BULGARIAN ACADEMY OF SCIENCES

CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **24**, No 3
Sofia • 2024 Print ISSN: 1311-9702; Online ISSN: 1314-4081
DOI: 10.2478/cait-2024-0031

Real-Time Hand Gesture Recognition: A Comprehensive Review of Techniques, Applications, and Challenges

Aws Saood Mohamed[1], *Nidaa Flaih Hassan*[1], *Abeer Salim Jamil*[2]

[1]*Department of Computer Science, University of Technology, Baghdad, Iraq*
[2]*Department of Computer Technology Engineering, Al-Mansour University College, Baghdad, Iraq*
E-mails: cs.21.06@grad.uotechnology.edu.iq 110020@uotechnology.edu.iq abeer.salim@muc.edu.iq

Abstract**:** *Real-time Hand Gesture Recognition* (*HGR*) *has emerged as a vital technology in human-computer interaction, offering intuitive and natural ways for users to interact with computer-vision systems. This comprehensive review explores the advancements, challenges, and future directions in real-time HGR. Various HGR-related technologies have also been investigated, including sensors and vision technologies, which are utilized as a preliminary step in acquiring data in HGR systems. This paper discusses different recognition approaches, from traditional handcrafted feature methods to state-of-the-art deep learning techniques. Learning paradigms have been analyzed such as supervised, unsupervised, transfer, and adaptive learning in the context of HGR. A wide range of applications has been covered, from sign language recognition to healthcare and security systems. Despite significant developments in the computer vision domain, challenges remain in areas such as environmental robustness, gesture complexity, computational efficiency, and user adaptability. Lastly, this paper concludes by highlighting potential solutions and future research directions trying to develop more robust, efficient, and user-friendly real-time HGR systems.*

Keywords**:** *Computer vision, Hand gesture recognition, Real-time systems, Deep learning, Transformers.*

1. Introduction

As a natural form of human communication, hand gestures can convey information and express emotions without words [1]. In the past few years, a novel technology has been explored and developed rapidly in the field of Human-Computer Interaction (HCI), known as Hand Gesture Recognition (HGR). HGR systems allow machines to comprehend gestures made by hand and interpret them, providing a more intuitive interface for humans to communicate with computers and other devices [2]. HGR technology is more than just a modern tool in computer vision; it represents a large opportunity to improve accessibility and inclusivity. HGR would eliminate communication barriers in the world of sign language translation for the deaf-hearing community.

Recent advancements in deep learning and techniques such as computer vision have resulted in far more accurate and robust systems that can take complex hand gestures, and translate them into either a spoken or written form of language in real-time [3]. HGR empowers natural interaction with virtual environments in the domain of Virtual and Augmented Reality. HGR technology has been integrated into gaming and virtual training programs which places the user right in the action for a more life-like experience [4]. This expanding use case of HGR in the medical field positively impacts patient care and rehabilitation [5].

The integration of HGR with AI is one of the important innovations in the industry sector. HGR systems are being leveraged in manufacturing and automation to control robotic arms and machinery. This would lead to increased accuracy, efficiency, and safety at the workplace and decrease human error [6]. These systems have to deal with dynamic environments and complex tasks, which represents the reason Reinforcement Learning (RL) being used together with real-time feedback mechanisms is critical for modern industrial operations. Real-time HGR is crucial because it requires identifying, recognizing gestures accurately, and continuously in actual time in the presence of changing lighting conditions or within complex backgrounds. The real-time performance of HGR demands access to effective algorithms and hardware platforms that can parse vast amounts of data accurately and rapidly [7].

While recent significant advancements in AI and computer vision have taken place, the future of HGR technology depends on overcoming the current challenges and improving system reliability. for that, the continuous scientific research in machine learning, deep learning, sensor technology, and human-computer interaction would lead to addressing these challenges. By taking advantage of interdisciplinary approaches and encouraging collaboration between academic and industry sectors, the capabilities of HGR can be fully realized, leading the way for more intuitive and efficient interactions between humans and machines.

This paper organized to covers various technologies used for real-time hand gesture recognition, categorizing them based on various technologies such as sensor-based HGR, Vision-based HGR, Gesture modes-based HGR, Recognition approaches, and Learning paradigms. In addition to the working principles, key features, advantages and disadvantages, performance limitations, and recent advancements of each technology and approach will be discussed in the following sections:

2. Hand gesture recognition technologies

HGR technologies can be categorized based on the types of sensors and methods used to capture and analyze hand movement data, as follows:

2.1. Sensor-based HGR

HGR systems can utilize various sensors to capture hand movement data. These types of sensors can be classified as contact sensors and non-contact sensors.

2.1.1. Contact sensors

These sensors are worn on the hand or arm and directly measure various aspects of hand movement. Common types of contact sensors used for HGR are a data gloves sensors and ElectroMyoGraphy sensor.

- **Data gloves sensors.** They represent an advanced approach to gesture recognition, utilizing multiple sensors to capture three-dimensional spatial information of hand postures. These sensors incorporate multiple types of sensors, including flex sensors and inertial sensors such as accelerometers and gyroscopes to provide detailed hand pose data, e.g. Inertial Measurement Units (IMUs) that consist of these sensors can be used to increase the accuracy of gesture detection due to its capability for tracking hand movement as well as orientation. Research [8] has demonstrated that a 3-layer Convolutional Neural Network (CNN) applied to raw IMU data achieved a high success rate of 97.5% in detecting dual-handed gestures. Another research [9] has proposed an Attention-based CNN-BiLSTM Network (A-CBLN) for dynamic gesture recognition using data gloves. In this study involving 32 subjects and seven dynamic gestures, the A-CBLN achieved an impressive 95.05% accuracy and 95.43% precision on the test dataset.

While data gloves offer high accuracy and environmental versatility, they present challenges in terms of user comfort, cost, and potential restriction of natural hand movements. These factors may impact their widespread adoption in certain applications.

- **ElectroMyoGraphy (EMG) sensor.** It is used to monitor the electrical activity of muscles to perform hand gesture recognition. Research [10] has obtained 95-100% accuracy for five gestures and only two EMG channels using the SVM classification. In [11], an RNN model with LSTM could be trained using just four EMG signals to identify five gestures. This model achieved 87±7 % accuracy in real-time testing. Recent publications have focused on minimalistic approaches to improve practicality and efficiency, making EMG-based gesture recognition a promising step towards embedded systems or real applications. While limitations such as optimal placement of electrodes, noise immunity, and muscle fatigue remain challenges, these improvements hold promise to overcome some of them, opening a path to enhance human-machine interaction on different platforms.

