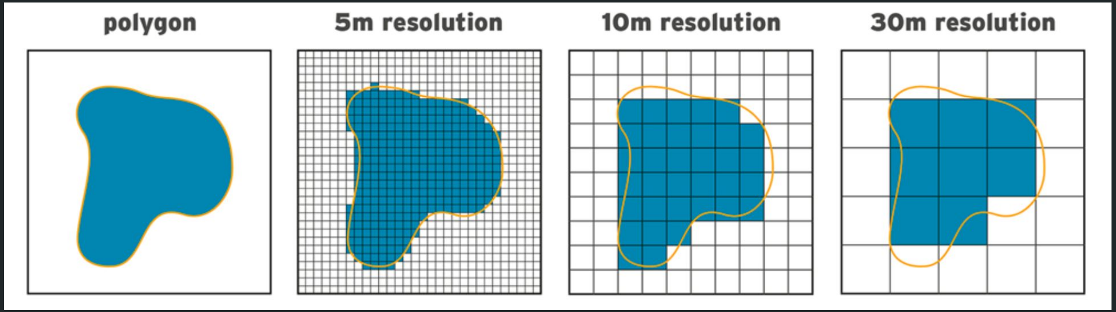
Producing Super-Resolution Digital Elevation Models

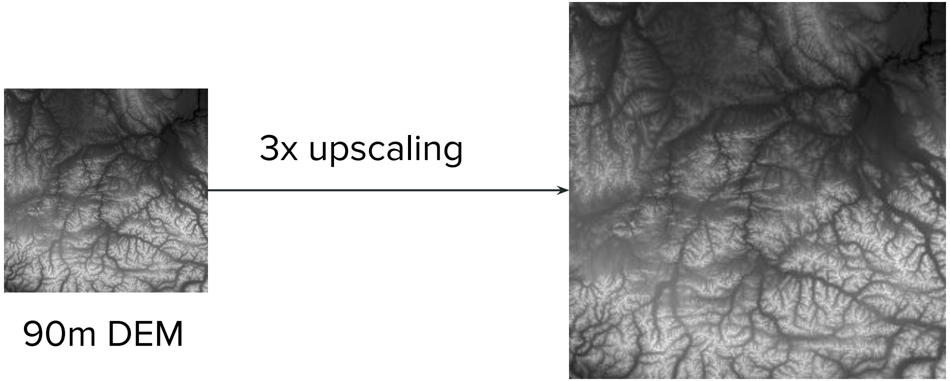
Team: sCUDA_Divers
PS ID: SS600



Problem

Resolution of a Digital Elevation Model (DEM) is defined as the size of a square grid whose average elevation value is depicted by a single pixel in the DEM.

The Task is to Increase the Resolution of a DEM using machine learning techniques.

