Hand Alphabet Recognition for Dactylology Conversion to English Print Using Streaming Video Segmentation

Manuel B. Garcia [a], Teodoro F. Revano, Jr. [a], Armi Cunanan-Yabut [b]

[a] College of Computer Studies, FEU Institute of Technology, Philippines
[b] College of Engineering, FEU Institute of Technology, Philippines

Abstract:

Assistive technologies gained traction in the medical field over the last few decades. Novel approaches have been developed to support people with disability to communicate effectively. However, little research has been conducted on the other side of the coin, that is, assistive technologies to help people who do not have a disability to understand the language of the disabled. This study describes the early development of a hand alphabet recognition that intends to accomplish a functioning dactylology conversion from sign language to English print in a live streaming video. Through video analysis, each frame is processed using a segmentation technique to partition it into different segments (e.g., pixels of hand gestures). The dactylology conversion algorithm was implemented in a mobile application where users can watch videos containing an on-screen sign language interpreter and understand fingerspelling used as communication by hearing- and speech-impaired people. Through the sample dataset of 13 videos of American Sign Language manually collected (n = 10) and recorded (n = 3), the application was tested for its accuracy in detecting the alphabet in a video (94.16%), and the correctness of conversion of the detected alphabet into English print (89.65%). This study contributes to the list of existing novel approaches that aim to promote social positive effects as well as improve the quality of life for both disabled and all the people they socialize with.

Keywords:

Dactylology, Assistive Technology, Sign Language, Video Segmentation, Fingerspelling, Hand Alphabet Recognition

Reference to this paper should be made as follows:

Garcia, M. B., Revano, T. F., Cunanan-Yabut, A. (2021). Hand Alphabet Recognition for Dactylology Conversion to English Print Using Streaming Video Segmentation. *2021 9th International Conference on Computer and Communications Management (ICCCM)*. DOI: 10.1145/3479162.3479169.

Corresponding Author:

Manuel B. Garcia, FEU Institute of Technology, Philippines. Email: **mbgarcia@feutech.edu.ph**

Disclaimer: This is a post-peer-review, pre-copyedit version of an article copied from https://manuelgarcia.info/publication/hand-alphabet-recognition and published in the proceedings of the *2021 9th International Conference on Computer and Communications Management (ICCCM)*. The final authenticated version is available online at: https://doi.org/10.1145/3479162.3479169

1. INTRODUCTION

In a world where millions of people are deaf-mute, Dactylology, or the science of communication using hands and fingers (e.g., one-handed alphabet, two-handed alphabet), is one of the, if not the only one, communication modalities that lets people with and without disability to express and send ideas and thoughts to each other. In fact, there are more than 120 distinctive sign lingos used in various nations such as American, French, German, Spanish, Filipino, Japanese, Indo-Pakistani, and more. The difficulty in establishing a universal communication modality between disabled people with different sign languages, and people with and without disabilities led to a stirring invitation of technology implementation. For instance, one study (Tao et al., 2018) developed a convolutional neural network model for American Sign Language alphabet recognition. The classification model was then combined with a multi-view augmentation strategy to exploit 3D information from depth images. On the other hand, an artificial neural network was utilized in another study (Oz & Leu, 2011) to design and develop an American Sign Language recognition system with a sensory glove and a three-dimensional motion tracker for extracting gesture data features. These technology-based communication advancements are some of the most novel contributions to the vast reaches of healthcare information, intelligent application systems, and even communication technology.

In a deeper look, there has been a widespread development of intelligent systems (Bayasut et al., 2011; Chiluisa-Castillo et al., 2018; Kawamura et al., 1995; Khayat et al., 2012) and studies (Khetarpal, 2014; Manzoor & Vimarlund, 2018; Zahid et al., 2013) that intend to establish and understand the usage of technology-based assistance to people with disabilities in performing basic communication tasks. Naves et al. (2012) developed an alternative communications system by uniting electromyographic (EMG) signals to the field of Human-Computer Interaction (HCI) to attend and serve patients severely disabled by amyotrophic lateral sclerosis. HCI and EMG were both incorporated into the EDITH system – a computer software package consisting of communication features designed for a multimedia environment. In the robotics field, a mobile robotic arm was developed by Gushi et al. (2017) for people with severe disabilities. The robotic arm can perform several tasks by using eye movements, which are detected by an image processing technique. In another field, Garcia (2019) developed a speech therapy game application to assist aphasic people in learning how to communicate again just like before their stroke occurred. Such technologies have been proven as an important tool and instrumental in promoting social positive effects as well as improving the quality of life not only for disabled but also of their family and relatives. Santos et al. (2017) confirmed the positive relationship between the quality of life of people and assistive technology (e.g., VISIMP; Garcia & Pilueta, 2020) making it a sought-after invention of our time. These stigmatized and marginalized social groups have now a way to establish their position and promote inclusion.

