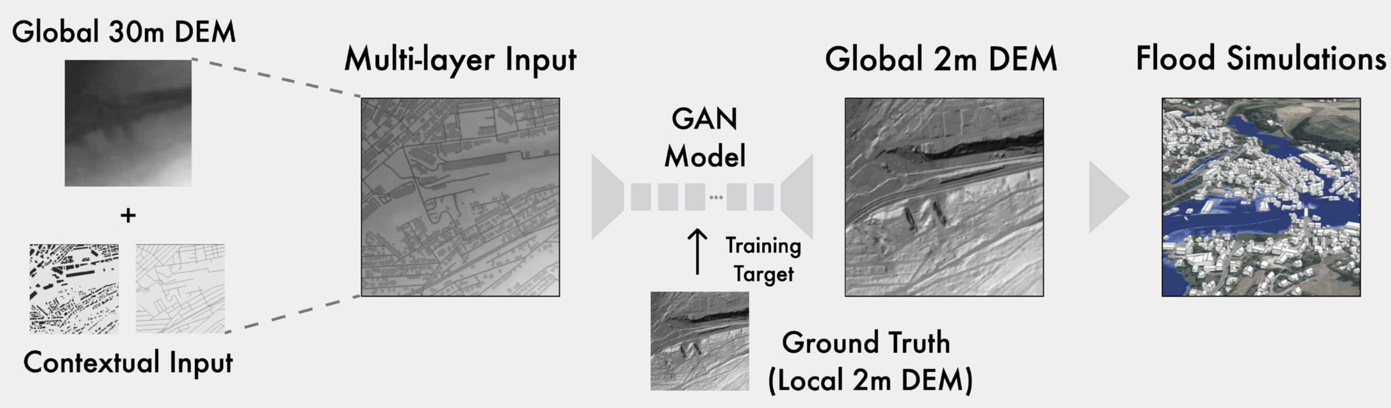
DEMGAN: Enhancing Digital Elevation Models with Deep Generative Models

The current landscape of open-access DEM data in the world trades of quality with quantity. On one hand, free, open-access Global DEMs are only available in 30 m resolution. On the other hand, high-resolution, sub-2m DEMs are either locked behind a hefty paywall or are only available in selected regions of the world.

Method

Our method generates high-resolution DEM images by leveraging the quality-quantity trade-off with a generative model based on Generative Adversarial Networks (GAN). Instead of using a classical super-resolution such as Bicubic upsampling, our method uses a GAN to translate low-resolution 30m Radar DEMs into high-resolution 2m DEMs without onsite LiDAR survey. This is achieved by training the deep learning model using a pair of low and high resolution DEM data with additional contextual input that demarcates the most up-to-date human infrastructure of a location.

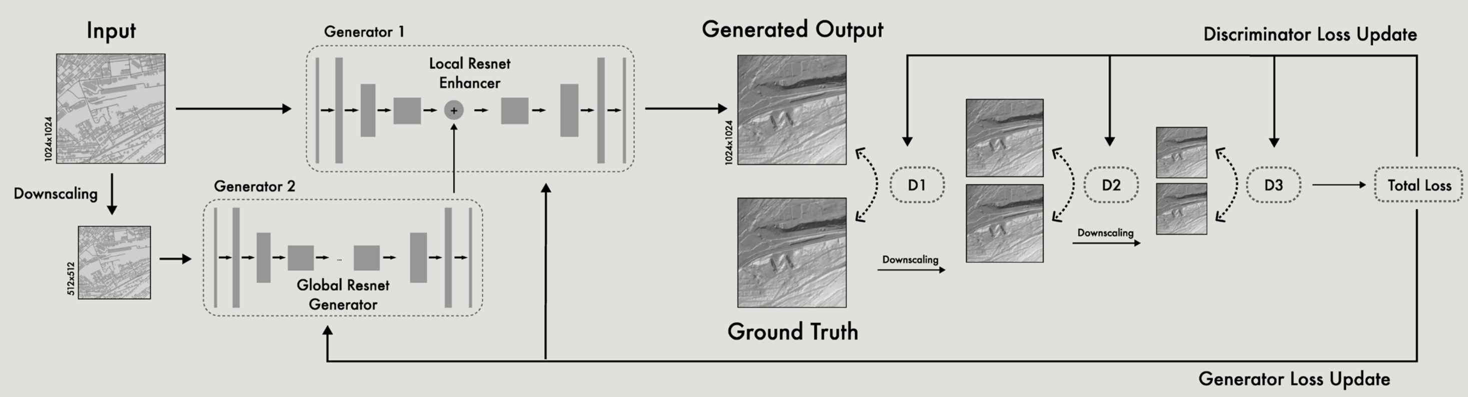
Passing the input image into the GAN network with a set of ground truth images, the model will be able to learn to generate synthetic high-resolution DEM that can be used in-lieu of LiDAR derived 2m DEMs. Trained in regions rich with high-resolution data, the model can then be used to synthesis 2m DEMs in regions that are lacking in high-resolution LiDAR data.



