

Workshop II: Basic Image Segmentation

Assignment Workshop 03

1. ให้นักศึกษาตอบคำถามต่อไปนี้

1.1 อธิบายเบื้องต้นเกี่ยวกับโมเดลที่เป็นโครงสร้างแบบ Autoencoder ว่าคืออะไร มีโครงสร้างอะไรบ้าง

1.2 Latent Space มีหน้าที่อะไรในโครงสร้างของ Autoencoder

1.3 ในโครงสร้างสถาปัตยกรรมแบบ U-Net จะมีส่วนสำคัญที่เรียกว่า Skip Connection ให้อธิบายว่าคืออะไร และช่วยอะไรบ้างใน U-Net

2. ให้นักศึกษาทำโจทย์ของไฟล์ Workshop_II-03_InageSegmentation.ipynb ให้เรียบร้อยพร้อมทั้งมาตอบคำถามต่อไปนี้

2.1 อธิบายวิธีการสร้างข้อมูลสังเคราะห์ให้มีการสร้างรูปแบบทั้งหมด 4 รูปแบบสำหรับการสร้างข้อมูลเพื่อนำมาเทรนโมเดล

2.2 คำสั่ง Conv2d กับ ConvTranspose2d ใน pytorch ใช้งานต่างกันยังไง

2.3 แสดงผลการทำนายของการทำโมเดล U-Net ที่เทรนด้วยข้อมูลสังเคราะห์แบบมี 4 วัตถุ

2. ให้นักศึกษาทำโจทย์ของไฟล์ Workshop_II-03_InageSegmentation.ipynb ให้เรียบร้อย พร้อมทั้งมาตอบคำถามต่อไปนี้

2.1 อธิบายวิธีการสร้างข้อมูลสังเคราะห์ให้มีการสร้างรูปแบบทั้งหมด 4 รูปแบบสำหรับการสร้างข้อมูลเพื่อนำมาเทรนโมเดล

2.2 คำสั่ง Conv2d กับ ConvTranspose2d ใน pytorch ใช้งานต่างกันยังไง

2.3 แสดงผลการทำงานของการทำโมเดล U-Net ที่เทรนด้วยข้อมูลสังเคราะห์แบบมี 4 วัตถุ

ตอบ

```
class SimpleDataset(Dataset):
    def __init__(self, size=250, image_size=128):
        self.size = size
        self.image_size = image_size

    def __len__(self):
        return self.size

    def __getitem__(self, idx):
        target = torch.zeros(4, self.image_size, self.image_size)

        # Generate circle
        target[0] = gen_circle_shape(self.image_size)
        # Generate rectangle
        target[1] = gen_square_shape(self.image_size)
        # Generate triangle
        target[2] = gen_triangle_shape(self.image_size)
        # Generate heart shape
        target[3] = gen_heart_shape(self.image_size)

        # Combine circle and rectangle
        combined_target = (target.sum(dim=0) > 0).float().unsqueeze(0)

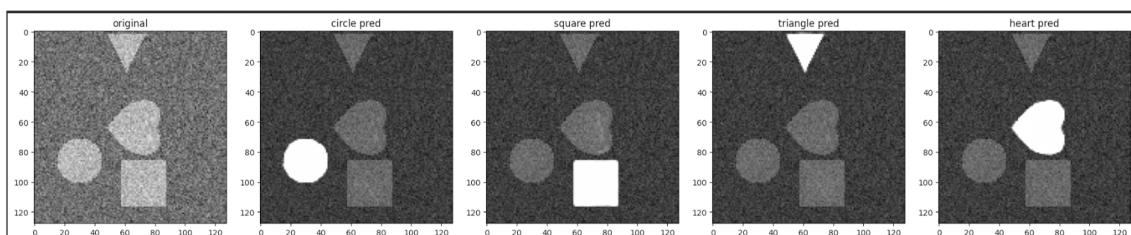
        # Add noise to create the input image
        noise = torch.randn(1, self.image_size, self.image_size) * 0.3
        input_image = combined_target + noise
```

2.1 ใน dataset จะเปรียบเสมือนว่ามีข้อมูลรูป default อยู่ 250 รูป ขนาด 128 pixel โดยจะ มีข้อมูล และ ตัวlabel โดย ข้อมูลจะเป็นรูปภาพช่วงเทา ที่มีทุกรูปร่างรวมอยู่ ได้แก่ สี่เหลี่ยม วงกลม สามเหลี่ยม หัวใจ และมีการเพิ่มการรบกวนของภาพ และตัวlabel จะเก็บเป็นอาเรย์ที่ประกอบด้วย mask ของแต่ละรูปร่าง

2.2 Conv2d ใช้ในการ convolution ภาพ ในโมเดล autoencoder เปรียบได้กับส่วน encode ส่วนใหญ่ใช้ในการ downsampling

ConvTranspose2d ใช้ในการ Transpose Convolution (Deconvolution) ในโมเดล autoencoder เปรียบได้กับส่วน decode ส่วนใหญ่ใช้ในการ upsampling

2.3



1. ให้นักศึกษาตอบคำถามต่อไปนี้

1.1 อธิบายเบื้องต้นเกี่ยวกับโมเดลที่เป็นโครงสร้างแบบ Autoencoder ว่าคืออะไร มีโครงสร้างอะไรบ้าง

1.2 Laten Space มีหน้าที่อะไรในโครงสร้างของ Autoencoder

1.3 ในโครงสร้างสถาปัตยกรรมแบบ U-Net จะมีส่วนสำคัญที่เรียกว่า Skip Connection ให้อธิบายว่าคืออะไร และช่วยอะไรบ้างใน U-Net

ตอบ

1.1 Autoencoder เป็น unsupervised learning และเป็น neural network ซึ่ง

ประกอบด้วยกัน 3 ส่วน ได้แก่ encoder (เปลี่ยนมิติข้อมูล), latent space, decoder(คืนมิติข้อมูล)

1.2 Laten Space คือ พื้นที่ของเวกเตอร์หลังจากการ encode(เปลี่ยนมิติข้อมูล) ซึ่งลักษณะของข้อมูลจะเปลี่ยนไปจากเดิม แต่ยังเก็บคุณลักษณะข้อมูลที่ใกล้เคียงเดิมไว้อยู่

1.3 เป็นเทคนิคที่ใช้ในการนำข้อมูลจาก layer ก่อนที่จะผ่าน latent space มาเพิ่มในการฝึก layer มาเพิ่มหลัง latent space ช่วยให้โมเดลเรียนรู้การ segment ได้ดีขึ้นเนื่องจาก layer หลังยังเรียนรู้กับข้อมูลที่ยังไม่สูญเสียรายละเอียดของรูปภาพ