accuracy by the tilt sensor.

In general, articles that discussed sensory DataGlove are limited due to several aspects, including the complexity of DataGlove development and the high cost of commercial DataGlove. In addition, requirement analysis is developed based on academic literature rather than exploring SL.

Therefore, there are several barriers to making these attempts limited. The needs of intensive analysis on SL allow researchers to understand the behavior and patterns of a particular SL. Expected technical literacy in such studies is high, in which researchers are required to have a background in SL, electrical engineering, and sensor technology, as well as programming language skills.

To simplify the DataGlove problems, issues need to be broken down into blocks. The first block is discussed in the SL context in which researchers are required to explore and study several aspects related to hand behavior, hand attributes, and gesture patterns toward developing DataGlove requirements. The second block discusses the concerns with the available sensors and technologies that can enable DataGlove to channel the required data representing the entire gestures. The last block discusses evaluating the proposed DataGlove to ensure the capacity of this proposal. This process can be done by identifying the signs systematically with a wider range of signs and larger scale of testing attempts. Neither the work proposed in ASL nor other SLs established a comprehensive study toward this end. Problems such as the gender of the nouns and similar sign recognition are either not discussed or excluded from the test. Finally, benchmarking the proposed DataGlove in the academic literature with a large-scale dataset would reduce the accuracy of these proposals. Therefore, this study discusses a generic framework to reduce the obstacles of sensory-based DataGlove for MSL recognition development.

3 Proposed framework

The proposed framework cycle is divided into three phases (Fig. 2). Module I (Investigation on SL and MSL Analysis) describes three involved information sources (i.e., investigating the preliminary scholarly research, expert, and observation) to gain knowledge about the SL field, and identifying and viewing the research problem. In addition, MSL analysis aims to examine the MSL from a different view point to extract useful features that may assist in drawing recognition system requirements. Module II (Design and Development of DataGlove) is proposed in several stages to develop an optimal design for DataGlove. Module III (Development of Recognition System) involved developing a proposed solution for recognition of the static gestures of MSL based on DataGlove.

3.1 Analysis module

The main goal of this module is to highlight the importance of human hand attitudes in SL and extracting critical attributes of gestures to facilitate the process of

Fig. 2 Block diagram of proposed MSL recognition framework
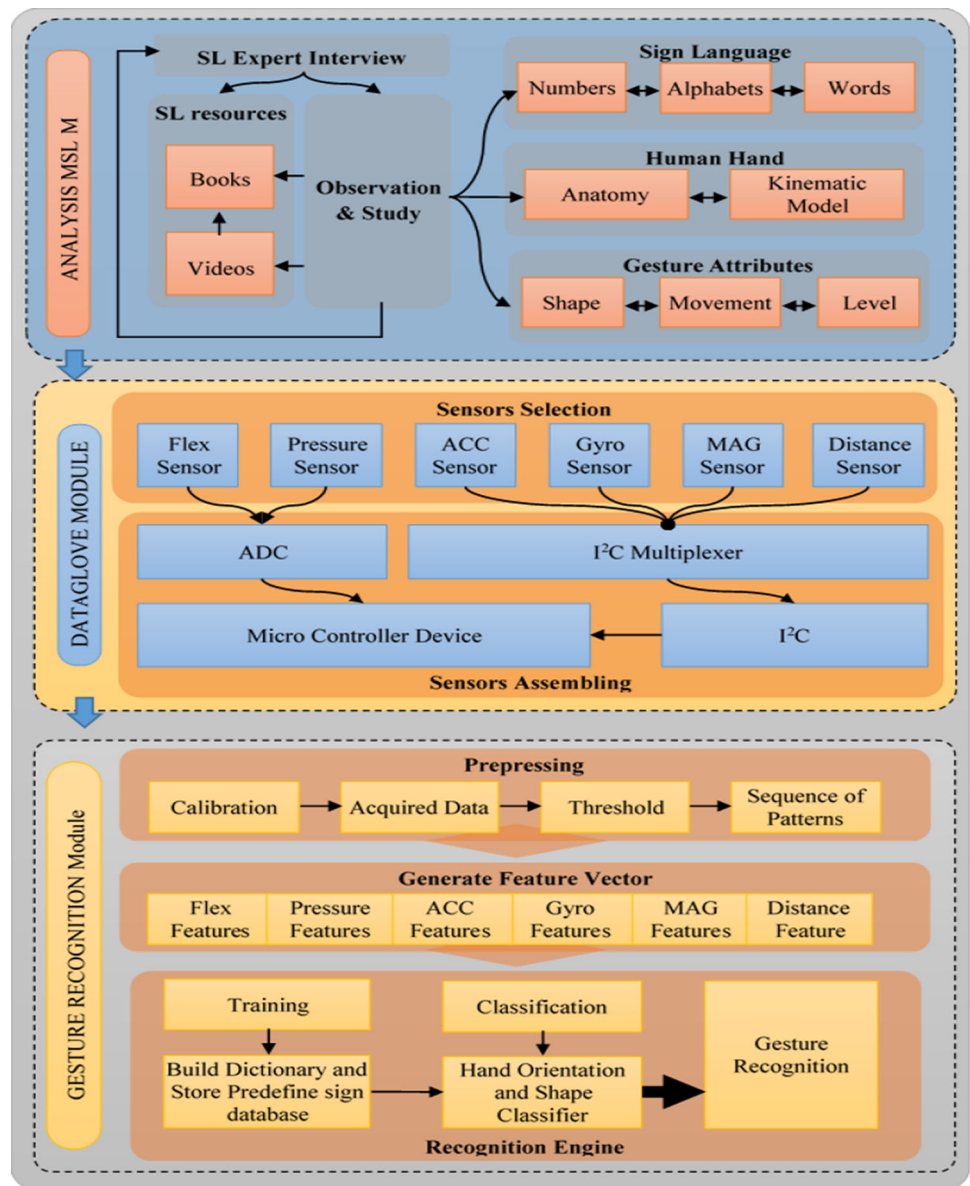
distinguishing between convergent gestures. These attributes would contribute, in later work, to the development of an expert system eligible for automatically interpreting

Fig. 3 Foremost resources in gaining knowledge

signs of MSL. To achieve this goal, the discussion is divided into a twofold investigation on SL and DataGlove SLR, and MSL analysis (see Fig. 3).

3.1.1 Investigation on SL and DataGlove SLR direction

Three investigation types were adopted in this study to gather relevant knowledge about the SL area. The first type involved the academic literature scientific articles were searched on the reliable electronic databases to study the relevant studies on DataGlove-based SLR. The second trend is gaining knowledge by conducting a personal interview with an expert in the field of SL, and the final stage is based on the observation process.
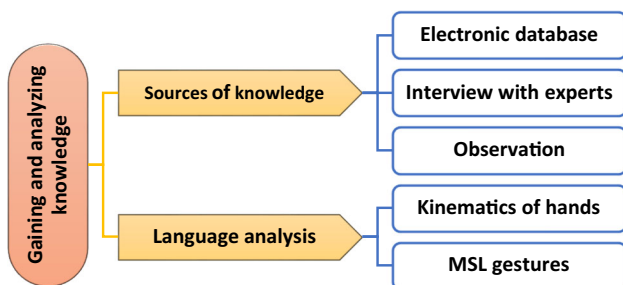
3.1.1.1 Gathering scholarly data The initial step was to carry out a primary search for related articles to SLR generally to constitute a complete picture of the technologies adopted in this field. Subsequently, the collected articles were examined from the primary research by narrowing down the scope to SLR based on the sensory glove. For this reason, proper keywords to initialize articles search with the scope of sensory glove SLR were identified by surveying the keywords used in previous research works. The next step is to identify high-quality electronic databases to be a major resource of scientific material such as IEEE Xplore, Web of Science, and Science Direct. Finally, the collected relevant articles were fully read resulting in a deep understanding of SLR and highlighting the most notable information.

3.1.1.2 Interview with SL expert The interview is considered one of the appropriate and constructive methods in acquiring knowledge and collecting data on a particular topic. Thus, the interview is conducted with a person that has the required experience in this field. Furthermore, straightforward contact with the person interviewed is likely to lead to particular and constructive proposals. Therefore, this study proposes that an interview be conducted with a person who has experience in teaching SL in the early stages of system development.

3.1.1.3 Observation Observation was adopted as one of the methods in investigation and elicitation. To maximize the benefit, we adopted three observation modes:

1. Observation of the learning materials of the SL that is available on the internet such as websites, videos, pictures, and mobile applications.
2. Observation of annotated pictures in syllabus books for MSL.
3. Observation videos of a native SL volunteer while performing MSL signs.

First, a specific SL must be selected and electronic learning materials must be found, which explain how the static and dynamic signs of selected SL are performed. In addition, the material should provide a description of the sign in terms of shape, orientation, and motion of the hand, including the display of several samples of a given sign. Second, we choose the textbooks that provide details about each sign supported by a pictorial representation. Subsequently, we determine the appropriate gestures for recognition. Finally, all the selected gestures, which are listed in the book, were recorded in five sessions. The singer is sitting in front of the video camera in a well-lit room. The camera was set to capture the upper body of the volunteer and was approximately 500 cm from the signer. The main goal of the observation cycle was to observe how the signer

performed gestures and monitored hand attributes, facial expressions, lip events, and body language. The knowledge gathered from the observation was used to understand MSL.

3.1.2 Language analysis

The literature reported the importance of analyzing the SL, which is the second direction in this module. SL communication is a symbolic nature and highly structured involves manual signing (MS, i.e., hand/arm gestures). In addition, non-manual signals are used to decipher the full meaning of the sentence through facial expressions, movements of head/lips, torso movements, and body postures. This trend is focused on examining the meaning of sign gestures by hand. Gesture analysis resulting from this phase is used in the development of DataGlove and related algorithm of MSL recognition.

Due to the importance of the human hand in the SL, it was salutary to look at the hand dissection to form a perception of the all-presumable movements that can be produced. Ideally, the analysis of hand anatomy helps to describe the scope of most potential postures that the hand can conduct. Therefore, the investigation was twofold: the first concern is to study the kinematics of the hand and the second concern is to study the MSL gestures. Figure 4 illustrates the main points of the analysis process to achieve the objective of this module.