2.1.2. Non-contact sensors

These sensors capture hand gestures from a distance without requiring direct contact with the hand. They are utilized in HGR systems in some cases where contact sensors are not feasible or desirable to apply in some applications. The common types of non-contact sensors are radar and Wi-Fi sensors.

- **Radar sensors.** Where short-range radar systems like Soli demonstrate real-time recognition of fine hand gestures. These systems offer advantages over optical cameras, functioning in low-light and occluded environments. Radar-based systems have achieved high accuracy in gesture recognition without requiring gloves. Using machine learning and signal processing techniques, these systems can generate range-Doppler maps and efficiently extract hand gestures, with some studies reporting classification accuracies up to 99.10% for 10 different gestures [12].

- **Wi-Fi sensors.** Wi-Fi signals have shown potential for gesture recognition, by analyzing changes in signal strength and phase caused by hand movements. This approach leverages existing Wi-Fi infrastructure, offering a non-intrusive and cost-effective method for gesture recognition. In addition, it is applied to two-hand gesture recognition and reaches a recognition accuracy of 95% [13]. While promising, this method requires sophisticated algorithms to extract gesture information from noisy Wi-Fi signals.

2.2. Vision-based HGR

Vision-based methods for hand gesture recognition involve utilizing cameras to capture and analyze hand gestures in real-time video streams. The most well-known cameras that are used in vision-based HGR are RGB cameras and Depth cameras. RGB cameras are commonly used for capturing visual data, enabling systems to extract features, recognize gestures, and support navigation tasks, which computer vision algorithms can use to extract features and recognize gestures [14]. RGB cameras are widely available and relatively cost-acceptable, but their performance can be affected by lighting conditions and background clutter. For that, depth cameras, offer a complementary approach by providing 3D spatial information. The combination of RGB and depth information enables more sophisticated gesture recognition systems [15]. Depth cameras, such as Microsoft Kinect, ASUS Xtion, and Mesa SwissRanger utilize RGB cameras and specialist technologies such as Time of Flight (ToF) and stereoscopic imaging, leading to improved performance within low-light and complex environments compared to traditional video cameras. The depth data; which is extracted from depth cameras, can be fed to deep learning algorithms and present high accuracy of gesture segmentation and recognition [16].

The analysis of captured images or video streams employs various techniques such as skin color detection, background subtraction for hand-palm identification, template-based tracking, and deformable contours for tracking hand regions across frames [14]. These techniques can be applied to a wide range of applications, including sign language recognition, digit recognition, and even healthcare systems enhancement through gesture recognition [17].

While combining computer vision frameworks with innovative strategies like iterative polygonal shape approximation and chain-coding schemes might achieve acceptable accuracy rates for recognizing hand gestures corresponding to different symbols and digits, more recent advancements in machine learning and deep learning have introduced new methods that may outperform these traditional computer vision techniques in wide scenarios [18].

Researchers in a paper [19] proposed a face-authenticated hand gesture-based human-computer interaction system for desktops. Their approach uses the Viola-Jones Algorithm for face recognition and authentication, followed by a CNN for HGR. The system achieved high accuracy in recognizing various hand gestures and was able to perform basic operations on a laptop. This demonstrates the potential of combining facial recognition for security with gesture control for intuitive device interaction.

3. Gesture modes in HGR

The technologies described in Sections 2.1 and 2.2 are utilized for collecting data that is regarded as a first step toward building HGR systems. These systems can be designed depending on the types of gesture modes, which can be categorized into static, dynamic, and hybrid modes.

3.1. Static gestures

In hand gesture recognition, static gestures refer to analyzing still hand positions for classification. This method is computationally less expensive compared to dynamic gestures which need continuous image sequence processing. Artificial Neural Network (ANN) paired with image segmentation techniques can produce highly accurate results for static gestures [20]. Static hand gestures are important for several applications such as simple sign recognition the complex robotic surgery. This can improve the accuracy and robustness of the classification in static hand gestures using new data collection methods and innovative training techniques for CNNs [21]. The HGR systems based on static gesture mode have an important role in human-computer interactions, robotics, and so on, indicating a high necessity to develop workable recognition algorithms.

3.2. Dynamic gestures

Dynamic gestures refer to analyzing continuous hand positions for classification. The deep learning approaches have propelled dynamic hand gesture recognition to a remarkable extent, though traditional methods such as Hidden Markov Models (HMA) and dynamic time warping are still applicable [22]. The domain has largely moved on to more sophisticated neural network architectures. Regarding capturing both spatial and temporal aspects of dynamic gestures, Convolutional Neural Networks (CNNs) paired with recurrent structures such as Long Short-Term Memory (LSTM) have achieved state-of-the-art results. CNN-LSTM hybrids have been proven to be effective at learning complex spatiotemporal features directly from raw input data [23].

3.3. Hybrid gestures

Hybrid gestures mode-based HGR systems attempt to combine different types of gestures (static and dynamic), e.g.,: a sign language gesture will need the user to create a defined shape with her fingers static part and then exactly move their hands [24]. Hybrid gestures typically demand systems that effectively handle both spatial (tracking ability) and temporal information. It could also leverage multiple modalities for better discriminative and resilient recognition [25].

The choice of gesture type depends on the specific application and its requirements for expressiveness and complexity. Static gestures are suitable for simple commands and interactions, while dynamic gestures can be used for more complex and nuanced communication.

4. Recognition approaches and learning paradigms

Hand Gesture Recognition (HGR) systems employ various approaches and learning paradigms to effectively interpret and classify gestures. This section explores the interconnection between recognition methods and the learning paradigms used to train and optimize these systems.

4.1. Handcrafted feature-based methods and supervised learning

Handcrafted feature-based methods in hand gesture recognition involve extracting features manually to represent gestures effectively [26]. These methods encompass global features and local features.

4.1.1. Global features

Global features describing overall hand morphology and movement of the hand over time, such as hand trajectory, velocity, and acceleration. These features are essential for recognizing dynamic gestures. Global features include techniques like color histograms, grayscale histograms, and Gabor filters.

4.1.2. Local features

Local features focus on specific hand areas such as fingertip positions, palm center location, hand orientation, and finger angles. These features are relatively easy to extract and can be effective for recognizing simple gestures. Examples of local features are Histogram Oriented Gradients (HOG), Local Binary Patterns (LBP), and Scale-Invariant Feature Transform (SIFT).

