Uncovering of Aged Sanskrit/Devanagari Documents Utilizing Generative Adversarial Networks and Tomography to Multidimensionally Reconstruct Missing Elements

Krishnamurthy, Srivatsa (School: Rock Ridge High School)

Sanskrit/Devanagari, what may be known as a dead language in modern times, at one point was a rich and advanced language that was spoken by many people who lived as far back as 1700 BCE. Ancient Devanagari scriptures are a manifestation of some of the most raw and original forms of language that were created with the intention for not only daily use, but also being functional and serving as works of art. Although many of these scriptures were written precisely for the future, they have not stood the test of time. While preservation attempts are active on many aged Devanagari scriptures, current technology does not put forth the effort of trying to reconstruct these documents to retain information. This research puts forth a methodology for reconstructing and inpainting 2D and 3D Himalayan Birch (Betula utilis) Devanagari documents/inscriptions. 2D (scanned documents) reconstruction and inpainting occur by implementing the Deep Generative Prior Model (Pan et al., 2020), which exploits a Generative Adversarial Networks (Goodfellow et al., 2014) image prior to inpaint and reconstruct missing portions, thus GAN-DGP. By specifically implementing GAN-DGP for versatile document restoration, it shows that the GAN-DGP model can be applied to the field of Digital Archaeo-Linguistics. For 3D (point clouds) reconstruction and image restoration, a novel convolutional operator, pointwise sparse convolution operator (“sparse3D”), was developed to work on sparse 3D point cloud data. Further, both U-Net (Brox et al., 2015) and PU-Net (Yu-et al., 2018) architectures were applied to focus on visual and geometric inpainting of the point cloud. Since Sanskrit/Devanagari is a language where minute details matter, both of these architectures will assist in inpainting with precision.

Awards Won:
Third Award of $1,000