3.1.2.1 Hand configurations Understanding finger kinematics is significant in various research fields, particularly medicine, biomechanics, and other scientific disciplines. A human hand moves with multiple degrees of freedom (DOFs) due to its articulated nature. Observably, the human hand contains 27 DOFs on which 21 DOFs are contributed by the five finger joints for local movements and the other six for global hand movements (Kortier et al. 2014; Lee and Kunii 1995). The human hand is highly articulated but constrained at the same time.

In the hand, 27 bones of different sizes are connected by flexible joints, which are responsible for different hand movements (Bullock et al. 2012). Figure 5 displays an overview of the bones in the hand. The pinky, ring, middle, and index fingers have the same joint configuration, whereas slight differences were observed when the thumb is compared with the other fingers. These 19 movable bones are connected by 14 joints with different intricacies, which make finger bending and other bones dependencies movements.

Finger joints are the connection areas between two different bones in the hand and fingers. Some people believe that several hand joints, particularly the joint that connects finger bones, have trivial functionality, which is
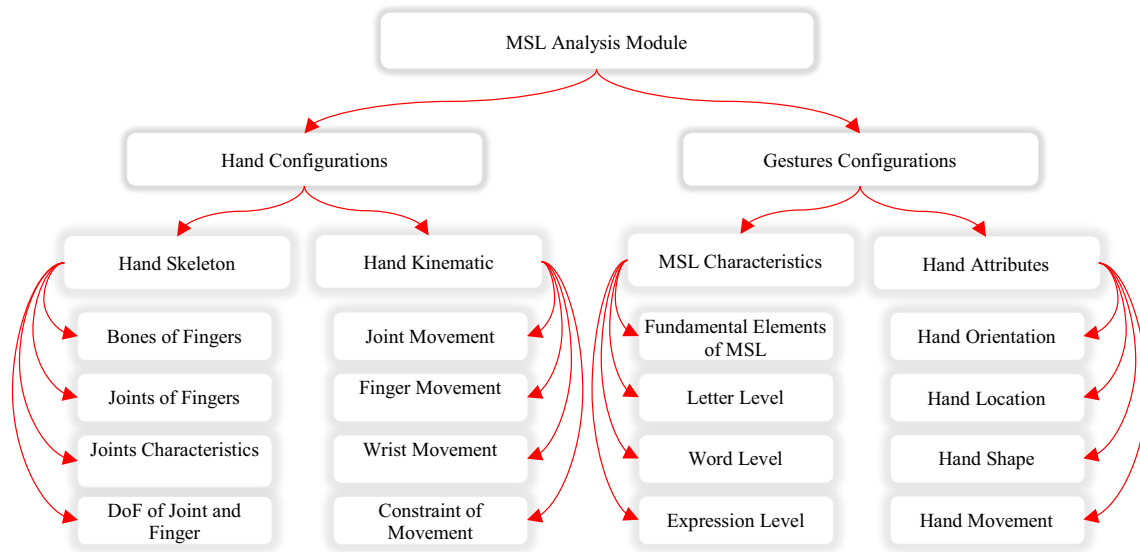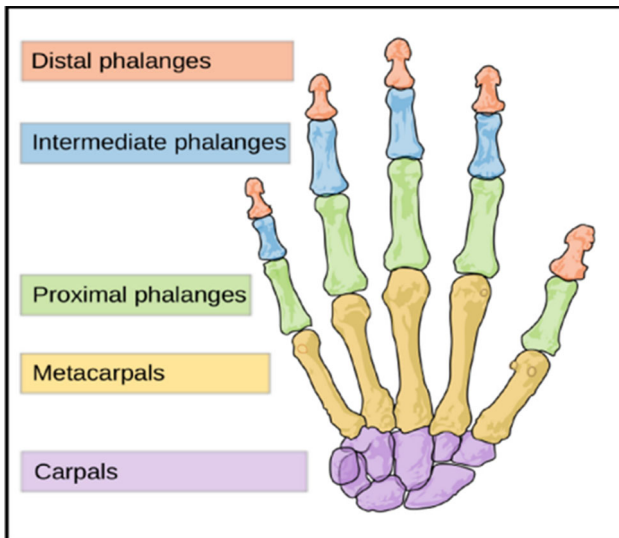
Fig. 4 Steps in MSL analysis
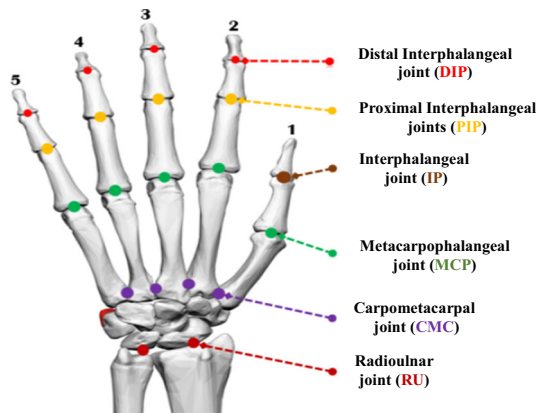

Fig. 5 Human hand bones (Ahmed et al. 2021)

Fig. 6 Joints of left hand, dorsal view (Ahmed et al. 2021)

true and complex to understand. Few constraints are associated with joint movement because every two bones are connected by a specific joint. In view of this fact, hand joints do not produce a complete rotation axis, and thus, only two axes are involved in the figure movement (see Fig. 6).

Each joint can produce a particular movement or rotation restricted by the bone size, joint laxity, bone position, and other restrictions. Perhaps most joints can perform two types of movement: adduction–abduction (AA) and flexion–extension (FE). Abduction refers to a structured movement away from the midline, whereas adduction points to the movement toward the center of the object. The object's center can be described as the midsagittal plane (Buczek et al. 2011) (see Fig. 7).

Radioulnar wrist joints produce both AA and FE motions in addition to pronation–supination (PS) motion. Pronation refers to rotations of the wrist so that the palm faces backward or downward, while supination describes the rotation of the wrist in which the palm faces forward or upward (see Fig. 8). These moments seem similar to the transverse plane.

3.1.2.2 Gesture configuration To analyze SL on a sound and reliable basis, the authors have identified trustworthy sources to serve as main references in selecting appropriate signs for this study. Two courseware books used to teach MSL have been adopted. The first book is *Bahasa Isyarat Malaysia* (translated as "Malaysian Sign Language"), published by the Malaysian Federation of the Deaf (Qiu et al. 2000) and the second is *Bahasa Malaysia Kod Tangan Jilid 1*, published by the school division of the Ministry
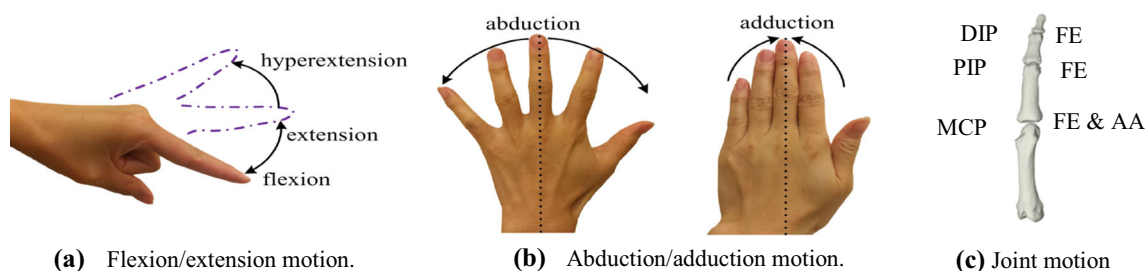
(a) Flexion/extension motion.   **(b)** Abduction/adduction motion.   **(c)** Joint motion

Fig. 7 Mechanics of fingers and digits: **a** flexion/extension motion of a finger concerning x-axis, **b** abduction/adduction movement of fingers concerning z-axis, and **c** biomechanical rotation axes for index finger joint (Ahmed et al. 2021)
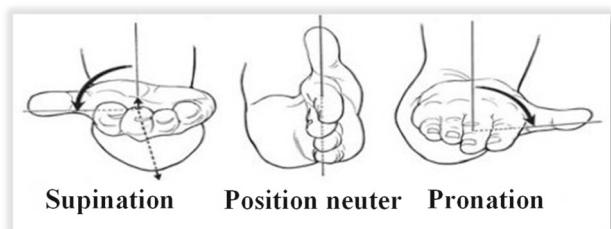

Fig. 8 Wrist PS motion

of Education of Malaysia in cooperation with the National Communications Committee of the whole communications working committees, 1985 (Malaysia 1985).

MSL is not a spoken language and does not require a voice to convey knowledge or communicate between people. Movements of arm, hand, palm, fingers, and even head with the facial expressions were used to replace the sound to convey information between individuals with hearing and speech disabilities. Each movement carries a certain meaning that enables the person who views through the eye of the other speaker and understands the meaning of the performed signs. Therefore, MSL is a visual language that is structured into three fundamental levels (i.e., ingredients, character, word, and expression).

- *Character level* Each language in the world has its own alphabet, which distinguishes a particular language from other languages. In SL, the alphabet is composed of hand postures instead of letters. MSL consists of 26 postures (each posture presents a special character of BIM) representing the alphabet.
- *Word level* MSL includes postures that hold the meaning of a single word or more. The speaker uses postures as aforementioned for real conversation in the majority of communication.
- *Expression level* Apart from words, in some cases, the movements mentioned in the preceding point are not sufficient to convey meaning. In the sense that, it is

resorted to utilizing, some expressions to be another fraction added to the essential components of MSL. These expressions can be facial expressions and movements of the lips or tongue and body position. Occasionally, the speaker moves his or her lips or draws a particular expression on his or her face while performing a certain sign.