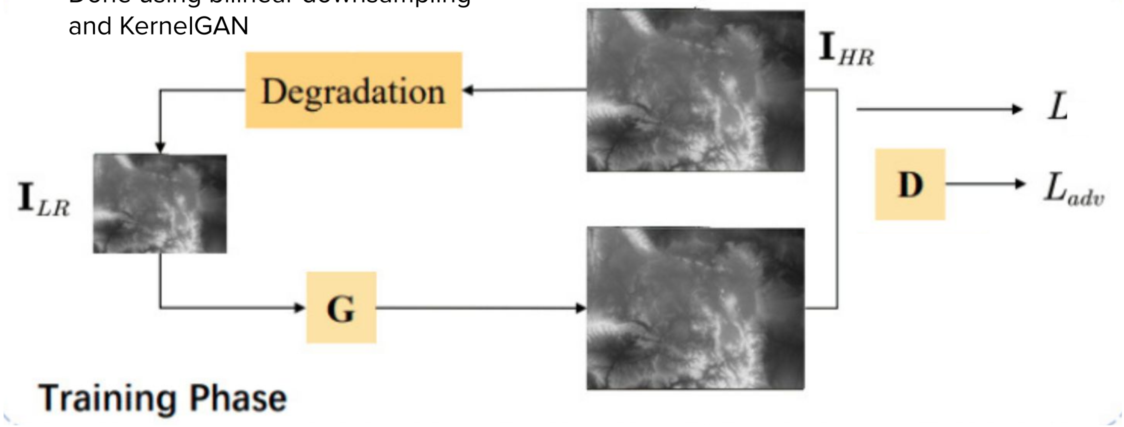


90m DEM

30m DEM

What’s our ML technique?: Overview

Done using bilinear downsampling
and KernelGAN



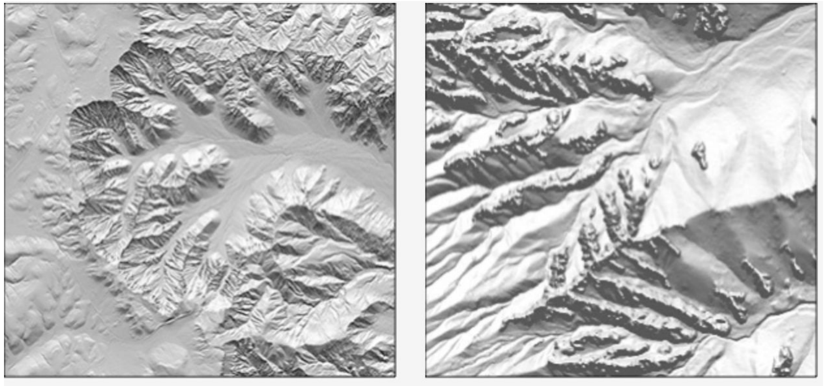
We use 2 deep convolutional neural networks i.e. generator G and discriminator D.

Generator G takes in a degraded DEM image i.e. a low-resolution image and upsamples it to a super-resolution DEM of the same resolution as original DEM.

The training of G is guided by a discriminator network that criticizes how good the outputs of generator are and also by a reconstruction loss.

What’s our ML technique?: Data

Every Machine Learning technique needs lots of data.



Samples of SRTM dataset

We initially used 2 datasets to pretrain our ML model and then fine tuned on CartoSAT data:

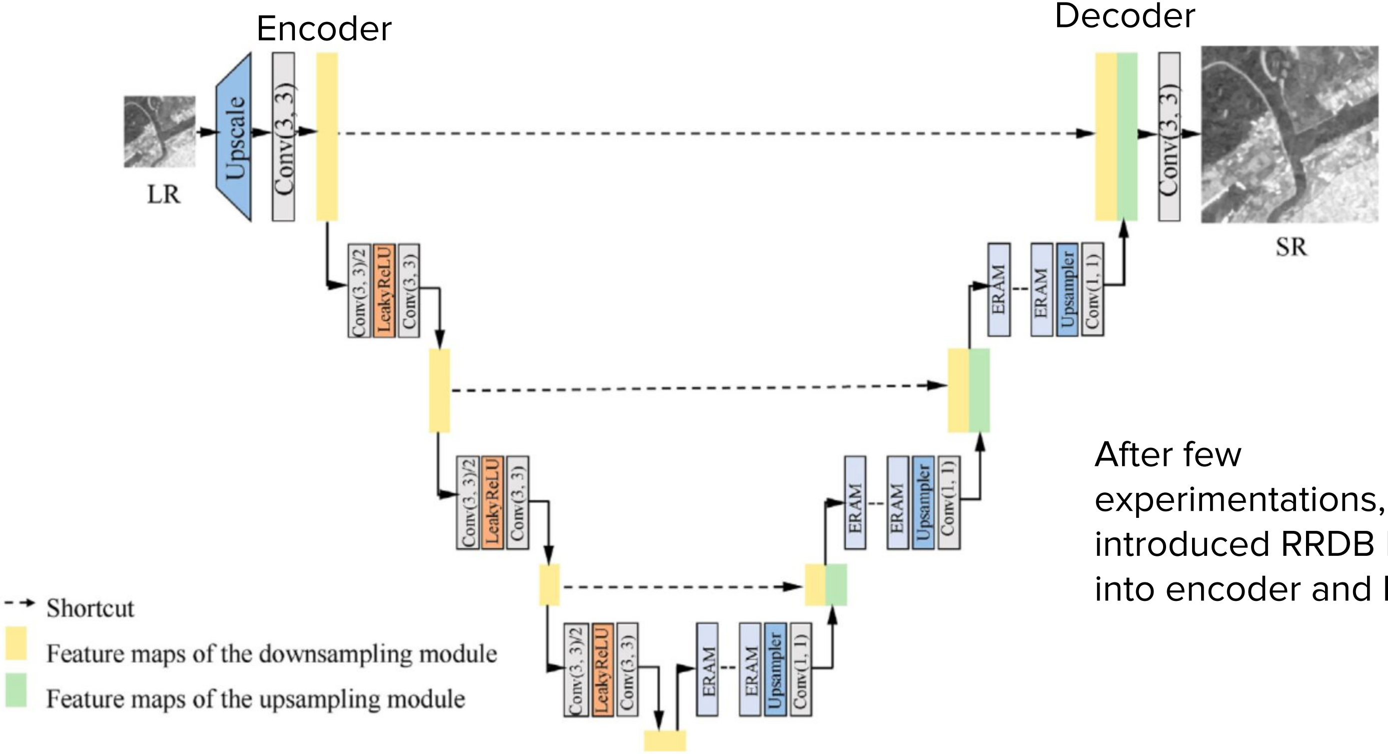
1. SRTM DEM dataset:
   90m Global DEM

2. ASTER dataset:
   104 30m DEMs collected from NASA’s portal for regions of North India.

3. CartoSAT DEM:
   30m DEM

4. CartoDEM:
   10m single scene

Generator Architecture: U-Net



After few experimentations, we've introduced RRDB Blocks into encoder and Decoder

