Sofian Khabbich, Jens Christian Will, Michael Hötter, and Hanno Homann

Individualized Hand Pose Control: A Patient Study

Abstract: Individuals with motor impairments often face challenges in interacting with digital devices. Video-based pose detection has emerged as a promising approach for creating accessible human-machine interfaces (HMIs). However, due to the diverse and highly individualized movement capabilities among this population, generic pose recognition systems are often inadequate. This study investigates the feasibility of a personalized hand-pose recognition system for individuals with motor impairments using minimal training data.
We implemented a video-based pose classification model using a webcam and the MediaPipe Hands framework to extract keypoints from hand poses. Each participant, regardless of their level of impairment, selected four individual hand poses according to their capabilities, which were then used to train a neural network. The trained models were evaluated in a real-time test application where participants controlled an animated figure using their hand poses.
Results from 14 participants demonstrated that most were able to achieve effective control, even in cases of severe motor limitations. Individuals with restricted finger mobility successfully adapted by utilizing wrist and elbow movements. However, participants with spasticity experienced higher misclassification rates due to difficulties in maintaining stable poses. Overall, our findings highlight the importance of individualized pose recognition systems for assistive technology. Future work should explore additional adaptations, such as head or eye-based controls, to further improve accessibility for users with severe motor impairments.

Keywords: motor disability, individualized hand pose control, few shot learning

1 Introduction

Individuals with physical disabilities often face significant challenges in daily life, including difficulties in communication and in controlling computers or assistive devices. Recent ai-based advancements in assistive technology, especially in video-based pose and gesture detection, offer new possibilities for assistive tools. Digital motion exercises consistently show evidence of physical improvements in motor control [1]. However, due to the highly individual and often restricted movement capabilities among this population [2, 3] generic solutions are often not applicable to each individual patient. This highlights the need for individually adaptable human-machine interfaces (HMIs).

Pose detection for individuals with disabilities has received much attention in this context: Some studies have explored the use of human pose estimation to improve safety and effectiveness of human-robot interactions within rehabilitation settings. For instance, joint angle measurements during upper-limb exercises can be used for responsive robotic platforms [4]. Other systems translate hand gestures into text or speech, facilitating communication for individuals with impairments. These systems employ various devices like: data gloves, video cameras, surface electromyography and other modalities [5].

Video cameras are powerful contactless and affordable sensors for capturing human motion. Their widespread availability and ease of integration make them ideal for developing accessible HMIs. Webcams have been used for hand gesture recognition systems to enable individuals with disabilities to control machines and complete tasks based on their hand movements and detected poses [6].

In a previous work, we have investigated one-shot learning on webcam images to interpret head-poses of healthy volunteers for a personalized computer input device [7]. In the present study, we evaluate the methodology on a group of people with disabilities for individualized control using hand poses. As the physical capabilities among the group differed considerably, each person was asked to demonstrate four different hand poses according to their personal abilities. The hand poses were captured using a video camera and used to train a pose classification network. Finally, the trained model was evaluated in a real-time test application.

The contributions of this work are (1) a feasibility study of individualized hand-pose recognition for accessibility with minimal training effort and (2) highlighting the variability in suitable hand poses chosen by participants, reinforcing the need of personalized gesture recognition rather than relying on generic solutions for assistive technology.

Sofian Khabbich, Jens Christian Will, Michael Hötter, Hanno Homann, Hannover University of Applied Sciences and Arts, Ricklinger Stadtweg 120, 30459 Hannover, Germany, e-mail: hanno.homann@hs-hannover.de

https://doi.org/10.1515/cdbme-2025-0208

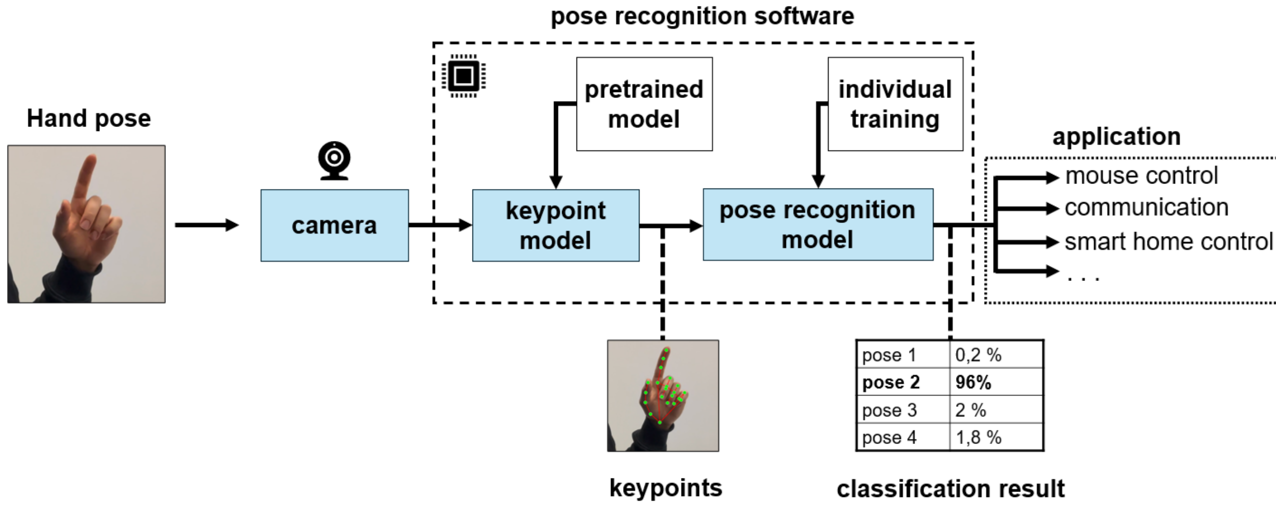


Fig. 1: Flow chart of the pose detection model

2 Methods

Data processing and model

Our objective was to capture individualized hand poses using a webcam which are then classified by a neural network for an HMI input. The workflow is illustrated in Fig. 1. Images are captured using a webcam and then processed by a pose recognition algorithm, similar as presented in [7]. First, hand pose keypoints are obtained through a keypoint extraction network. In this work, we deployed the pre-trained MediaPipe Hands model [8] by Google AI as a fast alternative to the previously used OpenPose model. The model computes a total of 21 three-dimensional keypoints for each detected hand.

The keypoint coordinates are then used as the input to a second neural network, which is individually trained to classify the presented poses. This second model is trained with a small amount of data specifically for the individual poses of each participant. Considering the strictly limited number of training data, we chose a minimal model design consisting of a single fully connected layer with softmax activation. In this work, we used the four principal directions (left, right, up, and down) as the model output classes.

Study design

The study involved people with motor impairments and working in a local day-care workshop. All participants were offered to take part in the study on a voluntary basis. A total of 14 persons took part in the study. The participants had a wide range of impairments, from minor motor impairments to paralysis and spasticity. Most of the participants had no fine motor skills in their hands.

Before the start of the study, the participants were informed in detail about the procedure. All participants were able to understand the task and procedure of the study and their consent was obtained. All types of personal data, including photos and video recordings were deliberately avoided throughout the study; instead only the keypoint coordinates were recorded anonymously.

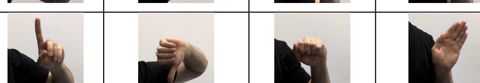
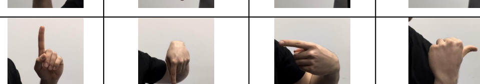
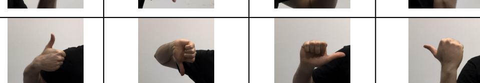
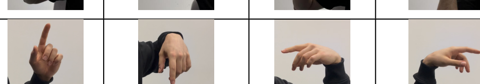
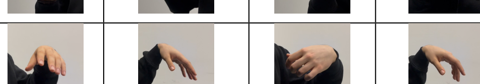
The implementation of the study was divided into three phases. In phase 1, the individual training dataset of each participant's poses was recorded. This dataset was then used to train the model in phase 2.

