Information Fusion 71 (2021) 97–98

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Information fusion for affective computing and sentiment analysis

1. Introduction

Emotions are intrinsically part of our mental activity and play a key role in communication and decision-making processes [1,2]. Emotion, cognition, and action interact in feedback loops and emotion can be viewed in a structural model tied to adaptation. Besides being important for the advancement of AI, detecting and interpreting emotional information is key in multiple areas of computer science, e.g., human-agent, -computer, and -robot interaction, but also e-learning, e-health, domotics, automotive, security, user profiling and personalization.

In recent years, emotion and sentiment analysis has become increasingly popular also for processing social media data on social networks, online communities, blogs, Wikis, microblogging platforms, and other online collaborative media [3,4]. The distillation of knowledge from such a big amount of unstructured information, however, is an extremely difficult task, as the contents of today's Web are perfectly suitable for human consumption, but remain hardly accessible to machines. The opportunity to capture the opinions of the general public about social events, political movements, company strategies, marketing campaigns, and product preferences has raised growing interest both within the scientific community, leading to many exciting open challenges, as well as in the business world, due to the remarkable benefits to be had from marketing [5] and financial market prediction [6].

Most of existing approaches to affective computing and sentiment analysis are still based on the syntactic representation of text, a method that relies mainly on word co-occurrence frequencies. Such algorithms are limited by the fact that they can only process information they can 'see'. As human text processors, we do not have such limitations as every word we see activates a cascade of semantically related concepts, relevant episodes, emotions, and sensory experiences [7], all of which enable the completion of other complex NLP tasks such as subjectivity detection [8], anaphora resolution [9], personality recognition [10], and more. Information fusion can aid to mimic the way humans process and analyze text and, hence, overcome the limitations of standard approaches to affective computing and sentiment analysis.

2. Contents of the special issue

We received 41 valid paper submissions for this special issue. After several rounds of rigorous reviews and revisions, we decided to publish only 7 of them in this special issue.

The article "Foundations of Multimodal Co-learning" by Zadeh et al. [11] empirically and theoretically studies the phenomenon of Multimodal Co-learning (MCl). First, the authors empirically observe that models trained using multimodal data perform better in unimodal downstream tasks than their unimodal trained counterparts. Furthermore, they use information-theoretic arguments to show the possibility of MCl as well as the required conditions for MCl to happen in practice. Finally, they show that MCl is an information entropy lower-bound to the unimodal learning, achieved by universal approximation capabilities of deep neural networks.

In "End-to-end multimodal affect recognition in real-world environments", Tzirakis et al. [12] propose an emotion recognition system that utilizes raw text, audio and visual information in an end-to-end manner. Authors propose a novel transformer-based architecture for the text modality that can robustly capture the semantics of sentences. They develop an audio model to process the audio channel, and adopt a variation of a high resolution network to process the visual modality. To fuse the modality-specific features, they propose novel attention-based methods.

"Quantum-inspired multimodal fusion for video sentiment analysis" by Li et al. [13] presents a novel quantum-theoretic multimodal fusion framework. The framework aims to recognize the sentiment of a multimodal sentence, which consists of word-aligned features of visual, acoustic and textual modalities. Because of the word-aligned format of data, the intra-modal interactions mainly consist of relations between different words in different modalities, while the inter-modal interactions are the interactions between visual, acoustic and textual modalities.

The work by Raheel et al. [14] entitled "DEAR-MULSEMEDIA: Dataset for emotion analysis and recognition in response to multiple sensorial media" created a new dataset for emotion analysis and recognition in response to MULtiple SEnsorial media (mulsemedia) consisting of physiological signals, e.g., electroencephalography, galvanic skin response, and photoplethysmography. Such signals were recorded while presenting a stimulus to the user and simultaneously engaging vision, auditory, olfactory, and haptic/tactile senses. Authors also analyzed the impact of mulsemedia content on viewer's emotions.

In "Deep learning based emotion analysis of microblog texts", Xu et al. [15] propose a microblog emotion classification model based on convolutional neural networks. Such a model introduces a word2vec neural network model to train distributed word embeddings on every single word. The trained word vectors are used as input features for the model to learn microblog text features through parallel convolution layers with multiple convolution kernels of different sizes. Experiment results show that the overall accuracy rate of the proposed model is 7.0% higher than that achieved by current mainstream methods, such as recurrent neural networks.

The review paper entitled "A survey on empathetic dialogue systems" by Ma et al. [16] focuses on the literature of empathetic dialogue

https://doi.org/10.1016/j.inffus.2021.02.010

Available online 9 February 2021

systems, whose goal is to enhance the perception and expression of emotional states, personal preference, and knowledge. Accordingly, authors identify three key features that underpin such systems: emotion-awareness, personality-awareness, and knowledge -accessibility. The main goal of this review is to serve as a comprehensive guide to research and development on empathetic dialogue systems and to suggest future directions in this domain.

