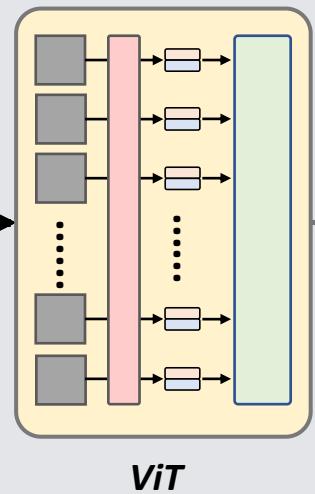
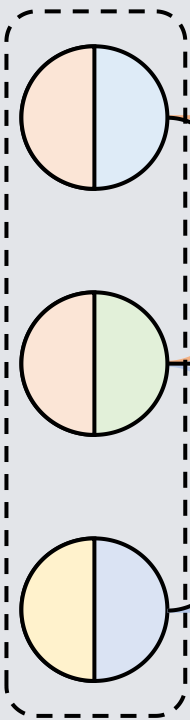


visual embedding



Keys: ρ - disentangler, ϕ - classifier, \mathcal{C} - decoupler, \rightarrow - first forward, \dashrightarrow - second forward



ρ_{attr}

ρ_{obj}

\leftrightarrow "black cat"
 \leftrightarrow "black dog"
 \leftrightarrow "orange cat"
 \mathcal{L}_{comp}

\mathcal{C}_{attr}

\mathcal{C}_{obj}

ϕ_{attr}

ϕ_{obj}

EMD

γ

γ

\mathcal{L}_{reg}

\mathcal{L}_{cls}

shared embedding space

A

B

C

.....

reasonable

\tilde{A}

\tilde{B}

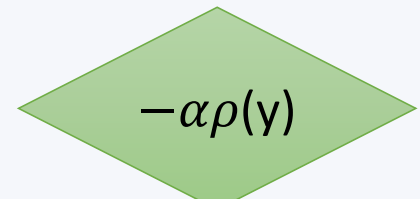
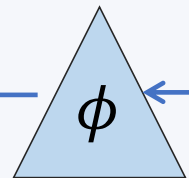
\tilde{C}

.....

unreasonable

cosine similarity

composition embedding



\mathcal{L}_{cls}^o

Open World

A

B

C

\tilde{A}

\tilde{B}

\tilde{C}

training

A

B

C

\tilde{A}

\tilde{B}

\tilde{C}

\times

\times

A

B

C

\tilde{A}

\tilde{B}

\tilde{C}

\checkmark

A

B

C



reasonable compositions

\tilde{A}

\tilde{B}

\tilde{C}



unreasonable compositions

Closed World