In general, Human Activity Recognition (HAR) applications have been using handcrafted feature-based methods to provide an informative set of features for best model performance. Nonetheless, choosing and verifying features by hand is both time-consuming and ineffective when compared to automatic methods of feature extraction. Recent work in image processing has shown promising results in color-based feature extraction that could be adapted for hand gesture recognition. Researchers in the paper [27] demonstrated the effectiveness of color histograms in extracting and differentiating color distributions in images. They applied various clustering techniques, including Fuzzy c-Means and a hybrid approach combining agglomerative hierarchies and k-Means, before using color histograms to identify color features. Moreover, similar approaches could be used for extracting color-based features of hands, which might lead to improving gesture detection and classification, especially in scenarios with varying skin tones or lighting conditions. This method could complement existing handcrafted feature extraction techniques like HOG or SIFT, especially for vision-based HGR systems.

The extracted features via handcrafted feature-based methods are typically utilized by supervised learning techniques, where algorithms are trained on labeled data to classify gestures accurately [28]. These algorithms such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), k-Nearest Neighbors (kNN), and Hidden Markov Models (HMM) are essential for interpreting the semantics conveyed by hand gestures, whether they are static postures or dynamic movements

[29]. That is why supervised learning serves as the cornerstone for training models to accurately detect and classify hand gestures in real-life settings.

4.2. Deep learning-based methods and transfer learning

Deep learning techniques, especially Convolutional Neural Networks and Recurrent Neural Networks (CNNs & RNNs) have been extensively used in HGR research due to their ability to feature generation from the raw data [30]. Deep learning models are typically trained on large datasets of labeled hand gestures, allowing them to learn features that best discriminate between different gestures. These models can be computationally intensive and often require significant amounts of training data. For this issue, deep transfer learning provides a solution for this issue where autonomous feature extraction and better performance can be achieved. Transfer learning can be used for hand gesture recognition applications to provide higher accuracy and efficiency. It can be applied to learn various features from gesture images which significantly increase recognition rates [31]. This way operates without manual feature extraction which may result in more accurate and robust recognition systems. In research [32], adaptive deep transfer learning is developed for gesture recognition with soft e-skin patches, reducing training data requirements and time while maintaining high accuracy levels. Moreover, these models show the effectiveness of transfer learning in enhancing hand gesture recognition systems and also demonstrate that transfer learning has enabled models to generalize better with limited training data.

The common deep learning algorithms used in HGR can be classified as Convolutional Neural Networks (CNNs), lightweight convolutional neural networks, and RNNs and sequence learning.

4.2.1. Convolutional Neural Networks (CNNs)

CNNs have shown excellent results for image data because of their ability to capture spatial information present in hand gesture images [33]. This proves that the power of CNN has also been in other domains like face detection, which has achieved an accuracy of 99.5% [34].

Furthermore, researchers in paper [35] proposed a 3D-CNN model designed for recognizing drivers' hand gestures in complex depth and intensity data. later enhancing it with a recurrent mechanism for dynamic gesture detection and classification. The 3D-CNN is used for the extraction of spectral and spatial features [36]. Another research [37], proposed three types of very deep 3D CNNs for gesture recognition, which can directly model the spatiotemporal information with their inherent hierarchical structure. The proposed method is evaluated on three challenging datasets, Ego Gesture, Jester, and Chalearn-IsoGD, and achieves state-of-the-art performance on all of them.

4.2.2. Lightweight convolutional neural networks

Recent years have seen the development of lightweight CNNs, offering performance comparable to heavier models while being more hardware-friendly. Researchers in a paper [17] proposed a lightweight, robust, and fast CNN for manual gesture recognition by image classification. The proposed system achieved 99.96% accuracy.

A hybrid network structure combining a lightweight VGG16 model and a random forest was presented for visual input-based gesture recognition.

Researchers in another paper [38] proposed an efficient DenseNet model that utilizes a fusion of channel and spatial attention for facial expression recognition. This approach combines densely connected convolutional layers with attention mechanisms to enhance feature extraction while reducing model complexity. The model achieved high accuracy (99.94-100%) on several lab-controlled datasets while maintaining a relatively low parameter count of 4.27 million. Such techniques show the capability for developing more efficient and accurate CNN models for real-time hand gesture recognition.

4.2.3. RNNs and sequence learning

RNNs are designed to process sequential data, demonstrating remarkable efficiency in capturing temporal dependencies within hand gesture sequences [39]. While traditional RNNs faced challenges with long-term dependencies, modern variants such as Long Short-Term Memory (LSTM) networks have largely overcome these limitations [40]. LSTM networks, excel at sequence learning, making them well-suited for capturing temporal dependencies in dynamic hand gestures. LSTMs have shown remarkable performance in handling extended sequences, making them particularly valuable in gesture recognition tasks where gestures may span varying durations [41].

4.3. Vision transformer-based methods and self-supervised learning

Vision Transformer (ViT) models are revolutionizing hand gesture recognition by leveraging self-attention mechanisms, allowing the model to capture complex relationships between image patches. This approach enhances accuracy in recognizing gestures despite variations in pose, lighting, and background. Originally designed for image classification tasks. ViT offers a novel approach compared to traditional CNNs [42]. Recent research has proposed specialized models like HGR-ViT, demonstrating exceptional performance across various hand gesture datasets [42]. In some applications, ViT can be trained using self-supervised learning techniques, where the model learns useful representations from unlabeled data before fine-tuning on a specific task [43].

Researchers aim to enhance human-computer interaction by employing ViT in hand gesture recognition systems, particularly in educational settings and sign language recognition. The combination of ViT with other models, such as CNNs, shows promise in improving the recognition of fine-grained details [44].

4.4. Unsupervised and adaptive learning in HGR

The role of unsupervised learning is paramount in hand gesture recognition systems specifically when it comes to identifying and labeling the fingers as open without any additional manual annotation tags. It can use the depth data of RGB-D sensors for the identification of hand shapes, based on finger extension and makes phase-agnostic recognition in real-time. The unsupervised learning techniques utilize temporal feature extraction from motion profile sequences in hand gesture recognition, which helps to improve the performance of any real-time systems. With unsupervised learning algorithms, these systems are well suited to process spatial and temporal data in an efficient way that allows reasonably accurate hand gesture classification without any human interaction or laborious supervised training steps [45].

In addition to the contributions of unsupervised learning, adaptive learning in HGR involves systems that can dynamically adjust and improve their recognition capabilities based on changing conditions. Various systems have been proposed to enhance gesture recognition accuracy [46]. Wearable biosensing system with in-sensor adaptive learning for real-time gesture classification and model updates. The feature adaptive learning in part designed to accommodate the continuous changes of electrode configurations, offers sEMG-based gesture-recognition with high density and evolves accuracy over evolving environments [28].

Table 1 shows the most powerful techniques of HGR on well-known datasets and the accuracy achieved for each of them.

