Medical Image Report Generation and Beyond
Zhiting Hu, Carnegie Mellon University

A tsunami of healthcare data has emerged in recent year. It is challenging yet highly valuable to leverage the power of machine learning to assist healthcare, e.g., by helping physicians and hospitals improve quality and efficiency. This talk presents our recent work on medical image report generation. Different from traditional image-to-text that generates individual labels or short single sentences, medical image reports are required to describe all radiology findings, include rich visual and textual context, and be precise in describing abnormality. We develop a set of useful techniques for the challenges, including multi-task learning, hierarchical text generation, retrieval-generation combination, visual-semantic co-attention, and cross-modality transformation, which yield usable results in practice. These techniques are also generally applicable to other contexts, such as generating paragraph descriptions of natural images, and controlled text generation in general. Our open-source general-purpose text generation toolkit, Texar, supports fast and standardized development of the above rich applications.