Generator Architecture: Modules

Modules:

Enhanced Residual Attention Module (ERAM)

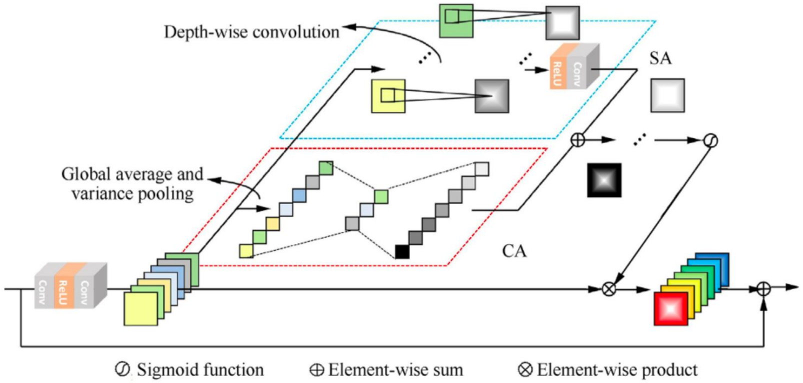
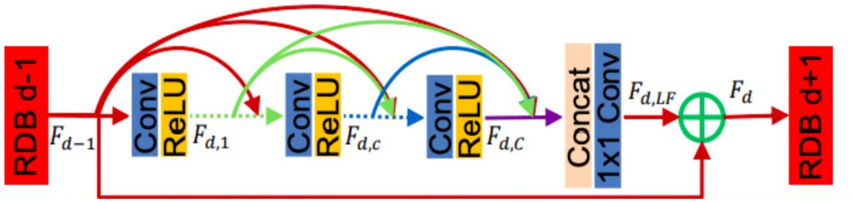


Fig. 3 Structure of ERAM.

Attention helps model make more discriminative features by attending to more important regions.

Residual in Residual Dense Blocks (RRDB)



Residual or skip-connections across different convolutional layers in a network ensures that information is not lost during the forward propagation.

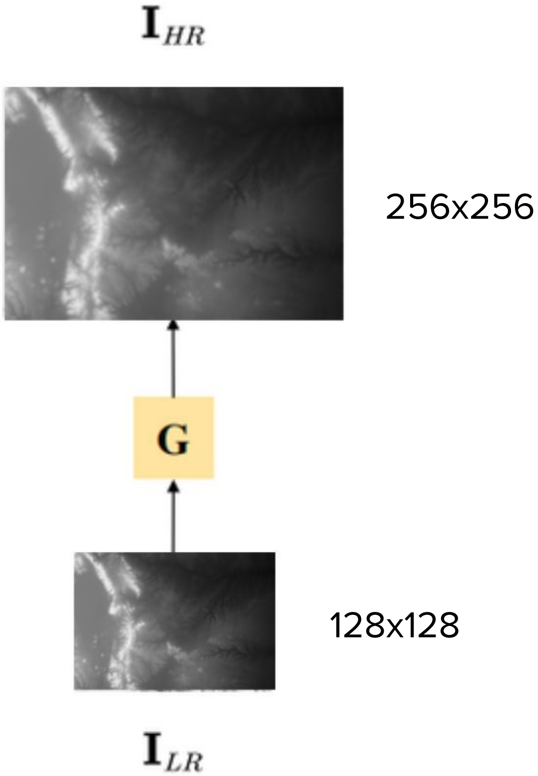
GAN Network: Generator Architecture

Details:

Training:
The model is trained to take in a 128x128 input DEM image to produce a 2 times upscaled i.e. 256x256 super resolution DEM.
(Takes 14 hours to train)

Prediction:
For any given DEM of any size, we divide the DEM into 128x128 patches and perform inference one each patch and stitch them back together to form a 2x upscaled DEM.

To upscale an image to 3x, we upscale the DEM 2 times using our 2x upscaler network to give a 4x DEM which is then downsampled 1.5 times to give a final 3x DEM.



