Abstract

Generative adversarial networks (GAN) have become prominent in the field of machine learning. Their premise is based on a minimax game in which a generator and discriminator “compete” against each other until an optimal point is reached. The goal of the generator is to produce synthetic samples that match that of real data. The discriminator tries to classify the real data as real and the generated data as not real. Together, the generator improves to the point where the fake data and real data are identical to the discriminator. GAN has been successfully applied in the image processing field over a large range of GAN variant architectures. Although not as prominent, the audio enhancement and synthesis field has also benefitted from GAN in a variety of different forms. In this survey paper, different techniques involving GAN will be explored relative to speech synthesis, speech enhancement, music generation, and general audio synthesis. Strengths and weaknesses of GAN will be looked at including variants created to combat those weaknesses. Also, a few similar machine learning architectures will be explored that may help achieve promising results.
References


Index Terms

Computer Science

Networks

Keywords

Audio, synthesis, generative adversarial networks, survey, enhancement.