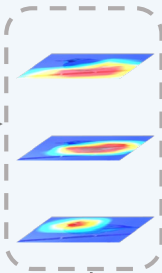


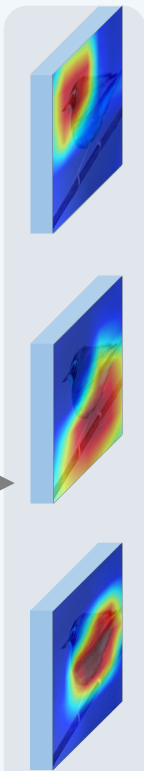
Backbone

CASL

SA



$\otimes$



DIV

CPT

RFC

$\mathcal{L}_{RFC}$

DIV

Divergence component

CPT

Compactness component

GAP

Global Average Pooling

CA

Channel Attention

AT

Adaptive Thresholding

SA

Spatial Attention

SCL

ACAM

GAP

CA

AT

Attribute Mask

ASE

brown eye

plain head

$\mathcal{L}_{ace} + \mathcal{L}_{ce}$