Machine Vision and Applications (2019) 30:807–811
https://doi.org/10.1007/s00138-019-01031-x

EDITORIAL

Guest editorial: special issue on human abnormal behavioural analysis

Gholamreza Anbarjafari[1,2,3] · Sergio Escalera[4] · Kamal Nasrollahi[8] · Hugo Jair Escalante[6] · Xavier Baro[5] · Jun Wan[7] · Thomas B. Moeslund[8]

Received: 1 May 2019 / Accepted: 7 May 2019 / Published online: 22 May 2019

An abnormality or anomaly[1] can be defined as the event or phenomenon that deviates from a common rule, type, arrangement and/or form. Without making any qualitative judgement and/or taking any posture on the validity of the sense of abnormality,[2] it is under no doubt that abnormalities represent entities of interest in several health, manufacturing, leisure and security scenarios: whether to apply corrective actions (e.g., on industrial applications, surveillance) and preventive treatments (e.g., in health, security), to perform in-depth exploration studies (e.g., sciences, like genomics, astrophysics, bio medicine, mines research), to invest resources (e.g., oil research, advertising campaigns) or to simply identify abnormalities and track their progress (e.g., patient monitoring) [2,4,5,7]. This applies for computer vision and its applications, where abnormalities have been the implicit motivation and subject of research in many studies and research fields.

The subject of anomalies is of particular interest to the looking at people (LaP) field. LaP focuses on the visual analysis of human behaviour, action/gesture/activity recognition, facial analysis and pose estimation. Since humans are involved in LaP tasks, it is clear that the study of abnormalities is critical in this context. This is particularly true in LaP applications like health and security, where uncommon patterns/behaviours are the most interesting ones. Despite its relevance, it is remarkable that so far the study of abnormalities in LaP, and more generally in computer vision, has not been characterised. Questions such as what makes an event/situation/behaviour abnormal, or how abnormal behaviour in images or video has been analysed, cannot be answered without reading a dozen of articles.

We edited a special issue that aims to compile the latest research human abnormal behavioural analysis. The focus of this issue is on all forms of behaviours and all kinds of atypical activities, hence comprising a broad scope across several fields of study. Including computer vision and machine learning, and benefiting from others are partially related to the issue in the aforementioned multi-disciplinary context. We received a total of twenty-two submissions that were subject to a strict reviewing process. Despite the high quality of all of the submissions, we only accepted fourteen outstanding contributions. In the following, we briefly summarise these contributions, grouped in thematic subjects.

Action and gesture recognition Abnormalities have been widely studied in the context of action and gesture recognition. Several contributions on this topic are part of the

✉ Gholamreza Anbarjafari
shb@icv.tuit.ut.ee

Sergio Escalera
sergio@maia.ub.es

Kamal Nasrollahi
kn@create.aau.dk

Hugo Jair Escalante
hugojair@inaoep.mx

Xavier Baro
xbaro@uoc.edu

Jun Wan
jun.wan@ia.ac.cn

Thomas B. Moeslund
tbm@create.aau.dk

1 iCV Lab, University of Tartu, Tartu, Estonia

2 Hasan Kalyoncu University, Gaziantep, Turkey

3 IDT, Loughborough University London, London, UK

4 Universitat de Barcelona and Computer Vision Center, Barcelona, Catalonia, Spain

5 Computer Vision Center and Universitat Oberta de Catalunya, Catalonia, Spain

6 Instituto Nacional de Astrofísica, Óptica y Electrónica, and Centro de Investigación y Estudios Avanzados del IPN, Mexico, Mexico

7 NLPR- CASIA, Beijing, China

8 University of Aalborg, Aalborg, Denmark

[1] Throughout this editorial, we use indistinguishably abnormality or anomaly.
[2] The authors emphasise that they are not using the term abnormality in any discriminatory or disrespectful sense.

