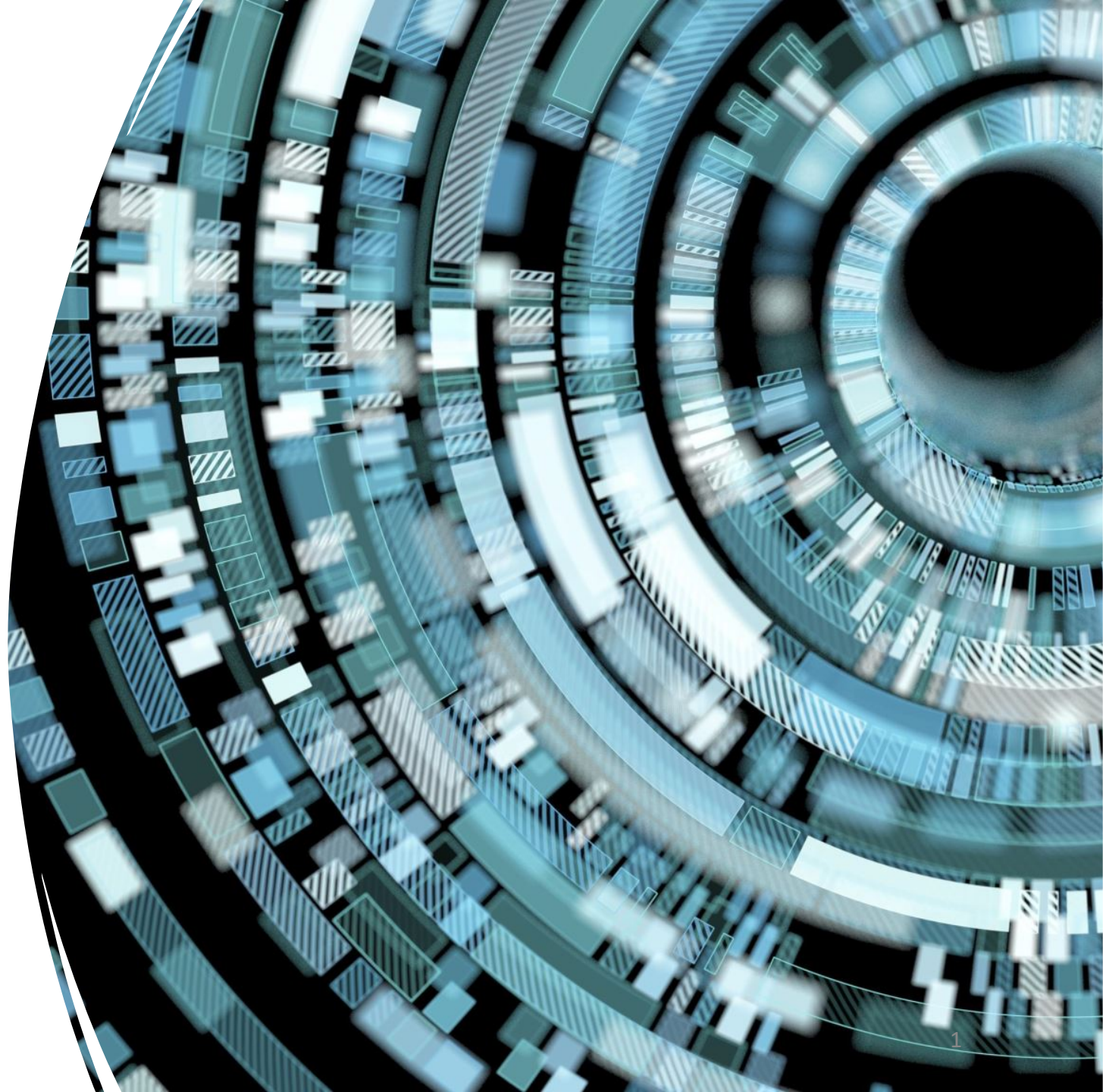


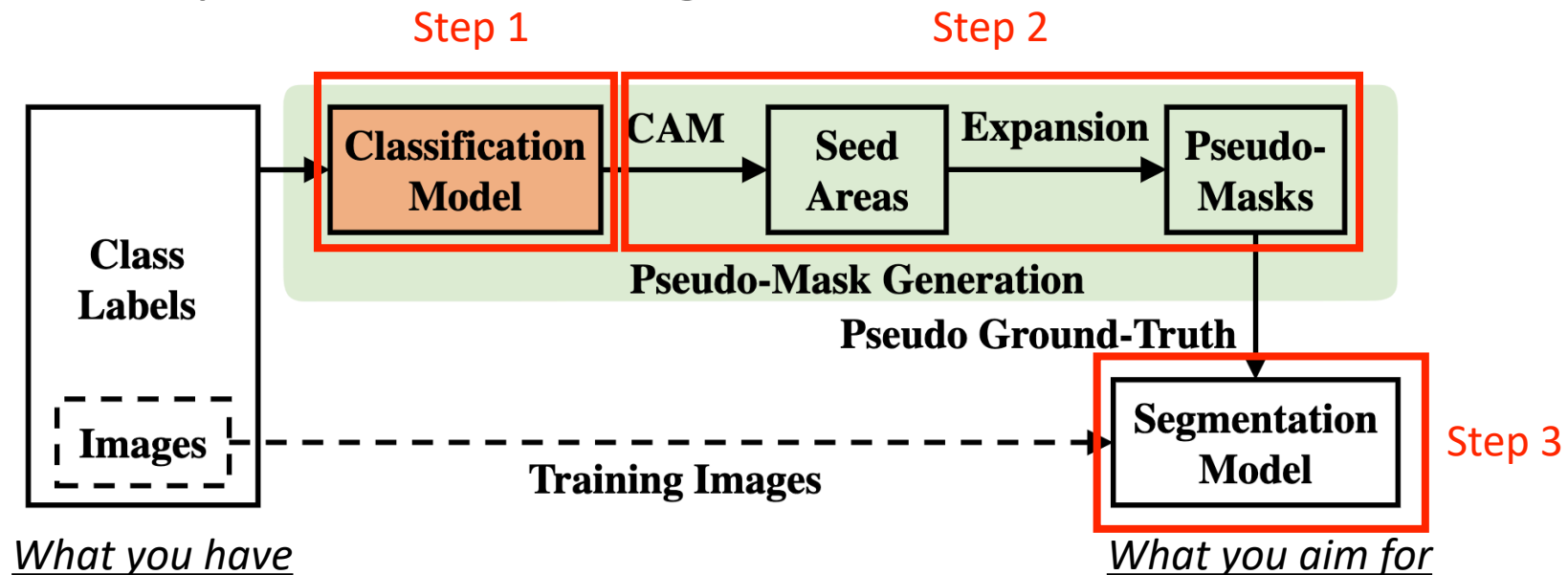
Deep Learning and Vision

Project



Task

- Weakly-Supervised Semantic Segmentation (WSSS)
 - Step 1: Multi-Label Classification
 - Step 2: Pseudo-Mask Generation
 - Step 3: Semantic Segmentation



Step 1 - Multi-Label Classification

- What you have: Images and Image-level labels (training data)
- What to do: Train a multi-label classification model
- What to submit (week11 Mon 23:59): multi-label classification accuracy on the validation set
- What to submit (week12 Mon 23:59): multi-label classification accuracy on the testing set

Step 2 - Pseudo-Mask Generation

- What you have: Images, Image-level labels and the classification model obtained in Step 1
- What to do: Generate pseudo-masks for training images
- What to submit (week11 Mon 23:59): pseudo-masks of all training images

Step 3 - Semantic Segmentation

- What you have: Images and the pseudo-masks obtained in Step 2
- What to do: Train a semantic segmentation model
- What to submit (week11 Mon 23:59): segmentation results on the validation set
- What to submit (week12 Mon 23:59): segmentation results on the testing set

Data

- Food images with image-level labels
- 103 classes
- Number of images:
 - 4,005 for training set
 - 505 for public test set (i.e., the validation set)
 - 1,000 for private test set (i.e., the testing set)