Table 1. The most powerful techniques of HGR

Reference	Technique / Algorithm	Accuracy	Dataset	Year
[35]	Recurrent 3D CNN	83.8 %	Custom Dataset captured with depth, color, and stereo-IR sensors	2016
[56]	3-D convolution and convolutional LSTM	98.89 %	Sheffield Kinect gesture (SKIG) data set	2017
[66]	C3D+LSTM+RSTTM	92.2 %	EgoGesture data (RGB-D)	2018
[21]	Lightweight CNN+ ResNeXt-101	94.04 %	EgoGesture dataset	2019
[37]	Deep 3D CNNs	98.5 %	EgoGesture dataset	2020
[3]	Lightweight CNN (MobileNetV2)	99.96 %	Custom dataset	2021
[18]	Lightweight VGG16 + Random Forest	99.98 %	The American Sign Language (ASL) dataset	2022
[42]	HGR-ViT	99.85 %	National University of Singapore (NUS) hand gesture	2023
[67]	GmTC (Graph and General Deep-Learning Network)	99.10 %	LSA64 dataset	2024

Tables 2 and 3 summarize the technologies and approaches which have been covered in the previous sections, as shown in the following context.

Table 2. The recent technologies of HGR with pros and cons

Technology	Sensor type	Gesture type	Pros	Cons
Data Gloves and EMG	Contact-based	Static, dynamic, hybrid	High accuracy, Direct measurement of hand/muscle movements	Inconvenient, Requires wearable hardware
RGB Cameras	Vision-based	Static, dynamic, hybrid	Low-cost, Unobtrusive	Sensitive to lighting conditions, Background clutter
Depth Cameras: (ToF, Structured Light)	Vision-based	Static, dynamic, hybrid	Robust to lighting variations, 3D information	More expensive than RGB, Computational overhead
Radar	Contactless	Dynamic	Not affected by lighting/occlusions, Can detect through obstructions	Limited gesture vocabulary, Signal processing complexity
Wi-Fi	Contactless	Dynamic	Leverages existing infrastructure, Low-cost	Noisy signals, Limited gesture recognition

Table 3. The recent approaches of HGR with advantages and limitations

Approach	Applicable to	Key features	Advantages	Limitations
Handcrafted features	Various sensors	Manual feature design	Interpretable features, Computationally efficient	Effort-intensive, May not generalize well
Deep learning: (CNNs, RNNs)	Vision-based sensors	Automatic feature learning	Automatic feature, High accuracy, Adaptable	Requires large training data, Computationally intensive
Vision transformers	Vision-based sensors	Self-attention mechanism	Captures long-range dependencies, State-of-the-art accuracy	Higher computational cost, Requires large datasets

5. Applications of real-time hand gesture recognition

Real-time hand gesture recognition has become important and highly useful for many spheres of life, because it allows human-machine interaction to be more natural and easier. The following part covers the major application areas with their respective use cases, where it also provides context to requirements and critical scenarios imposed on real-time HGR.

5.1. Human-Computer Interaction (HCI) and devices control

HGR is a well-developed sub-field of Human-Computer Interaction (HCI), aiming to understand and interpret human hand movements to control computers and various devices such as robots, smartphones, tablets, laptops, and smart home appliances, without the need for physical contact. This can be particularly useful for situations where touch-based interaction is inconvenient or unhygienic [22]. HGR can be used also to control assistive devices for disabled people like wheelchairs, and prosthetic limbs. This technology has the capacity to provide more natural and grounded interactions between human systems, spawning new paradigms beyond traditional computer input metaphors (e.g., keyboards and mice) [47]. Some of the common interesting HCI applications are Virtual Reality (VR) and Augmented Reality (AR).

VR and AR are two revolutionary technologies in computer vision that offer human-computer interaction through immersive experiences and intuitive interfaces. VR such as those incorporating hand gesture recognition and 3D stereoscopic projection, enhance user engagement in tasks like virtual manipulation and surgical operations [48]. AR and VR technologies benefit from object recognition technology, seamlessly integrating virtual content with the real world for interactive experiences. The integration of hand gesture recognition technology in Virtual Reality (VR) and Augmented Reality (AR) gaming applications, such as through devices like Knuckles controllers and MYO armbands, further exemplifies the significance of HCI in gaming [49]. These advancements in VR and AR technologies are reshaping entertainment, training simulations, and various other fields reliant on immersive virtual experiences.

5.2. Sign language recognition and translation

HGR is essential for developing sign language recognition and translation systems, which can bridge the communication gap between deaf and hearing individuals.

These systems capture and interpret sign language gestures, translating them into spoken or written language in real-time. HGR systems can improve the real-time performance and mutual understanding between sign language experts and spoken/written users by combining sensor-based with vision-based approaches [13].

For other works, HGR is improved using different technologies including smart wristbands with gesture recognition [48], vision-based systems [50], and deep learning approaches for sign language applications. They use sensors, machine learning algorithms, and neural networks to identify/manipulate hand gestures matching with right letters or words. With CNN and transfer learning techniques, researchers have obtained high accuracy in the recognition of gestures helping to communicate efficiently for sign language proficient person [51].

5.3 Healthcare and medical applications

Several health applications of HGR have been explored in research [15, 52], including medical image navigation and rehabilitation. Surgeons navigate and control the sterile environment during surgery using hand gestures with which they manage medical images. HGR interprets gestures in real time to manipulate objects within medical data visualization environments using vision-based systems with CNN models. Such systems provide a non-touch input method enabling more natural interactions with sensor-enabled devices that could used also to enable remote physical examination and robotic-assisted surgery as they achieve increased precision and control.

Another health application of HGR enables users to track and monitor their rehabilitation activities which can be used for physiotherapy providing the patient a real-time feedback. Additionally, hand movements detected by HGR are being used to monitor therapeutic exercises for stroke patients or arm rehabilitation solutions for motor-impaired people during Rehab programs. HGR can be used to control assistive devices for disabled people like wheelchairs, and prosthetic limbs.

5.4. Security and authentication

HGR is a viable form of user authentication and identification, proving to be more secure as well as changing the way in which one thinks about HGR compared with traditional passwords or PINs [53]. Gesture-based user authentication is one of the major applications in gesture recognition and it refers to a scenario where users perform predefined hand gestures that are then analyzed against enrolled templates for their corresponding approvals [54]. Many researchers have worked on gesture modalities like static hand gestures, dynamic hand gestures or even 3D-depth sensor-based three-Dimensional (3D) hand motion [55, 56].

Apart from conventional authentication setups, HGR has also been explored for continuous authentication, i.e., whenever a user interacts with the device in a normal way coverage Task [57]. It can improve security by catching fake imposters or potential unauthorized accesses in real-time. Further, HGR can be complemented with other biometric factors, i.e., facial recognition or voice based on different modalities, to assure full security for multimodal authentications, which will enhance the overall accuracy and durability.

