Recognition of Daily Living Activities Using Convolutional Neural Network Based Support Vector Machine

B.G.D.A. Madhusanka, R. Sureswaran

Abstract— The vision-based activity recognition moves toward the chance of restoring degrees of freedom to the elder and impaired in Activities of Daily Living (ADL). People can communicate their intentions in various manners however gaze-based implicit intention communication correspondence for elder people who are deaf and disable but retain visual capability were not discussed any longer in past research. Therefore, this paper proposes a strategy to improve the implicit intention communication correspondence for elder people on ADL utilizing the Convolutional Neural Network-based Support Vector Machine (CNN-SVM). To infer the intention of the user in ADL, the gaze-based implicit intention communication framework was introduced for tracking and analyzing. The system will be able to identify the necessary activities / requirements from the home environment after recognizing the user's implicit intention that will help the elder with the eye-gaze movement point them out to fulfill his/her intent. Finally, inferred intention may then be used to command caregivers to provide proper service.

Index Terms— activities of daily living, convolutional neural network, gaze-based communication, implicit intention inference, support vector machine

1 Introduction

According to more than 65% of human communication is non-verbal. In our everyday life we respond to several non-verbal cues and behaviors like postures, eye gaze, facial expression, and tone of voice. The first scientific research on nonverbal communication and behavior starts with “The Expression of the Emotions in Man and Animals of Charles Darwin” and nowadays there are an abundance of research on the types, effects and expression of unspoken communication and behavior. As stated in [1] research has identified several different types of nonverbal communication [2] such as facial expression, gestures, paralinguistic, body language and posture, proxemics, eye gaze, haptics, and appearance.

Consideration of the intention of others is an important aspect of human communication. Human communicates primarily by speech with others, but it can also be used to highlight specific features such as facial expression and gestures. Thus, a lot of recent research has focused on interpreting explicit human intent. Human intent shows by many of the above-mentioned implicit ways; therefore, it is important to make machine learn techniques how to interpret human explicit intention. Implicit knowledge, however, may not be sufficient to understand what human intent really is. Human intention is typically revealed through a combination of different expressing ways; sometimes for the purpose of emphasizing the intention, but sometimes naturally. Moreover, whether this is intended or by accident, the implicit expression of the human being is not always the same with his true intention. Thus, it is necessary to understand implicit human intention to improve the current system of interacting human computers.

Human speech ability has been reduced and the ability to speak is lost due to ageing and serious injury or diseases such as spinal trauma, paralysis, stroke, amputation, and disease of Parkinson. To these men, it would be difficult or impossible to carry out regular everyday activities such as standing up, dressing, cooking, eating, etc. independently [3]. While assistive technology, such as assistive devices , smart home systems, offer the optimism that illuminates the everyday independence of the elderly and the disabled, there is one barrier between a human user and assistive machines that can dramatically weaken the acceptance of assistive technologies. This challenge is the lack of an effective means of communication that helps elderly and disabled persons to communicate with the assistive system effectively and naturally. This problem will become even more serious with the development of increasingly assistive systems with higher function capabilities. To operate an assistive system, the user now needs to create specific requests for operation, most of which require broad motors motions [4].

Persons with hearing loss may have major adverse effects on all developmental fields, including language, cognitive , social , emotional, and behavioral functioning [5]. Because of these possible delays in development , people who are deaf or hard of hearing are at greater risk of experiencing mental health issues relative to their normally developing peers. Researchers evaluating people's emotional and behavioral issues have traditionally relied on questionnaires from home mates [6].

Also, increasing number of elderly people, particularly those living alone and apart from the rest of their family, has become a serious problem. Monitoring those people is effective but difficult and a burden on their family [7]. The deaf people reveal significant challenges in communication with caregivers since communication barriers in life are linked to poor physical and mental health of them. Also, society is getting complex and due to that need to pay high wages for caregivers, difficulty in finding well trained caregivers, and unlikeliness of caregiving tasks [8]. In addition to that, shortage of caregivers will become a serious problem in near future.

To prevent that, novel method is recognized the implicit intention of a human user, by using nonverbal communication to identify intention of the user to identify the necessary activities/requirements from the domestic area that is going to help the elder with the eye-gaze movement point them out to fulfill his/her intention [9]. Movement of the eye gaze requires relatively little human effort, and the capacity to regulate the gaze persists in most elderly and disabled individuals with movement impairments, making the gaze a very effective way of communicating with an assistive device for the elderly and disabled [10]. This concept is expected to work with the elderly people who are in deaf and disabled people to uplift their living standards.