Granted, these novel approaches have been developed to aid people with disability to communicate effectively. Notwithstanding, little research has been conducted on the other side of the coin, that is, assistive technologies to assist people who do not have a disability to also

2021 9th International Conference on Computer and Communications Management (ICCCM)
DOI: 10.1145/3479162.3479169

understand and comprehend the language of the disabled. In fact, most people do not clearly understand sign language. Therefore, aside from the research gap, there is also a communication gap between the deaf communities and the public. This study describes the early development of a hand alphabet recognition that intends to achieve a functioning dactylology conversion from sign language to English print in a live streaming video. Through video analysis, each frame is processed using a video segmentation technique to partition it into different segments (e.g., pixels of hand gestures). The dactylology conversion algorithm was implemented in a mobile application where users could watch a video containing an on-screen sign language interpreter and understand fingerspelling used as communication by hearing and speech-impaired people. Not only does the mobile application provide a new communication modality, but it also sensitizes and offers awareness on how to communicate with deaf and mute people.

2. RELATED WORKS

The growth of multimedia information led to an extensive interest in a video indexing of media information and video retrieval for accessing the acquired information stored in a database. For a better performance video indexing and retrieval system, a proper video segmentation algorithm must be applied (Thounaojam et al., 2014). According to Dhiman and Dhanda (2016), video segmentation refers to the process of decaying video data into meaningful segments that have a strong correlation with the real world. In the computer science field, there are several schemes to use when performing video segmentation and every algorithm has different advantages and disadvantages. Figure 1 shows an example of a segmented single video frame.

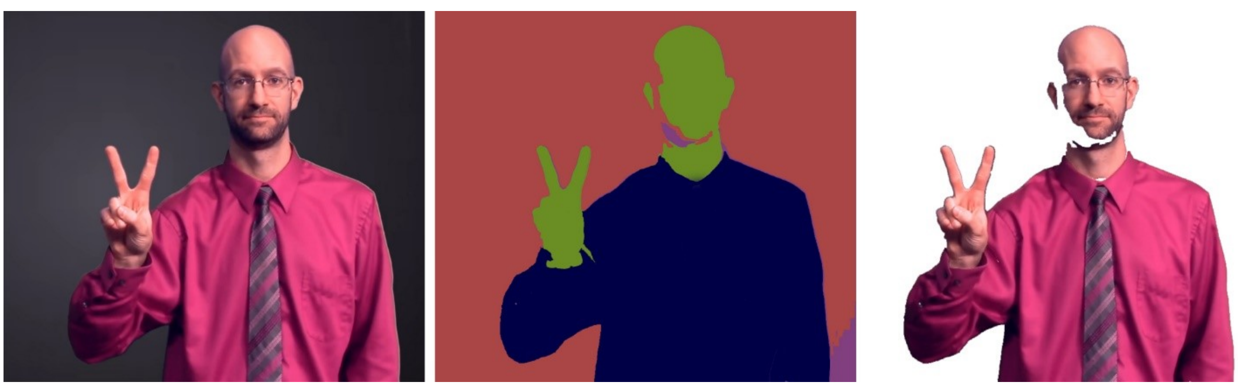

Figure 1: Image Segmentation in a Video Sequence

Beevi and Natarajan (2009) proposed a video segmentation algorithm for MPEG-4 encoding systems to build segmentation results with low computation load by using baseline, shadow cancellation, and adaptive threshold modes. Similarly, a frame-by-frame technique with computationally efficient results was proposed by Vora and Raman (2017) through clustering of visually similar generic object segments in a video via extraction using top-k region proposal to generate preliminary masks of a foreground object. Alternatively, El Hassani et al. (2008) managed to use a region merging process for spatial and motion information to implement their