3.1.2.3 Hand attributes Not surprisingly, the hand's state description at a particular piece of sign plays an important role in shaping the gestures of the SL. The hand performs a series of complex movements in multiple forms to provide a wide range of gestures. These gestures are used in daily communication among deaf people. The complexity of the hand is due to a combination of several aspects (i.e., hand shape, hand orientation, and others). These aspects are taxonomized in this study and explained individually under the label of hand attribute. The developed classification of the gesture based on analyzed hand attribute is used in this thesis to group the sign and formulate the requirements for glove development. Attributes of the hand (see Fig. 9), such as number, movement, location, orientation, shape, and movement stages, are described in detail in the following sections.

3.2 DataGlove module

3.2.1 Design constraints

Hardware design is crucial to the development of an auto sign-language recognition device to ensure that the data are recorded accurately. Thus, certain specifications and constraints were set to achieve high-quality results (Fig. 10). The design constraints are summarized as follows:

3.2.1.1 Physical glove specifications

1. User-friendly: The system ought to be simple and easy; therefore, no restrictions per use are expected.
2. Close-packed: The system ought to be compact, lightweight, and portable.

Fig. 9 Characteristics of hand attributes
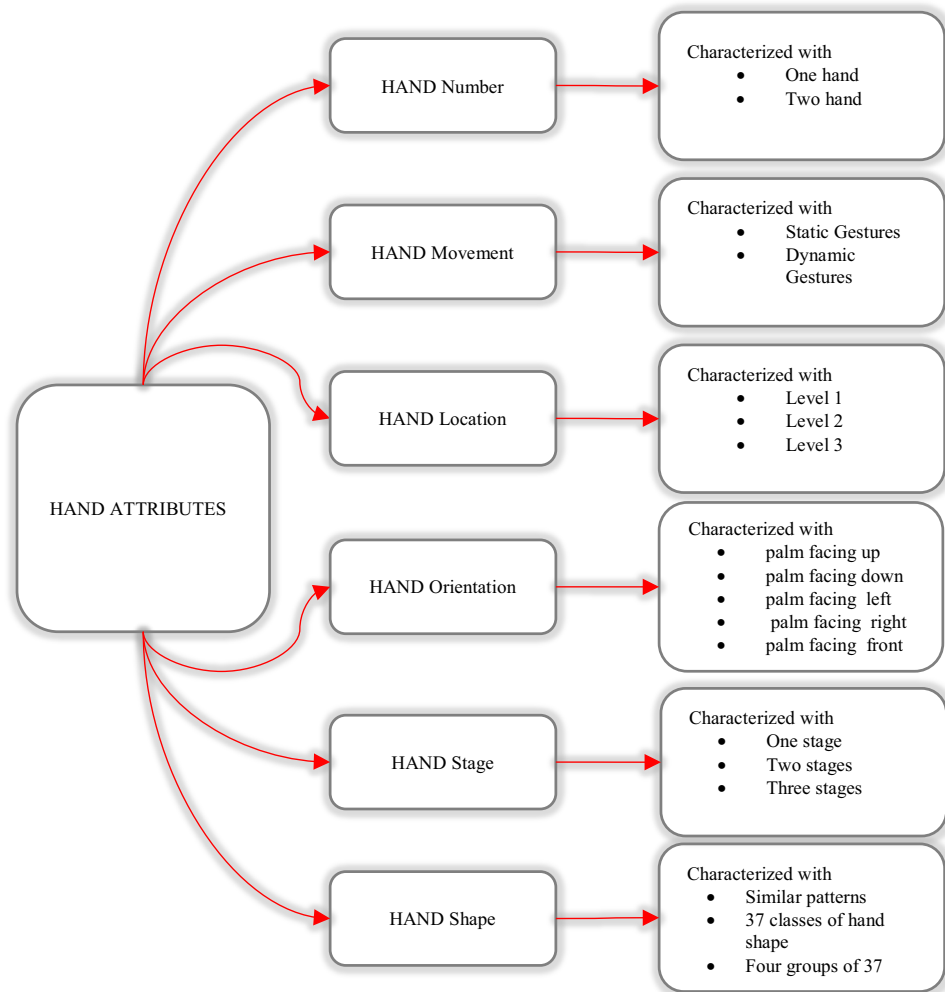
3. **Affordable:** The system ought to be inexpensive financially.
4. **Adaptable:** The system ought to be eligible for integration with existing gadgets such as smartphones.
5. **Comfortable:** The glove must be breathable, easy to wear and take off, and not restricted to the user's hand.
6. **Safe:** No exposed wires carry a current of 1 mA or greater, and the glove does not cause harm to the user.

3.2.1.2 Data specifications

1. Identifying sign resources, such as educational books, websites, or other learning methods, where numerous sources for teaching SL are available.
2. Identifying the number of hands, which affects the number of signs to be targeted as well as the system layout.
3. Identifying the type of gesture, which is either a static gesture drawn without any movement or a dynamic gesture.

4. Identifying sign samples as the result of the application of all the preceding points.
5. Identifying participants who will perform signs that have already been identified.

3.2.1.3 Acquiring and processing of data

1. Selection of acquisition gadgets to be mounted on the sensory glove to capture hand motion and finger bend.
2. Selection of enough sensors in terms of sensor type and number to accurately record static gestures.
3. DataGlove can expand to capture a dynamic gesture and continuous sentences.
4. The design of the DataGlove should provide high-resolution data based on the best currently available technologies.
5. The system can be calibrated for different users, is simple to use, and quick to set up.
6. A suitable method should classify the signs with fast recognition and good accuracy.
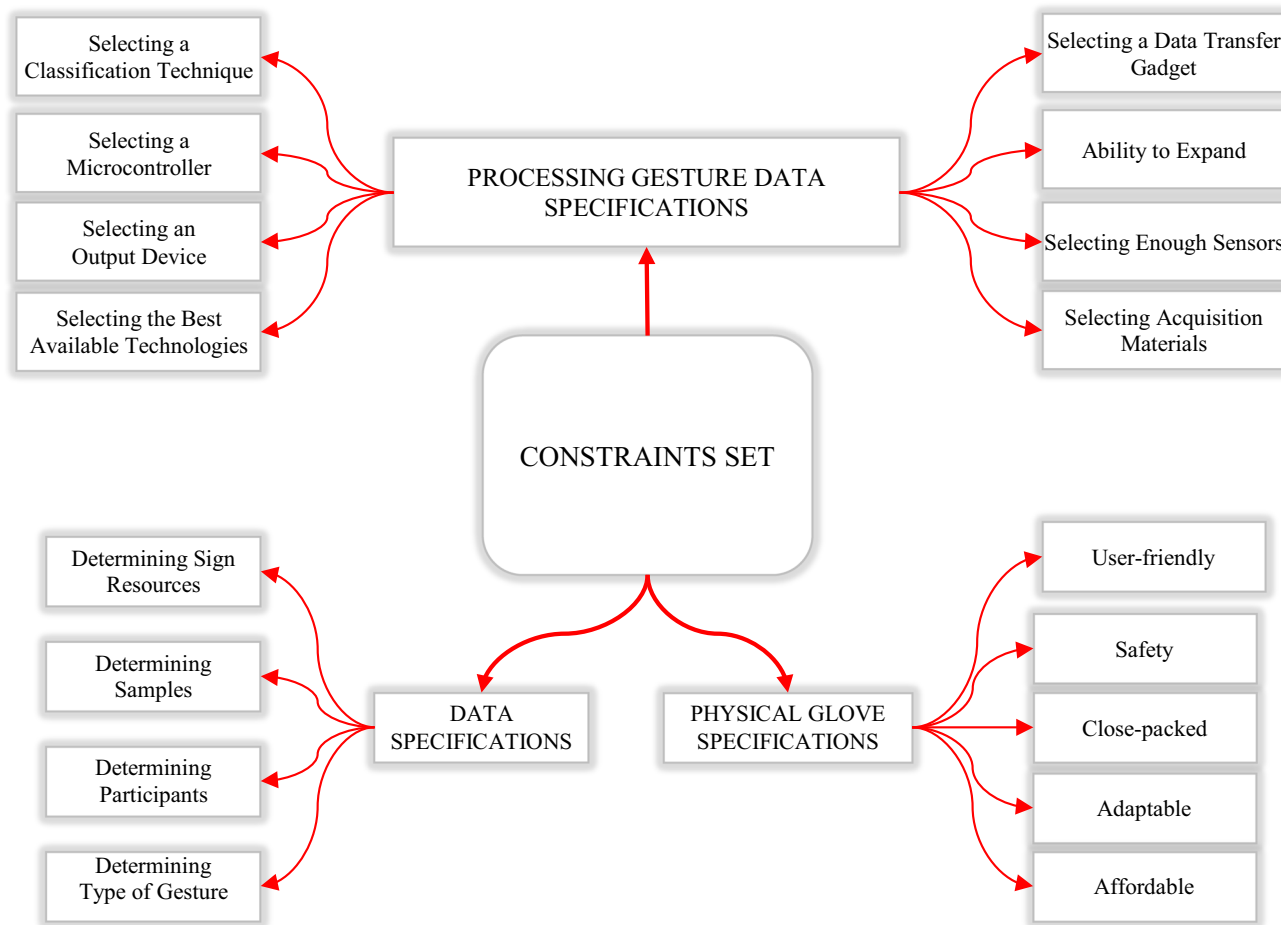
Fig. 10 Constraints set and primary keys for proposed framework

7. Selecting a processing gadget represented by the microcontroller, which is designated as an intelligent subsystem that is responsible for reading, interpreting the signs, and producing the output.

8. Selecting an output gadget represented by a loud-speaker to generate uttered words or display a system to produce visual words, such as a connected LCD, directly to the microcontroller to present the consequence.

3.2.2 Appropriate glove material selection

Several important features must be considered when selecting a suitable glove for the task of detecting the sign. For example, the material of the glove must be made of an elastic material so as not to impede the movement of the hand and provide sufficient flexibility to allow the fingers to bend comfortably. Another issue is related to the dimension of the hand. The size of the hand generally

varies among individuals. The size of a young person's hand is different from that of an adult; furthermore, the hand of the male is larger than that of the female, not to mention the differences among individuals from the same class and similar age in the size of hands.

3.2.2.1 Sensor selection Many sensor technologies are available in the market. The outputs of different sensors may be integrated to achieve a particular task. The nature of SL is needed to measure the bend of figures, the orientation of the hand, and other features such as figure AA motion to distinguish the signs.

a) Finger Bending and Flexion-related Sensors

Observing the academic literature is suggested to apply the capabilities of flex sensors in measuring a finger band rather than other sensors because the flex sensor is ideal for measuring repetitive bending, acceleration, or range of motion while providing high-speed measurements. In the literature, several sizes of flex sensors are suggested,

specifically 2.2 and 4.5 inches. However, researchers have neither justified the usage of different sizes nor paid intention to the position of the sensors. Therefore, a ready design where the difference of data resulting from each sensor and position of the flex sensor is adjusted with the hand kinematic has yet to be discussed in the literature. These issues raised several questions that the literature has not answered, such as the following: What is the optimal size of the flex sensor to be used in SLR gloves? Where is the best location to assemble the flex sensor at the finger? Does the number of joints covered by flex sensor play an important role in the data acquired by flex? To what extent is the data important in the recognition process? To answer these questions, an experiment needs to be designed to understand the capacity of flex sensors, the differences between flex sizes, and its relation to the data quality. Furthermore, the optimal position that produces better information for the number of joints covered and the hand kinematic has to be identified.

b) Hand/Wrist and Finger Orientations-related Sensors
Inertial measurement unit (IMU) has been used to measure the orientation of the hand or wrist. However, few researchers have suggested using IMU to illustrate the finger orientation instead of finger flexion. Finger orientation is yet to be the best alternative to finger flexion. However, with some signs, finger orientation is valuable information to distinguish signs. A design that combines IMU and flex has not been discussed in the literature and thus a question might be asked with respect to the feasibility and usefulness of making this combination. We believe that the combination of both sensors is experimentally possible and deserves an attempt to ensure data with better quality. Thus, both IMU and flex are proposed to be allocated over fingers to increase the benefit of having a better features vector.

Apart from this, the IMU sensor (also called micro-electromechanical system) is needed due to its capability to measure the rotation and motion of objects. Indeed, no abundant alternative sensors can estimate the movement or rotation of the hand rather than using IMU. Three different types of IMU are used in the academic literature (3, 6, and 9 DOFs). Similar to the case of the flex sensor, no experimental comparison can confirm the usage of each. Therefore, we propose to use and test different IMUs to decide which is suitable for the MSLR process with the case of finger orientation and hand/wrist orientation.

c) Abduction–Adduction-related Sensors
From the observation of SL shapes and the knowledge provided in the academic literature, several signs are required to touch or twist a figure, which has the same shape from different angles. Such signs require additional information to be distinguished (e.g., the case of U and V).

A touch sensor or force-sensing resistor (FSR) is used to provide the desired information in the literature. Few researchers have suggested sparking two wires to detect the touch between fingers. Sparking wires might cause the burning of the circuit that leaves no choice but avoid such a solution. Size, comfort, and capability of sensors to detect different cases are the main selection criteria in the abduction–adduction process. Other procedures adopted in commercial gloves used the flex sensors to identify the angle between fingers. However, a glove that costs \$22,000 to understand one sign behavior is not a viable solution. In addition, we believe that the flex sensor is capable of handling the angle, not the figure twisting. Therefore, the flex sensor is excluded from the process of abduction–adduction measurement. Nevertheless, the lack of information for the best sensor of measuring abduction–adduction leaves no way but to compare the previous solution experimentally. Therefore, the selection of the touch sensor or FSR is subject to the experimental confirmation.

d) Hand Location-related Sensors

The above-mentioned sensors are utilized in the academic literature. Due to some reasons (not reported in the literature but we believe are due to the lack of SL analysis), some signs are identical (similar shapes, orientations, bend, and finger angle) but in different positions are difficult to distinguish and not discussed in the literature (e.g., Baba/Father and Emak/Mother). Therefore, the position of the hand needs to be investigated. At first glance, IMU seems to be a good indicator when things are related to the arm direction or angle. However, this option requires complex processing and calibration per user or per use. Therefore, we propose using sensors that can measure the distance between the elbow and ground. Three sensors are proposed to measure this feature, namely, time of flight (TOF), ultrasonic sensor, and barometric sensor. Comparison among these sensors experimentally can help to distinguish identical signs.

Based on sensor numbers, type, and position (finger bending and flexion, hand/wrist and finger orientations, abduction–adduction, and hand location), the newly designed glove can help to have a wider range of words in the classification module.

e) Microcontroller Selection

The microcontroller, which is similar to the system's mind process the collected data from the sensors to recognize the gesture. Different kinds of microcontrollers are employed in recognition systems such as ATmega, MSP430G2553, ARM7, and ARM9. Furthermore, previous researchers have used Raspberry Pi, Arduino, and Odroid XU4 as an

electronic platform. The Arduino is a common choice for SLR projects. Arduino boards are microcontrollers that execute written code as their firmware interprets it. The Arduino board is interfaced with sensors through analogue and digital ports. A microcontroller called Arduino Mega 2560 with 54 digital input/output pins, 16 analog inputs, 4 UARTs (hardware serial ports), and 16-bit ADC is used in the present study.

3.2.2.2 Output unit selection The final stage of the recognition system is displaying what has been distinguished: either recognized gestures converted into a text or pronunciation of what has been signed or represented in the form of animation. One of the following devices is used to achieve the above: a personal computer, an LCD screen, a loudspeaker, or even a smartphone. In this study, the intercepted sign is displayed as text on the PC screen using the Arduino serial monitor.

3.3 Proposed intelligent sensory glove

The intelligent sensory glove comprises a glove equipped with sensors interfaced with an Arduino ATMega328 microcontroller. A flex sensor, a pressure sensor, a TOF sensor, and 9-DOF IMUs are all included in the glove. Each sensor serves a particular function, such as measuring finger bends with flex sensors. To detect finger abduction and adduction, the pressure sensor is used to measure the contact between adjacent fingers. Finally, fingers and hand orientation are determined by the IMU sensor, and the hand position is measured with a TOF laser sensor. Each sensor is positioned in the proper position to achieve the sensor's purpose. Five flex sensors are located on the dorsal side of the fingers using pockets. Four pressure sensors are sewn to the side of the fingers. Five IMUs are placed on the fingertips (one for each fingertip), and one IMU is attached to the back of the palm, while the distance sensor is located at the elbow. Figure 11 shows the main components of the data glove for this work.

The flex and touch sensors are connected to the microcontroller through resistors to the analog port. IMUs and distance communicate with the Arduino through the I^2C protocol. The recognized gestures are then displayed as text on the monitor. Figure 12 shows the system component, connection circuit, and position of each sensor. Table 1 presents the items used in prototype hardware for the SLR system in general and MSLR in particular.

3.4 Gesture recognition module

The sign translator design commences with the glove, the heart of the system. Several fundamental actions must be observed while commencing with establishing the glove.

3.4.1 Case study

In accordance with expert advice, *Bahasa Isyarat Malaysia (Malaysian Sign Language)* and *Bahasa Malaysia Kod Tangan Jilid 1* were adopted as major references in selecting MSL postures. Approximately 1708 signs listed in both books have been investigated. These gestures are typically used in daily life and classified into three main categories: numbers, letters, and words. Table 2 presents a list of gestures selected in this study.

3.4.1.1 Case study I: fingerspelling and alphabet In the context of SL, reading and writing involve different techniques, for instance, reading can be spelled letter by letter (typically technical terms or names). The MSL alphabet consists of 26 letters (i.e., A to Z) performed by a single hand. Most of the characters are performed with static movement except for the letters J and Z that require movement. Therefore, the selected sample in this case study includes 24 letters (all MSL letters except J and Z). Figure 13 presents the hand postures of 26 letters of the MSL alphabet.

3.4.1.2 Case study II: numbers Numbers in MSL have their own postures. Numbers from 0 to 10 are formed using one hand and a static gesture (Fig. 14), whereas numbers greater than 10 are formed using dynamic gestures. Thus, signs of numbers from 0 to 10 are selected due the scope of this study.

3.4.1.3 Case study III: MSL isolated words Besides numbers and letters, MSL consists of a considerable set of words. A considerable number of works performed by hand posture execution are available. These words are either simple (i.e., composite of one static form) or complex (i.e., consisting of more than one form). Within the scope of this study, words performed with one hand were selected along with static gestures. Thus, 60 is the total number of static signs among the total of 1708 signs (i.e., signs listed in the two selected books) (see Table 3).