Data example



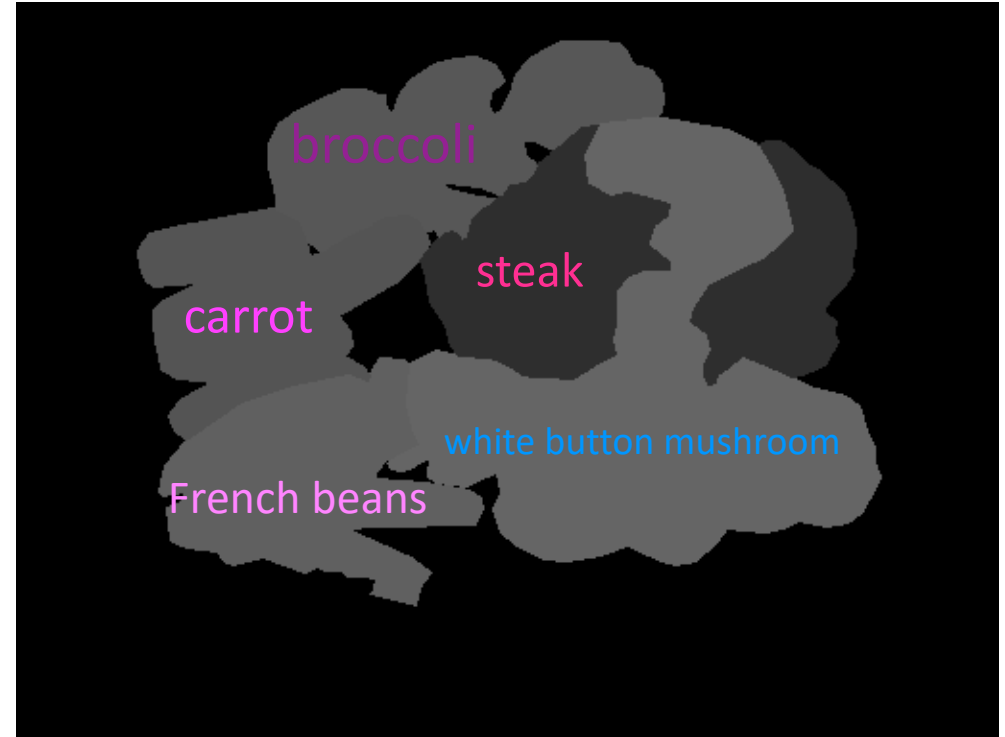
- Image-Level Labels
 - pork
 - corn
 - asparagus

Data example



- Image-Level Labels
 - steak
 - carrot
 - broccoli
 - French beans
 - white button mushroom

Data example



- Pixel-level Labels
(hidden ground truth)

Evaluation (the example code will be given)

- Task1: F_1 score

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

How many objects are found and correctly recognised?

- Task2: Mean Intersection over Union (mIoU)
 - Task3: Mean Intersection over Union (mIoU)
- The quality of the mask (how much coverage)

Two Phases

- Phase1: Public Data (Validation Set) has 505 images
 - 27 Sep 2021, 20:00 SGT to 25 Oct 2021, 23:59 SGT
 - Open leaderboard visible to all
 - Each team has up to 10 submission trials per day

Two Phases

- Phase2: Private Data (Testing Set) has 1,000 images
 - 25 Oct 2021, 23:59 SGT to 1 Nov 2021, 23:59 SGT
 - Hidden leaderboard
 - Each team has up to 2 submissions **in total** (The final result will be depended on the last submission)
 - You need to submit the results of Step 1 and Step 3

Competition Site

- The competition site will be released later.
- Create an account for your team
 - Your account ID must be “**smu_cs701_21term1_Tx**” where **$x \in [1,2,\dots,10]$** is your team number
 - Only one account per team

Limitations

- Input Size: 512*512
- Params Size: up to 500M each model
- If pretrain: only ImageNet is allowed (torchvision.models)
- Here are two example trainings:

	FPN-R50	CCNet-R50
Input Size	512*512	512*512
Batch Size	4	4
GPU RAM	~8G	~11.5G
Params Size	~220M	~380M
Time (80k step)	~9h	~18h

Timeline

- Week 6 (weekend): Optional workshop.
- 27 Sep 20:00 SGT (Week 7): Dataset release. Public leaderboard opens
- 25 Oct 23:59 SGT (Week 11): Public leaderboard closed.

Private test data released at 23:59.

- 1 Nov 23:59 SGT (Week 12): Private test submission is due.
- 8 Nov 19:00 (Week 13): Teams make presentation on final results.
- 15 Nov 23:59 SGT (Week 14): Project report due.

Grading (25%)

- Performance on public leaderboard (5% — 2%,1%,2% for step 1,2,3)
- Performance on private leaderboard (10%— 5%,5% for step 1,3)
- Project report (≤ 6 pages) & presentation (≤ 10 mins) (10%)
- Bonus score for interesting extension of the problem or significantly novel idea of the solution (3%)