In phase 3, the quality of the pose prediction was evaluated in close collaboration with the participant using a real-time test application. In this application an animated figure could be moved on the screen of a notebook in four directions: up, down, left, and right. Before the participants recorded their own datasets, the control principle was demonstrated by the study leader to achieve an intuitive understanding of the animated character following the hand pose direction.

The participants were then asked to chose four individual hand poses and were completely free to chose positions matching their capabilities. Either the left or the right hand could be used. After recording the dataset and training the model, a repeated demonstration of the four poses was used to evaluate the success of the pose recognition model using the real-time application. The success in controlling the animated figure was rated subjectively by the study leader during the evaluation phase 3 on a 10-step scale from 0 (no effective control) to 1 (full intentional control).

3 Results

The participants presented their four hand poses to the camera. As no pictures were taken, Tab. 1 displays re-enactments of the participants' chosen poses. As the majority of participants were unable to fully control their fingers, the poses often mainly differed by wrist rotation (e.g. participant 10). For subjects with pronounced spasticity the four poses were difficult to distinguish visually (e.g. participant 5). Some participants, on the other hand, were able to move their fingers more clearly and thus to form more distinctive poses (e.g. participant 2).

Participant	Pose Up	Pose Down	Pose Left	Pose Right
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				

Tab. 1: Re-enactment of the individually chosen poses

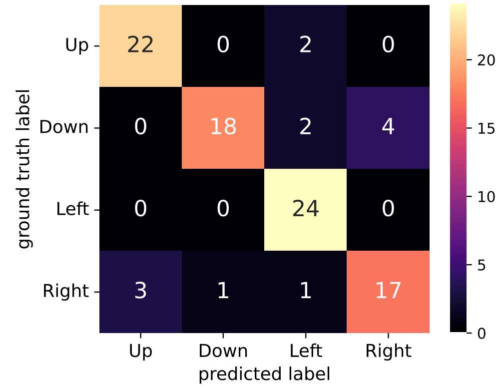


Fig. 2: Confusion matrix of the test data of all participants

The actual number of images acquired per pose varied, as the attention of the volunteers influenced the number of feasible recordings. Depending on this, between 6 and 11 recordings were taken per pose. In total, there were between 3 and 6 training recordings plus 1 to 3 validation and test set recordings per pose.

This comparatively high proportion of the validation data was chosen deliberately, as the pose recognition model based on few-shot learning should work reliably even with a small number of examples per pose. The dataset splits were strictly disjoint, ensuring that no sample used for training was reused in the validation or test sets. Importantly, test data were held back and not exposed to the model during training or validation.

Fig. 2 shows the aggregated confusion matrix of the test set across all models and subjects. Some misclassifications occur as values off the diagonal. In addition to misclassification, it is also possible to analyze which hand poses are difficult to separate due to their similarity and are therefore more error-prone for the pose recognition model. Summed over all participants, the confusion matrix shows that 86.2% of the test set data was correctly classified by the pose recognition model. Most of the misclassifications in Fig. 2 occurred for participants 5, 6, and 14. For these participants, almost all test recordings were incorrectly classified. This agreed well with the success rating of the intentional control of the animated character in the real-time test application, which is shown in Fig. 3.

A key observation here is that participants with severe limitations in finger motor skills were nevertheless able to achieve significant separability of the keypoints for the different hand poses by rotating other joints, particularly the wrist and elbow joints (e.g. participant 10). In contrast, participants with spasticity showed great difficulties in controlling the animated figure (e.g. participants 5 and 14).

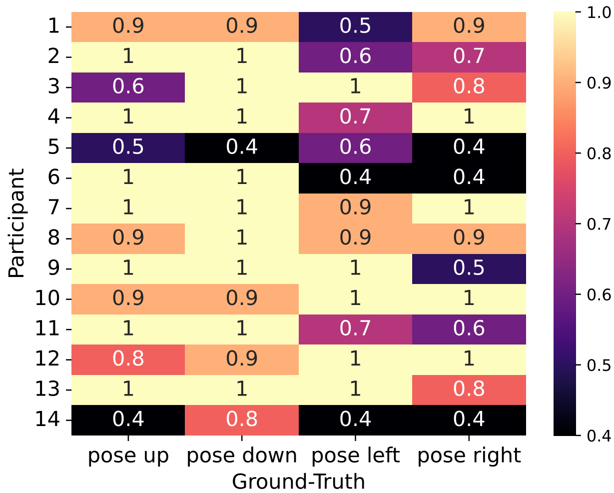


Fig. 3: Success rating for control of the animated figure in the real-time test application: 10-step scale from 0 (no effective control) to 1 (full intentional control)