3.4.2 Participant selection

An effective participant selection process is important because of inappropriate procedures that may seriously affect the findings and outcomes of a study. Results of academic literature analysis showed that the number of participants that perform a gesture range between 1 and 10. Therefore, this work involves 10 adult volunteers (five males and five females between 20 and 35 years old) who perform the signs.
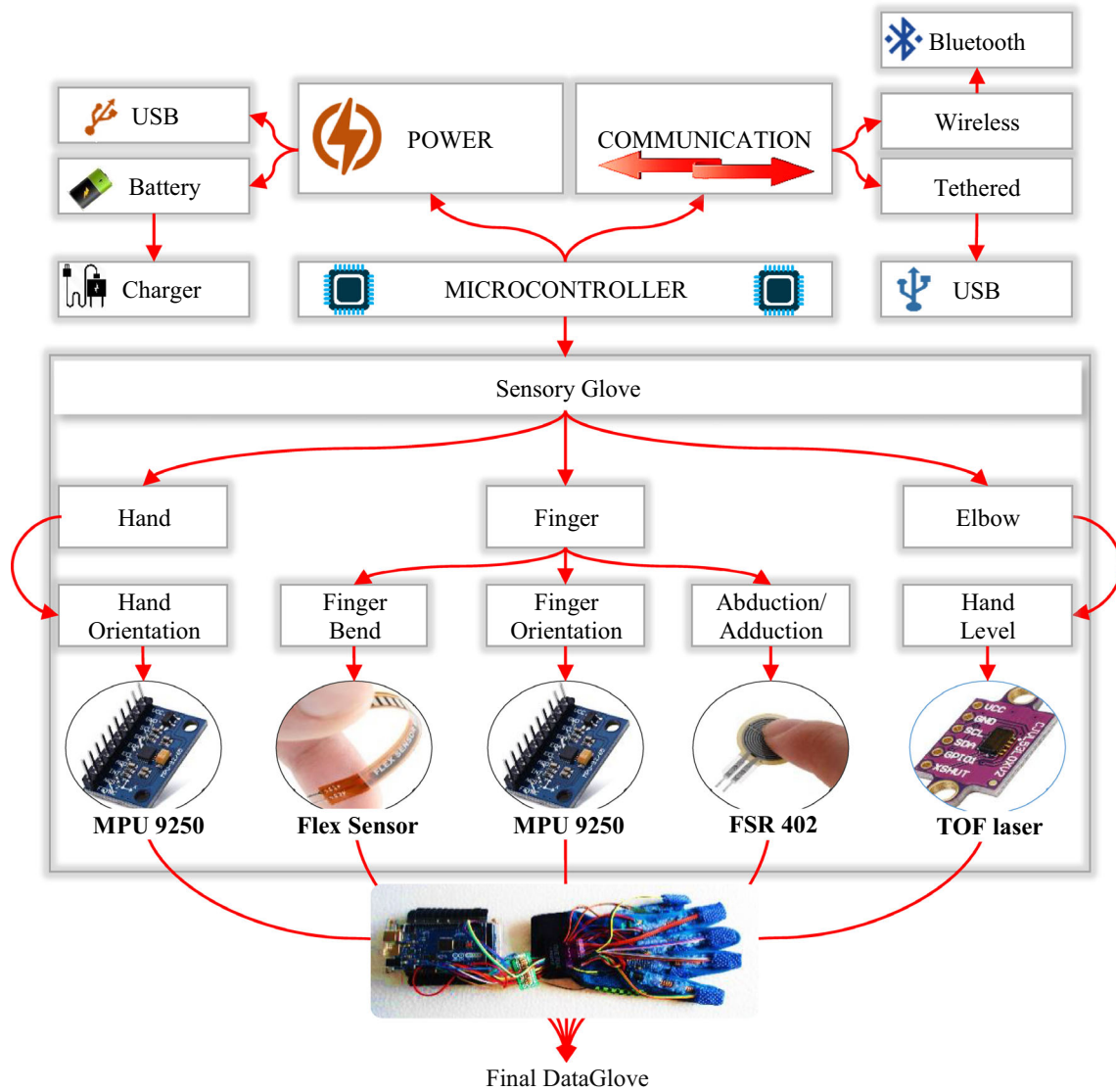
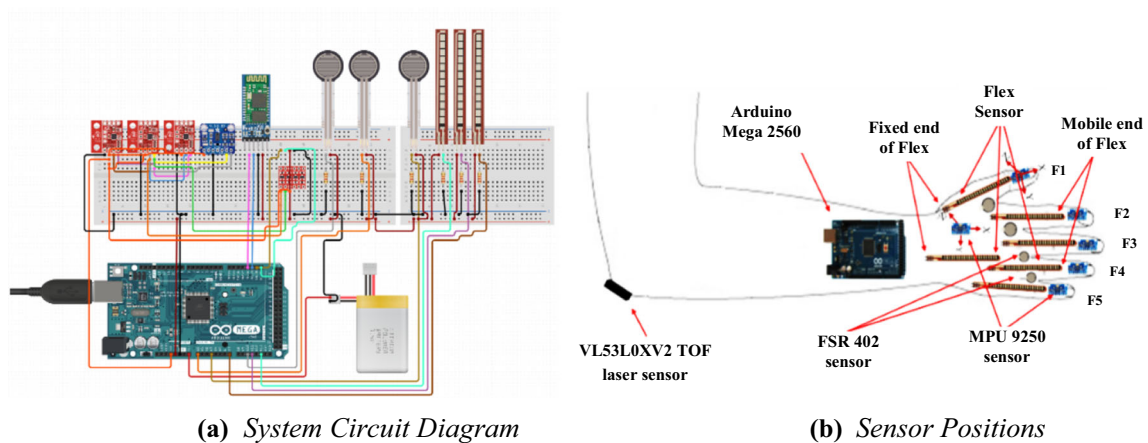**Fig. 11** Block diagram of proposed intelligent sensory glove

(a) *System Circuit Diagram*   **(b)** *Sensor Positions*

Fig. 12 New design of proposed data glove: **a** System circuit, **b** Sensor positions

Table 1 Bill of materials used to design flex-FSR-TOF and multi-IMU-based Data Glove (equipment prices quoted on 03/08/2020)

Name	Model	Quantity	Position	Cost
Flex sensor	Flex Sensor 2.2 Inches	6	Fingers and wrist	RM 32.00
IMU	MPU 9250	6	Fingers and wrist	RM 24.40
FSR	Force Sensing Resistor 402	4	Between four fingers	RM 29.30
TOF	VL53L0X	1	Elbow	RM 29.00
MCU	Arduino Mega 2650	1	Forearm	RM 40.60
Bluetooth module	HC-05	1	Arduino mega	RM 15.60
Multiplexer	TCA9548A I2C 8 Channel Multiplexer	1	Wrist glove	RM 18.00
Battery	1500 mAh Li-ion Battery	1	Arduino mega	RM 46.10

Table 2 Selected sample description

Sample	Description	Number of samples
Number	All numbers from 0 to 10 only	11
Alphabet	All letters of the MSL except j and z because they have dynamic motion	24
Word	Signs are generated using the right hand and the type of movement is static	60
Total number		95
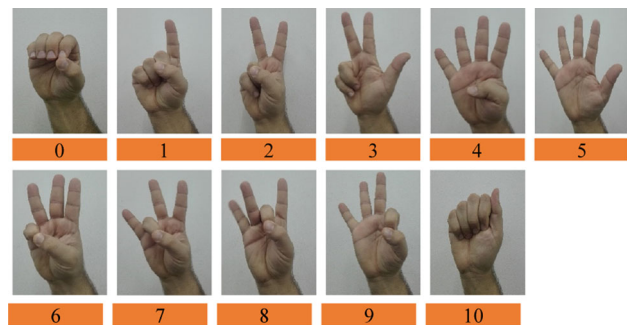
Fig. 14 Hand gestures of MSL numbers from 0 to 10

Fig. 13 26 Letter postures of MSL

3.4.3 Recognition technique selection

The desired outcome of developing a translation system is an effective and competent tool that provides support to people who need such a device to overcome difficulties and barriers in communicating with others. Accordingly, any system that is able to interpret SL in real time with fool-proof accuracy and reliability is an essential way to eliminate these issues. Accordingly, we select a template-matching method for static posture recognition, which operates on the basis of statistics to determine the closest match of acquired information values with predefined training samples called templates. In fact, this method is characterized by the lack of a need for complex training processes or computations, thereby increasing its speed. Consequently, template matching is the most suitable method for the recognition of the static gesture in real time.

The SLR software receives the values given by the flex, FSR, and TOF sensors, and IMUs through an Arduino Mega 2650 Microcontroller Board. To this end, recognition system software consists of two essential parts: the first part concerns system operation management, including hardware configuration and quantization of the captured sensory data. The second part concerns the artificial learning process of MSL gestures. The proposed system

Table 3 60 Static word gestures of MSL

ID	W1	W2	W3	W4	W5	W6
G						
M	Bapa	Emak	Jantan	Betina	Dato'	Datin
E	Father	Mother	Male	Female	Dato' (native chief)	Datin (wife of the dato')

ID	W7	W8	W9	W10	W11	W12
G						
M	Putera	Puteri	Nasi	Tebing	Babi	Mangkuk
E	Prince	Princess	Rice	Cliff	Pig	Bowl

ID	W13	W14	W15	W16	W17	W18
G						
M	Badak Sumbu	Salah	Selatan	Genggam	Batuk	Peniti
E	Rhinoceros	Wrong	South	Grasp	Cough	Pin

ID	W19	W20	W21	W22	W23	W24
G						
M	Leah	Bisu	Geraham	Sulit	Pekak	Leher
E	Leah (weary)	Mute	Molar	Difficult	Deaf	Neck

ID	W25	W26	W27	W28	W29	W30
G						
M	Vitamin	Siung	Buta	Tahu	Taring	Cekak
E	Vitamin	Clove	Blind	Know	Fangs	Short

ID	W31	W32	W33	W34	W35	W36
G						
M	Mata	Gigi	Bahu	Puntung	Senyap	Utara
E	Eyes	tooth	Shoulder	Thorns	Quite	North

ID	W37	W38	W39	W40	W41	W42
G						
M	Minum	Air	Makan	Sakit	Demam	Juruawat
E	Drink	Water	Eat	Sick	Fever	Nurse

ID	W43	W44	W45	W46	W47	W48
G						
M	Salah	Tidur	Dengar	Saya	Bulan	Babi
E	Wrong	Sleep	Hear	I	Moon	Pig

ID	W49	W50	W51	W52	W53	W54
|G						
M	Polis	Kenyang	Dengar	Pingat	Sirap	Ketuk
E	Police	Absolutely	Listen	Medal	Syrup	Knock

ID	W55	W56	W57	W58	W59	W60
G						
M	Terpengaruh	Tengah	Waktu	Boss	Kerajaan	Mangkuk
E	Influenced	Noon	Time	Boss	Government	Bowl

ID label of gesture, *G* gesture, *M* meaning in Malay language, *E* meaning in English language, *W* word

software is based on the statistical template-matching model. The entire model can be divided into three parts: calibration of the sensors, training of the model, and gesture recognition. Figure 16 shows the main steps of gesture recognition.

3.4.3.1 Initialization setup

Software initialization steps are coded in the setup function. The early-stage involved serial communication initialization where data rate (i.e., bits per second) for serial data transmission is set to create the communication channel with the PC. In the following stage, initialization of sensors, including sensor parameterization, registration of addresses, sensor initialization, and proper scaling are set and applied.