Loss Functions

1. **L1 Loss** - Mean Absolute Error between Super Resolved DEM and Ground Truth DEM
2. **SSIM** - Structural Similarity Index is a perceptual metric to measure similarity of two images.
3. **Edge Loss** - Sobel Filter detects all the edge in the DEM. We apply Sobel Filter on the generated SRDEM and on HRDEM. MAE is computed.

Loss Functions

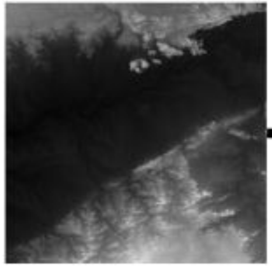
1. **L1 Loss** - Mean Absolute Error between Super Resolved DEM and Ground Truth DEM

2. **SSIM** - Structural Similarity Index is a perceptual metric to measure similarity of two images.

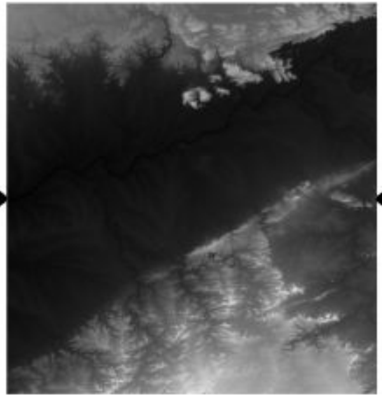
3. **Edge Loss** - Sobel Filter detects all the edge in the DEM. We apply Sobel Filter on the generated SRDEM and on HRDEM. MAE is computed.

We used a combination of all the above 3 losses.