special issue. Li et al. describe *a spatiotemporal attention-based ResC3D model for large-scale gesture recognition* with an emphasis on abnormal gesture recognition from RGBD videos [9]. They conducted an experimental evaluation using the Chalearn IsoGD Dataset [15] showing the effectiveness of the attention model. In the same line, Lin et al., in their *Abnormal Gesture Recognition Based on Multi-Model Fusion Strategy* article [10], proposed a novel refined fused model that combines a masked (RGB and Depth) ResC3D networks with a skeleton LSTM for abnormal gesture recognition in RGBD videos. Based on their experimental results, their proposed method could distinguish the abnormal gesture samples effectively achieving state-of-the-art performance in the IsoGD dataset. The IsoGD dataset is becoming a useful resource for studying abnormal behaviour [14,15].

On the other hand, Raman et al. focused on predicting the walking directions of pedestrians, a topic that can have a direct impact on abnormal behavioural analysis. In their work entitled *Beyond Estimating Discrete Directions of Walk: A Fuzzy Approach* [11], the authors proposed a type 1 fuzzy approach over apposite inter- and intra-frame locomotion features of pedestrian to predict eight different directions of walk. Their exhaustive experimental results showed that their proposed methodology overcome the existing state-of-the-art techniques in direction of walk prediction.

Instead of working with videos, Hu et al. focused on still images to study the problem of abnormal behavioural detection in the driving context [6]. In their *Driving behavior Recognition from Still Images by Using Multi-stream Fusion CNN* article, Hu et al. used multi-stream CNNs to extract multi-scale features by filtering images with receptive fields of different kernel sizes and then used fusion of the multi-scale information to generate their final prediction. The conducted experiments indicated that their method achieves the significant performance improvements compared to the state of the art. Likewise, Ding et al. dealt with the problem of posture assessment, a topic that also can have a direct impact on abnormal behavioural analysis [3]. In their *A Real-Time Webcam-Based Method for Assessing Upper-Body Postures* paper, Ding et al. presented a vision-based method for real-time assessment of upper-body postures of a subject whose lower body is not visible to the camera. They introduced a database of upper-body postures that cover the various scenarios when a subject is sitting in front of a desk.

Applications with potential societal impact As previously mentioned, abnormalities can be found in any domain and their detection could be used develop systems that can have an impact on society. A couple of submissions in this line are part of the special issue.

On the one hand, Sun et al. report a study for automatically monitoring infants in the context of health clinics [13]. In their paper entitled *Video-Based Discomfort Detection for Infants*, Sun et al. proposed a video-based method for automated detection of discomfort/pain of newborn infants based on a two-phase classification workflow, where Phase 1 was subject-independent (considering generic visual descriptors) and Phase 2 was subject-dependent (including subject templates). Experimental results in a database with 12 infants revealed almost perfect performance for the subject-dependent setting with lower performance for the subject-independent phase. On the other hand, Lashkar et al. introduced, in their paper *A motion-based waveform for the detection of breathing difficulties during sleep* [8], a method based on motion to detect breathing movements of sleeping patients from infrared videos. The detected movements are mapped into a waveform where different sleeping–breathing patterns can be detected. Experimental results showed the effectiveness of their proposed technique, motivating further research on automatic recognition of breathing events.

In general terms, both previous papers addressed quite novel scenarios in applications that could have a direct impact on society. We foresee these works will motivate further research on abnormal behavioural detection in health.

Surveillance applications Perhaps the most common application for abnormality detection methodologies from visual information is surveillance/security. Because of this, it is of no surprise that the special issue is formed by several articles in such topic.

Zhang et al. introduced a method for reducing the effect of perspective distortion present in real surveillance scenarios [16]. In their article *Detection of Abnormal behavior in Narrow Scene with Perspective Distortion*, the authors introduced a mechanism for making up the distorting effect in the region of interest extraction followed by improved pyramid L-K optical flow method with perspective weight and disorder coefficient. A set of rules are defined based on the optical flow monitored across frames. Their experimental results indicated that their algorithm can effectively compensate for the distortion effect and improve the accuracy of abnormal behavioural detection.

Instead of focusing on subjects' behaviours, Shean et al. focused on crowd behaviour in their paper entitled *Statistical detection of a panic behavior in crowded scenes* [12]. They proposed a sparse representation that accounts for temporal variation of motion magnitudes and orientations, the crowd density and interactions. Their goal was characterising and detecting mass abnormal behaviour; they achieved this goal by characterising the distribution of normal crowd behaviour and identifying outliers of such distribution as *panic* behaviour. The experimental evaluation showed the superiority of their proposed method over the existing state-of-the-art techniques. The same problem was studied by Xu et al. in their work *Dual-Channel CNN for Efficient*

Abnormal Behavior Identification through Crowd Feature Engineering [16]. The authors proposed a deep learning-based solution that is feed with raw video and with a motion-based spatiotemporal descriptor, resulting in a two-channel convolutional network that learns to detect abnormal behaviour in crowds. Their proposed method showed substantial improvements in terms of detection accuracy and efficiency for online crowd abnormal behavioural detection.

The special issue also includes methods that study abnormal behaviour for surveillance from other perspectives. Camina et al. describe in their paper *Bagging RandomMiner: A one-class classifier for file accesses-based masquerade detection* an ensemble classifier for file access-based masquerade detection by using a one-class classification approach. Based on their work, it has been shown that one-class classification attains better classification results than many state-of-the-art multi-class classifiers for such important problem.

In terms of security applications, the special issue includes relevant articles covering subject-level and crowd-level solutions. In addition, solutions that go beyond video surveillance are also considered.

Unconscious behavioural analysis Last but not least, the special issue received several submissions on the detection of abnormal unconscious behaviours. This sort of behaviours refers to aspects that go beyond the conscious/standard ones, for example personality and emotion analysis, in which the sole visual information is not enough for an effective analysis.

Avots et al. focused on recognising emotions from audio–visual information by considering a cross-corpus evaluation. In their paper entitled *Audio-Visual Emotion Recognition in Wild*, the authors perform an extensive evaluation of multimodal features when recognising emotions across datasets. Interestingly, they provide a comparison between automatic emotion recognition using automatic techniques and human decision-makers. At a broader scale, Favaretto et al. introduced a methodology to detect personality and basic emotion characteristics of crowds in video sequences. In their work *Detecting Personality and Emotion Traits in Crowds from Video Sequences*, Favaretto et al. explore the feasibility of inferring personality (OCEAN model) and emotion from groups of people. In their experimental results, they showed that mapping emotion characteristics to OCEAN dimensions generates coherent information when compared to data provided in the available literature.

From a different perspective, Bablani et al. studied the problem of deception detection in their work *An Efficient Concealed Information Test: EEG Feature Extraction and Ensemble Classification for Lie Identification* [1]. They approached the problem of detecting lies from subjects by using EEG signal data. Ten subjects who participated in the study were shown images that should cause a stimuli. Signals were represented with traditional features, and an ensemble model was adopted for classification.

Although it is clear that characterising unconscious behaviours is a complicated task that is far from being solved. Efforts like those described above will motivate further research on this relevant topic.

From the above survey, it is clear that the special issue covers a wide diversity in terms of methodologies, applications and even information sources. This evidences the broadness of the field of abnormal behavioural analysis. All of the articles in this issue address important research questions and propose novel solutions that resulted quite competitive. More importantly, these articles pose additional questions that will surely motivate further research on the subject in the forthcoming years.

As guest editors of this special issue, we carefully examined the feedback and discussions provided by the authors and reviewers throughout the revision cycle. This allowed us to filter out a number of high-level research contributions on human behavioural analysis which are focusing on addressing some abnormality detection. We were delighted to see that many researchers have started working on this particular challenge, which this SI was designed for.