3.4.3.2 Sensor calibration

Calibration is the process of identifying the initial values and adjusting these data into similar or close values to reducing measurement errors. This process can be achieved by taking the minimum and maximum sensor values and then processing them (i.e., normalization and quantizing the values). Calibration of a particular sensor aims to convert the read sensor values into a predefined range of discrete datasets according to a scaled-down factor as in

$$N_i = (S_i - S_{i\min}) \times \frac{(R_{\max} - R_{\min})}{(S_{i\max} - S_{i\min})} + S_{i\min} \tag{1}$$

where N_i is the normalized value for the ith sensor, $R_{\max}$ and $R_{\min}$ define the required range, $S_{i(\min)}$ and $S_{i(\max)}$ are the read minimum and maximum sensor values, and S_i is the actual value of the ith sensor.

To address the difference in hand anatomy, before starting to record or perform any signs, the participant is asked to perform several actions for the calibration. The first action is stretching the hand (i.e., all fingers) as straight as possible for a few seconds, and then reading the maximum flex values of the straight fingers. Then, the second action is performed, which is closing the hand and bending all fingers as much as possible for a few more seconds. At this point, the minimum values of the flex sensor are recorded and saved. This value represents the finger bending. Similarly, other sensor data are collected for calibration.
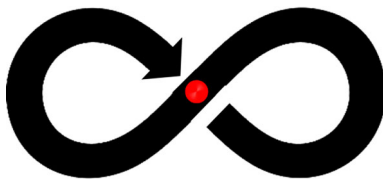

Fig. 15 8-shape rotation for magnetometer calibration

An IMU calibration procedure is necessary before starting to record raw data to guarantee improved IMU performance. In this study, the palm is placed horizontally to calibrate the ACC and Gyro. An average of 100 raw data measurements are recorded to extract the ACC_{offset} and $Gyro_{offset}$ for the ACC and Gyro, respectively. For the magnetometer calibration, the subject is asked to conduct an 8-shaped rotation (which means rotating the hand around a fixed point in the form of an infinity symbol; the shape remains exactly the same, but its position in space varies, as shown in Fig. 15) to obtain the minimum (Mag_{min}) and maximum (Mag_{max}) values per axis. The magnetometer offset (Magoffset) can be declared by calculating the average of Mag_{min} and Mag_{max}. The Acc', Gyro', and Mag' are the calibrated data that have been calculated by applying the equations using the offset inputted.

$$ACC' = ACC - ACC_{offset} \tag{2}$$

$$Gyro' = Gyro - Gyro_{offset} \tag{3}$$

$$Mag' = Mag - Mag_{offset} \tag{4}$$

3.4.3.3 Feature extraction

Concerning the flex sensor, by using the max and min values obtained from the calibration for each flex sensor, the new values of the sensors are transformed into three values: 0, 1, and 2. The sensor reading converts to 0 when the finger is fully flexed and the angle of the PIP joint (refer Fig. 7) is approximately 90°. Sensor reading turns to 1 when the finger is half-folded and at a joint angle of approximately 45°. While the finger is completely straight, the feature value is 2.

Regarding the pressure sensor, the purpose of using the sensor is to determine the touch between the fingers. Accordingly, when the sensor value is equal to 0 (no pressure is applied), the value of the feature is equal to 0. When the sensor detects pressure, an increment on the sensor value declares the touch. In this case, the value of the feature is equal to 1.

The triple component (ACC, Gyro. and magnetometer) of the IMU sensor serves to estimate the trajectory and motion of the hand to recognize the dynamic gestures and hand orientation. However, the ACC and Gyro values are sufficient to estimate the hand orientation as long as the process of recognition takes place on static gestures only. Therefore, the ACC and Gyro features are adopted in this study to determine the orientation of the hand. As a result of several experiments performed on the IMU sensor, we found that the best threshold value for this work is 0.35 for ACC and 1.25 for Gyro. Accordingly, the IMU features are calculated using the following equations:

Table 4 Feature descriptions

| Features	Descriptions
F1–F6	Flex features of thumb to pinky finger and wrist
C1–C4	Contact features of thumb-index, index-middle, middle-ring, and ring-pinky fingers
ax, ay, az, gx, gy, gz	IMU features of wrist include 3 axes ACC, 3 axes Gyro
Tax, Tay, Taz, Tgx, Tgy, Tgz,	IMU features of thumb include 3 axes ACC, 3 axes Gyro
Iax, Iay, Iaz, Igx, Igy, Igz,	IMU features of index finger include 3 axes ACC, 3 axes Gyro
Max, May, Maz, Mgx, Mgy, Mgz,	IMU features of middle finger include 3 axes ACC, 3 axes Gyro
Rax, Ray, Raz, Rgx, Rgy, Rgz,	IMU features of ring finger include 3 axes ACC, 3 axes Gyro
Pax, Pay, Paz, Pgx, Pgy, Pgz,	IMU features of pinky finger include 3 axes ACC, 3 axes Gyro
D1	Distance features for elbow
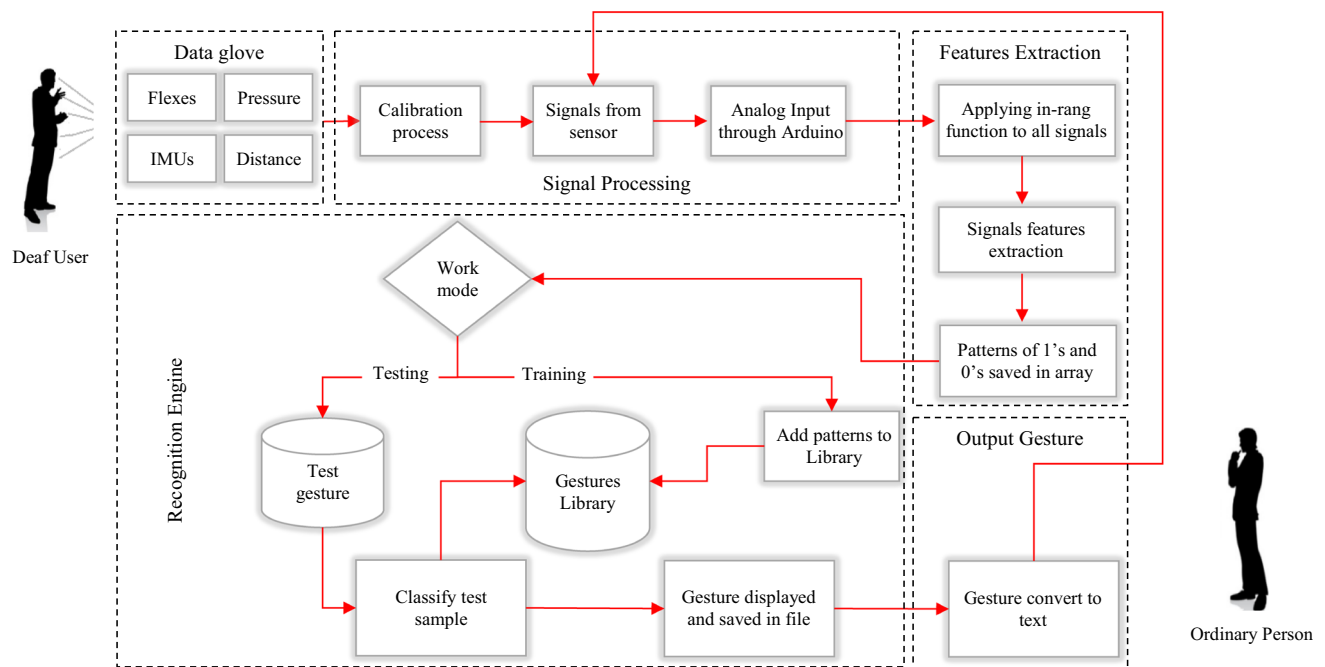
Fig. 16 Diagram blocks of MSL recognition

$$ax = \begin{cases} 1 & \text{if} x > 0.35 \\ 2 & \text{if } x < -0.35 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

$$mx = \begin{cases} 1 & \text{if} x > 1.25 \\ 2 & \text{if } x < -1.25 \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

The same procedure is followed for "ay" and "az" features. Table 4 lists the extracted features from various sensors (Fig. 16).

3.4.3.4 Gesture classification Template matching method based on pattern recognition is used in this study on

Fig. 17 Flowchart of gesture testing procedure
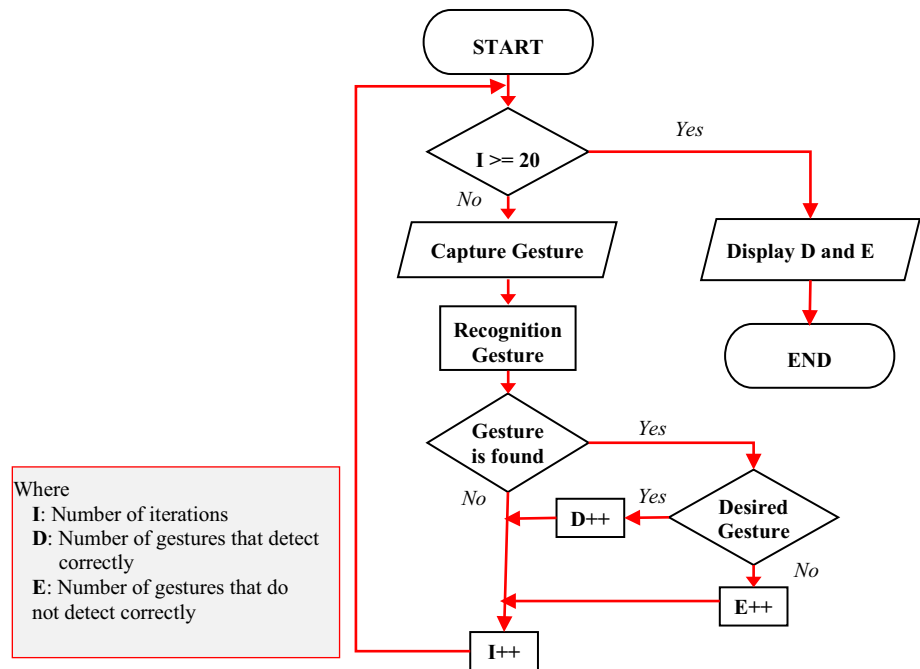
recognition approach. The pattern is represented by a collection of hand attributes as a feature vector. The recognition method operates in two modes: training and classification. In the training mode, the features are extracted from captured raw data. The next step in the training process is to store mean values for each reading sensor when a certain posture is signed. The classifier assigns the input pattern in the classification model to one of the pattern classes under consideration based on the extracted features. If the sensor values match the values that are already stored in the system, then the corresponding letter appears. Otherwise, the device continues to read the sensor values until a valid character is recognized.

