

Description:

In this project you'll be using machine learning to do SuperResolution. Superresolution is a technique to enhance image resolution:

https://en.wikipedia.org/wiki/Super-resolution_imaging

Basics:

Images usually have two spatial dimensions (height and width) as well as a third (spectral) dimension, often called depth, bands, or channels. A black and white image is only 2D and has no 3rd dimension so its resolution can be written as $H*W$. An RGB image has three bands for red, blue, and green frequencies of the spectrum and its resolution can be written as $H*W*3$. A multispectral image captures many more frequency bands and has a larger third dimension, so for example its resolution can be $H*W*15$. Hyperspectral can have resolutions of $h*w*30$, or even larger.

Different types of imaging use different types of cameras with different trade offs. RGB cameras have high spatial resolution ($H*W$) but low spectral resolution (3). Hyperspectral cameras have low spatial resolution ($h*w$) [low quantity represented by lowercase letters] while high spectral resolution (30 or higher). An ideal image would have high resolution in both spatial ($H*W$) as well as spectral (D) dimensions, but usually cameras don't offer both.

Superresolution:

Sometimes we get two images of the same scene from two different types of cameras: RGB (with high spatial resolution and low spectral one ($H*W*3$)) as well as hyperspectral (with low spatial resolution but high spectral one ($h*w*30$)). Superresolution is a technique which takes two such images of the same scene and produces an ideal image (with high spatial as well as high spectral resolution ($H*W*30$)).

Task:

Your task is to achieve superresolution using ML/DL models, i.e., train a model that takes two images (one RGB ($H*W*3$) and one hyperspectral ($h*w*31$)) and produce an image that contains the best of both ($H*W*31$).

Dataset:

You will use the CAVE data set by Columbia University:

<https://www1.cs.columbia.edu/CAVE/databases/multispectral/>

You can download the Single zip file (complete_ms_data.zip) from **Access Instructions** and unzip it. In it you will find many folders each containing a hyperspectral image. For example in

the folder `./balloons_ms/balloons_ms/` you will find 31 images of resolution $H*W=512*512$ labeled from `balloons_ms_01.png` to `balloons_ms_31.png`. You can consider each of these images as one band of a hyperspectral image, i.e., if you stack these 31 images of resolution $512*512$ one after the other, you will obtain a hyperspectral image of resolution $512*512*31$. You will then crop this hyperspectral image across the height and width dimensions into smaller images of resolution $64*64*31$ each. Each time you crop a portion of the image, you advance by 32 pixels in one dimension, that way you get more images.

You do this for each folder and the set of these smaller images ($H*W*D=64*64*31$ each) will be your ground truth (Y).

Input generation:

From each of these images in this ground truth you'll then generate two images:

1. LowResHSI ($h*w*D$)

One will simulate an image with low spatial dimensions and high spectral dimension. You'll obtain this by applying a $8*8$ averaging filter on each band of your image of resolution $64*64*31$ to obtain an image of resolution of $8*8*31$. Let's call this one a low (spatial) resolution HSI image.

2. HiResRGB ($H*W*d$)

From each of the same $64*64*31$ images you will then generate another image having high spatial resolutions and low spectral resolution. You'll basically have three $1*1$ filters on this image (one will average the bands 1-10, second will average bands 11-20, and the third will average bands 21-31 in the spectral dimensions). This will give you an image of resolution $64*64*3$. This will simulate the RGB image (it's not exactly RGB). Let's call it the high (spatial) resolution RGB image. These LowResHSI and HiResRGB images will form the training and testing dataset of your model (X).

Model:

You can use any ML model you like as long as you can explain (the what, why, and how) of every choice you make (model, hyperparameters, etc.).

Output:

When given two images of the same scene (one LowResHSI and one HiResRGB) your model should output a high (spatial) resolution and high spectral resolution HiResSHI image ($64*64*31$).

Cost function:

Use the mean square error function to measure the different between your predicted output (\hat{y}) and the actual ground truth that you had ($64*64*31$ image).

I have attached a paper that performs superresolution using ML that you can use for inspiration. You can also google about superresolution and stuff to discover information on your own.

https://ieeexplore.ieee.org/abstract/document/9449622?casa_token=yeBZYUJZPxEAAAAA:sSxiBqaDhQa-mHI74saqOQTPaW3l6FbAwzZlft04Pv5rBVWxo1pmuB1nh6hl2vVtzmb8mb2bEw

<https://ieeexplore.ieee.org/abstract/document/9650906>

Grading:

10% marks for data generation (X and Y).

30% marks for the working of your model and your explanations in viva/presentation.

Note:

The terms multispectral and hyperspectral can have blurred boundaries, i.e., at how many bands the hyperspectral starts, etc. In this document, I have used them interchangeably. 31 bands, I've called hyperspectral.

Thank you and best of luck :)