Our Results



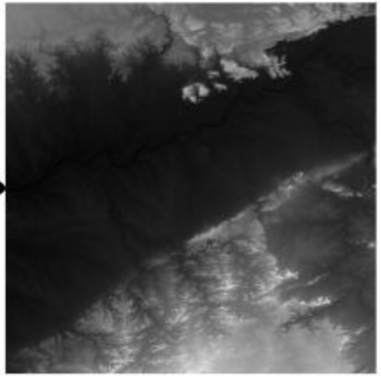
SRTM
90m



Generated by network
30m

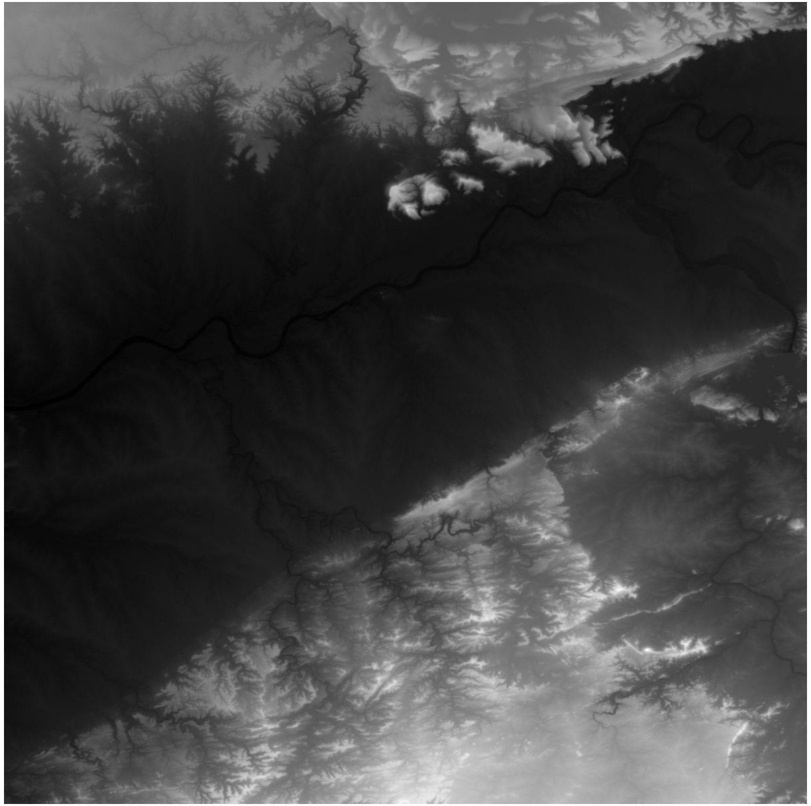
Comparison with GT

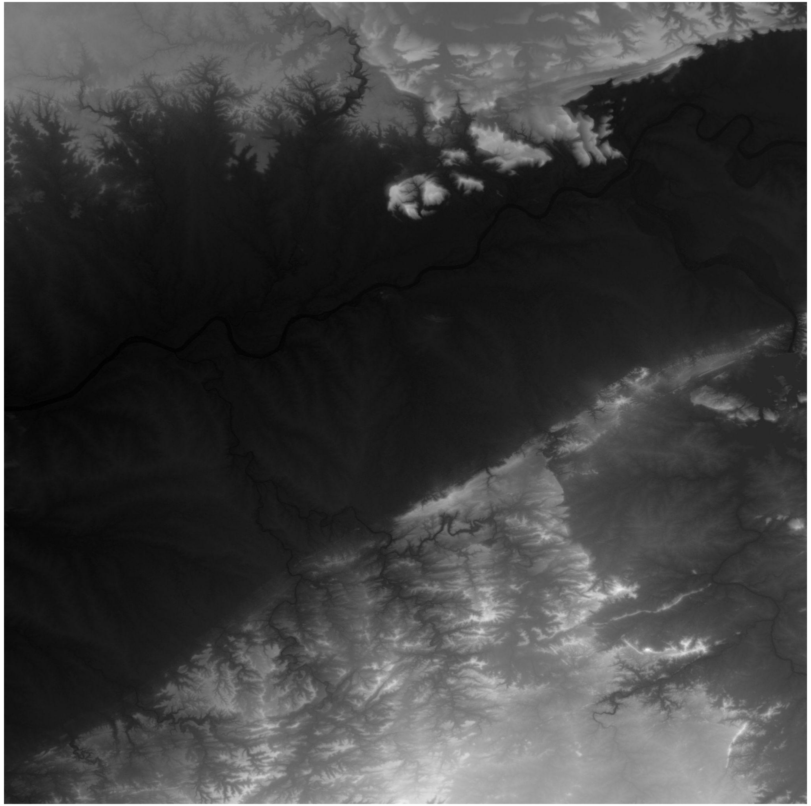
MAE: 394.04589
MSE: 165977.05
SSIM: 0.853477
PSNR: 28.22330



