

Homework 3 Report – Visual Recognition using Deep Learning

313551113 王唯誠

[GitHub - 108062237/CVDL_HW3](https://github.com/108062237/CVDL_HW3)

1. Introduction

In this assignment, we perform instance segmentation on four types of cells in medical images. The task requires using Mask R-CNN to detect each cell's bounding box and predict its class. Additionally, we compare the performance of different backbones to select the most suitable architecture.

2. Method

a. Data Preprocessing

1. Load original images and masks.

- According to the folder structure, read the color image from `train/[image_id]/image.tif`, then sequentially load `class1.tif` through `class4.tif` in the same directory.
- In each `classX.tif`, the pixel values represent instance IDs for that class (0 = background; >0 = the Nth cell).

2. Binarization and target field generation

- For each `classX.tif`, first take the unique non-zero pixel IDs to identify each cell instance.

```
for inst_id in np.unique(mask_arr):
    if inst_id == 0:
        continue
    binary_mask = (mask_arr == inst_id)
    masks.append(torch.tensor(binary_mask, dtype=torch.uint8))
    labels.append(cls)
```

- Then, for each instance's binary mask, compute the minimum and maximum x and y coordinates to form its bounding box.

```
for m in masks:
    pos = torch.nonzero(m)
    ymin = torch.min(pos[:, 0]).item()
    xmin = torch.min(pos[:, 1]).item()
    ymax = torch.max(pos[:, 0]).item()
    xmax = torch.max(pos[:, 1]).item()
    boxes.append([xmin, ymin, xmax, ymax])
```

- Finally, package this information into a target dictionary.

```
target = {
    'boxes': boxes,
    'labels': labels,
    'masks': masks,
    'image_id': torch.tensor([idx])
}
```

3. Data Splitting & Loader Construction

Randomly shuffle all image_id and split them into training and validation sets with an 80%/20% ratio.

4. Transform Pipeline

Since the chosen transforms library already provides built-in normalization and resizing [1], there is no need to reimplement these steps in the code.

b. Model Architecture

Base Architecture: Mask R-CNN integrated with a Feature Pyramid Network (FPN) to enhance multi-scale feature representation [1].

Backbone: By default, ResNet50-FPN is used; switchable to ResNet101, ResNet152, ResNeXt50_32x4d, or ResNeXt101_32x8d.

RPN Anchors: Utilizes the default anchor configuration provided by FPN (multiple scales and three aspect ratios).

ROI Heads:

1. Box Head: Replaces the default classification and regression layers with FastRCNNPredictor, producing class scores and bounding-box offsets for the specified num_classes.
2. Mask Head: Replaces the default mask branch with MaskRCNNPredictor.