Model Details

The input to the model is created by concatenating an OSM vector layer (Building polygons and street networks) with the 30m Global DEM layer. The result is a multichannel image that is fed into the GAN network based a Conditional Image-to-Image GAN that translates input image data into a target image where the input image acts as conditions to guide the generation of the final output.

In this case, the lower resolution global 30m DEM is translated into high-resolution 2m DEM with vector building and street data as conditions. The model is created using two generator networks based on the Resnet architecture with three discriminator PatchGANs that updates the loss of the entire network in every epoch until the generator's output would pass the scrutiny of the discriminator as real images 50% of the time, converging to the best possible performance with no one winning against the other.

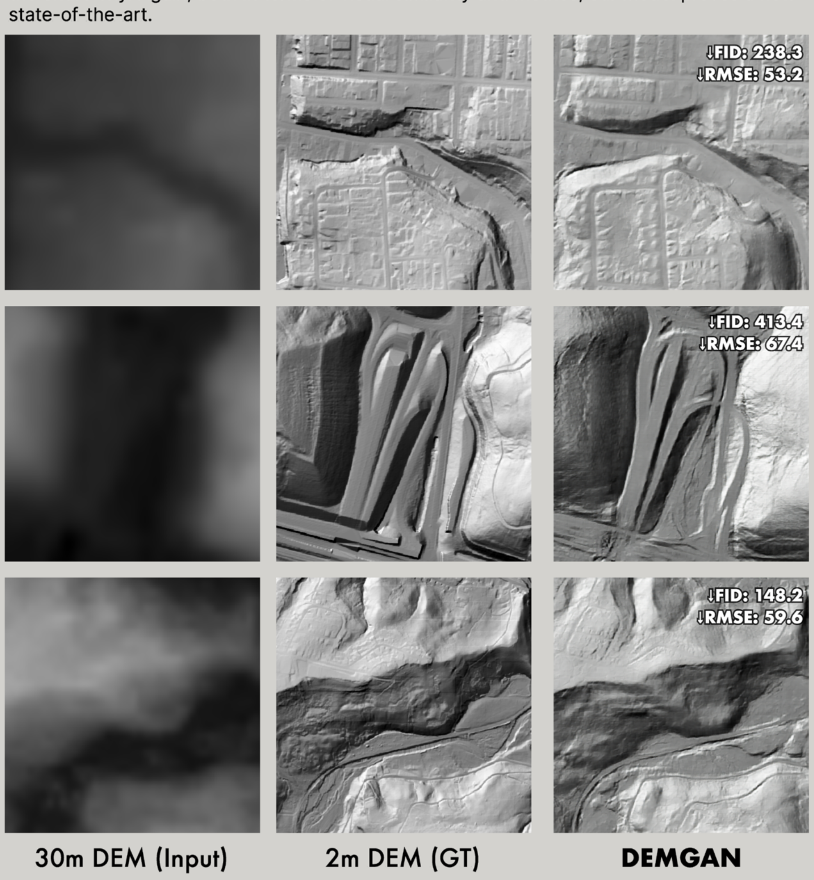


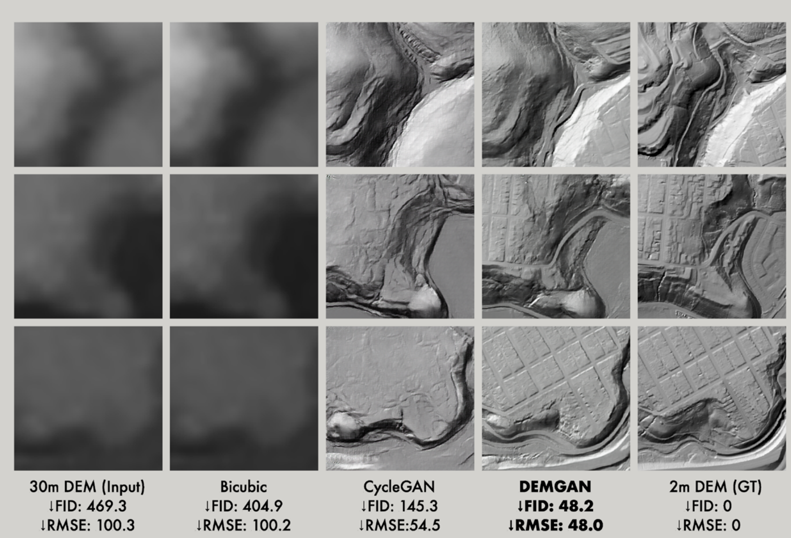
Results

The following results uses the same methods described above but are captured under different experiment setups to test the feasibility and scalability of the model. The model is trained with 50,000 pairs of data. For all locations and experiments, the 30m DEM model is obtained from CopernicusDEM published by the European Space Agency. Location-specific high-resolution DEMs are extracted from different sources such as the U.S. Geological Survey (USGS), Vancouver City Government and SwissTopo.

1. A New State-of-the-Art

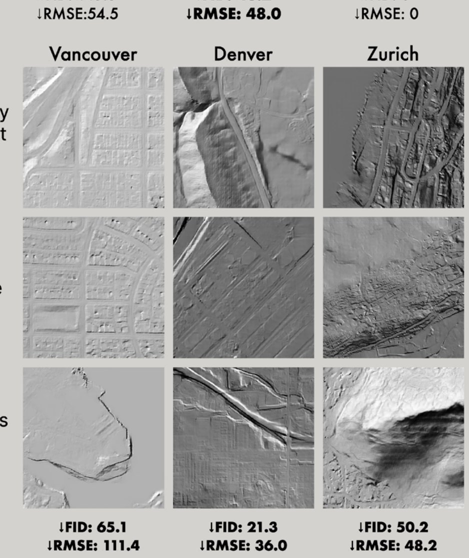
Two metrics, Frechet Inception Distance (FID) and Root Mean Square Error (RMSE) are used to measure the accuracy of the results. The smaller the number, the truer the generated results are to the ground truth. We randomly selected three samples in the dataset of the city of Pittsburg for the image below. Visually, the