2021 9th International Conference on Computer and Communications Management (ICCCM)
DOI: 10.1145/3479162.3479169

proposed time-consistent video segmentation algorithm for real-time application. Another technique in partitioning video information for further analysis is the graph-based hierarchical video segmentation (de Souza et al., 2015). This method used four main steps that start with generating a graph for a *k*-sized frame-block followed by a calculation of hierarchical scales. Then, a calculation for the inference of video segmentations through a thresholding process is accomplished. Lastly, temporal coherence video segments are calculated by merging two consecutive segmented blocks. Li et al. (2016), on the other hand, utilized an algorithm called suboptimal low-rank decomposition (SOLD) to decompose the representation coefficient matrix into sub-matrices of low ranks. The efficiency analysis revealed that this method is faster and more effective than HGB and SHGB. Further, intelligent approaches are likewise proposed for recognizing hand gestures in a natural manner. Chaudhary et al. (2013) grouped these approaches into fuzzy logic, genetic algorithm, and artificial neural networks. Verma and Dev (2009) used fuzzy clustering-based finite state machines to recognize hand gestures efficiently. Nolker and Ritter (2002), on the other hand, used a neural network to distinguish fingertips transformable into finger joint angles of a hand model. This allowed a full reconstruction of a three-dimensional hand shape, with 16 segments and 20 joint angles. Hu et al. (2000) used the extraction of a human parametric 2D model to estimate posture and recognize human activity. In their system, the genetic algorithm was applied to make a model with human silhouette.

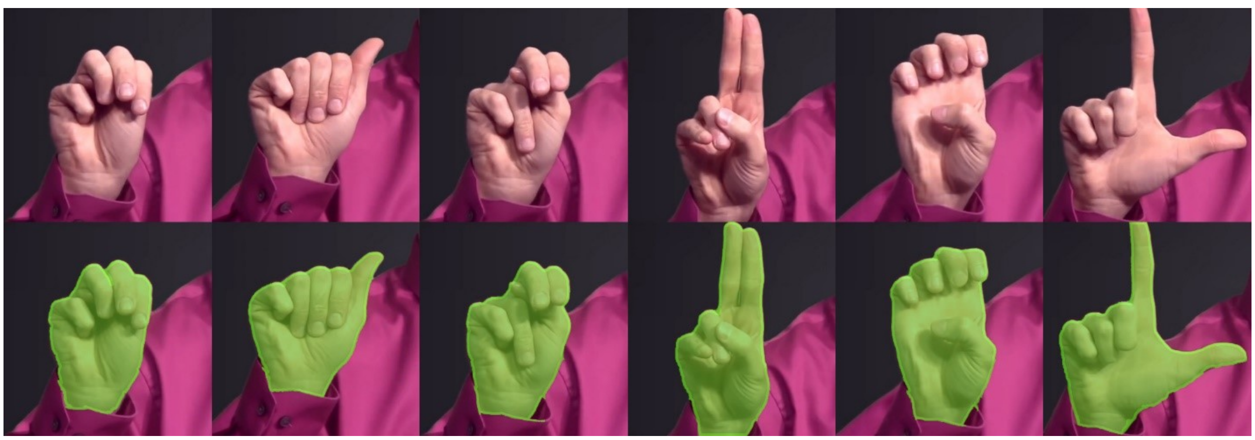

Figure 2: Frame-by-Frame Segmentation of Hand Gesture with Sign Language Alphabet Recognition

Another area of related works that needs to be reviewed is the hand recognition system where the recognized gestures could be utilized as a part of a more intelligent system (Elakkiya et al., 2012), or for controlling a robot (Faudzi et al., 2012). To produce gesture recognition systems, different approaches have been proposed from using additional hardware, such as gloves and color markers, to the use of skin-based segmentation for feature extraction (Cicirelli et al., 2015; Koh et al., 2019; Oprea et al., 2018; Rautaray & Agrawal, 2012; Shi et al., 2010; Sidek & Abdul Hadi, 2014; Zengeler et al., 2018). The growing importance of gesture recognition in society (Khan & Ibraheem, 2012) led to a stimulating revolution among systems inventors and gave birth to essential applications in numerous areas like surveillance systems, robotics, HCI, healthcare, education, etc. In addition, sign language recognition has been likewise a beneficiary

DOI: 10.1145/3479162.3479169

and received special attention from all the advancements of gesture recognition. Nevertheless, the development of gesture recognition systems has formed many lessons based on the drawbacks of the existing and completed prototype. For instance, a neural network classifier is too time-consuming (e.g., learning ten words in four days; Murakami & Taguchi, 1991) to make progress in computer hardware. An orientation histogram method becomes problematic when dealing with similar gestures but different histograms or different gestures with similar histograms (Freeman & Michal, 1995).