c. Training Setup

Optimizer: AdamW

Learning rate: 1e-4

Epochs: 30

Batch size: 2

Device: RTX4090 & RTX3060

3.Experiments & Results

Backbone	loss curve																																																																																													
Resnet50	<div>Training vs Validation Loss over Epochs</div> <table border="1"><thead><tr><th>Epoch</th><th>Train</th><th>Validation</th></tr></thead><tbody><tr><td>1</td><td>2.65</td><td>2.00</td></tr><tr><td>2</td><td>1.90</td><td>1.95</td></tr><tr><td>3</td><td>1.75</td><td>1.85</td></tr><tr><td>4</td><td>1.80</td><td>1.90</td></tr><tr><td>5</td><td>1.82</td><td>2.02</td></tr><tr><td>6</td><td>1.78</td><td>1.88</td></tr><tr><td>7</td><td>1.78</td><td>2.05</td></tr><tr><td>8</td><td>1.78</td><td>1.90</td></tr><tr><td>9</td><td>1.80</td><td>1.80</td></tr><tr><td>10</td><td>1.82</td><td>1.95</td></tr><tr><td>11</td><td>1.68</td><td>1.85</td></tr><tr><td>12</td><td>1.65</td><td>1.75</td></tr><tr><td>13</td><td>1.60</td><td>1.70</td></tr><tr><td>14</td><td>1.58</td><td>1.68</td></tr><tr><td>15</td><td>1.55</td><td>1.68</td></tr><tr><td>16</td><td>1.55</td><td>1.65</td></tr><tr><td>17</td><td>1.55</td><td>1.65</td></tr><tr><td>18</td><td>1.55</td><td>1.65</td></tr><tr><td>19</td><td>1.52</td><td>1.65</td></tr><tr><td>20</td><td>1.52</td><td>1.62</td></tr><tr><td>21</td><td>1.52</td><td>1.65</td></tr><tr><td>22</td><td>1.52</td><td>1.60</td></tr><tr><td>23</td><td>1.52</td><td>1.62</td></tr><tr><td>24</td><td>1.50</td><td>1.62</td></tr><tr><td>25</td><td>1.50</td><td>1.60</td></tr><tr><td>26</td><td>1.50</td><td>1.58</td></tr><tr><td>27</td><td>1.50</td><td>1.60</td></tr><tr><td>28</td><td>1.48</td><td>1.55</td></tr><tr><td>29</td><td>1.48</td><td>1.58</td></tr><tr><td>30</td><td>1.48</td><td>1.58</td></tr></tbody></table>	Epoch	Train	Validation	1	2.65	2.00	2	1.90	1.95	3	1.75	1.85	4	1.80	1.90	5	1.82	2.02	6	1.78	1.88	7	1.78	2.05	8	1.78	1.90	9	1.80	1.80	10	1.82	1.95	11	1.68	1.85	12	1.65	1.75	13	1.60	1.70	14	1.58	1.68	15	1.55	1.68	16	1.55	1.65	17	1.55	1.65	18	1.55	1.65	19	1.52	1.65	20	1.52	1.62	21	1.52	1.65	22	1.52	1.60	23	1.52	1.62	24	1.50	1.62	25	1.50	1.60	26	1.50	1.58	27	1.50	1.60	28	1.48	1.55	29	1.48	1.58	30	1.48	1.58
Epoch	Train	Validation																																																																																												
1	2.65	2.00																																																																																												
2	1.90	1.95																																																																																												
3	1.75	1.85																																																																																												
4	1.80	1.90																																																																																												
5	1.82	2.02																																																																																												
6	1.78	1.88																																																																																												
7	1.78	2.05																																																																																												
8	1.78	1.90																																																																																												
9	1.80	1.80																																																																																												
10	1.82	1.95																																																																																												
11	1.68	1.85																																																																																												
12	1.65	1.75																																																																																												
13	1.60	1.70																																																																																												
14	1.58	1.68																																																																																												
15	1.55	1.68																																																																																												
16	1.55	1.65																																																																																												
17	1.55	1.65																																																																																												
18	1.55	1.65																																																																																												
19	1.52	1.65																																																																																												
20	1.52	1.62																																																																																												
21	1.52	1.65																																																																																												
22	1.52	1.60																																																																																												
23	1.52	1.62																																																																																												
24	1.50	1.62																																																																																												
25	1.50	1.60																																																																																												
26	1.50	1.58																																																																																												
27	1.50	1.60																																																																																												
28	1.48	1.55																																																																																												
29	1.48	1.58																																																																																												