5.5. Other emerging applications

HGR can also be explored in other emerging applications [48], such as smart home and office environments, where hand gestures can be used to operate appliances or peripherals. In the automotive industry, in-vehicle systems can be controlled through hand gestures, enhancing driver safety and convenience. The education and entertainment sectors are also leveraging HGR for interactive learning experiences and innovative entertainment applications. These diverse applications demonstrate how HGR is breaking down barriers in human-machine interaction across various domains, each with its unique requirements and challenges that need to be addressed for successful implementation.

6. Future challenges and directions

Real-time hand gesture recognition has come a long way over the past years. However, there are still many challenges in creating and deploying practical systems. This part will highlight these challenges and potential solutions as well as future research directions.

6.1. Environmental factors

Environmental factors such as illumination changes, complex backgrounds, and occlusions can significantly impact the performance of vision-based HGR systems. Changes in lighting can affect hand appearance and feature extraction [58]. Complex backgrounds make it difficult to segment the hand from its surroundings, potentially causing false positives and misclassifications [59]. Occlusions, whether by other objects or self-occlusion, can hinder gesture tracking and recognition [60].

Potential solutions are:

- **Feature invariance.** Creating features that are immune to changes in lighting, background clutter, and occlusions.
- **Multi-sensor fusion.** That is combining the information from several sensors (such as RGB and Depth cameras) would enable more robustly solving open problems.
- **Adaptive background modelling.** This means the background model will dynamically adapt to changes in the environment.

6.2. Gesture complexity and diversity

The high Degrees Of Freedom (DOF) of the human hand allows for a wide range of possible gestures, making it challenging to build models that can recognize all conceivable representations [60]. Inter-user variability, where different people perform the same gesture differently, and intra-user variability, where a single user may perform the same gesture differently over time or in various contexts, can impair recognition accuracy [61].

Potential solutions are:

- **Size & heterogeneity of datasets.** Training models on many varied examples that span the heterogeneity between one user to another (e.g., as well as variations within users over time).

- **Adaptive Learning** is when you create models that learn from individual users, and learning improves over time as the conditions or system changes.
- **Context-aware recognition.** Incorporating context such as the user's environment or recent activities can be useful to disambiguate gestures with one another and contribute towards better recognition accuracy.

6.3. Computational constraints and real-time performance

Deep learning models for HGR can be computationally expensive, posing challenges for deployment on resource-constrained devices such as wearable sensors and mobile phones [62]. Real-time HGR systems must process and recognize gestures quickly enough to avoid perceptible latency during user interaction, which is particularly challenging for complex gestures or resource-limited devices [63].

Potential solutions are:

- **Building of lightweight models.** Frameworks for creating efficient and compact deep learning models that don't require as many computational resources Hardware.
- **Acceleration.** Utilizing specialized hardware available in modern computers such as GPUs, TPUs, or FPGAs to accelerate the process of using deep learning model implementations.
- **Edge computing.** Some computation can be offloaded to edge devices like wearable sensors, which in turn reduces latency and improves real-time performance.

6.4. Data scarcity and overfitting

Deep learning models require large amounts of labeled data for training. The collecting process and labeling of hand gesture data can be time-consuming and expensive [11]. Also, when trained on limited data, models are prone to overfitting, leading to poor generalization for unseen gestures [58].

Potential solutions are:

- **Data augmentation.** Utilizing techniques like image transformations and synthetic data generation to grow the training datasets both in size and diversity.
- **Transfer learning.** Relying on pre-trained models to lower the data we need for training.
- **Regularization techniques.** Regularisation techniques such as dropout and weight decay can also protect against overfitting.

6.5. User independence and adaptability

Developing HGR systems that can generalize to new users without requiring user-specific training, In addition to creating models capable of adapting to changes in user behavior and the real world are significant challenges [61].

Potential solutions are:

- **User-independent feature extraction.** Producing features that are invariant to the sizing, shape, and form of the hand motion.
- **Adaptive learning.** Needing to develop models that can learn and adapt themselves continuously with emerging data points as well as changing conditions.

- **Personalization techniques.** This refers to providing options for users who wish to configure the system based on their gesture preferences in detail.

For all mentioned above of challenges, the suggested solutions if implemented in professional approaches will help in building more robust, efficient, and user-friendly real-time hand gesture recognition systems that can serve several applications.

7. Experimental evaluation

In this section, common evaluation metrics for gesture recognition and segmentation are explored.

7.1. Accuracy

Accuracy is the percentage of the total number of correctly classified samples out of the total number of samples available. Accuracy rate is a metric that can be used to measure the entire classifier's performance in gesture recognition and segmentation. The calculation formula is

$$\text{Accuracy} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}, \tag{1}$$

where TP refers to True-Positive instances, TN to True-Negatives, FP to False-Positives, and FN indicating for False-Negatives.

7.2. Precision

It is the fraction of relevant instances among all mined positive instances A precision rate is a measure that can be used to characterize the accuracy of gesture identification/classification. The calculation formula is

$$\text{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}. \tag{2}$$

7.3. Recall

The recall rate is what proportion of samples that are in positive classes that were detected by the classifier. In gesture recognition and segmentation, recall can be used to measure the comprehensiveness of the classifier. The formula is

$$\text{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}. \tag{3}$$

7.4. F1-score

Using accuracy and recall in calculating values, the F1-score is a sum of them (average) so it evaluates both binary data classification systems where evaluating with only one criterion could provide no clue about completeness. In gesture tracking and segmentation, for instance, the F1-score can serve as an evaluation-pursuit metric to determine which classifier is best. The calculation formula, as in the next equation,

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{4}$$

7.5. Intersection over Union (IoU)

The overlapping area of the predicted region and the real one to total size. IoU has been applied to measure the segmentation performance of a model for gesture segmentation. The calculation formula is

$$\text{IoU} = \frac{\text{Intersection}}{\text{Union}}, \tag{5}$$

where the intersection of the predicted frame and true frame, is divided by the merging of said frames. The larger the intersection, the closer the applied result is to the expected outcome.

Most of the mentioned papers in this comprehensive review have utilized these metrics to evaluate the adopted models in their research. For that, to illustrate how to use these metrics, consider the following example.

Example. Suppose a hand gesture recognition system evaluated on a test set of 1000 gestures across 10 classes, and suppose the system has correctly identified 920 gestures, with the following breakdown:

True positives TP = 920, false positives FP = 50, false negatives FN = 80, and true negatives TN = 7950.