2 SYSTEM OVERVIEW

Recognizing the human purpose from the objects deliberately studied by an individual, the framework needs knowledge about the relation between objects and purpose. This knowledge depicts how objects are linked to a certain form of purpose, or how objects may fulfill a kind of intent. This awareness is conceived as a model of Nave Bayesian Graphic Probabilistic (NBGP) [11]. NGBP model that can be used in a wide variety of classification tasks. Since it is a probabilistic model, the algorithm can be coded up easily and the predictions made real-time quick. Because of this, it is easily scalable and is traditionally the algorithm of choice for real-world applications that are required to respond to user's requests instantaneously. Though NBGP model performs well in prediction accuracy, it assumes strong independence among features given class variable and assumption may reduce prediction accuracy when two or more features are dependent given class variable. To improve prediction accuracy, the model can relax NBGP assumption and allow dependencies among features given class variable. Capturing feature dependencies more likely improves prediction accuracy for classification applications in more than two features have some correlation given class variable. The above Probabilistic model can be encoded using Bayesian Network as shown in Fig. 1 [11].

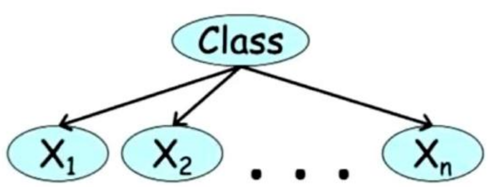

Fig. 1: Probabilistic graphical model for Naïve Bayes classifier

Experimentally identified objects in the kitchen scenario. The intention knowledge is developed to identify the objects in the kitchen. The experiment focused on visual attention identification and inference of intention which are essential aspects of the home environment. Thereafter, modules are combined to form the entire structure outlined in Fig. 2. The user views the live kitchen scene fed back from the Human Machine Interface (HMI). During the purpose expression procedure, the HMI is assumed to hold a stable scene image which contains objects for the user. By clicking on the button, once the user activates the inference engine, the device starts analyzing the data from the eye gaze to extract the objects being viewed and to infer the user's intent. The assumed result will be shown to consumer on the monitor for confirmation. The position of the visualized object can be determined from the location of the object in the scene image as a pointing vector from the scene camera to the target.

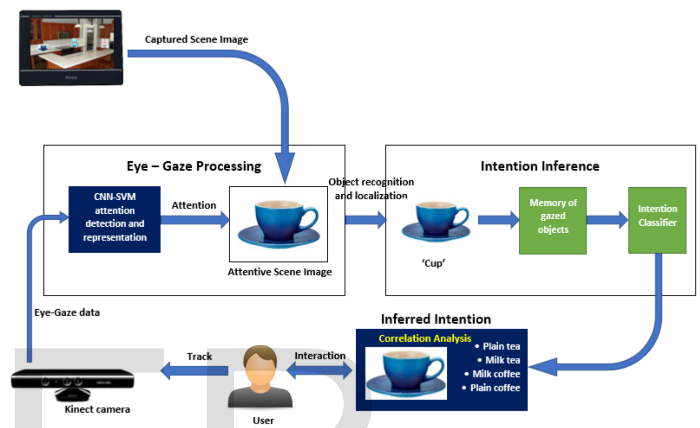

Fig. 2: Gaze-based implicit intention communication framework

3 CONVOLUTIONAL NEURAL NETWORK BASED SUPPORT VECTOR MACHINE

Care identification using natural visual habits does not allow users to prepare and memorize specific instructions or execute artificial actions relative to the conventional long-sighted method and deliberate blink, thus minimizing the user's deliberate effort. A combined CNN-SVM model [12] is used for cluster gaze points on the visualized object after increasing focus is observed, and an analogous circle is determined to represent the cluster of gaze-points. Finally, the visualized objects can be identified and recognized using the picture on the scene from the overlap of the area of human attention.

The intentional gaze detection classifier for visual-behavior is based on a CNN-SVM, classification procedure is shown in the Fig. 3 [12, 13]. This classifier uses eye-gazing features derived from natural imagery, which do not compel users to perform any abnormal visual behaviour, such as prolonged gazing or a deliberate blink. The features used for this classification include gaze dwelling time, pupil size/gradient variation, gaze velocity and gaze concentration. These characteristics were selected based on the literature review and the experimental observation [13]. The classifier is conditioned until he can differentiate between the purposeful gaze and the intention-free gaze. The extracted eye-gaze features are inserted into the classifier during use, and the output is its best guess of which visualization state the user is in. And the visual attention

ISSN 2229-5518

sensed is when one detects the intentional focus. The extracted eye-gaze features are entered in the classifier, and the output is its best guess on which visualization state the user is in.

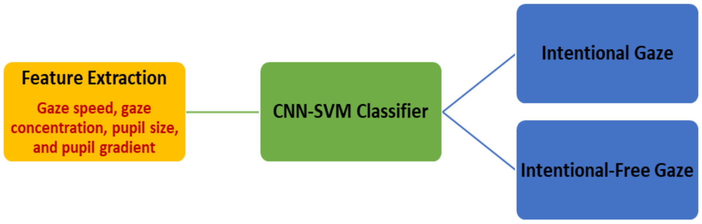

Fig. 3: Detection of intentional gaze using a CNN-SVM classifier

After the CNN classifier is well equipped, the SVM classifier conducts the classification by replacing the fully linked layers. The data, feature maps and kernels are typically organized in square matrixes since the CNN is mainly designed for image processing. The process for implementing the combined CNN-SVM model is shown in Fig. 4 [12].