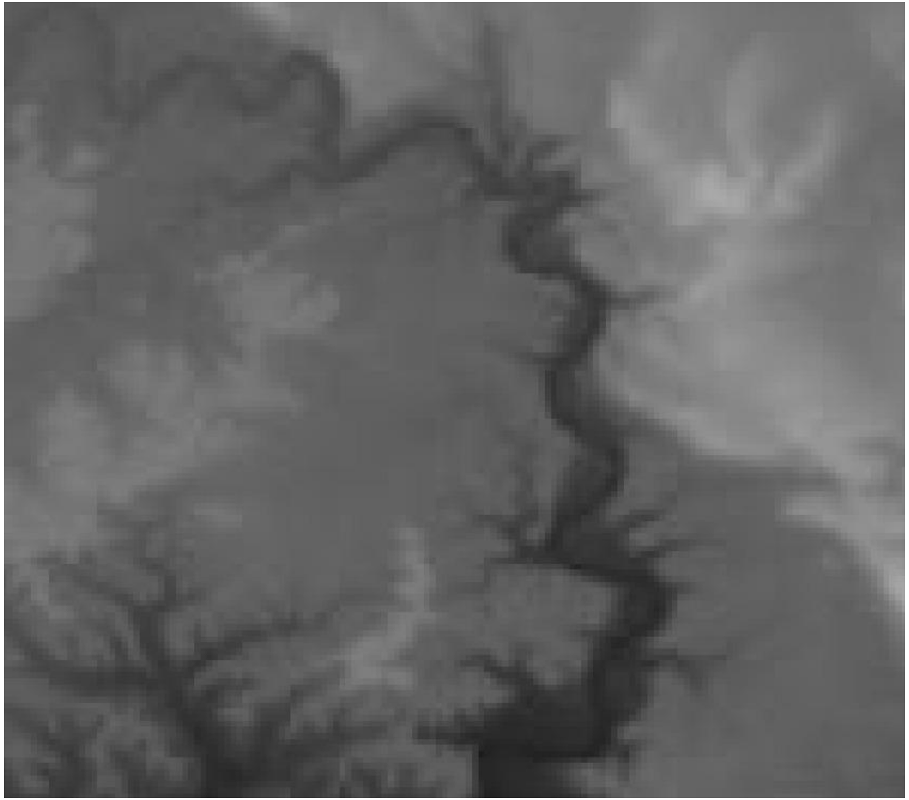
CartoSAT
30m

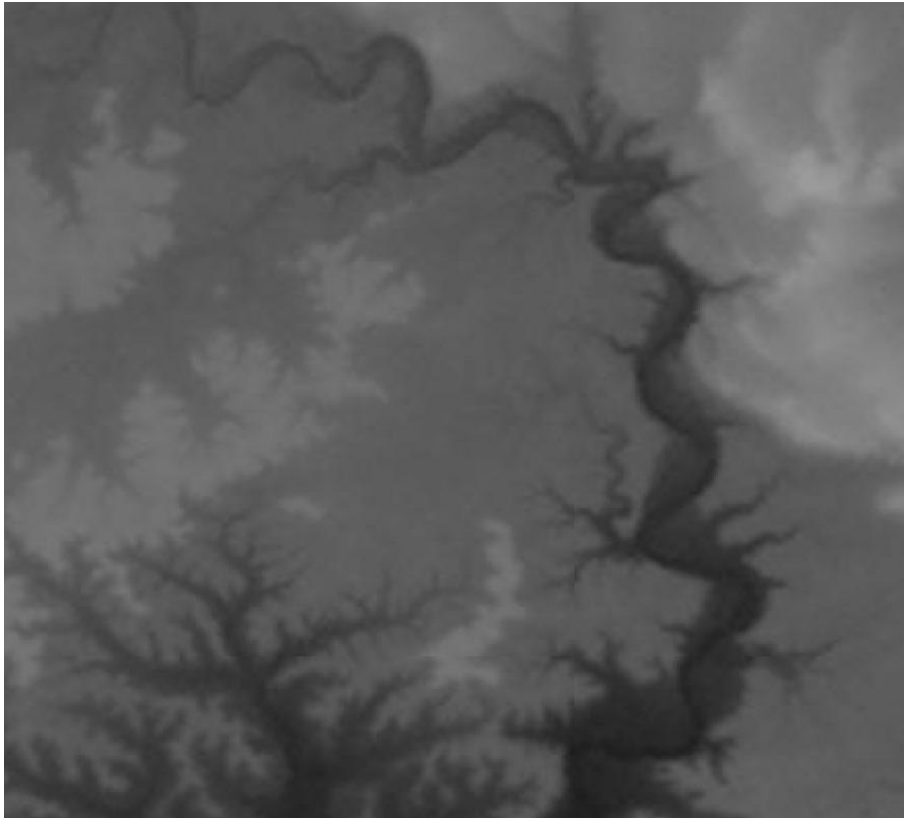
Comparison: Complete Image (90m vs 30m Model)