Then the results would be as follows:

- Accuracy = (TP + TN)/(TP + TN + FP + FN) =
  = (920 + 7950)/(920 + 7950 + 50 + 80) = 0.987 = 98.7%.
- Precision = TP/(TP + FP) = 920/(920 + 50) = 0.948 = 94.8%.
- Recall = TP/(TP + FN) = 920/(920 + 80) = 0.920 = 92.0%.
- F1-score = 2×(Precision×Recall)/(Precision + Recall) =
  =2×(0.948×0.920)/(0.948 + 0.920) = 0.934 = 93.4%.

Moreover, when evaluating real-time hand gesture recognition systems, it is crucial to consider both accuracy and processing speed. As researchers in paper [64] showed in their study of real-time face detection algorithms, there is a trade-off between these factors. They found that while the Viola-Jones algorithm was faster, YOLO v3 achieved higher accuracy, especially in challenging conditions. For hand gesture recognition, similar considerations could apply. A system might need to make a trade-off between high accuracy among various gesture types and environmental conditions with the ability to process input quickly enough for real-time interaction.

Therefore, when evaluating the HGR systems, it's important to measure not only overall accuracy but also performance metrics such as processing time per frame/Frames Per Second (FPS). This comprehensive evaluation approach can provide a more complete picture of a system's suitability for real-world applications, where both accuracy and responsiveness are important.

8. Conclusion

While various advancements have been made in sensing technologies, recognition algorithms and learning paradigms, including CNNs, RNNs, and the Vision Transformer, have proven to be state-of-the-art in real-time HGR. However, much of this progress still faces unsolved challenges, such as the issue due to environmental conditions like vision occlusions, wearables around the body, gesture complexity, and computational constraints, especially in a multi-modal convolutional RNN scenario which requires heavy repeated processing for classification, or data scarcity when trying to use generative learning on an instance-level at recognition. This leads to the need of necessitated robust feature extraction, multi-sensor fusion, efficient model design, and adaptive learning approaches to address the challenges.

Future research should focus on novel sensing modalities, reliable & efficient models, contextual information incorporation, and personalization techniques. Moreover, addressing these challenges and taking advantage of advancements in related fields as research develops, achieving more intuitive human-machine interactions might be anticipated, enhancing the quality of life and productivity.

References

1. Bavelas, J. B. The Social Life of Hand Gestures. – In: Face-to-Face Dialogue: Theory, Research, and Applications. Oxford Academic, 2022.
2. Athavale, S., M. Deshmukh. Dynamic Hand Gesture Recognition for Human-Computer Interaction. A Comparative Study. – International Journal of Engineering Research and General Science, Vol. **2**, 2014, No 2, pp. 38-55.
3. Baumgartl, H., D. Sauter, C. Schenk, C. Atik, R. Buettner. Vision-Based Hand Gesture Recognition for Human-Computer Interaction Using Mobile Net V2. – In: Proc. of 45th IEEE Annual Computers, Software, and Applications Conference (COMPSAC'21), IEEE, 2021, pp. 1667-1674.
4. Griffiths, D., J. Boehm. A Review on Deep Learning Techniques for 3D Sensed Data Classification. – Remote Sensing, Vol. **11**, 2019, No 12, 1499.
5. Dhande, A., S. Mantri, H. Pande. Comparative Analysis of Human Hand Gesture Recognition in Real-Time Healthcare Applications. – In: Proc. of International Conference on Expert Clouds and Applications, Springer, Singapore, 2022, pp. 461-475.
6. Ameur, S., A. B. Khalifa, M. S. Bouhlel. A Novel Hybrid Bidirectional Unidirectional LSTM Network for Dynamic Hand Gesture Recognition with Leap Motion. – Entertainment Computing, Vol. **35**, 2020, 100373.
7. Han, D., B. Mulyana, V. Stankovic, S. Cheng. A Survey on Deep Reinforcement Learning Algorithms for Robotic Manipulation. – Sensors, Vol. **23**, 2023, No 7, 3762.
8. Wu, J., P. Ren, B. Song, R. Zhang, C. Zhao, X. Zhang. Data Glove-Based Gesture Recognition Using CNN-BiLSTM Model with Attention Mechanism. – Plos One, Vol. **18**, 2023, No 11, e0294174.
9. Jiang, S., B. Lv, W. Guo, C. Zhang, H. Wang, X. Sheng, P. B. Shull. Feasibility of Wrist-Worn, Real-Time Hand, and Surface Gesture Recognition via sEMG and IMU Sensing. – IEEE Transactions on Industrial Informatics, Vol. **14**, 2017, No 8, pp. 3376-3385.
10. Chen, X., Y. Li, R. Hu, X. Zhang, X. Chen. Hand Gesture Recognition Based on Surface Electromyography Using Convolutional Neural Network with Transfer Learning Method. – IEEE Journal of Biomedical and Health Informatics, Vol. **25**, 2020, No 4, pp. 1292-1304.
11. Jaramillo-Yánez, A., M. E. Benalcázar, E. Mena-Maldonado. Real-Time Hand Gesture Recognition Using Surface Electromyography and Machine Learning: A Systematic Literature Review. – Sensors, Vol. **20**, 2020, No 9, 2467.
12. Choi, J. W., S. J. Ryu, J. H. Kim. Short-Range Radar Based Real-Time Hand Gesture Recognition Using LSTM Encoder. – IEEE Access, Vol. **7**, 2019, pp. 33610-33618.
13. Dang, L. M., K. Min, H. Wang, M. J. Piran, C. H. Lee, H. Moon. Sensor-Based and Vision-Based Human Activity Recognition: A Comprehensive Survey. – Pattern Recognition, Vol. **108**, 2020, 107561.
14. De, O. P., S. Mukherjee, S. Nandy, T. Chakraborty, S. Saha. Computer Vision Based Framework for Digit Recognition by Hand Gesture Analysis. – In: Proc. of 7th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON'16), IEEE, 2016, pp. 1-5.