30	1.48	1.58																																																																																												
Resnet101	<div>Training vs Validation Loss over Epochs</div> <table border="1"><thead><tr><th>Epoch</th><th>Train</th><th>Validation</th></tr></thead><tbody><tr><td>1</td><td>1.95</td><td>1.70</td></tr><tr><td>2</td><td>1.78</td><td>1.72</td></tr><tr><td>3</td><td>1.75</td><td>1.62</td></tr><tr><td>4</td><td>1.72</td><td>1.65</td></tr><tr><td>5</td><td>1.75</td><td>1.65</td></tr><tr><td>6</td><td>1.70</td><td>1.65</td></tr><tr><td>7</td><td>1.72</td><td>1.75</td></tr><tr><td>8</td><td>1.70</td><td>1.68</td></tr><tr><td>9</td><td>1.68</td><td>1.68</td></tr><tr><td>10</td><td>1.68</td><td>1.68</td></tr><tr><td>11</td><td>1.68</td><td>1.60</td></tr><tr><td>12</td><td>1.68</td><td>1.58</td></tr><tr><td>13</td><td>1.68</td><td>1.60</td></tr><tr><td>14</td><td>1.65</td><td>1.62</td></tr><tr><td>15</td><td>1.65</td><td>1.65</td></tr><tr><td>16</td><td>1.62</td><td>1.58</td></tr><tr><td>17</td><td>1.62</td><td>1.62</td></tr><tr><td>18</td><td>1.62</td><td>1.65</td></tr><tr><td>19</td><td>1.60</td><td>1.58</td></tr><tr><td>20</td><td>1.60</td><td>1.55</td></tr><tr><td>21</td><td>1.58</td><td>1.55</td></tr><tr><td>22</td><td>1.58</td><td>1.52</td></tr><tr><td>23</td><td>1.55</td><td>1.58</td></tr><tr><td>24</td><td>1.52</td><td>1.50</td></tr><tr><td>25</td><td>1.52</td><td>1.55</td></tr><tr><td>26</td><td>1.50</td><td>1.48</td></tr><tr><td>27</td><td>1.52</td><td>1.50</td></tr><tr><td>28</td><td>1.52</td><td>1.58</td></tr><tr><td>29</td><td>1.50</td><td>1.55</td></tr><tr><td>30</td><td>1.50</td><td>1.40</td></tr></tbody></table>	Epoch	Train	Validation	1	1.95	1.70	2	1.78	1.72	3	1.75	1.62	4	1.72	1.65	5	1.75	1.65	6	1.70	1.65	7	1.72	1.75	8	1.70	1.68	9	1.68	1.68	10	1.68	1.68	11	1.68	1.60	12	1.68	1.58	13	1.68	1.60	14	1.65	1.62	15	1.65	1.65	16	1.62	1.58	17	1.62	1.62	18	1.62	1.65	19	1.60	1.58	20	1.60	1.55	21	1.58	1.55	22	1.58	1.52	23	1.55	1.58	24	1.52	1.50	25	1.52	1.55	26	1.50	1.48	27	1.52	1.50	28	1.52	1.58	29	1.50	1.55	30	1.50	1.40
Epoch	Train	Validation																																																																																												
1	1.95	1.70																																																																																												
2	1.78	1.72																																																																																												
3	1.75	1.62																																																																																												
4	1.72	1.65																																																																																												
5	1.75	1.65																																																																																												
6	1.70	1.65																																																																																												
7	1.72	1.75																																																																																												
8	1.70	1.68																																																																																												
9	1.68	1.68																																																																																												
10	1.68	1.68																																																																																												
11	1.68	1.60																																																																																												
12	1.68	1.58																																																																																												
13	1.68	1.60																																																																																												
14	1.65	1.62																																																																																												
15	1.65	1.65																																																																																												
16	1.62	1.58																																																																																												
17	1.62	1.62																																																																																												
18	1.62	1.65																																																																																												
19	1.60	1.58																																																																																												
20	1.60	1.55																																																																																												
21	1.58	1.55																																																																																												
22	1.58	1.52																																																																																												
23	1.55	1.58																																																																																												
24	1.52	1.50																																																																																												
25	1.52	1.55																																																																																												
26	1.50	1.48																																																																																												
27	1.52	1.50																																																																																												
28	1.52	1.58																																																																																												
29	1.50	1.55																																																																																												
30	1.50	1.40																																																																																												
