



Multimodal Generative AI and Machine Learning

Raghuveer M. Rao
DEVCOM Army Research Laboratory, Adelphi MD 20783, USA
e-mail: raghuveer.m.rao.civ@army.mil

Since data-driven machine learning (ML) models, such as deep-learning neural networks, require domain-specific training data of good quality, and in good quantity, synthetic data generation has been explored, especially in application domains where such data are hard to come by in sufficient quantity. Image and scene synthesis methods such as Generative Adversarial Networks (GANs) and diffusion models have shown exceptional capability in generating realistic data. However, a key issue is whether such synthesized data contribute to improved performance of the ML models when pressed into application. Other questions that arise would be related to the importance of photorealism – or perceptual realism in non-imaging modalities to ML training, methods and tools available for data synthesis, measures for assessing closeness of synthesized data to real data, and the relationship between the synthesis methods and the ML models themselves. The DEVCOM Army Research Laboratory has undertaken multiple projects addressing these issues. Over a diverse range of imaging domains, it has been found that using synthetic data results in improved performance of the ML models in data classification tasks. Examples include drone-based detection of humans and vehicles, infrared image synthesis and classification, and human activity recognition. Commercial off the shelf scene synthesis techniques in combination with domain and style transfer methods have been effective in several instances. The investigations have also touched upon problems such as network dissection, interpretability and foundation models. The generative models have been particularly useful in identifying neural network behavior affecting specific attributes of the data. The talk will provide details of these projects along with key results and thoughts for the future.