3.5 Recognition evaluation and validation

For evaluation, the system performance measures are made concerning the reading error rate at the first stage. The flowchart in Fig. 17 demonstrates the approach used to measure the system performance metrics. Each letter is tested individually over five participants, where 20 iterations are applied to each letter to measure the frequency of recognition. Therefore, the performance of the proposed system can be measured by calculating the recognition accuracy of each gesture followed by the total accuracy for the entire system. However, such an erroneous result might be either "misclassification" or "gestured is not recognized," in other words, wrong detection or no detection, respectively. The accuracy and error rates are calculated using the following equations:

$$Accuracy\% = \frac{detected\ right}{Num.\ of\ itration} * 100 \tag{7}$$

$$Err\ of\ wrong\ reading\% = \frac{detected\ wrong}{Num.\ of\ itration} * 10 \tag{8}$$

$$Err\ of\ not\ detected\% = \frac{not\ detected}{Num.\ of\ itration} * 100 \tag{9}$$

The MSL recognition system should be validated to guarantee its quality. Therefore, the system performance is measured by different individuals. Three other participants are asked to perform the same signs in the same situation to compare the system performance with those of different individuals. Each participant performs each sign 20 times. Then, the accuracy and error rates of the system are calculated.

4 Comparative analysis with academic literature

Table 5 summarizes the main issues of the relevant research, including our study. These studies use three types of sensors (flex sensors, touch sensors, and IMU) to capture hand gestures of different SL. (Arif et al. 2016) proposed an ASL sign recognition system based on the sensory approach. The alphabet of ASL was collected using a glove equipped with five flex sensors that sense the finger actions, a three-axis ACC to differentiate among the static and dynamic tokens, a touch sensor, and a microcontroller to process the collected data. However, the system is evaluated by alphabet only. (Sharma et al. 2015) designed a data

Table 5 Benchmark summary of previous works and the present study

Reference	Language	Gesture	Finger Bend	Finger Touch	Hand orientation	Hand location	Total sensors	Channel
Shukor et al. 2015)	MSL	Nine gestures	Five tilt sensor		3-Axis ACC		6	8
Elmahgiubi et al. 2015)	ASL	20 Letters	Five flex sensors	Five contact sensors	One MPU6050	–	11	13
Sharma et al. 2015)	ASL	26 Letters	Nine 2.2- Inch Flex Sensors	Five contact sensors	One ACC ADXL335	–	15	17
Arif et al. 2016)	ASL	26 Letters	Five flex sensors	Two conductive foils	One ACC ADXL335	–	7	9
Vijayalakshmi and Aarthi 2016)	English alphabets	Eight Letters	Three 4.5-Inch and Two 2.2-Inch	One tactile sensor	One ACC	–	7	9
Tanyawiwat and Thiemjarus 2012)	ASL	26 Letters and two gestures	Five flex sensors	Five fabric contact sensors	3D ACC	–	11	13
Proposed system	MSL	95 Static gestures	Six flex sensors	Five pressure sensors	Six IMU MPU 9250	One TOF laser Sensor	17	65

glove consisting of five flex sensors, four contact sensors, and one ADXL335 ACC sensor for ASL letter recognition. However, this work provided a detailed explanation of the electronic system components rather than reporting the results of the system test, the method of evaluating the data glove or even indicating the accuracy of the system. (Tanyawiwat and Thiemjarus 2012) proposed a fused sensory concept. Five fabric contact sensors, five flex sensors, and one 3D ACC were utilized in constructing a data glove. However, this system was not able to recognize several characters with acceptable accuracy, where the accuracy of E, G, and N letters were 47.47, 48.55, and 42.82, respectively. (Vijayalakshmi and Aarthi 2016) developed a data glove equipped with five flex sensors, one three-axis ACC, and tactile (contact) sensor, and the system was evaluated using the English alphabet. However, the study selected eight letters from ASL, from A to H, for recognition. In Elmahgiubi et al. (2015), the system hardware consisted of five flex sensors, five contact sensors, and one 6-DoF IMU sensor, MPU6050. However, the system was evaluated by recognizing only 20 letters of ASL. (Shukor et al. 2015) developed a sensor to recognize MSL signs. The capture device was embedded with 10 tilt sensors to measure finger bending. Only nine gestures were used to evaluate the system including three letters A, B, C, and numbers from 1 to 3, besides three isolated words that are "saya," "Apa," and "makan". However, the recognition system was unable to solve the ambiguity of similar gesture problems concerning MSL such as U, R, and V characters and others.

In previous studies, the design of the sensory capture sign device was not based on SL analysis. Moreover, researchers have yet to come up with definitive solutions to similar gesture issues. In addition, 95 different gestures used daily are performed by five males and five females who have been selected for recognition. Our proposed system can capture all hand gesture information through 65 channels of data from 17 sensors as efficiently as possible and recognize similar gestures effectively.

5 Conclusion

This study aimed to eliminate the communication barrier between people and those with hearing disabilities. Therefore, a framework for MSL recognition is discussed with regard to the design of sub-modules alongside the details for each module. The framework discussed consists of three main sub-modules: first module (Analysis of MSL Modules) is related to extracting useful features to help identify system requirements and recognizing complex signs with high accuracy. The second module (DataGlove Modules) is concerned with developing an optimal Data-Glove based on sensor testing and system experimental results. The data collected by DataGlove are preprocessed and classified in the third module (Gesture Recognition Module). This development discussion can contribute to bridging the gap between hearing impaired and other people. Below is a summary of this research that describes current knowledge about the topic and the additional knowledge from the present study, alongside the related research objective and research questions.

In future research, an SL translation system can be extended to recognize complex signs such as dynamic gestures by taking advantage of the abundance of data generated by 17 built-in sensors and 64 data channels of the

DataGlove. In addition, the accuracy is improved and the time complexity for a real-time system application is minimized by deploying an efficient method of dimensionality reduction. Often, the conversation in SL consists of a series of continuous gestures. Thus, a mechanism is needed to enable the segment of the continuous signs to isolate gestures for recognition. Further research can be conducted to consider the recognition of SL based on deep learning techniques at static, dynamic, and sentence levels.

Summary points

What was already known about the topic	What this study added to our knowledge
Data suggest different approaches of gesture recognition for SL	This study focused on the recognition of SL gestures based on the sensory approach, which involves several issues and challenges not yet solved perfectly
	A comprehensive overview of recognition technology for SL gestures regarding the sensory approach based on (systematic) qualitative research
	Taxonomy of collected articles into groups, analysis of data on articles, highlights of important issues (i.e., motivation, challenges, and recommendations), and presentation of gaps in this area
Understanding the features of SL is essential for picture-perfect recognition. However, analyses of MSL have not been reported yet	First analysis process to MSL concerning gesture recognition based on the attributes and kinematics of the human hand
Information on finger bending is obtained by using either a flex sensor or IMU sensor. A combination of using both flex sensor and IMU sensor has not been presented yet	New design of the DataGlove enables this technique to capture all the hand attributes that are indispensable for the formation of any sign through 65 channels of data. The design of the DataGlove is based on the outcome of the language analysis and systematic sensor tests
Gestures of SL are formed based on hand attributes. Previous studies discussed and solved problems related to finger bending, finger touch, and orientation (or trajectory) of hand. However, issues related to hand position according to the signer's body have not been discussed yet	The DataGlove presents a solution for ambiguous gestures related to hand position issues

Declarations

Conflict of interest The authors declare no conflict of interest.

References

Abdulateef SK, Abdali T-AN, Alroomi MDS, Altaha MAA (2020) An optimise ELM by league championship algorithm based on food images. Indones J Electr Eng Computer Sci 20:132–137

Abhishek KS, Qubeley LCF, Ho D. (2016) Glove-based hand gesture recognition sign language translator using capacitive touch sensor. In: Electron devices and solid-state circuits (EDSSC), 2016 IEEE international conference on, pp. 334–337

Abualola H, Al Ghothani H, Eddin AN, Almoosa N, Poon K. (2016) Flexible gesture recognition using wearable inertial sensors. In: Circuits and systems (MWSCAS), 2016 IEEE 59th international midwest symposium on. pp. 1–4.

Adnan NH, Wan K, Shahriman A, Zaaba S, Basah SN, Razlan ZM et al (2012) Measurement of the flexible bending force of the index and middle fingers for virtual interaction. Procedia Eng 41:388–394

Aguiar S, Erazo A, Romero S, Garcés E, Atiencia V, Figueroa JP. (2016) Development of a smart glove as a communication tool for people with hearing impairment and speech disorders. In: Ecuador technical chapters meeting (ETCM), IEEE. pp. 1–6.