Resnext101	<div>Training vs Validation Loss over Epochs</div> <table border="1"><thead><tr><th>Epoch</th><th>Train</th><th>Validation</th></tr></thead><tbody><tr><td>1</td><td>2.05</td><td>1.88</td></tr><tr><td>2</td><td>1.78</td><td>1.58</td></tr><tr><td>3</td><td>1.70</td><td>1.45</td></tr><tr><td>4</td><td>1.68</td><td>1.55</td></tr><tr><td>5</td><td>1.68</td><td>1.60</td></tr><tr><td>6</td><td>1.70</td><td>1.60</td></tr><tr><td>7</td><td>1.68</td><td>1.55</td></tr><tr><td>8</td><td>1.65</td><td>1.52</td></tr><tr><td>9</td><td>1.65</td><td>1.45</td></tr><tr><td>10</td><td>1.65</td><td>1.48</td></tr><tr><td>11</td><td>1.62</td><td>1.48</td></tr><tr><td>12</td><td>1.62</td><td>1.40</td></tr><tr><td>13</td><td>1.62</td><td>1.50</td></tr><tr><td>14</td><td>1.62</td><td>1.58</td></tr><tr><td>15</td><td>1.62</td><td>1.45</td></tr><tr><td>16</td><td>1.60</td><td>1.40</td></tr><tr><td>17</td><td>1.60</td><td>1.52</td></tr><tr><td>18</td><td>1.58</td><td>1.40</td></tr><tr><td>19</td><td>1.58</td><td>1.42</td></tr><tr><td>20</td><td>1.58</td><td>1.40</td></tr><tr><td>21</td><td>1.55</td><td>1.42</td></tr><tr><td>22</td><td>1.55</td><td>1.42</td></tr><tr><td>23</td><td>1.52</td><td>1.42</td></tr><tr><td>24</td><td>1.52</td><td>1.42</td></tr><tr><td>25</td><td>1.55</td><td>1.42</td></tr><tr><td>26</td><td>1.52</td><td>1.40</td></tr><tr><td>27</td><td>1.52</td><td>1.42</td></tr><tr><td>28</td><td>1.52</td><td>1.45</td></tr><tr><td>29</td><td>1.50</td><td>1.40</td></tr><tr><td>30</td><td>1.52</td><td>1.40</td></tr></tbody></table>	Epoch	Train	Validation	1	2.05	1.88	2	1.78	1.58	3	1.70	1.45	4	1.68	1.55	5	1.68	1.60	6	1.70	1.60	7	1.68	1.55	8	1.65	1.52	9	1.65	1.45	10	1.65	1.48	11	1.62	1.48	12	1.62	1.40	13	1.62	1.50	14	1.62	1.58	15	1.62	1.45	16	1.60	1.40	17	1.60	1.52	18	1.58	1.40	19	1.58	1.42	20	1.58	1.40	21	1.55	1.42	22	1.55	1.42	23	1.52	1.42	24	1.52	1.42	25	1.55	1.42	26	1.52	1.40	27	1.52	1.42	28	1.52	1.45	29	1.50	1.40	30	1.52	1.40
Epoch	Train	Validation																																																																																												
1	2.05	1.88																																																																																												
2	1.78	1.58																																																																																												
3	1.70	1.45																																																																																												
4	1.68	1.55																																																																																												
5	1.68	1.60																																																																																												
6	1.70	1.60																																																																																												
7	1.68	1.55																																																																																												
8	1.65	1.52																																																																																												
9	1.65	1.45																																																																																												
10	1.65	1.48																																																																																												
11	1.62	1.48																																																																																												
12	1.62	1.40																																																																																												
13	1.62	1.50																																																																																												
14	1.62	1.58																																																																																												
15	1.62	1.45																																																																																												
16	1.60	1.40																																																																																												
17	1.60	1.52																																																																																												
18	1.58	1.40																																																																																												
19	1.58	1.42																																																																																												
20	1.58	1.40																																																																																												
21	1.55	1.42																																																																																												
22	1.55	1.42																																																																																												
23	1.52	1.42																																																																																												
24	1.52	1.42																																																																																												
25	1.55	1.42																																																																																												
26	1.52	1.40																																																																																												
27	1.52	1.42																																																																																												
28	1.52	1.45																																																																																												
29	1.50	1.40																																																																																												
30	1.52	1.40																																																																																												

4. Conclusion

In this assignment, we used Mask R-CNN to perform instance segmentation on cells and gained an understanding of the mAP evaluation metric and the COCO format. Ultimately, we achieved a score of 0.22 in the competition.

5. Reference

[1] https://github.com/pytorch/vision/blob/main/torchvision/models/detection/mask_rcnn.py