Pre-training stage

Base data

$x \rightarrow f_\phi \rightarrow$

Fine-tuning stage

Without Base data

Novel data

$x \rightarrow f_\phi \rightarrow$

θ_{base} θ_{novel}

$\mu_{base} < \mu_{novel}$

$\sigma_{base} > \sigma_{novel}$

0

Mismatch of base and novel weight distributions

w\ Classifier Normalization

$\mu_{base} = \mu_{novel}$

$\sigma_{base} = \sigma_{novel}$

0

Balance the weight distributions