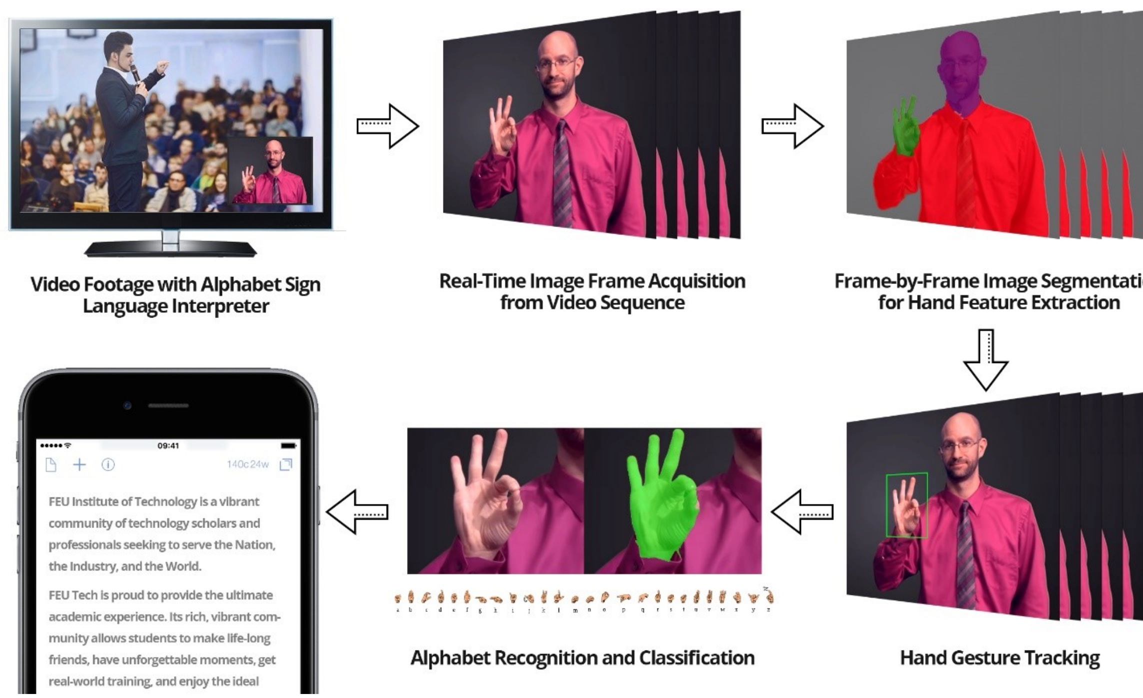

Figure 3: System Block Diagram of Hand Alphabet Sign Language Conversion to English Print

3. METHODS

The main purpose of this study is to translate the hand alphabet of the American Sign Language into English print to bridge a communication gap between people with and without disabilities when using dactylology as a communication modality. Towards the realization of this goal, various image and video processing techniques were utilized based on the experimental results of other existing studies. First, the strategy of Vora and Raman (2017) in video object segmentation was the basis of the core processes of hand gesture recognition. The skin detection algorithm of Garcia et al. (2018) was slightly mimicked particularly on image processing techniques to enhance the accuracy result. For this application, each video frame undergoes several processes such as (1) color illumination restoration to recover details of a frame, (2) histogram equalization to redistribute color intensities, (3) lighting correction to re-adjust dark areas, and (4) noise reduction to remove unwanted pixels. For this to become possible, each

DOI: 10.1145/3479162.3479169

frame must be extracted from the video media file. Afterward, the extracted frames proceed to the core processes of the algorithm, such as (1) hand detection and extraction via an image segmentation with the mixture of background subtraction and three-frame-difference method, (2) object tracking based on a motion consistency algorithm (He et al., 2018), and (3) hand alphabet recognition using a convolutional neural network model for the classification process. Figure 2 shows segmentation and tracking of hand gestures frame-by-frame to automatically detect and recognize the sign language alphabet. Meanwhile, the system block diagram that exhibits how the whole process works is illustrated in Figure 3.

