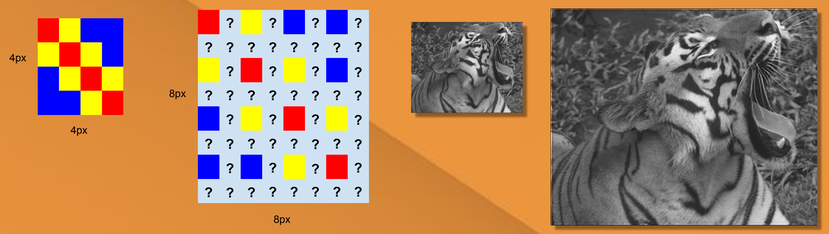
Infrared Super Resolution

Connor Northend | Jose Jimenez | Julian Castro

Background

Super resolution is the process of scaling a photo beyond its original dimensions, e.g. turning a 320x256 photo into a 1024x768 photo. Super resolution has applications that range from enhancing old movies to exploring the far reaches of space via satellite photos. This poses a question: A 1024x768 photo has more information than a 320x256 one, so how can the extra data be synthesized?



Overview

IRSR is an machine learning framework that was designed specifically for upscaling infrared images and videos. It uses information obtained from consecutive frames in a video to super resolve each image or frame. Through a neural network training, the computer can learn a set of kernel weights to convolve with an input image, and a mapping can be determined that takes low resolution inputs and generates high resolution outputs. The goal of this project was to run the super resolution algorithm directly on a camera so a balance between computation time and image clarity was designed for.

FLIR Boson

Camera features:

- 12 core processor built for speed and with machine learning in mind
- Small thermal sensor (320x256) but can potentially run a super resolution algorithm in real time to produce an HD image

Network Structures

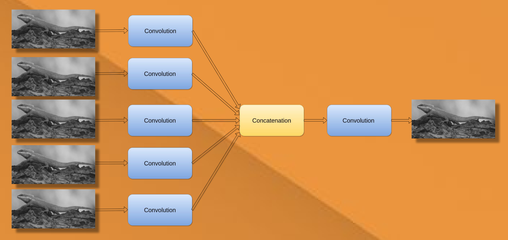
All of the networks were trained by taking a high definition image or video and downsampling them by a factor of 3. The data was then re-upsampled in order to generate a data pair that had the original HD image and a corresponding less detailed version. The less detailed version was then sent through the network, and the output of the network was compared to the original HD data. The values of the convolution kernels were then adjusted in order to minimize the loss between the network's output and the original HD data. The input images are resized bicubically, and the networks performs enhancement of the images. The training was completed on a Titan X Pascal due to the large number of training images and iterations required.

SRCNN

- Uses one frame to super resolve an image
- Based on paper by Chao Dong et al.

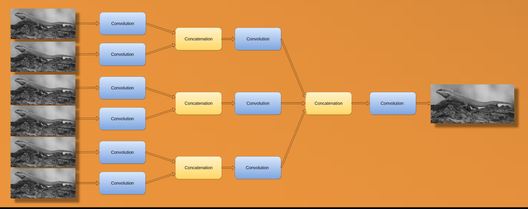
VSRnet

- Uses multiple frames to reconstruct super resolved image
- Based on paper by Armin Kappeler et al.

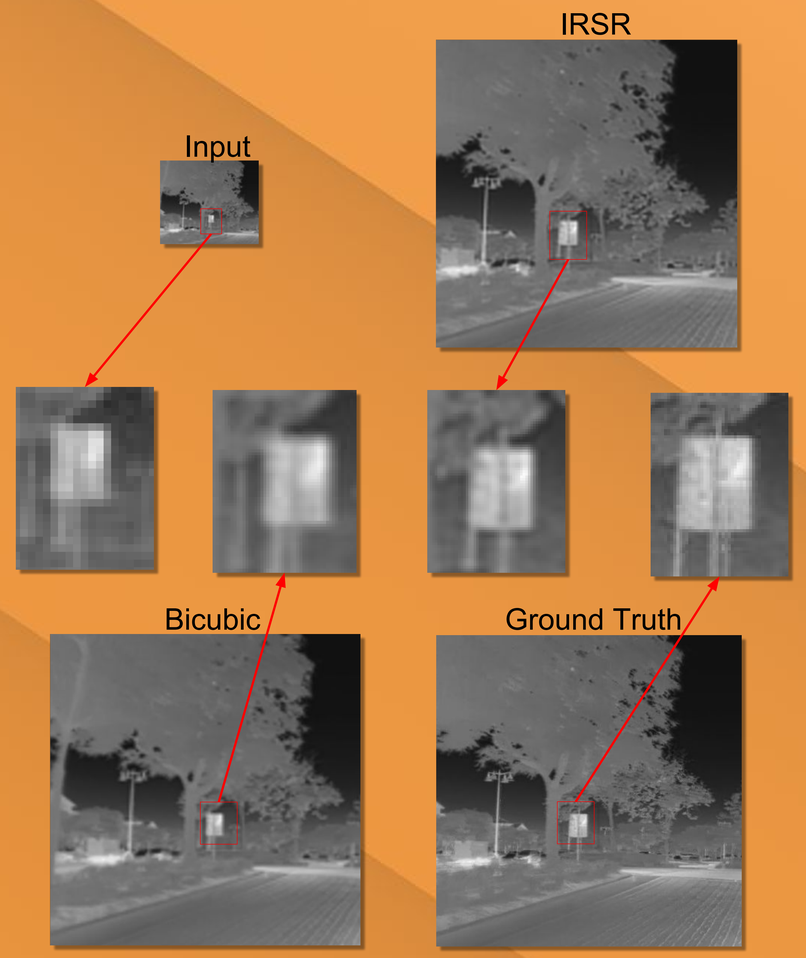


IRSR

- Uses multiple frames to reconstruct super resolved image
- Based on VSRnet with extra layers added in the center for performance



Super Resolved Thermal Image



Left image: Super Resolved using SRCNN
Middle image: Super Resolved using VSRnet
Right Image: Super Resolved using IRSR

Acknowledgements:
We would like to thank our professor Ilan Ben-Yaacov, teaching assistant Ekta Prashnani, and advisor B.S. Manjunath. We would also like to thank Louis Tremblay and Jim Klingshirn from FLIR, as well as the creator of VSRnet Armin Kappeler.