Guest Editorial: Information Fusion for Medical Data: Early, Late, and Deep Fusion Methods for Multimodal Data

THIS special issue addresses important current topics on multimodal data fusion in the medical context. All clinical data, including genomic and proteomic, play a role in the diagnosis and in particular in the treatment planning and follow-up. This is true for all types of data analyses whether in classification, regression, retrieval, clustering, or other. The interaction between several types of information is not always well understood. Experienced clinicians automatically and even unconsciously add multiple sources of information into their decision process, but machine learning tools often concentrate on single information sources. This special issue presents five examples where several data sources are fused. The papers give several examples of fusion techniques and also the results obtained in quite different application scenarios.

In the following we present the accepted articles:

Alzheimer's Disease (AD) is the most common type of neurodegenerative disease in the world. However, its origins are still unknown, and the underlying process of neurodegeneration is a matter of research. Clinical diagnosis of AD is based on the examination of many variables that range from clinical history, brain imaging, blood tests or many other biological data, although the main part is the most easily noticeable: cognitive decline. The appreciation of cognitive decline is a subjective task, with many confounding variables and co-morbidities, and it poses some new questions: How does structural changes affect cognition? Is it possible to predict cognitive decline based solely on structural MRI imaging? The research presented in [item 1) in the Appendix] suggest that it is. This work study a new way to characterize neurodegeneration, using the latent features derived of a denoising convolutional autoencoder. This neural network is composed of two main parts: an encoder, that transforms the thousand-voxel images to tens of features, and a decoder, that takes those features and tries to reconstruct the original image. This network is trained by minimizing the mean squared error between the reconstructed image and the input. To make the system more robust, the input images are corrupted by adding white noise on top. Some of the latent features (an encoding of the original images on the manifold space defined by the autoencoder) show high correlation with cognitive test scores such as Mini-Mental State Examination (MMSE) or the Alzheimer's Disease Assessment Scale (ADAS) with 11 questions (ADAS-11), with a Pearson correlation higher than 0.6. These features were also used in a classification environment, where they could differentiate between AD-affected subjects and controls with high accuracy (>80%). This proves that the system can be used to explore the process of neurodegeneration and cognitive decline, as well as providing new tools for an easier and more objective diagnosis.

The work presented in [item 2) in the Appendix] is on the use of activity data such as wrist bands and waist bands to predict classes of users regarding their general activity. Instead of using a single data set, the work uses five data sets and allows thus not only to compare the evaluation within a particular data set, but also to train with some sets and evaluate on unseen data sets. The results show that a testing in unseen data sets has a much lower performance than training and testing on the same type of data. When training with several resources the quality slightly improved, meaning that the generalization performance is better in this case. As a conclusion, the authors propose to always use heterogeneous resources for training to reach optimal performance on new data if variations can be expected in the data. The article clearly shows how many biases are present in articles using machine learning on medical data. Most approaches being evaluated on retrospective data and will likely not hold in the same way on prospective data, where changes in the data acquired occur regularly.

Medical image synthesis problem has important applications in scenarios where some imaging modalities may be of limited access or missing due to various reasons such as cost, radiation, or the utilization of intravenous contrast. To tackle this issue a new, CNN-based, architecture is proposed in [item 3) in the Appendix], the Residual Inception Encoder-Decoder Net (RIED-Net). The residual inception block, where one additional inception path with a 1×1 convolution is adapted for feature map resizing, improves upon the U-Net by reserving more pixel-level information and makes the architecture robust to overfitting and the gradient vanishing problem. Furthermore, it addresses the problem when the input feature maps have different channels from the output feature maps. Results are illustrated in two case studies, Breast cancer and Alzheimer's disease. In the first example, "virtual" recombined images are rendered from contrast-enhanced digital mammography (CEDM) low energy (LE) images. In the second, synthetic Positron Emission Tomography (PET) images are rendered from Magnetic Resonance Images (MRI).

With an aging population, it is vital to analyze different events that have a direct correlation with quality of life. Events related to the autonomy and independence of the population must have a

Digital Object Identifier 10.1109/JBHI.2019.2958429

relevant place in modern societies. Falling is one of those events, intimately related to mobility and independence, especially of the elder. State of the art studies in the area present two major drawbacks: the use of small datasets (less than 150 people) and less than 30% of incidence in older population. To reduce this gap, in [item 4) in the Appendix], the authors present an approach capable of predicting falls in elders (people over 65 years of age) based on a dataset containing 281 participants. To achieve that, a multifactorial screening protocol was used, and different data types have been enrolled, such as clinical and self-reported data and data from instrumental functional tests. The data were later fused in three stages: early fusion – data fusion before classification phase; late fusion – after the classification phase; and slow fusion which combines the two earliest strategies. The overall results indicate that there is no significant difference in the evaluation metrics used between the prediction in each of the three stages. However, two important points have emerged: recall becomes a more important metric than specificity for the prediction, and information from sensor instrumentation played a relevant role in the prediction task.