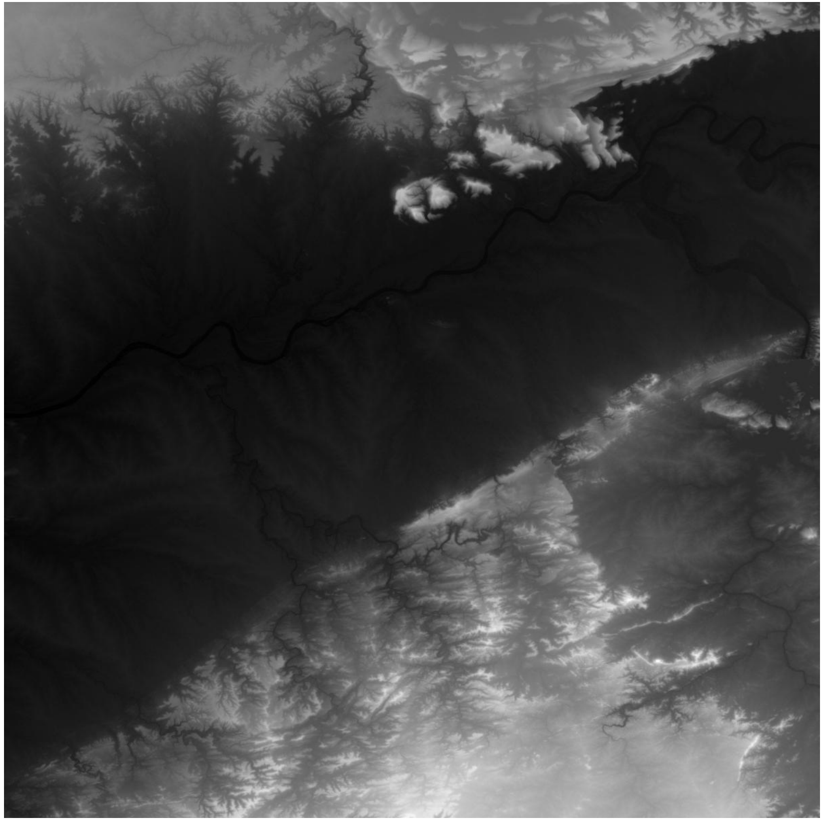


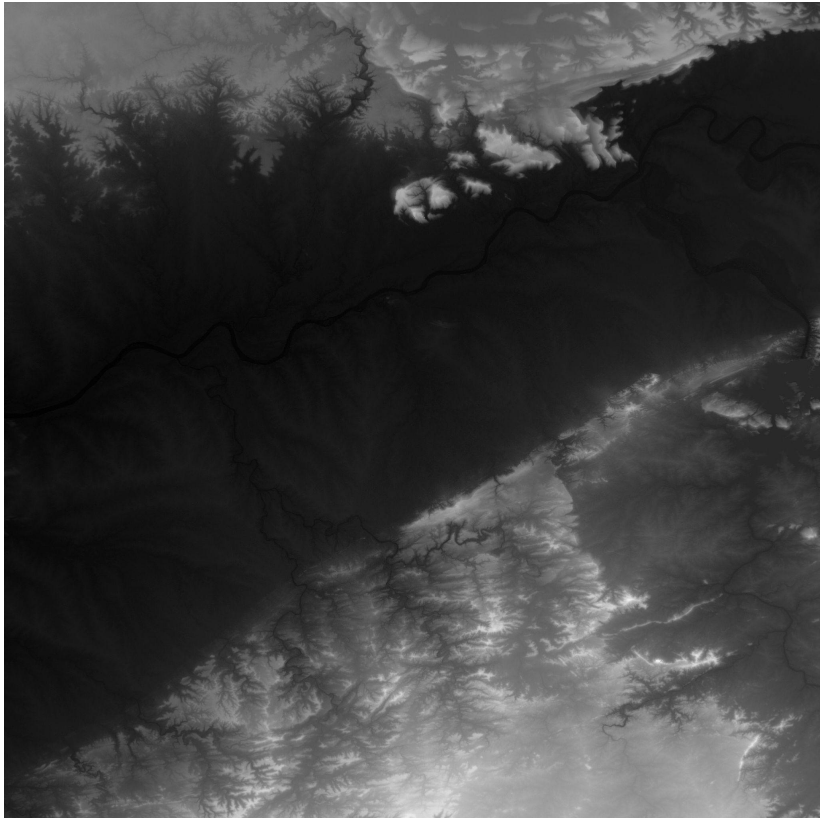
Comparison: Zoomed in (90m vs 30m Model)





Comparison: Complete (90m vs 90m->30m Bicubic)

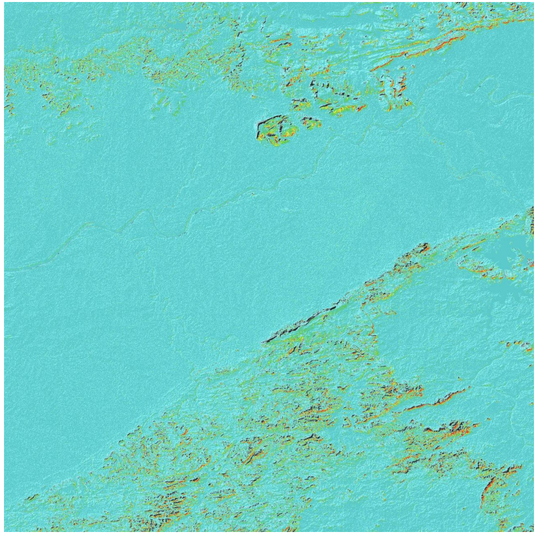




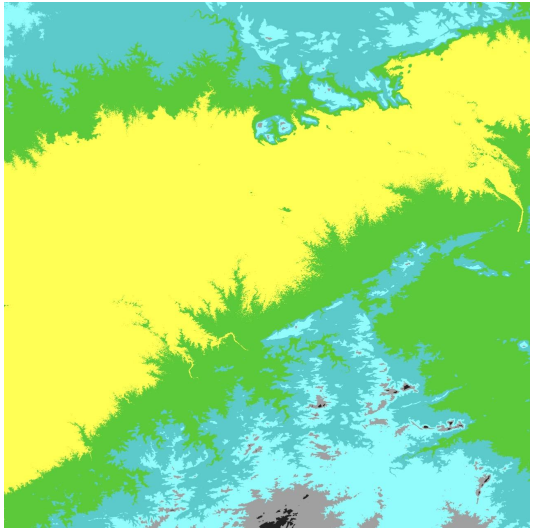
Comparison: Cropped (30m Bicubic vs 90m->30m Model)