15. G u p t a, A., S. K u m a r. Review for Optimal Human-Gesture Design Methodology and Motion Representation of Medical Images Using Segmentation from Depth Data and Gesture Recognition. – Current Medical Imaging, Vol. **20**, 2024.
16. L e o n, D. G., J. G r ö l i, S. R. Y e d u r i, D. R o s s i e r, R. M o s q u e r o n, O. J. P a n d e y, L. R. C e n k e r a m a d d i. Video Hand Gestures Recognition Using Depth Camera and Lightweight CNN. – IEEE Sensors Journal, Vol. **22**, 2022, No 14, pp. 14610-14619.
17. D o s o v i t s k i y, A., et al. An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale. – arXiv preprint arXiv:2010.11929, 2020.
18. E w e, E. L., C. P. L e e, L. C. K w e k, K. M. L i m. Hand Gesture Recognition via Lightweight VGG16 and Ensemble Classifier. – Applied Sciences, Vol. **12**, 2022, No 15, 7643.
19. M u k t h i n e n i, V., R. M u k t h i n e n i, O. S h a r m a, S. J a m j a l a N a r a y a n a n. Face Authenticated Hand Gesture Based Human Computer Interaction for Desktops. – Cybernetics and Information Technologies, Vol. **20**, 2020, No 4, pp. 74-89.
20. N g u y e n, T. N., H. H. H u y n h, J. M e u n i e r. Static Hand Gesture Recognition Using Artificial Neural Network. – Journal of Image and Graphics, Vol. **1**, 2013, No 1, pp. 34-38.
21. K ö p ü k l ü, O., A. G u n d u z, N. K o s e, G. R i g o l l. Real-Time Hand Gesture Detection and Classification Using Convolutional Neural Networks. – In: Proc. of 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG'2019), IEEE, 2019, pp. 1-8.
22. L e i, Q., H. Z h a n g, Z. X i a, Y. Y a n g, Y. H e, S. L i u. Applications of Hand Gestures Recognition in Industrial Robots: A Review. – In: Proc. of 11th International Conference on Machine Vision (ICMV'2018), SPIE, Vol. **11041**, 2019, pp. 455-465.
23. S i r i a k, R., I. S k a r g a-B a n d u r o v a, Y. B o l t o v. Deep Convolutional Network with Long Short-Term Memory Layers for Dynamic Gesture Recognition. – In: Proc. of 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS'19), IEEE, 2019, pp. 158-162.
24. A l o n s o, D. G., A. T e y s e y r e, L. B e r d u n, S. S c h i a f f i n o. A Deep Learning Approach for Hybrid Hand Gesture Recognition. – In: Proc. of 18th Mexican International Conference on Artificial Intelligence (MICAI'2019), Advances in Soft Computing, Xalapa, Mexico, 27 October – 2 November 2019, Proceedings 18, Springer International Publishing, 2019, pp. 87-99.
25. A b d u l r a z z a q, H. I., N. F. H a s s a n. Modified Siamese Convolutional Neural Network for Fusion Multimodal Biometrics at Feature Level. – In: Proc. of 2nd Scientific Conference of Computer Sciences (SCCS'19), IEEE, 2019, pp. 12-17.
26. Y a s e n, M., S. J u s o h. A Systematic Review on Hand Gesture Recognition Techniques, Challenges and Applications. – PeerJ Computer Science, Vol. **5**, 2019, e218.
27. B h u v a n y a, R., M. K a v i t h a. Image Clustering and Feature Extraction by Utilizing an Improvised Unsupervised Learning Approach. – Cybernetics and Information Technologies, Vol. **23**, 2023, No 2, pp. 3-19.
28. M o i n, A., A. Z h o u, A. R a h i m i, A. M e n o n, S. B e n a t t i, G. A l e x a n d r o v, J. M. R a b a e y. A Wearable Biosensing System with In-Sensor Adaptive Machine Learning for Hand Gesture Recognition. – Nature Electronics, Vol. **4**, 2021, No 1, pp. 54-63.
29. S i, G., Z. G u, H. Z h e n g. Duet of ViT and CNN: Multi-Scale Dual-Branch Network for Fine-Grained Image Classification of Marine Organisms. – Intelligent Marine Technology and Systems, Vol. **2**, 2024, No 1.
30. H a q, M. A., M. R i d l w a n, I. N a i l a. Leveraging Self-Attention Mechanism for Deep Learning in Hand-Gesture Recognition System. – In: E3S Web of Conferences. Vol. **500**. 2024.
31. K u m a r a n, N., M. S. A n u r a g, M. S a m p a t h. Hand Gesture Recognition Using Transfer Learning Techniques. – Journal of Current Research in Engineering and Science (JCRES), Vol. **4**, 2021, No 1.
32. R o n g, Y., G. G u. Deep Transfer Learning-Based Adaptive Gesture Recognition of a Soft e-Skin Patch with Reduced Training Data and Time. – Sensors and Actuators. A: Physical, Vol. **363**, 2023, 114693.
33. A g e i s h i, N., F. T o m o h i d e, A. B. A b d a l l a h. Real-Time Hand-Gesture Recognition Based on Deep Neural Network. – In: SHS Web of Conferences. EDP Sciences, 2021, 04009.