3.1. Hand Detection and Extraction

After the preprocessing of video frames, the first step towards the recognition of hand features is segmentation. A common cue when segmenting body parts like a hand is the skin color (Garcia et al., 2018) since it is invariant to scale and rotation changes (Stergiopoulou & Papamarkos, 2009). However, the result of the segmentation is affected by illumination conditions. As such, segmented skin-colored regions might not be a skin but of another region with a similar color. Fortunately, a sign language interpreter customarily conveys the information using hand movements with a stationary body. Consequently, detecting a moving object in a video sequence will probably result in a moving hand. Weng et al. (2010) proposed a new interframe difference algorithm for detecting a moving object in a video through the combination of background subtraction and three frame-difference methods. The first process is to subtract the current frame to the previous and next frames separately and add together the results to generate a grayscale image. Afterward, another grayscale image will be created by subtracting the current frame from the background image. The final output is a binary image made from the sum of two previously generated grayscale images, which helps in adding a bounding box on a region that has a constantly moving object.

3.2. Hand Gesture Tracking

Once the target object is segmented and the feature is extracted, a hidden bounding box stays on that object for tracking purposes to perform the same process on the next frames until the end of a video sequence. He et al. (2018) proposed an object tracing based on motion consistency (MCT), which serves as the basis of the algorithm used in tracking the segmented hand. MCT states that the object that needs tracking must be known in the first frame. The segmented region during the hand detection and extraction processes determines the "*known target object*" prior to shifting to the state transition model, which is used to select the candidate samples in the current video frame. Subsequently, the target state prediction estimates the target position including motion directions and distances based on a motion consistency. Hence, the tracking result is determined by the position factor with the holistic responses of each candidate.

3.3. Hand Alphabet Recognition

To classify the hand alphabet sign language, a convolutional neural network model for classification was tested, trained, and evaluated through the combination of various technologies

DOI: 10.1145/3479162.3479169

such as Python programming language, OpenCV for real-time computer vision, and TensorFlow for machine learning with the inclusion of PIP packages such as matplotlib, NumPy, OpenCV, and TensorFlow. After classification, the alphabet sign language is converted into an English print and displayed on the mobile application. To recognize a word and/or to separate a string (converted sign language) into actual English words, a spellchecker and autocorrect algorithm was used through the comparison of translated sign language to an actual English word dictionary.

4. EXPERIMENTAL RESULTS

Through the sample dataset of 13 videos of American Sign Language manually collected (n = 10) and recorded (n = 3), the application was tested for its accuracy in detecting the alphabet in a video, and the correctness of conversion of the detected alphabet into English print. Upon the calculations, the detection accuracy of 94.16% and conversion accuracy of 89.65% was obtained as shown in Table 1. Manual labeling was performed in all videos to determine the number of alphabets made by the sign language interpreter. A common error in the detection of the alphabet is the recognition of fingerspelled letters that must be traced in the air such as "*Z*" and "*J*'. In addition, "*K*" and "*P*" have a similar handshape, where the former is palm down while the latter is palm forward, that confuses the algorithm. Notwithstanding, the spellchecker and autocorrect algorithm was able to translate the sign language into the correct English words, especially for those words with problematic letters.

Table 1: Detection and Conversion Accuracy Results of Sample Sign Language Videos

Video Index	Number of Frames	Number of Alphabet	Detected Alphabet	Detection Accuracy	Correct Character	Conversion Accuracy
1	3744	145	144	99.31	121	84.03
2	3576	124	120	96.77	110	91.67
3	4176	169	151	89.35	140	92.72
4	3336	139	131	94.24	118	90.08
5	4104	171	169	98.83	149	88.17
6	5136	214	201	93.93	192	95.52
7	1824	76	72	94.74	66	91.67
8	2664	111	100	90.09	92	92.00
9	2352	98	92	93.88	89	96.74
10	3384	141	124	87.94	115	92.74
11	3096	129	120	93.02	103	85.83
12	1416	45	43	95.56	34	79.07
13	1512	56	54	96.43	46	85.19
Mean:	**3102**	**124**	**117**	**94.16**	**106**	**89.65**

