Human behavior understanding from motion and bodily cues using deep neural networks

IMPACT PARAGRAPH

In this addendum, the scientific as well as social impact of the presented work is discussed. The paragraph below addresses the drafted four questions given in the "Regulations for obtaining the doctoral degree Maastricht University" [1].

(Research) What is the main objective of the research described in the thesis and what are the most important results and conclusions?

This thesis investigated how automatic human behavior understanding from video data can support real-world applications in the fields of Ambient Assisted Living (AAL), Surveillance and Affective Computing (AC). Unlike language, which has well studied semantic and syntactic structure, human behaviors are lacking a generic underlying architecture [2]. As a matter of fact, human behaviors are generated in different forms and levels of complexity. In this thesis, we aim to study human behaviors in different environments by interpreting how body motion and body postures evolve over time. Specifically, in Chapter 3 and Chapter 4, we build two novel frameworks that extract and interpret spatio-temporal motion features using trajectory data. We demonstrate that by learning the spatial as well as the temporal distribution of the trajectory points, we can detect abnormal behaviors and predict in near real-time what will happen next. These models showed to be beneficial in surveillance applications, such as detecting abnormal events in public environments like train stations or public squares, and in AAL applications, such as detecting dangerous events in elderly affected by dementia and Alzheimer's disease. In Chapter 5, Chapter 6, and Chapter 7, we build novel frameworks to encode body postures and their interaction with context information over time. Our main objective in these chapters is to learn stable behavioral patterns defined as *Personality patterns*. Our research highlights that integrating computational models with psychological theories, such as the Big-5 personality traits, can help the interpretation of human behaviors in social as well as nonsocial environments.

(Relevance) What is the (potential) contribution of the results from this research to science, and, if applicable, to social sectors and social challenges?

We are in the midst of a wave of technological innovations that are revolutionizing several sectors of our society. However, the integration and the automation of new technologies in our society remains a fundamental challenge to be tackled. In this thesis, we study the automatic understanding of human behaviors and its applications in the fields of Ambient Assisted Living, Surveillance and Affective Computing.

As the average population age in Europe, and in the world in general, is steadily increasing, new challenges surrounding the economic burden of having more old people than young are emerging. In this context, smart and automatized healthcare applications made with low-cost sensors could reduce the economic impact while improving the living conditions of the old population [3]. In this thesis, we investigate the detection of abnormal events such as confusion and repetitive behaviors that impact elderly affected by dementia and Alzheimer's disease. In Chapter 3, we propose a dataset which

encourages the analysis of human behaviors in unrestricted settings for the discovery of abnormal patterns from spontaneous activities. Datasets containing abnormal agitation and confusion behaviors from video data are rare to find as well as difficult to collect. Hence, with the help of medical professionals, in this dataset, we designed certain activities which might provoke trajectory data similar to the ones created by people in confusion states. The dataset is used to design and train a machine learning model that distinguishes between normal and abnormal activities. Finally, by applying a transfer learning strategy, we test the model directly in hospitals with four patients affected by Alzheimer's disease. This work is part of the ICT4Life European project, which aimed to implement a platform integrating a series of innovative services, targeting aging people with cognitive impairments.

Affects and emotions are fundamental human experiences that have a great impact on humans' lives, choices, well-being, and so forth. Consequently, the ability to automatically recognize and interpret affective attributes and emotional patterns can have a huge impact on several sectors in our society. Specifically, in this thesis, we study how human body postures, interactions, and behaviors can be automatically mapped to personality labels. As human beings, we are able to interpret affective states of other individuals from little information. Therefore, the obtained results are of great importance for future interdisciplinary behavioral studies, aiming to combine data-driven approaches with psychological studies to enhance the understanding of human behavior from machines' perspective.

(Target group) To whom are the research results interesting and/or relevant? And why?

The primary target group of the presented studies is the researchers in Computer Vision and the field of AI in general. Five novel frameworks have been introduced, from Chapter 3 to Chapter 7. Quantitative as well as qualitative experiments are carried out to investigate the efficacy of our methods on public datasets against state-of-the-art computer vision algorithms.

The secondary target group is the potential users of the AI applications presented in this thesis and are described below.

8

For the users in the healthcare field, we presented the ICT4Life platform (Chapter 3), which aimed at providing innovative ICT services targeting the aging population. The ICT4Life platform includes AI algorithms to monitor, detect, and prevent dangerous events supporting the old generation to live independently for as long as possible. Nevertheless, the aging process brings several difficulties that affect not only the old generation, but also the elder's family and healthcare professionals. In this context, the ICT4Life platform focuses on the integration of different information (sensory information as well as medical files) to provide high level insights on the patient's health condition. These insights are accessible to all the users of the platform, i.e. patient, carers, medical staff.

For the users in the surveillance field, long and unconstrained security footage is hard to monitor due to limited human resources. The ability to hold attention and to react in case of rarely suspicious events is demanding and prone to human error [4]. Thus, the presented work in Chapter 3 and Chapter 4 which differentiate important events versus unimportant events is critical to optimize the limited attention of surveillants, and, at the same time, to alleviate the costs of surveillance systems.

For the users in the affective computing field, we presented three novel frameworks

(Chapter 5, Chapter 6, and Chapter 7) which link body postures and context interactions to personality attributes. The detection and the recognition of human affective cues can be applied in several applications for industries as well as for education. For example, companies can understand their customers better and provide more satisfactory services by pulling emotional strings that are difficult to reach with the standard marketing strategies. Schools and universities can personalize the educational contents by looking and understanding the involvement of the students.

(Activity) In what way can these target groups be involved in and informed about the research results, so that the knowledge gained can be used in the future?

In this thesis, we presented five works that have been published to peer-reviewed international conferences as well as high-impact journals. The work published in peer-reviewed conferences has been displayed in presentation form (Chapter 3 and Chapter 7) as well as poster form (Chapter 5). Furthermore, the work in Chapter 3 was part of the ICT4Life European project [2], and it has been presented during international consortium meetings as well as revisions to the European commission.

REFERENCES

[1] UM, "Regulations for obtaining the doctoral degree Maastricht University," *Regulation Governing the Attainment of Doctoral Degrees*, 2018.

[2] L. C. G. F. d. Santos, *Laban Movement Analysis: A Bayesian Computational Approach to Hierarchical Motion Analysis and Learning.* PhD thesis, 2014.

[3] R. Li, B. Lu, and K. D. McDonald-Maier, "Cognitive assisted living ambient system: a survey," *Digital Communications and Networks*, vol. 1, no. 4, pp. 229–252, 2015.

[4] A. Hampapur, L. Brown, J. Connell, S. Pankanti, A. Senior, and Y. Tian, "Smart surveillance: applications, technologies and implications," in *Fourth International Conference on Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint*, vol. 2, pp. 1133–1138, IEEE, 2003.

[2] https://cordis.europa.eu/project/id/690090