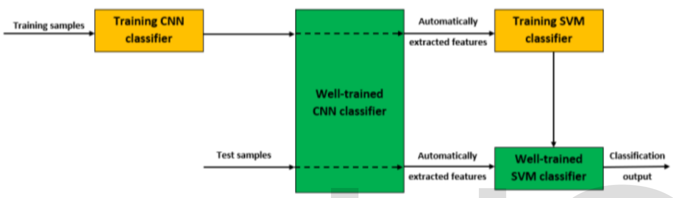

Fig. 4: Implementation process of the CNN-SVM combined model

4 Conclusion

In this paper, an implicit novel, gaze-based communication of intent for elderly people who are deaf and disabled but retain visual ability. To infer the user's intent in ADL, the gaze-based implicit purpose communication system for monitoring and analyzing has been introduced. In the experiment, CNN-SVM was used to enhance the implicit communication of wishes for the elderly on ADL. The awareness of the system was interpreted to understand the human purpose of how objects are connected to a certain form of purpose, or how objects may fulfill a form of intent. Based on this information as NBGP basis. Implicit purpose gaze-based contact modules are combined to form the entire system, where the consumer looks at the live scene from the kitchen feedback the HMI. During the intent expression process, the HMI is assumed to hold a stable scene image that contains objects for user. The goal of this paper is to promote the ADL in deaf and disabled people, check the proposed structure for engagement with target users, and the author's next step is to get results and validation.

References

[1] Dong, S. Y., Kim, B. K., & Lee, S. Y. (2016). Implicit agreeing/disagreeing intention while reading self-relevant sentences: a human fMRI study. Social neuroscience, 11(3), 221-232.

[2] Fornaser, A., De Cecco, M., Mizumoto, T., & Yasumoto, K. (2019, February). Reliability assessment on human activity recognition. In 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC) (pp. 363-368). IEEE.

[3] Ni, Q., Patterson, T., Cleland, I., & Nugent, C. (2016). Dynamic detection of window starting positions and its implementation within an activity recognition framework. Journal of biomedical informatics, 62, 171-180.

[4] Moladande, M. W. C. N., & Madhusanka, B. G. D. A. (2019, March). Implicit Intention and Activity Recognition of a Human Using Neural Networks for a Service Robot Eye. In 2019 International Research Conference on Smart Computing and Systems Engineering (SCSE) (pp. 38-43). IEEE.

[5] Chen, L. B., Tsai, C. W., Chang, W. J., Cheng, Y. M., & Li, K. S. M. (2016, January). A real-time mobile emergency assistance system for helping deaf-mute people/elderly singletons. In 2016 IEEE International Conference on Consumer Electronics (ICCE) (pp. 45-46). IEEE.

[6] Yousaf, K., Mehmood, Z., Saba, T., Rehman, A., Rashid, M., Altaf, M., & Shuguang, Z. (2018). A novel technique for speech recognition and visualization based mobile application to support two-way communication between deaf-mute and normal peoples. Wireless Communications and Mobile Computing, 2018.

[7] Kuenburg, A., Fellinger, P., & Fellinger, J. (2016). Health care access among deaf people. The Journal of Deaf Studies and Deaf Education, 21(1), 1-10.

[8] Smith, S., Doyle, J., Hannigan, C., Murphy, E., & Dinsmore, J. (2017). Supporting Older People with Multimorbidity: Experiences of Caregiver Burden among Paid Care Workers in Ireland. Age and Ageing, 46(3), iii13-iii59.

[9] Salimu, S., Kinshella, M., Marianne, M., Banda, M., Newberry, L., Dube, Q., ... & Nyondo-Mipando, A. L. (2020). "Those women have hope in us and if you explain properly, they understand": Health workers' views on factors affecting caregiver engagement with bubble CPAP.

[10] Li, S., & Zhang, X. (2017). Implicit intention communication in human–robot interaction through visual behavior studies. IEEE Transactions on Human-Machine Systems, 47(4), 437-448.

[11] Qian, W., Lauri, F., & Gechter, F. (2019). A Probabilistic Approach for Discovering Daily Human Mobility Patterns with Mobile Data. arXiv preprint arXiv:1911.09355.

[12] Wu, H., Kitagawa, Y., Wada, T., Kato, T., & Chen, Q. (2007, November). Tracking iris contour with a 3D eye-model for gaze estimation. In Asian Conference on Computer Vision (pp. 688-697). Springer, Berlin, Heidelberg.

[13] Joseph, A. W., & Murugesh, R. (2020). Potential Eye Tracking Metrics and Indicators to Measure Cognitive Load in Human-Computer Interaction Research. Journal of Scientific Research, 64(1).