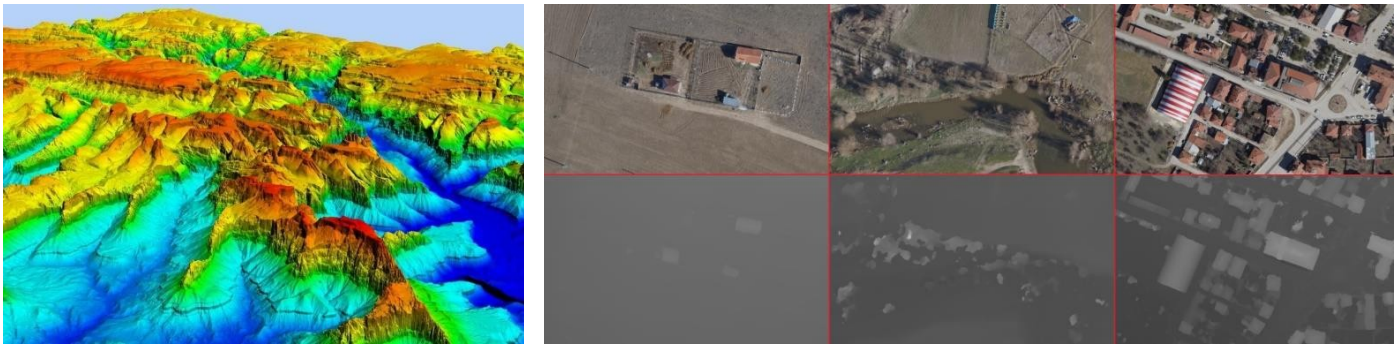
Neural Networks

Digital Surface Model (DSM) can be obtained using different sensors and techniques. Independent of how DSM is constructed, it is a 2-dimensional array holding floating values for each cell which represents the height for the location of that cell. For example, height value is zero in sea-level and 8848 meters in top of mountain Everest. So, you can think of DSM as a single-channel image with 32bits floating data type pixels. Note that pre-trained models (i.e., AlexNet, VGG, GoogLeNet, ResNet, DenseNet) generally accept RGB images which have 3 channels that are generally normalized between 0 and 1 (from 0 and 255). Given this information solve one of the below problems (solve only one):

1) Develop a classifier which will classify a 64x64 DSM patch with pixel resolution with 0.5 meters (32 meters x 32 meters area) as *unknown*, *terrain* (only terrain), *building* (has building but can contain terrain as well), and *vegetation* (has vegetation but can contain terrain as well).
2) Given DSM data (Figure 1b bottom) you want to generate orthophoto (raster) data (Figure 1b top).

Explain why you choose (1) or (2). Assume there exists no training data or very limited training data. Draw your architecture and explain it (including its input and output). What are your assumptions? What are the limitations and how do you overcome them (for training and/or inference)? How will you train it (give details such as optimizer, loss)? How will you make an inference using it?



a) 3D visualization of an example DSM b) Orthophoto (top) and corresponding DSM (bottom)

Figure 1. Illustration of an DSM

Figure 2. 3D scene of the Drone-derived Digital Surface Model (DSM) and orthophoto.