DOI: 10.1145/3479162.3479169

5. CONCLUSION AND RECOMMENDATION

In this study, a mobile application for hand alphabet recognition for dactylology conversion to English print was presented. Grounded from various algorithms and methodologies, the preliminary results of the experimental assessment specify a very encouraging outcome with a 94.16% detection accuracy and 89.65% conversion accuracy, at least based on the dataset supplied. One problem that needs to be addressed is the accuracy enhancement of the classification model particularly on letters with the same handshape and fingerspelled alphabet that must be traced in the air. As of writing, the mobile application is still limited because it is an ongoing and the first phase of the project that focuses on the conversion of sign languages. For future work and the next phase of the project, a classifier will be modeled to detect words in a sign language to extend the hand alphabet. This extension of the algorithm is expected to be useful when communicating and enhancing interactions with people with disabilities. Conversion of sign language to English print then the text to speech audio is also feasible to remove the hassle of reading textual information but this is only a recommendation for future authors for now. Overall, this evaluation study presented support to the list of existing novel approaches in promoting social positive effects as well as improving the quality of life for both disabled and all the people they socialize with.

ACKNOWLEDGMENTS

The authors would like to thank the FEU Institute of Technology for funding the conference presentation.

REFERENCES

Bayasut, B. L., Ananta, G. P., & Muda, A. K. (2011). Intelligent biometric detection system for disabled people. *2011 11th International Conference on Hybrid Intelligent Systems (HIS)*, 346-350. https://doi.org/10.1109/HIS.2011.6122130

Beevi, C. P. Y., & Natarajan, S. (2009). A Novel Video Segmentation Algorithm with Shadow Cancellation and Adaptive Threshold Techniques. 304-311. https://doi.org/10.1007/978-3-642-10546-3_36

Chaudhary, A., Raheja, J. L., Das, K., & Raheja, S. (2013). Intelligent Approaches to interact with Machines using Hand Gesture Recognition in Natural way: A Survey. *Human-Computer Interaction*.

Chiluisa-Castillo, D., Ortega-Barreto, F., Robles-Bykbaev, V., & Pesántez-Avilés, F. (2018). An Intelligent Platform to Design and Develop Low-Cost Assistive Technologies and Robotic Assistants for Children with Disabilities. *2018 IEEE XXV International Conference on Electronics, Electrical Engineering and Computing (INTERCON)*, 1-4. https://doi.org/10.1109/INTERCON.2018.8526413

Cicirelli, G., Attolico, C., Guaragnella, C., & D'Orazio, T. (2015). A Kinect-Based Gesture Recognition Approach for a Natural Human Robot Interface. *International Journal of Advanced Robotic Systems*, *12*(3), 22. https://doi.org/10.5772/59974

de Souza, K. J. F., Araújo, A. D. A., Guimarães, S. J. F., do Patrocínio, Z. K. G., & Cord, M. (2015). Streaming Graph-Based Hierarchical Video Segmentation by a Simple Label Propagation. *2015 28th SIBGRAPI Conference on Graphics, Patterns and Images*, 119-125. https://doi.org/10.1109/SIBGRAPI.2015.23

Dhiman, P., & Dhanda, M. (2016). A Review on Various Techniques of Video Segmentation. *International Journal for Innovative Research in Science & Technology*.

El Hassani, M., Jehan-Besson, S., Brun, L., Revenu, M., Duranton, M., Tschumperl, #233, , D., & Rivasseau, D. (2008). A Time-Consistent Video Segmentation Algorithm Designed for Real-Time Implementation. *VLSI Design*. https://doi.org/10.1155/2008/892370

Elakkiya, R., Selvamani, K., Kanimozhi, S., Velumadhava, R., & Kannan, A. (2012). Intelligent System for Human Computer Interface Using Hand Gesture Recognition. *Procedia Engineering*, *38*, 3180-3191. https://doi.org/10.1016/j.proeng.2012.06.369

Faudzi, A. A. M., Ali, M. H. K., Azman, M. A., & Ismail, Z. H. (2012). Real-time Hand Gestures System for Mobile Robots Control. *Procedia Engineering*, *41*, 798-804. https://doi.org/https://doi.org/10.1016/j.proeng.2012.07.246

Freeman, W. T., & Michal, R. (1995). Orientation Histograms for Hand Gesture Recognition. *IEEE International Workshop on Automatic Face and Gesture Recognition*.

Garcia, M. B. (2019). A Speech Therapy Game Application for Aphasia Patient Neurorehabilitation – A Pilot Study of an mHealth App. *International Journal of Simulation: Systems, Science & Technology*, *20*(2). https://doi.org/10.5013/IJSSST.a.20.S2.05

Garcia, M. B., & Pilueta, N. U. (2020). The VISIMP Portable Communications Device for Visually Impaired Individuals – Development and Feasibility Study of an Assistive Technology. *Journal of Critical Reviews*, *7*(15), 511-516.

Garcia, M. B., Revano, T. F., Habal, B. M., Contreras, J., & Enriquez, J. (2018). A Pornographic Image and Video Filtering Application Using Optimized Nudity Recognition and Detection Algorithm. *10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment aend Management*, 1-5. https://doi.org/10.1109/HNICEM.2018.8666227

Gushi, S., Higa, H., Uehara, H., & Soken, T. (2017). A Mobile Robotic Arm for People with Severe Disabilities: Evaluation of Scooping Foods. *International Conference on Intelligent Informatics and Biomedical Sciences*. https://doi.org/10.1109/ICIIBMS.2017.8279740

He, L., Qiao, X., Wen, S., & Li, F. (2018). Robust Object Tracking Based on Motion Consistency. *Sensors*, *18*(2), 572. https://doi.org/10.3390/s18020572

Hu, C., Yu, Q., Li, Y., & Ma, S. (2000). Extraction of Parametric Human Model for Posture Recognition Using Genetic Algorithm. *4th International Conference on Automatic Face and Gesture Recognition*. https://doi.org/10.1109/AFGR.2000.840683

Kawamura, K., Bagchi, S., Iskarous, M., & Bishay, M. (1995). Intelligent Robotic Systems in Service of the Disabled. *IEEE Transactions on Rehabilitation Engineering*, *3*(1), 14-21. https://doi.org/10.1109/86.372888

Khan, R., & Ibraheem, N. (2012). Survey on Gesture Recognition for Hand Image Postures. *International Journal of Computer And Information Science*(3). https://doi.org/10.5539/cis.v5n3p110

Khayat, G. A. E., Mabrouk, T. F., & Elmaghraby, A. S. (2012). Intelligent Serious Games System for Children with Learning Disabilities. *17th International Conference on Computer Games (CGAMES)*, 30-34. https://doi.org/10.1109/CGames.2012.6314547

Khetarpal, A. (2014). Information and Communication Technology (ICT) and Disability. *Review of Market Integration*, *6*(1), 96-113. https://doi.org/10.1177/0974929214560117

Koh, J. I., Cherian, J., Taele, P., & Hammond, T. (2019). Developing a Hand Gesture Recognition System for Mapping Symbolic Hand Gestures to Analogous Emojis in Computer-Mediated Communication. *ACM Transactions on Interactive Intelligent Systems*, *9*(1), 1-35. https://doi.org/10.1145/3297277

Li, C., Lin, L., Zuo, W., Wang, W., & Tang, J. (2016). An Approach to Streaming Video Segmentation With Sub-Optimal Low-Rank Decomposition. *IEEE Transactions on Image Processing*, *25*(5), 1947-1960. https://doi.org/10.1109/TIP.2016.2537211

Manzoor, M., & Vimarlund, V. (2018). Digital Technologies for Social Inclusion of Individuals with Disabilities. *Health and Technology*, *8*(5). https://doi.org/10.1007/s12553-018-0239-1

Murakami, K., & Taguchi, H. (1991). Gesture Recognition using Recurrent Neural Networks. *SIGCHI Conference on Human Factors in Computing Systems*, 237-242. https://doi.org/10.1145/108844.108900

Naves, E., Rocha, L., & Pino, P. (2012). Alternative Communication System for People with Severe Motor Disabilities Using Myoelectric Signal Control. *ISSNIP Biosignals and Biorobotics Conference: Biosignals and Robotics for Better and Safer Living (BRC)*, 1-4. https://doi.org/10.1109/BRC.2012.6222176

Nolker, C., & Ritter, H. (2002). Visual Recognition of Continuous Hand Postures. *IEEE Transactions on Neural Networks*, *13*(4), 983-994. https://doi.org/10.1109/TNN.2002.1021898

Oprea, S. O., Garcia-Garcia, A., Orts-Escolano, S., Villena-Martinez, V., & Castro-Vargas, J. A. (2018). A Long Short-Term Memory Based Schaeffer Gesture Recognition System. *Expert Systems*, *35*(2). https://doi.org/10.1111/exsy.12247

Oz, C., & Leu, M. C. (2011). American Sign Language Word Recognition with a Sensory Glove Using Artificial Neural Networks. *Engineering Applications of Artificial Intelligence*, *24*(7), 1204-1213. https://doi.org/10.1016/j.engappai.2011.06.015

Rautaray, S. S., & Agrawal, A. (2012). Real Time Gesture Recognition System for Interaction in Dynamic Environment. *Procedia Technology*, *4*. https://doi.org/10.1016/j.protcy.2012.05.095

Santos, R., Sampaio, P., Sampaio, R., Gutierrez, G., & Almeida, M. (2017). Assistive Technology and its Relationship to the Quality of Life of People with Disabilities. *Revista De Terapia Ocupacional Da Universidade De São Paulo*, *28*(1), 54-62. https://doi.org/10.11606/issn.2238-6149.v28i1p54-62

Shi, L., Wang, Y., & Li, J. (2010). A Real Time Vision-Based Hand Gestures Recognition System. 349-358. https://doi.org/10.1007/978-3-642-16493-4_36

Sidek, O., & Abdul Hadi, M. (2014). Wireless Gesture Recognition System using MEMS Accelerometer. *International Symposium on Technology Management and Emerging Technologies*, 444-447. https://doi.org/10.1109/ISTMET.2014.6936550

Stergiopoulou, E., & Papamarkos, N. (2009). Hand Gesture Recognition Using a Neural Network Shape Fitting Technique. *Engineering Applications of Artificial Intelligence*, *22*(8), 1141-1158. https://doi.org/10.1016/j.engappai.2009.03.008

Tao, W., Leu, M. C., & Yin, Z. (2018). American Sign Language Alphabet Recognition Using Convolutional Neural Networks with Multiview Augmentation and Inference Fusion. *Engineering Applications of Artificial Intelligence*, *76*, 202-213. https://doi.org/10.1016/j.engappai.2018.09.006

Thounaojam, D. M., Trivedi, A., Manglem Singh, K., & Roy, S. (2014). A Survey on Video Segmentation. *Intelligent Computing, Networking, and Informatics*, 903-912. https://doi.org/10.1007/978-81-322-1665-0_91

Verma, R., & Dev, A. (2009). Vision Based Hand Gesture Recognition Using Finite State Machines and Fuzzy Logic. *International Conference on Ultra Modern Telecommunications & Workshops*, 1-6. https://doi.org/10.1109/ICUMT.2009.5345425

Vora, A., & Raman, S. (2017). Flow-free Video Object Segmentation. *National Conference on Computer Vision, Pattern Recognition, Image Processing, and Graphics*, 34-44. https://doi.org/10.1007/978-981-13-0020-2_4

Weng, M., Huang, G., & Da, X. (2010). A New Interframe Difference Algorithm for Moving Target Detection. *3rd International Congress on Image and Signal Processing*, *1*, 285-289. https://doi.org/10.1109/CISP.2010.5648259

Zahid, M. J. A., Ashraf, M. M., Malik, B. T., & Hoque, M. R. (2013). Information Communication Technology (ICT) for Disabled Persons in Bangladesh: Preliminary Study of Impact/Outcome. *International Working Conference on Transfer and Diffusion of IT*. https://doi.org/10.1007/978-3-642-38862-0_48

Zengeler, N., Kopinski, T., & Handmann, U. (2018). Hand Gesture Recognition in Automotive Human–Machine Interaction Using Depth Cameras. *Sensors*, *19*(1), 59. https://doi.org/10.3390/s19010059

DOI: 10.1145/3479162.3479169