Finally, "Conversational transfer learning for emotion recognition" by Hazarika et al. [17] proposes an approach where authors pre-train a hierarchical dialogue model on multi-turn conversations (source) and then transfer its parameters to a conversational emotion classifier (target). In addition to the popular practice of using pre-trained sentence encoders, their approach also incorporates recurrent parameters that model inter-sentential context across the whole conversation. Based on this idea, the authors perform several experiments across multiple datasets and find improvement in performance and robustness against limited training data.

Acknowledgments

We thank INFFUS Journal Managers for their valuable support during the preparation of this special issue. In addition, we are deeply grateful to the reviewers for their contributions to the decision process.

References

[1] Z. Wang, S. Ho, E. Cambria, A review of emotion sensing: categorization models and algorithms, Multimed. Tools Appl. 79 (2020) 35553–35582.
[2] Y. Susanto, A. Livingstone, B.C. Ng, E. Cambria, The hourglass model revisited, IEEE Intell. Syst. 35 (5) (2020) 96–102.
[3] D. Camacho, A. Panizo-LLedot, G. Bello-Orgaz, A. Gonzalez-Pardo, E. Cambria, The four dimensions of social network analysis: an overview of research methods, applications, and software tools, Inf. Fusion 63 (2020) 88–120.
[4] E. Cambria, H. Wang, B. White, Guest editorial: big social data analysis, Knowl. Based Syst. 69 (2014) 1–2.
[5] E. Cambria, M. Grassi, A. Hussain, C. Havasi, Sentic computing for social media marketing, Multimed. Tools Appl. 59 (2) (2012) 557–577.
[6] F. Xing, E. Cambria, R. Welsch, Intelligent asset allocation via market sentiment views, IEEE Comput. Intell. Mag. 13 (4) (2018) 25–34.
[7] E. Cambria, Y. Li, F. Xing, S. Poria, K. Kwok, SenticNet 6: ensemble application of symbolic and subsymbolic AI for sentiment analysis. CIKM, 2020, pp. 105–114.
[8] I. Chaturvedi, E. Cambria, R. Welsch, F. Herrera, Distinguishing between facts and opinions for sentiment analysis: survey and challenges, Inf. Fusion 44 (2018) 65–77.
[9] R. Sukthanker, S. Poria, E. Cambria, R. Thirunavukarasu, Anaphora and coreference resolution: a review, Inf. Fusion 59 (2020) 139–162.
[10] Y. Mehta, N. Majumder, A. Gelbukh, E. Cambria, Recent trends in deep learning based personality detection, Artif. Intell. Rev. 53 (2020) 2313–2339.
[11] A. Zadeh, P.P. Liang, L.-P. Morency, Foundations of multimodal co-learning, Inf. Fusion 64 (2020) 188–193.
[12] P. Tzirakis, J. Chen, S. Zafeiriou, B. Schuller, End-to-end multimodal affect recognition in real-world environments, Inf. Fusion 68 (2021) 46–53.
[13] Q. Li, D. Gkoumas, C. Lioma, M. Melucci, Quantum-inspired multimodal fusion for video sentiment analysis, Inf. Fusion 65 (2021) 58–71.
[14] A. Raheel, M. Majid, S.M. Anwar, Dear-mulsemedia: dataset for emotion analysis and recognition in response to multiple sensorial media, Inf. Fusion 65 (2021) 37–49.
[15] D. Xu, Z. Tian, R. Lai, X. Kong, Z. Tan, W. Shi, Deep learning based emotion analysis of microblog texts, Inf. Fusion 64 (2020) 1–11.
[16] Y. Ma, K.L. Nguyen, F. Xing, E. Cambria, A survey on empathetic dialogue systems, Inf. Fusion 64 (2020) 50–70.
[17] D. Hazarika, S. Poria, R. Zimmermann, R. Mihalcea, Conversational transfer learning for emotion recognition, Inf. Fusion 65 (2021) 1–12.

Amir Hussain[a], Erik Cambria[b,*], Soujanya Poria[c], Ahmad Hawalah[d], Francisco Herrera[e]
[a] *Edinburgh Napier University, UK*
[b] *Nanyang Technological University, Singapore*
[c] *Singapore University of Technology and Design, Singapore*
[d] *Taibah University, Saudi Arabia*
[e] *University of Granada, Spain*

[*] Corresponding author.
E-mail address: cambria@ntu.edu.sg (E. Cambria).