Difference Maps

The ideal difference map would be light blue



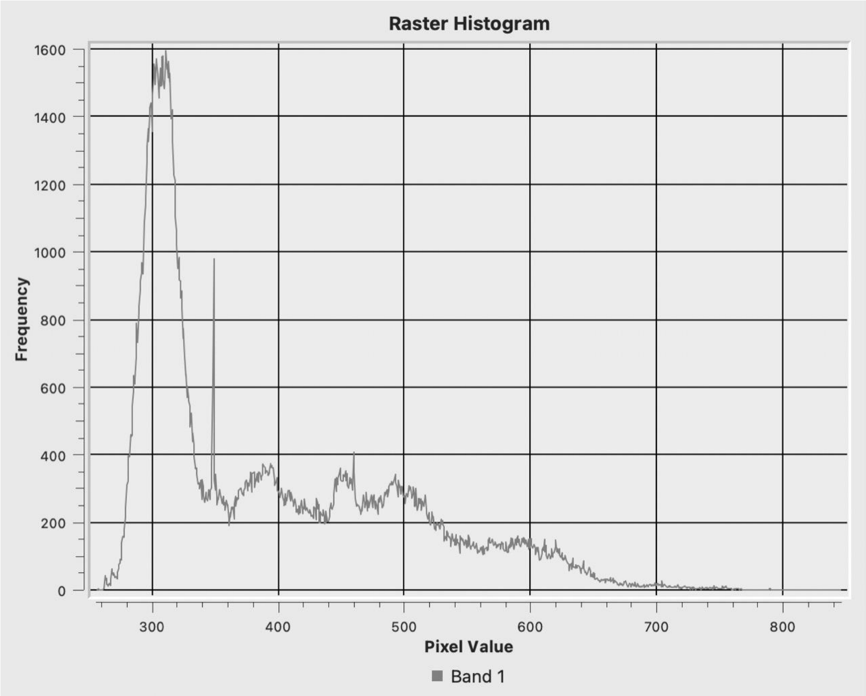
>15 - red
10-15 - yellow
5-10 - green
0-5 - blue



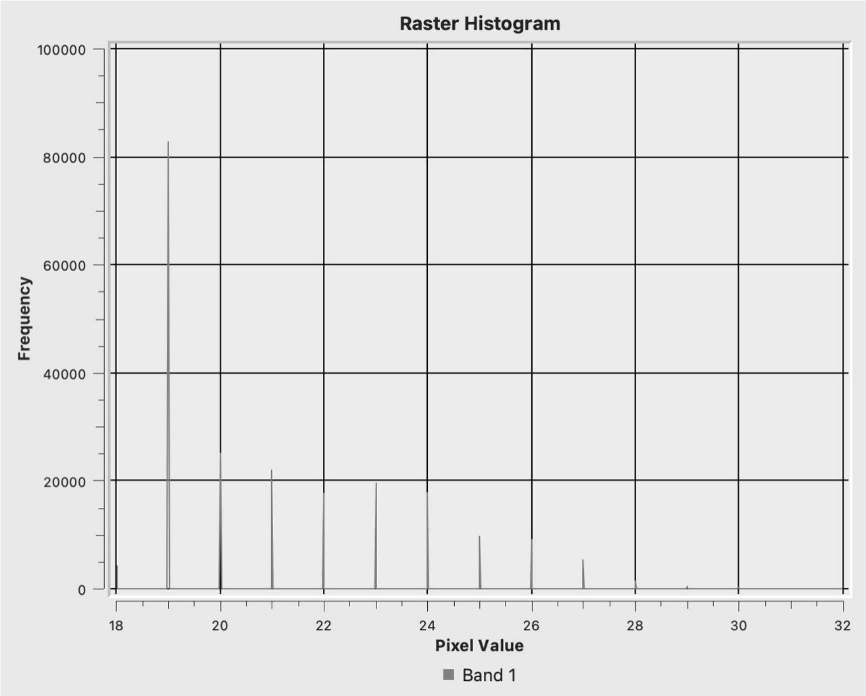
Difference between output by our model and groundtruth

Difference between output by our bilinear and groundtruth

Histogram



30m DEM Generated by Model



30m DEM Upsampled from 90m DEM by Bicubic interpolation

Conclusion

We have come up with a novel architecture that involves attention and is trained using a loss that preserves edges during super resolution of an image.

Our method beats traditional and several state of the art methods.

We have also been working on a prototype Web Interface for uploading DEMs and super resolving them.

With our work, we hope that we've moved towards producing more accurate DEMs thus giving more reliable DEM applications such as flood estimators, terrain analysis, etc.

Thank you

Mohammed Abbas Ansari
Uzma Firoz Khan
Mohammad Kashif
Md Haider Zama

Saquib Ali
Aaatir Nadim
Azhan Mohammed (Mentor