

# Computer Vision Training Video Collection Requirements and Guidelines

## Overview of the objective

The objective is to collect videos of multiple approaches towards a specific building within a scene. The approaches should vary in direction, angle, zoom, etc..

Proper video collection depends on the diversity and quality of the data used to train the model. The model's ability to function correctly is directly tied to the range of environments and scenarios it has been exposed to.

Data diversity is not simply about volume, but the collection of meaningfully different samples. A good location or "scene" is one that introduces distinct characteristics that improve the model's understanding of targets and their context. A location is considered unique when it adds new combinations of features, relationships, or environmental conditions.

Locations should include a wide range of structure types that are relevant to the target domain. These may include:

- Industrial facilities (factories, warehouses, refineries, power substations, wastewater treatment)
  - Structures with visible pipes, smokestacks, tanks
- Transportation infrastructure (rail yards, bridges, highways)
- Urban features (dense residential blocks, commercial centers, high-rise clusters)
- Abandoned or degraded structures
- Rural and remote structures (isolated buildings, agricultural sites, small settlements)

Priority should be given to structures that exhibit distinct geometric layouts and have recognizable functional components. For example, don't pick multiple scenes in a neighborhood with cookie cutter houses that look very similar.

Geographic variation is critical so try to vary this as much as possible:

- Different regions
- Varied climates
- Diverse terrains (coastal, flat, forested, etc)

Beyond the structure itself, the surrounding context plays a key role in localization. Locations should vary across:

- urban density (dense cities vs sparse rural areas)

- Background clutter and occlusion
- Seasonal and temporal conditions (snow cover, foliage, shadows)
- Lighting and atmospheric conditions (time of day, cloud cover)
- Dynamic elements within the scene (moving vehicles, transient objects, etc)

Example images of good scenes:



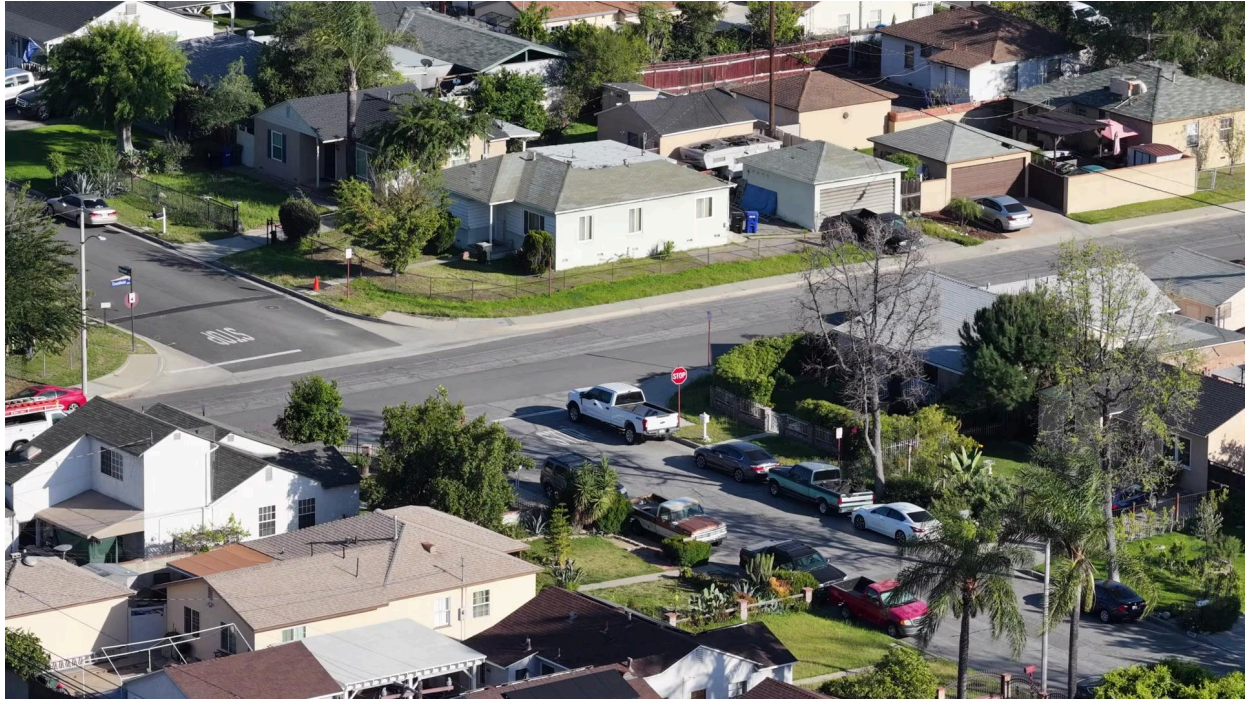














Requirements of collecting a good scene

**Variability is our friend, mix up types of shots, cameras, zooms, approach direction and angles, etc...**

### **We Need a minimum of the following shots per scene:**

1. Top down
  - a. start at max altitude and descend straight down on to the target building. Get as close as reasonably possible.
2. Top down with lateral arc
  - a. Start at max altitude and then fly in an arc going forward and down towards the target. (Imagine the second half of a thrown ball where it is descending towards the ground)
3. An approach towards target building with approximately 45 degree down angle in the approach
4. A second approach towards target building with approximately 45 degree down angle in the approach, coming from a different direction as shot #3.
5. Ground level cell phone photos of target building and any structures within center frame of we are flying towards:
  - a. We want 360 degree coverage of the buildings
  - b. Vary the distance between camera and target building between pics, also vertical and horizontal
  - c. If you cannot get ground level pics, take a close up still photo with a drone. Ensure the still photo does not look similar to any approach videos frames

### **Approach flights towards target building:**

- Length of approach: Rule of thumb here is start with the target building taking up around 5-10% of the total frame and then proceed to fly towards the target building
- Vary your approaches directions, stability levels, camera/zoom levels, exposure, etc... anything to add variability within shots

### **Capture the location of the target building (map pin or lat/long)**

- This allows us to find it on a map for post processing

### **Also submit srt files along with mp4**

- This is useful data for us and most drones should allow for it

### **We also want IR footage of the same shots if possible**

- If pilots have a drone with separate IR camera, or even a combination that does both simultaneously (like DJI mavic 3 thermal advanced) this is highly preferred.