4 Discussion

This feasibility study showed that the majority of our participants were able to control the animated figure satisfactorily despite their severe disabilities. When the fine motor control of the fingers was not given, the volunteers found individual hand poses to achieve the given task. Notedly, all persons enjoyed taking part and to experience themselves as self-effective through the gameified interaction.

However, it became also apparent that those participants with spasticity could hardly achieve a stable resting positions, which made the individual pose estimation more prone to error. This made it difficult for the human observers as well as for the classification network to clearly detect which of the four poses the participant was presenting during the real-time evaluation.

5 Conclusions

The study conducted showed that an individually trained pose recognition model based on few-shot learning can be a practical solution for hand control for people with severe motor impairments. Despite their limitations, most of the participants were able to use specific hand poses to successfully control the test application. It is particularly noteworthy that even people with severely limited finger mobility were able to achieve distinguishable pose control through targeted adjustments, such as the use of wrist or elbow movements.

However, the results also revealed challenges. In particular, participants with spasticity had difficulty maintaining stable poses, leading to a higher misclassification rate. This underlines the need for further research into alternative control methods, such as head or eye control, in order to provide these user groups with an effective interaction option. Overall,our study shows promising insights into the individualization of pose recognition models for people with motor impairments and potential for future developments. Further studies with extended training methods and additional control options are required to further optimize usability and recognition accuracy.

Author Statement
This work has been realized based on a cooperation between Hannover University of Applied Sciences and Arts and a workshop for persons with disablities at Annastift Leben und Lernen gGmbH, Hannover. Conflict of interest: Authors state no conflict of interest. Informed consent: Informed consent has been obtained from all individuals included in this study. Ethical approval: The research related to human use complies with all the relevant national regulations, institutional policies and was performed in accordance with the tenets of the Helsinki Declaration.

References

[1] J Corey, JM Tsai, A Mhadeshwar, S Srinivasan, and A Bhat. Digital motor intervention effects on motor performance of individuals with developmental disabilities: a systematic review. *Journal of Intellectual Disability Research*, 68(11):1221–1252, 2024.
[2] Arnold J Capute, Bruce K Shapiro, and Frederick B Palmer. Spectrum of developmental disabilities: continuum of motor dysfunction. *Orthopedic Clinics of North America*, 12(1):3–22, 1981.
[3] Casey J. Zampella, Leah A. L. Wang, Margaret Haley, Anne G. Hutchinson, and Ashley B. de Marchena. Motor skill differences in autism spectrum disorder: a clinically focused review. *Current Psychiatry Reports*, 23, 2021.
[4] Óscar G. Hernández, Vicente Morell, José L. Ramon, and Carlos A. Jara. Human pose detection for robotic-assisted and rehabilitation environments. *Applied Sciences*, 11(9), 2021.
[5] Lin Guo, Zongxing Lu, and Ligang Yao. Human-machine interaction sensing technology based on hand gesture recognition: A review. *IEEE Transactions on Human-Machine Systems*, 51(4):300–309, 2021.
[6] Rishi Sharma. Hand gesture recognition systems: As assistive technology to control machines. Master's thesis, California State University, Northridge, June 2023.
[7] Hanno Homann, Cedric Rohbani, and Jens Christian Will. One-shot learning hmi for people with disabilities. *Current Directions in Biomedical Engineering*, 10(4):319–323, 2024.
[8] Fan Zhang, Valentin Bazarevsky, Andrey Vakunov, Andrei Tkachenka, George Sung, Chuo-Ling Chang, and Matthias Grundmann. Mediapipe hands: On-device real-time hand tracking. *arXiv preprint arXiv:2006.10214*, 2020.