AI-POWERED DIAGNOSTICS

Self-attention Feature Fusion: A Platform For Multimodal Biomedical AI

Andrew Matas Battelle Memorial Institute **Collin Dunlap** Battelle Memorial Institute **Michael Riedl** Battelle Memorial Institute

Modern wearable sensors enable unprecedented access to a wide range of low-latency health information for the modern soldier. Full use of this information may enable rapid response to unknown threats and diagnosis of diseases before symptoms arise. Therefore, the time is ripe to develop a trusted and automated system capable of integrating multiple sensor inputs and providing accurate and reliable, real-time threat information to the warfighter so that they can adapt before symptoms arise. Such a system faces several technical challenges innate to any form of sensor fusion, such as dealing with sensor dropout, and combining different modalities such as image, text, spectral, and time series data. Additionally, many traditional approaches rely on methods such as extracting hand-crafted features from the data, which are time-consuming to develop and may be sub-optimal, or treating sensors as independent, which can fail to capture correlations between different sensors. In this talk, we will present Self-Attention Feature Fusion (SAFF), a novel platform based on the transformer architecture in the field of deep learning. SAFF can learn features for a variety of different sensor modalities in an unsupervised way, leveraging large amounts of pre-existing unlabeled data and combining them in a common latent space to perform sensor fusion. Importantly, SAFF is agnostic to sensor type and therefore can combine real-time information from wearable or environmental sensors including diagnostic assays, as well as contextual information such as intelligence data in the form of images or text. The predictions from this AI framework can used to integrate data from multiple devices and provide diagnosis and actionable instructions in response to unknown threats.

As a test case to demonstrate the SAFF platform we use a publicly available dataset from Tamneh et al, 2017, which includes multi-modal signals from 68 subjects during a driving task, to predict real-time cognitive states. We demonstrate that SAFF is robust to sensor dropout by showing that we obtain parity in performance when comparing with models from Mou et al, 2023, which were trained on a subset of the available sensors. Additionally, we prove the multimodal capability of SAFF by incorporating tabular and image data and showing that including this information boosts our performance. We will show how this approach can be applied to chemical and biological defense by outlining a case study where we use the model in an unsupervised manner as an anomaly detector to identify patients deviating from healthy baseline behavior.