**UCSD Capstone - Step 4: Survey Of Existing Research**

**Possible Approaches (**with code**):**

1. ***Semantic Segmentation of Aerial Imagery - DigitalSreeni*** (Video)
   * This video covers the creation of a multi-class semantic segmentation model which is trained on a small dataset of aerial imagery over Dubai. This video showcases some examples in dataset augmentation, data preprocessing, and how to train a simple U-Net for the segmentation task.
   * **Code:** <https://github.com/bnsreenu/python_for_microscopists/tree/master/228_semantic_segmentation_of_aerial_imagery_using_unet>
   * **Video:** <https://www.youtube.com/watch?v=jvZm8REF2KY>
2. ***The Multi-Temporal Urban Development SpaceNet 7 Dataset***
   * This is the original proposed baseline presented by SpaceNet for our challenge. This approach creates a binary classification model which performs semantic segmentation of 4m resolution aerial imagery for building footprint extraction. The authors propose a simple U-Net with a VGG16 backbone.
   * **Paper:** <https://arxiv.org/abs/2102.04420>
   * **Code:** <https://github.com/CosmiQ/CosmiQ_SN7_Baseline/tree/master>

**Additional Research:**

1. ***SpaceNet 7 - lxastro - Solution***
   * This is the winning solution’s approach to the SpaceNet 7 challenge. This approach trades the U-Net + VGG16 backbone for the HRNet architecture. The approach also differs in their post-processing of the prediction masks.
   * **Paper:** <https://github.com/SpaceNetChallenge/SpaceNet7_Multi-Temporal_Solutions/blob/master/1-lxastro0/report-final-lxastro0.docx>
   * **Code:** <https://github.com/SpaceNetChallenge/SpaceNet7_Multi-Temporal_Solutions/tree/master/1-lxastro0>
2. ***MeVIT: A Medium-Resolution Vision Transformer for Semantic Segmentation on Landsat Satellite Imagery for Agriculture in Thailand***
   * This paper takes a look at using Vision Transformers for semantic segmentation of Landsat Imagery. Landsat offers a similar resolution to Sentinel and I believe that it may be possible to apply their techniques to my task.
   * **Paper:** <https://www.mdpi.com/2072-4292/15/21/5124>
3. ***Satellite Image Semantic Segmentation***
   * This paper proposes the use of a Swin Transformer for the semantic segmentation of land use classes through the analysis of 50cm resolution imagery.
   * **Paper:** <https://ar5iv.labs.arxiv.org/html/2110.05812>
4. ***Benchmark for Building Segmentation on Up-Scaled Sentinel-2 Imagery***
   * This paper proposes a dual superresolution and semantic segmentation model ensemble for the use of building footprint extraction on upscaled sentinel 2 imagery. I believe that an approach like this may be a valuable consideration in my task.