The guest editors express their appreciation to the authors for their high-quality work and their contribution to the state of the art of this emerging domain. We would like to thank all reviewers for their helpful feedback and comments. We thank Professor Rahul Sukthankar for his support of this special issue and the Springer team for their support.

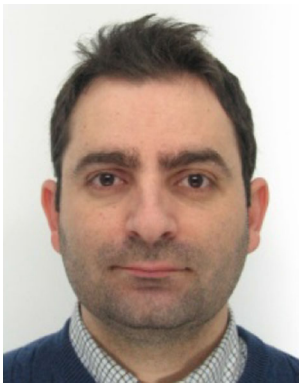
We hope you find this special issue both interesting and inspiring.

Acknowledgements This work has been partially supported by the Estonian Research Council Grant PUT638, the Spanish Projects TIN2015-65464-R and TIN2016-74946-P (MINECO/FEDER, UE), CERCA Programme/Generalitat de Catalunya and the Estonian Centre of Excellence in IT (EXCITE) funded by the European Regional Development Fund). This work is partially supported by ICREA under the ICREA Academia programme.

References

1. Bablani, A., Edla, D.R., Tripathi, D., Kuppili, V.: An efficient concealed information test: Eeg feature extraction and ensemble classification for lie identification. Mach. Vis. Appl. (2018). https://doi.org/10.1007/s00138-018-0950-y
2. Bellantonio, M., Haque, M.A., Rodriguez, P., Nasrollahi, K., Telve, T., Escalera, S., Gonzalez, J., Moeslund, T.B., Rasti, P., Anbarjafari, G.: Spatio-temporal pain recognition in cnn-based super-resolved facial images. In: Video Analytics. Face and Facial Expression Recognition and Audience Measurement, pp. 151–162. Springer (2016)
3. Ding, Z., Li, W., Ogunbona, P., Qin, L.: A real-time webcam-based method for assessing upper-body postures. Mach. Vis. Appl. (2019). https://doi.org/10.1007/s00138-019-01033-9

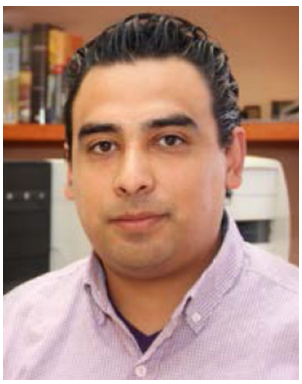
4. Gorbova, J., Colovic, M., Marjanovic, M., Njegus, A., Anbarjafari, G.: Going deeper in hidden sadness recognition using spontaneous micro expressions database. Multimedia Tools Appl (2019). https://doi.org/10.1007/s11042-019-7658-5
5. Haque, M.A., Bautista, R.B., Noroozi, F., Kulkarni, K., Laursen, C.B., Irani, R., Bellantonio, M., Escalera, S., Anbarjafari, G., Nasrollahi, K., et al.: Deep multimodal pain recognition: a database and comparison of spatio-temporal visual modalities. In: 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), pp. 250–257. IEEE (2018)
6. Hu, Y., Lu, M., Lu, X.: Driving behaviour recognition from still images by using multi-stream fusion CNN. Mach. Vis. Appl. (2018). https://doi.org/10.1007/s00138-018-0994-z
7. Kulkarni, K., Corneanu, C., Ofodile, I., Escalera, S., Baro, X., Hyniewska, S., Allik, J., Anbarjafari, G.: Automatic recognition of facial displays of unfelt emotions. IEEE Trans. Affect. Comput. 1–14 (2018)
8. Lashkar, S., Ammar, H.: A motion-based waveform for the detection of breathing difficulties during sleep. Mach. Vis. Appl. (2018). https://doi.org/10.1007/s00138-018-0980-5
9. Li, Y., Miao, Q., Qi, X., Ma, Z., Ouyang, W.: A spatiotemporal attention-based resc3d model for large-scale gesture recognition. Mach. Vis. Appl. (2018). https://doi.org/10.1007/s00138-018-0996-x
10. Lin, C., Lin, X., Xie, Y., Liang, Y.: Abnormal gesture recognition based on multi-model fusion strategy. Mach. Vis. Appl. (2018). https://doi.org/10.1007/s00138-018-0969-0
11. Raman, R., Boubchir, L., Sa, P.K., Majhi, B., Bakshi, S.: Beyond estimating discrete directions of walk: a fuzzy approach. Mach. Vis. Appl. (2018). https://doi.org/10.1007/s00138-018-0939-6
12. Shehab, D., Ammar, H.: Statistical detection of a panic behavior in crowded scenes. Mach. Vis. Appl. (2018). https://doi.org/10.1007/s00138-018-0974-3
13. Sun, Y., Shan, C., Tan, T., Long, X., Pourtaherian, A., Zinger, S.: Video-based discomfort detection for infants. Mach. Vis. Appl. (2018). https://doi.org/10.1007/s00138-018-0968-1
14. Wan, J., Escalera, S., Anbarjafari, G., Jair Escalante, H., Baró, X., Guyon, I., Madadi, M., Allik, J., Gorbova, J., Lin, C., Xie, Y.: Results and analysis of chalearn lap multi-modal isolated and continuous gesture recognition, and real versus fake expressed emotions challenges. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 3189–3197 (2017)
15. Wan, J., Zhao, Y., Zhou, S., Guyon, I., Escalera, S., Li, S.Z.: Chalearn looking at people rgb-d isolated and continuous datasets for gesture recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 56–64 (2016)
16. Xu, Y., Lu, L., Xu, Z., He, J., Zhou, J., Zhang, C.: Dual-channel cnn for efficient abnormal behavior identification through crowd feature engineering. Mach. Vis. Appl. (2018). https://doi.org/10.1007/s00138-018-0971-6



