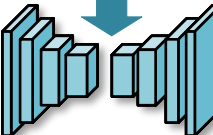
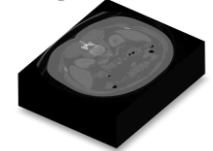
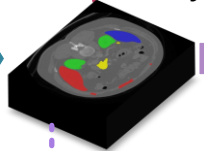


Original volume

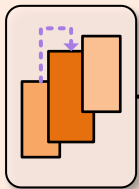


3D-Backbone  
e.g. 3D U-Net,  
UNETR, nnFormer

Coarse mask  $y^o$

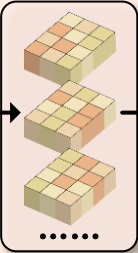


SAI Pipeline



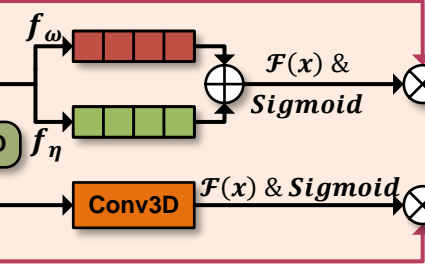
Preliminary  
feature  
extraction  
layer

Axis  
Grouping



$\{y_i^\lambda\}_{i=1}^G$

channel-wise groups



spatial-wise groups

$$\mathcal{F}(x) = Wx + b$$

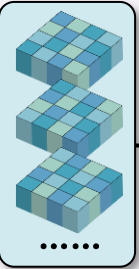
Spatial  
Redistribution

$y_s^\mu$

PDI Pipeline

Segmentation map  
enhancing

Permute



Merge

$\{y_i^\alpha\}_{i=1}^{N_s}$

$\{y_i^\delta\}_{i=1}^{N_s}$

Self-  
Attention

$\{y_i^r\}_{i=1}^{N_s}$

R

Permute



$\{y_i^r\}_{i=1}^{N_s}$

R Reshape

D Divide

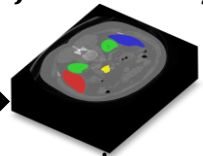
C Concat

Conv1x1+GN  
+RELU

Conv7x7+GN  
+RELU

Conv1x1+GN

Refined mask  $y^r$



$\mathcal{L}_{DiceCE}$

Ground truth