The use of deep learning approaches for integrating multimodal data is growing as well. In [item 5) in the Appendix] the authors leverage an autoencoder-based semi-supervised learning approach to tackle the complex issue of identifying drug-drug interactions. They do this via extracting relevant features from the food and drug administration (FDA) adverse event reports. Results show that the approach outperforms competing models and show the promise of such approaches for deriving useful information by connecting disparate information sources.

As can be seen from the above summaries, there are still some open questions and research lines to pursue in the field of Information Fusion in the medical domain.

From the data side, work has been developed either using similar type of data or with data from completely different types. In [item 2) in the Appendix], raw acceleration is used, but while some samples coming from the hip, while others originate from the wrist. Image from image synthesis if performed in [item 3) in the Appendix], albeit source and target being different types of medical images. [item 1) in the Appendix] found that data-driven decomposition of MRI is largely related with clinical variables such as age, tau protein deposits and also neuropsychological examinations. In [item 4) in the Appendix], different fusion strategies are used to combine information from personal data, inertial sensors, and pressure platform. In [item 5) in the Appendix] it is shown that features derived from adverse events contain effective information about the severity levels of drug-drug interactions.

For the future, the use of all of the available data to automatically aid the specialists in the decision process is going to be a reality in personalized medicine. To achieve this, algorithms need to be prepared not only to deal with the different types of data, but also with the existence of different data for each patient.

From the machine learning point of view, a good balance between the use of traditional techniques and deep learning methods is observed. In [item 2) in the Appendix], a shallow Neural Network with a single hidden layer with ten nodes was used, while in [item 4) in the Appendix] five classifiers were tested, including k-Nearest Neighbors, Decision Tree, Random Forest, Logistic Regression and Multilayer Perceptron. Leveraging deep learning methodologies, [item 1) in the Appendix] use latent features derived from a denoising convolutional autoencoder, while [item 3) in the Appendix] propose a Residual Inception Encoder-Decoder Net (RIED-Net). Interestingly, in [item 5) in the Appendix] deep learning is combined with traditional techniques by developing an algorithm with stacked autoencoders that identify reliable samples, followed by a weighted support vector machine.

We foresee the continuation of the use of deep learning techniques, maybe in combination with traditional methods and even hand-engineered features. Other issues such as interpretability of the models become even more significant when different sources of data are present. It is necessary to clearly explain to the specialist why the algorithm gave a certain result and which type of information was more relevant. The need to combine information, also translates into the need of information systems that gather all the data, raising data protection issues.

INÊS DOMINGUES, *Guest Editor*
Medical Physics, Radiobiology, and
Radiation Protection Group
IPO Porto Research Centre (CI-IPOP)
4200-072 Porto, Portugal
ines.domingues@isec.pt

HENNING MÜLLER, *Guest Editor*
HES-SO
3960 Sierre, Switzerland
henning.mueller@hevs.ch

ANDRES ORTIZ, *Guest Editor*
Universidad de Málaga
29016 Málaga, Spain
aortiz@ic.uma.es

BELUR V. DASARATHY, *Guest Editor*
Consultant–Decision Systems &
Information Fusion Technologies
fusion-consultant@ieee.org

PEDRO H. ABREU, *Guest Editor*
CISUC
Department of Informatics Engineering
University of Coimbra
3030-790 Coimbra, Portugal
pha@dei.uc.pt

VINCE D. CALHOUN, *Guest Editor*
Tri-Institutional Center for Translational
Research in Neuroimaging and Data
Science (TReNDS)
Georgia State University, Georgia Institute
of Technology, Emory University
Atlanta, GA 30303 USA
vcalhoun@gsu.edu

APPENDIX
RELATED WORK

1) F. J. Martinez-Murcia, A. Ortiz, J.-M. Gorriz, J. Ramirez, and D. Castillo-Barnes, "Studying the manifold structure of Alzheimer's disease: A deep learning approach using convolutional autoencoders," *IEEE J. Biomed. Health Informat.*, to be published, doi: 10.1109/JBHI.2019.2914970.
2) V. Farrahi, M. Niemela, P. Tjurin, M. Kangas, R. Korpelainen, and T. Jamsa, "Evaluating and enhancing the generalization performance of machine learning models for physical activity intensity prediction from raw acceleration data," *IEEE J. Biomed. Health Informat.*, to be published, doi: 10.1109/JBHI.2019.2917565.
3) F. Gao, T. Wu, X. Chu, H. Yoon, Y. Xu, and B. Patel, "Deep residual inception encoder-decoder network for medical imaging synthesis," *IEEE J. Biomed. Health Informat.*, to be published, doi: 10.1109/JBHI.2019.2912659.
4) J. R. Silva, I. Sousa, and J. S. Cardoso, "Fusion of clinical, self-reported, and multisensor data for predicting falls," *IEEE J. Biomed. Health Informat.*, to be published, doi: 10.1109/JBHI.2019.2951230.
5) N. Liu, C.-B. Chen, and S. Kumara, "Semi-supervised learning algorithm for identifying high-priority drug-drug interactions through adverse event reports," *IEEE J. Biomed. Health Informat.*, to be published, doi: 10.1109/JBHI.2019.2932740.