Gholamreza Anbarjafari heads the Intelligent Computer Vision (iCV) Research Lab in the Institute of Technology at the University of Tartu. He is also a Reader at Institute of Digital Technologies, Loughborough University London. He is an IEEE Senior Member and the Chair of the Signal Processing/Circuits and Systems/Solid-State Circuits Joint Societies Chapter of the IEEE Estonian section. He received the Estonian Research Council Grant (PUT638) and the Scientific and Technological Research Council of Turkey (Project 1001–116E097) in 2015 and 2017, respectively. He has been involved in many international industrial projects. He is expert in computer vision, human-robot interaction, affective computing, human behavior analysis and artificial intelligence. He is an associate editor of several journals such as SIVP and JIVP and has been a lead guest editor of several special issues on human behavior analysis. He has supervised over 10 MSc students and 7 PhD students. He has published over 100 scientific works. He has been in the organizing committee and technical committee of conferences such as ICOSST, ICGIP, SIU, SampTA, FG and ICPR. He is organizing a challenge and a workshop on FG17, CVPR17, ICCV17, and ECML19.

Sergio Escalera is an associate professor at the Department of Mathematics and Informatics, Universitat de Barcelona. He is an adjunct professor at Universitat Oberta de Catalunya, Aalborg University, and Dalhousie University. He obtained the PhD degree on multi-class visual categorization systems at the Computer Vision Center, UAB. He obtained the 2008 Best Thesis Award on Computer Science at Universitat Autònoma de Barcelona. He leads the Human Pose Recovery and Behavior Analysis Group at UB and CVC. He is also a member of the Computer Vision Center at Campus UAB. He is an expert in human behavior analysis in temporal series, statistical pattern recognition, visual object recognition, and HCI systems, with special interest in human pose recovery and behavior analysis from multi-modal data. He is vice-president of ChaLearn Challenges in Machine Learning, leading ChaLearn Looking at People events. He is Chair of IAPR TC-12: Multimedia and visual information systems.

Kamal Nasrollahi is currently employed as associate professor at Visual Analysis of People (VAP) Laboratory in Aalborg University where he is involved in several national and international research projects. He obtained his MSc and PhD degrees from Amirkabir University of Technology and Aalborg University, in 2007 and 2010, respectively. His main research interest is in facial analysis systems on which he has published more than 80 peer-reviewed papers on different aspects of such systems in several international conferences and journals. He has won three best conference-paper awards. He has served as a technical committee member for several conferences and reviewer for several journals.



Hugo Jair Escalante received the PhD degree in computer science, for which he received the Best PhD Thesis on Artificial Intelligence 2010 Award (Mexican Society in Artificial Intelligence). He is researcher scientist at Instituto Nacional de Astrofisica, Optica y Electronica, INAOE, Mexico. In 2017, he received the UANL research award in exact sciences. He is secretary and member of the board of directors of ChaLearn, a non-profit organization dedicated to organizing challenges, since 2011. He is information officer of the IAPR Technical Committee 12. Since 2017, he is editor of the Springer Series on Challenges in Machine Learning. He has been involved in the organization of several challenges in machine learning and computer vision, collocated with top venues, see http://chalearnlap.cvc.uab.es/. He has served as co-editor of special issues in the International Journal of Computer Vision, the IEEE Transactions on Pattern Analysis and Machine Intelligence, and the IEEE Transactions on Affective Computing. He has served as area chair for NIPS 2016-2018, is data competition chair of PAKDD 2018 and has been member of the program committee of venues like CVPR, ICPR, ICCV, ECCV, ICML, NIPS, IJCNN. His research interests are in machine learning, challenge organization, and its applications on language and vision.

Xavier Baro received his B.S. degree in Computer Science at the UAB in 2003. In 2005, he obtained his MS degree in Computer Science at UAB and in 2009 the PhD degree in Computer Engineering. At present, he is associate professor and researcher at the Computer Science, Multimedia and Telecommunications Department at Universitat Oberta de Catalunya (UOC).

Jun Wan received his B.S. degree from the China University of Geosciences, Beijing, China, in 2008 and the PhD degree from the Institute of Information Science, Beijing Jiaotong University, Beijing, China, in 2015. Since January 2015, he has been an assistant professor at the National Laboratory of Pattern Recognition (NLPR), Institute of Automation, Chinese Academy of Science (CASIA). He received the 2012 ChaLearn One-Shot-Learning Gesture Challenge Award, sponsored by Microsoft, ICPR 2012. He also received the 2013, 2014 Best Paper Award from the Institute of Information Science, Beijing Jiaotong University. His main research interests include computer vision, machine learning, especially for gesture and action recognition, facial attribution analysis (i.e., age estimation, facial expression, gender and race classification). He has published papers in top journals, such as JMLR, TPAMI, TIP, and TCYB. He has served as the reviewer on several top journals and conferences, such as JMLR, TPAMI, TIP, TMM, TSMC, PR, ICPR2016, CVPR2017, ICCV2017, FG2017.

Thomas B. Moeslund received his PhD from Aalborg University in 2003 and is currently head of the Visual Analysis of People lab at Aalborg University: www.vap.aau.dk. His research covers all aspects of software systems for automatic analysis of people. He has been involved in 14 national and international research projects, as coordinator, WP leader and researcher. He has published more than 300 peer-reviewed journal and conference papers. Awards include a Most Cited Paper Award in 2009, a Best IEEE Paper Award in 2010, a Teacher of the Year Award in 2010, and a Most Suitable for Commercial Application Award in 2012. He serves as associate editor and editorial board member for four international journals. He has co-edited two special journal issues and acted as PC member/reviewer for a number of conferences. Professor Moeslund has co-chaired the following eight international conferences/workshops/tutorials: ARTEMIS'12 (ECCV'12), AMDO'12, Looking at People'12 (CVPR12), Looking at People'11 (ICCV'11), Artemis'11 (ICCV'11), Artemis'10 (MM'10), THEMIS'08 (ICCV'09) and THEMIS'08 (BMVC'08).