34. O b a i d a, T. H., A. S. J a m i l, N. F. H a s s a n. Real-Time Face Detection in Digital Video-Based on Viola-Jones Supported by Convolutional Neural Networks. – International Journal of Electrical and Computer Engineering (IJECE), Vol. **12**, 2022, No 3, pp. 3083-3091.
35. M o l c h a n o v, P., X. Y a n g, S. G u p t a, K. K i m, S. T y r e e, J. K a u t z. Online Detection and Classification of Dynamic Hand Gestures with Recurrent 3D Convolutional Neural Network. – In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 4207-4215.
36. H u a n g, J., W. Z h o u, H. L i, W. L i. Attention-Based 3D-CNNs for Large-Vocabulary Sign Language Recognition. – IEEE Transactions on Circuits and Systems for Video Technology, Vol. **29**, 2018, No 9, pp. 2822-2832.
37. Z h a n g, Y., L. S h i, Y. W u, K. C h e n g, J. C h e n g, H. L u. Gesture Recognition Based on Deep Deformable 3D Convolutional Neural Networks. – Pattern Recognition, Vol. **107**, 2020, 107416.
38. L o n g, D. T. Efficient DenseNet Model with Fusion of Channel and Spatial Attention for Facial Expression Recognition. – Cybernetics and Information Technologies, Vol. **24**, 2024, No 1, pp. 171-189.
39. S u t s k e v e r, I., O. V i n y a l s, Q. V. L e. Sequence to Sequence Learning with Neural Networks. – Advances in Neural Information Processing Systems, Vol. **27**, 2014.
40. O r d ó ñ e z, F. J., D. R o g g e n. Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition. – Sensors, Vol. **16**, 2016, No 1, 115.
41. M i t t a l, A., P. K u m a r, P. P. R o y, R. B a l a s u b r a m a n i a n, B. B. C h a u d h u r i. A Modified LSTM Model for Continuous Sign Language Recognition Using Leap Motion. – IEEE Sensors Journal, Vol. **19**, 2019, No 16, pp. 7056-7063.
42. T a n, C. K., K. M. L i m, R. K. C h a n g, C. P. L e e, A. A l q a h t a n i. HGR-ViT: Hand Gesture Recognition with Vision Transformer. – Sensors, Vol. **23**, 2023, No 12, 5555.
43. X i e, Z., Y. L i n, Z. Y a o, Z. Z h a n g, Q. D a i, Y. C a o, H. H u. Self-Supervised Learning with Swin Transformers. – arXiv e-prints, arXiv:2105.04553, 2021.
44. T a v a k o l i, M., C. B e n u s s i, P. A. L o p e s, L. B. O s o r i o, A. T. de A l m e i d a. Robust Hand Gesture Recognition with a Double Channel Surface EMG Wearable Armband and SVM Classifier. – Biomedical Signal Processing and Control, Vol. **46**, 2018, pp. 121-130.
45. V e r m a, B., A. C h o u d h a r y. Unsupervised Learning Based Static Hand Gesture Recognition from RGB-D Sensor. – In: Proc. of 8th International Conference on Soft Computing and Pattern Recognition (SoCPaR 2016), Springer International Publishing, 2018, pp. 304-314.
46. Z h a n g, Y., Y. C h e n, H. Y u, X. Y a n g, R. S u n, B. Z e n g. A Feature Adaptive Learning Method for High-Density Semg-Based Gesture Recognition. – Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Vol. **5**, 2021, No 1, pp. 1-26.
47. S a r m a, D., M. K. B h u y a n. Methods, Databases and Recent Advancement of Vision-Based Hand Gesture Recognition for Hci Systems: A Review. – SN Computer Science, Vol. **2**, 2021, No 6, 436.
48. J i a n g, S., P. K a n g, X. S o n g, B. P. L o, P. B. S h u l l. Emerging Wearable Interfaces and Algorithms for Hand Gesture Recognition: A Survey. – IEEE Reviews in Biomedical Engineering, Vol. **15**, 2021, pp. 85-102.
49. P y u n, K. R., et al. Machine-Learned Wearable Sensors for Real-Time Hand-Motion Recognition: toward Practical Applications. – National Science Review, Vol. **11**, 2024, No 2, nwad298.
50. K a u r, H., J. R a n i. A Review: Study of Various Techniques of Hand Gesture Recognition. – In: Proc. of 1st IEEE International Conference on Power Electronics, Intelligent Control, and Energy Systems (ICPEICES'16), IEEE, 2016, pp. 1-5.
51. W a n g, F., R. H u, Y. J i n. Research on Gesture Image Recognition Method Based on Transfer Learning. – Procedia Computer Science, Vol. **187**, 2021, pp. 140-145.
52. S a i f u l I s l a m, M., W. R a h m a n, A. A b d e l k a d e r, P. T. Y a n g, S. L e e, J. L. A d a m s, E. H o q u e. Using AI to Measure Parkinson's Disease Severity at Home. – arXiv e-prints, arXiv-2303, 2023.
53. S i k k a n d a r, M. Y. Design a Contactless Authentication System Using Hand Gestures Technique in COVID-19 Panic Situation. – Annals of the Romanian Society for Cell Biology, 2021, pp. 2149-2159.

54. R a u t a r a y, S. S., A. A g r a w a l. Vision Based Hand Gesture Recognition for Human Computer Interaction: A Survey. – Artificial Intelligence Review, Vol. **43**, 2015, pp. 1-54.
55. S a t y b a l d i n a, D., G. K a l y m o v a. Deep Learning Based Static Hand Gesture Recognition. – Indonesian Journal of Electrical Engineering and Computer Science, Vol. **21**, 2021, No 1, pp. 398-405.
56. Z h u, G., L. Z h a n g, P. S h e n, J. S o n g. Multimodal Gesture Recognition Using 3-D Convolution and Convolutional LSTM. – IEEE Access, Vol. **5**, 2017, pp. 4517-4524.
57. A l-D o r i, A. S. M., J. M. K a d h i m. Touchscreen-Based Smartphone Continuous Authentication System (SCAS) Using Deep Neural Network. – Turkish Journal of Computer and Mathematics Education, Vol. **12**, 2021, No 11, pp. 2382-2391.
58. K e t a n C h a k r a b o r t y, B., D. S a r m a, M. K. B h u y a n, K. F. M a c D o r m a n. Review of Constraints on Vision-Based Gesture Recognition for Human-Computer Interaction. – IET Computer Vision (Wiley-Blackwell), Vol. **12**, 2018, No 1.
59. L a i, B. W., C. C. L i, E. J e n g. Dual-Handed Dynamic Gesture Recognition Using Inertial Sensors. – In: Proc. of 3rd IEEE International Conference on Electronic Communications, Internet of Things and Big Data (ICEIB'23), IEEE, 2023, pp. 223-225.
60. R a s t g o o, R., K. K i a n i, S. E s c a l e r a. Hand Sign Language Recognition Using Multi-View Hand Skeleton. – Expert Systems with Applications, Vol. **150**, 2020, 113336.
61. B a m b a c h, S., S. L e e, D. J. C r a n d a l l, C. Y u. Lending a Hand: Detecting Hands and Recognizing Activities in Complex Egocentric Interactions. – In: Proc. of IEEE International Conference on Computer Vision, IEEE, 2015, pp. 1949-1957.
62. Z e n g, W., C. W a n g, Q. W a n g. Hand Gesture Recognition Using Leap Motion via Deterministic Learning. – Multimedia Tools and Applications, Vol. **77**, 2018, pp. 28185-206.
63. K u m a r a n, N., M. S. A n u r a g, M. S a m p a t h. Hand Gesture Recognition Using Transfer Learning Techniques. – Journal of Current Research in Engineering and Science (JCRES), Vol. **4**, 2021, No 1.
64. O b a i d a, T. H., N. F. H a s s a n, A. S. J a m i l. Comparative of Viola-Jones and YOLO V3 for Face Detection in Real Time. – Iraqi Journal of Computers, Communications, Control & Systems Engineering, Vol. **22**, 2022, No 2, pp. 63-72.
65. Z h a n g, Y., C. C a o, J. C h e n g, H. L u. EgoGesture: A New Dataset and Benchmark for Egocentric Hand Gesture Recognition. – IEEE Transactions on Multimedia, Vol. **20**, 2018, No 5, pp. 1038-1050.
66. M i a h, A. S. M., M. A. M. H a s a n, Y. T o m i o k a, J. S h i n. Hand Gesture Recognition for Multi-Culture Sign Language Using Graph and General Deep Learning Network. – IEEE Open Journal of the Computer Society, Vol. **5**, 2024, pp. 144-155.

Received: 16.07.2024; Second version: 26.08.2024; Accepted: 28.08.2024 (fast track)