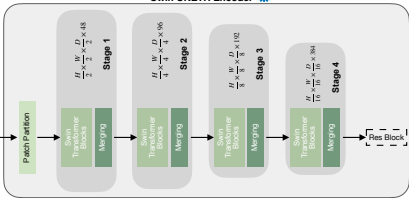


Swin UNETR Encoder



Imaging feature vector 1×256

Conv Block $\times 4$



$$\frac{H}{32} \times \frac{W}{32} \times \frac{D}{32} \times 768$$