Ahmad MA (2016) Artificial neural network vs. support vector machine for speech emotion recognition. Tikrit J Pure Sci 21:167–172

Ahmed SF, Ali SMB, Qureshi SSM. (2010) Electronic speaking glove for speechless patients, a tongue to a dumb. In: Sustainable utilization and development in engineering and technology (STUDENT), 2010 IEEE conference on. pp. 56–60

Ahmed MA, Zaidan BB, Zaidan AA, Salih MM, Lakulu MMB (2018) A review on systems-based sensory gloves for sign language recognition state of the art between 2007 and 2017. Sensors 18:2208

Ahmed M, Zaidan B, Zaidan A, Salih MM, Al-qaysi Z, Alamoodi A (2021) Based on wearable sensory device in 3D-printed humanoid: a new real-time sign language recognition system. Measurement 168:108431

Ani AIC, Rosli AD, Baharudin R, Abbas MH, Abdullah MF. (2014) Preliminary study of recognizing alphabet letter via hand gesture. In: Computational science and technology (ICCST), 2014 international conference on. pp. 1–5

Anupreethi H, Vijayakumar S (2012) MSP430 based sign language recognizer for dumb patients. Procedia Eng 38:1374–1380

Arif A, Rizvi STH, Jawaid I, Waleed MA, Shakeel MR. (2016) Techno-Talk: An American Sign Language (ASL) Translator. In: Control, decision and information technologies (CoDIT), 2016 international conference on. pp. 665–670

Bajpai D, Porov U, Srivastav G, Sachan N. (2015) Two Way Wireless data communication and American sign language translator glove for images text and speech display on mobile phone. In: Communication systems and network technologies (CSNT), 2015 fifth international conference on. pp. 578–585

Bhatnagar VS, Magon R, Srivastava R, Thakur MK. (2015) A cost effective sign language to voice emulation system. In: Contemporary computing (IC3), 2015 eighth international conference on. pp. 521–525

Borghetti M, Sardini E, Serpelloni M (2013) Sensorized glove for measuring hand finger flexion for rehabilitation purposes. IEEE Trans Instrum Meas 62:3308–3314

Buczek FL, Sinsel EW, Gloekler DS, Wimer BM, Warren CM, Wu JZ (2011) Kinematic performance of a six degree-of-freedom hand model (6DHand) for use in occupational biomechanics. J Biomech 44:1805–1809

Bui TD, Nguyen LT (2007) Recognizing postures in Vietnamese sign language with MEMS accelerometers. IEEE Sens J 7:707–712

Bullock IM, Borràs J, Dollar AM. (2012) Assessing assumptions in kinematic hand models: a review. In: Biomedical robotics and biomechatronics (BioRob), 2012 4th IEEE RAS & EMBS international conference on. pp. 139–146.

Das P, De R, Paul S, Chowdhury M, Neogi B. (2015) Analytical study and overview on glove based Indian sign language interpretation technique

Dipietro L, Sabatini AM, Dario P (2008) A survey of glove-based systems and their applications. IEEE Trans Syst, Man, Cybern Part C (applications and Reviews) 38:461–482

Elmahgiubi M, Ennajar M, Drawil N, Elbuni MS. (2015) Sign language translator and gesture recognition. In: Computer & information technology (GSCIT), 2015 global summit on. pp. 1–6

P. O. P. Malaysia, "Bahasa Isyarat Malaysia," *Penerbit Persekutuan Orang Pekak Malaysia,* 2000.

Fu Y-F, Ho C-S (2008) Development of a programmable digital glove. Smart Mater Struct 17:025031

Gałka J, Mąsior M, Zaborski M, Barczewska K (2016) Inertial motion sensing glove for sign language gesture acquisition and recognition. IEEE Sens J 16:6310–6316

Gupta D, Singh P, Pandey K, Solanki J. (2015) Design and development of a low cost electronic hand glove for deaf and blind. In: Computing for sustainable global development (INDIACom), 2015 2nd international conference on. pp. 1607–1611.

Ibarguren A, Maurtua I, Sierra B (2009) Layered architecture for real-time sign recognition. Comput J 53:1169–1183

Ibarguren A, Maurtua I, Sierra B (2010) Layered architecture for real time sign recognition: hand gesture and movement. Eng Appl Artif Intell 23:1216–1228

Iwasako K, Soga M, Taki H (2014) Development of finger motion skill learning support system based on data gloves. Procedia Computer Sci 35:1307–1314

Jebali M, Dakhli A, Jemni M (2021) Vision-based continuous sign language recognition using multimodal sensor fusion. Evolv Syst. https://doi.org/10.1007/s12530-020-09365-y

Jianbin Q, Ji W, Lam H-K, Wang M (2020) Fuzzy-affine-model based sampled-data filtering design for stochastic nonlinear systems. IEEE Transactions on Fuzzy Systems

Kadam K, Ganu R, Bhosekar A, Joshi S (2012) American sign language interpreter. In: Technology for education (T4E), 2012 IEEE fourth international conference on. pp. 157–159

Kanwal K, Abdullah S, Ahmed YB, Saher Y, Jafri AR. (2014) Assistive Glove for Pakistani Sign Language Translation. In: Multi-topic conference (INMIC), 2014 IEEE 17th international. pp. 173–176.

Kau L-J, Su W-L, Yu P-J, Wei S-J. (2015) A real-time portable sign language translation system. In: Circuits and systems (MWSCAS), 2015 IEEE 58th international midwest symposium on. pp. 1–4

Kausar S, Javed MY (2011) A survey on sign language recognition. In: 2011 Frontiers of information technology. pp. 95–98

Khan S, Gupta GS, Bailey D, Demidenko S, Messom C. (2009) Sign language analysis and recognition: a preliminary investigation. In: Image and vision computing New Zealand, 2009. IVCNZ'09. 24th international conference. pp. 119–123.

Kong W, Ranganath S (2014) Towards subject independent continuous sign language recognition: a segment and merge approach. Pattern Recogn 47:1294–1308

Kortier HG, Sluiter VI, Roetenberg D, Veltink PH (2014) Assessment of hand kinematics using inertial and magnetic sensors. J Neuroeng Rehabil 11:70

Kosmidou VE, Hadjileontiadis LJ (2009) Sign language recognition using intrinsic-mode sample entropy on sEMG and accelerometer data. IEEE Trans Biomed Eng 56:2879–2890

Lee J, Kunii TL (1995) Model-based analysis of hand posture. IEEE Comput Graphics Appl 15:77–86

Lee CK, Ng KK, Chen C-H, Lau HC, Chung S, Tsoi T (2021) American sign language recognition and training method with recurrent neural network. Expert Syst Appl 167:114403

Lei L, Dashun Q. (2015) Design of data-glove and Chinese sign language recognition system based on ARM9. In: Electronic measurement & instruments (ICEMI), 2015 12th IEEE international conference on. pp. 1130–1134.

Maarif H, Akmeliawati R, Gunawan TS (2018) Survey on language processing algorithm for sign language synthesizer. Int J Robot Mech 4:39–48

K. P. Malaysia, "Komunikasi Seluruh Bahasa Malaysia Kod Tangan: Jilid 1," ed: Kuala Lumpur: Dewan Bahasa dan Pustaka, 1985.

Mehdi SA, Khan YN. (2002) Sign language recognition using sensor gloves. In: Neural information processing, 2002. ICONIP'02. proceedings of the 9th international conference on. pp. 2204–2206.

Oz C, Leu MC (2007) Linguistic properties based on American sign language isolated word recognition with artificial neural networks using a sensory glove and motion tracker. Neurocomputing 70:2891–2901

Oz C, Leu MC (2011) American sign language word recognition with a sensory glove using artificial neural networks. Eng Appl Artif Intell 24:1204–1213

Pezzuoli F, Corona D, Corradini ML (2021) Recognition and classification of dynamic hand gestures by a wearable data-glove. SN Computer Sci 2:1–9

Pradhan G, Prabhakaran B, Li C (2008) Hand-gesture computing for the hearing and speech impaired. IEEE MultiMed 15:20–27

Qiu J, Ji W, Rudas IJ, Gao H (2020) Asynchronous sampled-data filtering design for fuzzy-affine-model-based stochastic nonlinear systems. IEEE Trans Cybern

Ramli S (2012)"GMT feature extraction for representation of BIM sign language." In: Control and System Graduate Research Colloquium (ICSGRC), IEEE, 2012, pp. 43-48

Rastgoo R, Kiani K, Escalera S (2020) Sign language recognition: a deep survey. Expert Syst Appl 164:113794

Sharma D, Verma D, Khetarpal P (2015) LabVIEW based sign language trainer cum portable display unit for the speech impaired. In: India conference (INDICON), 2015 annual IEEE. pp. 1–6

Shukor AZ, Miskon MF, Jamaluddin MH, Ali FB, Asyraf MF, Bahar MBB (2015) A new data glove approach for Malaysian sign language detection. Proc Computer Sci 76:60–67

Sruthi C, Lijiya A. (2019) Signet: a deep learning based indian sign language recognition system. In: 2019 International conference on communication and signal processing (ICCSP), pp. 0596–0600

Swee TT, Ariff A, Salleh S-H, Seng SK, Huat LS. (2007) Wireless data gloves Malay sign language recognition system. In: Information, communications & signal processing, 2007 6th international conference on. pp. 1–4.

Swee TT, Salleh S-H, Ariff A, Ting C-M, Seng SK, Huat LS. (2007) Malay sign language gesture recognition system. In: Intelligent and advanced systems, 2007. ICIAS 2007. international conference on. pp. 982–985

Tang J, Cheng H, Zhao Y, Guo H (2018) Structured dynamic time warping for continuous hand trajectory gesture recognition. Pattern Recogn 80:21–31

Tanyawiwat N, Thiemjarus S (2012) Design of an assistive communication glove using combined sensory channels. In: Wearable and implantable body sensor networks (BSN), 2012 ninth international conference on. pp. 34–39

Tubaiz N, Shanableh T, Assaleh K (2015) Glove-based continuous Arabic sign language recognition in user-dependent mode. IEEE Trans Human-Machine Syst 45:526–533

Tyagi A, Bansal S (2021) Feature extraction technique for vision-based indian sign language recognition system: a review. Comput Methods Data Eng 1:39–53

Vijay PK, Suhas NN, Chandrashekhar CS, Dhananjay DK (2012) Recent developments in sign language recognition: a review. Int J Adv Comput Eng Commun Technol 1:21–26

Vijayalakshmi P, Aarthi M. (2016) Sign language to speech conversion. In: Recent trends in information technology (ICRTIT), 2016 international conference on pp. 1–6

Zamani M, Kanan HR. (2014) Saliency based alphabet and numbers of American sign language recognition using linear feature extraction. In: 2014 4th International conference on computer and knowledge engineering (ICCKE), pp. 398–403

Zhang X, Chen X, Li Y, Lantz V, Wang K, Yang J (2011) A framework for hand gesture recognition based on accelerometer and EMG sensors. IEEE Trans Syst, Man, Cybernetics-Part a: Syst Hum 41:1064–1076

Zhou Y, Jiang G, Lin Y (2016) A novel finger and hand pose estimation technique for real-time hand gesture recognition. Pattern Recogn 49:102–114