Take-Home Questionnaire – Geospatial Instance Segmentation with TorchGeo + Mask R-CNN

Name: _____ Date: _____

1. Instance vs Object vs Semantic Segmentation

In your own words, explain the difference between:

- **Object detection**
- **Semantic segmentation**
- **Instance segmentation**

Give **one geospatial example** for each.

Answer:

2. Why TorchGeo?

Why do we use **TorchGeo** instead of writing a raw PyTorch `Dataset` for VHR-10 and other remote sensing datasets?
Mention at least **two advantages** that are specific to geospatial data.

Answer:

3. Variable Image Sizes & `collate_fn`

In the workshop, we used a custom `collate_fn` that returns a list of samples (not a stacked tensor).

a) Why is this necessary for VHR-10 images?
b) What might go wrong if we used the default collate function?

Answer:

4. Transfer Learning Intuition

We started from a **COCO-pretrained Mask R-CNN** and only changed the prediction heads.

a) What kind of features learned on COCO might still be useful for overhead remote sensing images?
b) What might *not* transfer well?

Answer:

5. `num_classes` and Background

In the code, we set:

```
num_classes = len(base_ds.categories)
```

a) Why do we need to include the **background** class in `num_classes` ?
b) What do you think might happen if we forgot to include the background class?

Answer:

---

6. Experiment with Score Thresholds

Re-run inference and visualize predictions with:

- Score threshold = **0.3**
- Score threshold = **0.9**

Compare visually.

a) What changes do you see in the **number** of predicted instances?
b) How would you describe the trade-off between **precision** and **recall** as you increase the threshold?

Answer:

---

7. Effect of Training Data Size

Train the model on:

- Only **20 images**
- Then on **100 images**

For each setting:

- Note the training loss trend
- Run the small evaluation and compare precision/recall

What do you observe about model performance when you increase the training data?

Answer:

---

8. Class Imbalance in VHR-10

Using the dataset, count how many objects of each class exist in the 100-image subset you trained on.

a) Which classes are **rare** in your subset?
b) How might this class imbalance affect Mask R-CNN's performance in a real project?

Answer:

---

9. Failure Cases Gallery (Qualitative Analysis)

Create a "failure gallery":

- Pick **5 images** where the model clearly misses objects or misclassifies them.

For one example, briefly describe:

- What is the failure?
- Why do you think the model might have failed (e.g., object size, occlusion, confusing background)?

Answer:

---

10. Design a Real Geospatial Instance Segmentation Project

Imagine you are designing a real-world project (e.g., detecting ships, storage tanks, buildings, or vehicles).

Answer briefly:

a) What dataset would you need? (imagery type, spatial resolution, labels)
b) What pre-processing or **data augmentation** would you add beyond what we did?
c) How would you evaluate the model in a way that makes sense for that real-world use case (not just a single number)?

Answer: