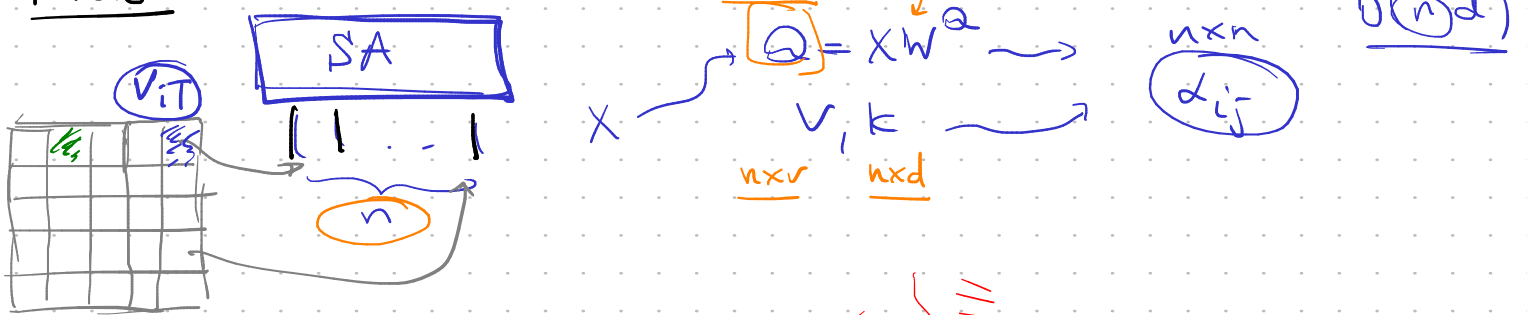
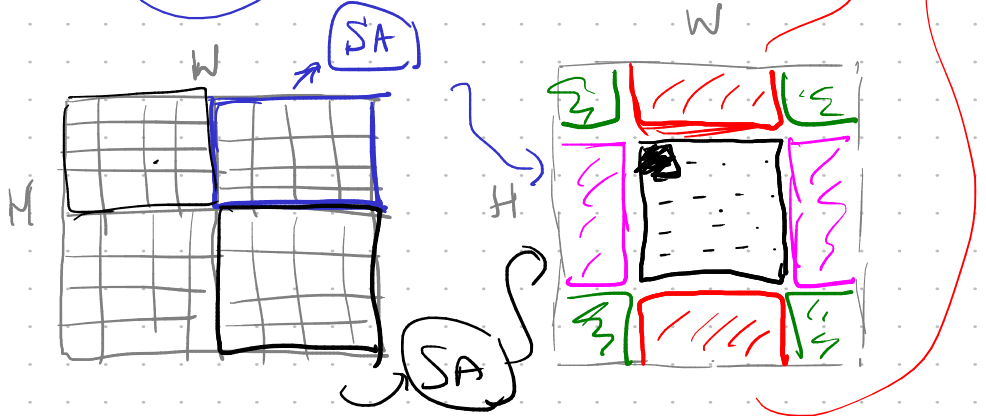


PROBLEM



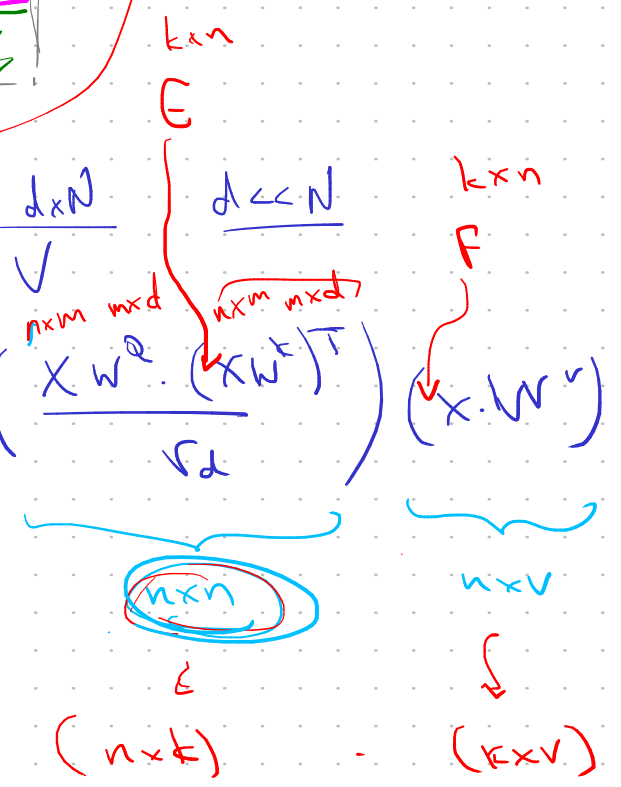
① Swin Transformer



② Linformer

$$\frac{N \times N}{X} \approx \frac{N \times d}{U} \frac{d \times N}{V}$$

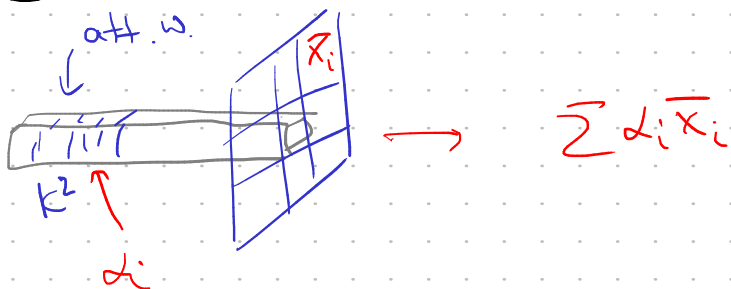
$$Att = \text{Softmax}\left(\frac{QK^T}{\sqrt{d}}\right) V = \text{Softmax}\left(\frac{XW^Q \cdot (XW^K)^T}{\sqrt{d}}\right) (XW^V)$$



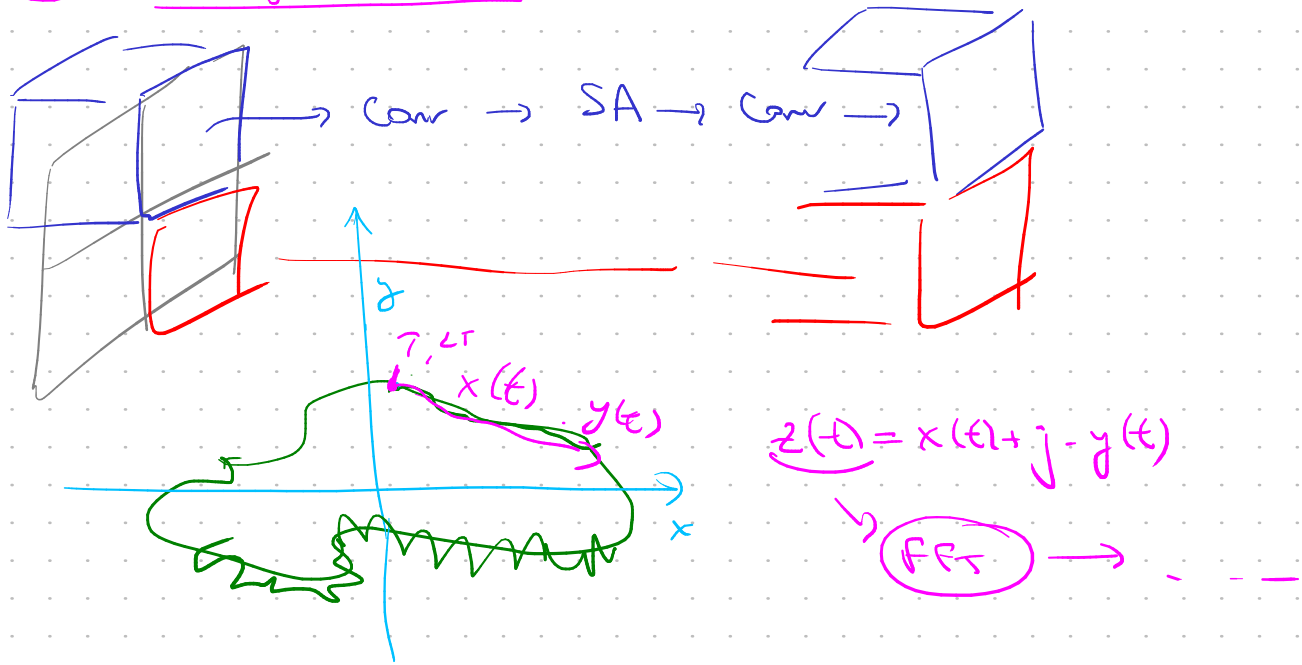
$$Att = \text{Softmax}\left(\frac{XW^Q \cdot (EXW^K)^T}{\sqrt{d}}\right) (FXW^V)$$

$n \times k$

③ Outlook attention



4) Group Transformer



$$IoU = \frac{|X \cap \hat{X}|}{|X \cup \hat{X}|}$$

$$BCE = \frac{1}{N} \sum_n \dots$$

5) Instance segmentation

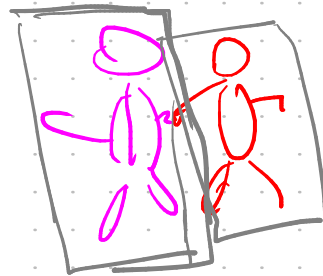
Semantic segm.



instance segm

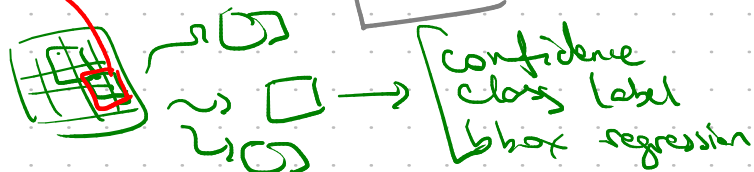


Object detection



Two-stage

RPN



⑥ DETR

