Background Research

In previous research done by Goodfellow, he proposed a framework for creating things that imitate some origin material. It utilizes 2 different networks, a discriminator and generator. The discriminator attempts to distinguish between real data and data generated by the generator. The generator attempts to generate data that will fool the discriminator. The discriminator will then train to better distinguish between the generated and real data. This creates an arms race, where each network attempts to get better at recognizing or creating data. Eventually, the generator is able to create data that imitates real data (Goodfellow et al., 2014).

In previous research done by Gadelha, he proposed a model that utilized GANs to learn the 3D structure of objects from multiple 2D images. His model trains the discriminator to distinguish between real photos of the object and the photos of the model generated by the generator network. The generator generates a model of an object from random vectors and outputs a 3D model. In order to obtain an image, a projection module is used to create a silhouette of the 3D model at a random position. This is then fed back into discriminator, and the networks train based off the result of the discriminator. The network is able to generate models to a reasonable degree of accuracy. However, the model has a few key limitations. It isn’t able to account for any structure that doesn’t change the silhouette. For example, the network fails to find concave interiors since they don’t change the silhouette of the object. The network also only generates models for objects (Gadhela, Maji, & Wang, 2016).

In previous research done by Wu, he also uses GANs to produce 3D models. However, he uses a different generator. He utilizes a convolutional neural network to generate the 3D models from latent vectors. He also utilizes an image encoder which takes an image as input, then outputs a latent vector. Using this architecture, he is able to recreate objects reasonably accurately from 2D images. However, it still has some limitations. The network is only designed to handle singular objects, not large rooms or landscapes (Wu, Zhang, Xue, Freeman, & Tenenbaum, 2016).

In previous research done also done by Wu, he uses a neural network to estimate heatmaps of 2D images, which are then used to estimate a 3D structure. In order to get training data, he used a projection layer, which takes a 3D model and produces a 2D image, which can be input into another network which will label key points in the picture using the heat map. With these key points, another network attempts to recreate the 3D structure of the objects. However, this approach does have limits. The network only produces a skeleton of the model (Wu et al., 2016).

In previous research done by Wu, he uses a 2.5D image, which contain other data like depth information, to estimate 3D objects. His network attempts to estimate a 2.5D image from the input image. From this image, he reconstructs a 3D shape using another network. In order to keep consistency between the model and 2.5D image, he enforces constraints on the network that are based upon the 2.5D image (Wu et al., 2017).

This research attempts to use these existing techniques on a different set of data. Instead of using it on objects, it will attempt to instead estimate the structure of land from 2D images.

This is important because it has applications in space. On other planets, it could be used to create an accurate map of the landscape. With this map, better navigation could be employed on rovers, making it easier to navigate complex terrain with less risk of damage.

References

Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., … Bengio, Y. (2014). General Adversarial Nets. Neural Information Processing Systems Conference, 27, Retrieved from https://papers.nips.cc/paper/5423-generative-adversarial-nets

Gadhela, M., Maji, S., & Wang, R. (2016). 3D Shape Induction from 2D Views of Multiple Objects. Retrieved from https://arxiv.org/abs/1612.05872

Wu, J., Zhang, C., Xue, T., Freeman, W. T., & Tenenbaum, J. T. (2016). Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling. Neural Information Processing Systems Conference, 29. Retrieved from https://papers.nips.cc/paper/6096-learning-a-probabilistic-latent-space-of-object-shapes-via-3d-generative-adversarial-modeling

Wu, J., Zhang, C., Xue, T., Freeman, W. T., Sun, X., & Tenenbaum, J. T. (2017). MarrNet: 3D Shape Reconstruction via 2.5D Sketches. Neural Information Processing Systems Conference, 30. Retrieved from http://papers.nips.cc/paper/6657-marrnet-3d-shape-reconstruction-via-25d-sketches

Wu, J., Xue, T., Freeman, W. T., Lim, J. J., Tian, T., Torralba, A., & Tenenbaum, J. T. (2016). Single Image 3D Interpreter Network. European Conference on Computer Vision, 14, 365-382. https://doi.org/10.1007/978-3-319-46466-4\_22.