generated results from the 30m DEM input looks almost identical to the ground truth. In the image to the right, we compare the performance of our model with two other types of techniques, bicubic interpolation (Bicubic) and CycleGAN for the entire city. Again, both in terms of visual similarity and metrics, our model proves a new state-of-the-art.





We also investigated the adaptability of the proposed method on different types of urban morphologies and elevation. 3 different cities are chosen for a comparative study besides Pittsburg. We see that in most cases, our proposed method is able to score as well, if not better than the Pittsburgh model which we tested extensively in Experiment 1.

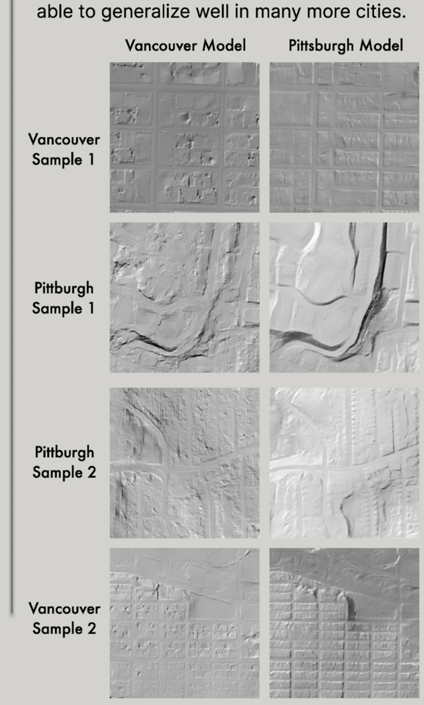
We conclude that the model is able to learn from the original 2m DEM data and generate additional information from the input as long as the training dataset contains sufficient number of images regardless of contextual difference.



3. Cross City Translation

Lastly, the ability of the model to generalise it's learned params into other locations with different terrain patterns is crucial in the wide-spread application of our model. We investigated this using Pittsburgh and Vancouver as examples.

From the two figures, we see that both the Vancouver and Pittsburgh model are able to generate terrain details at 2 meters resolution. It can be hypothesized that a much more powerful model trained on hundreds of thousands of images would be able to generalize well in many more cities.



DEMGAN showed that the generated terrains are much more accurate than using traditional super-resolution algorithms because additional information is generated during the process. With much lower FID and RMSE compared to traditional super-resolution techniques, the results from DEMGAN could be used as substitutions where LiDAR-derived DEMs are unavailable.