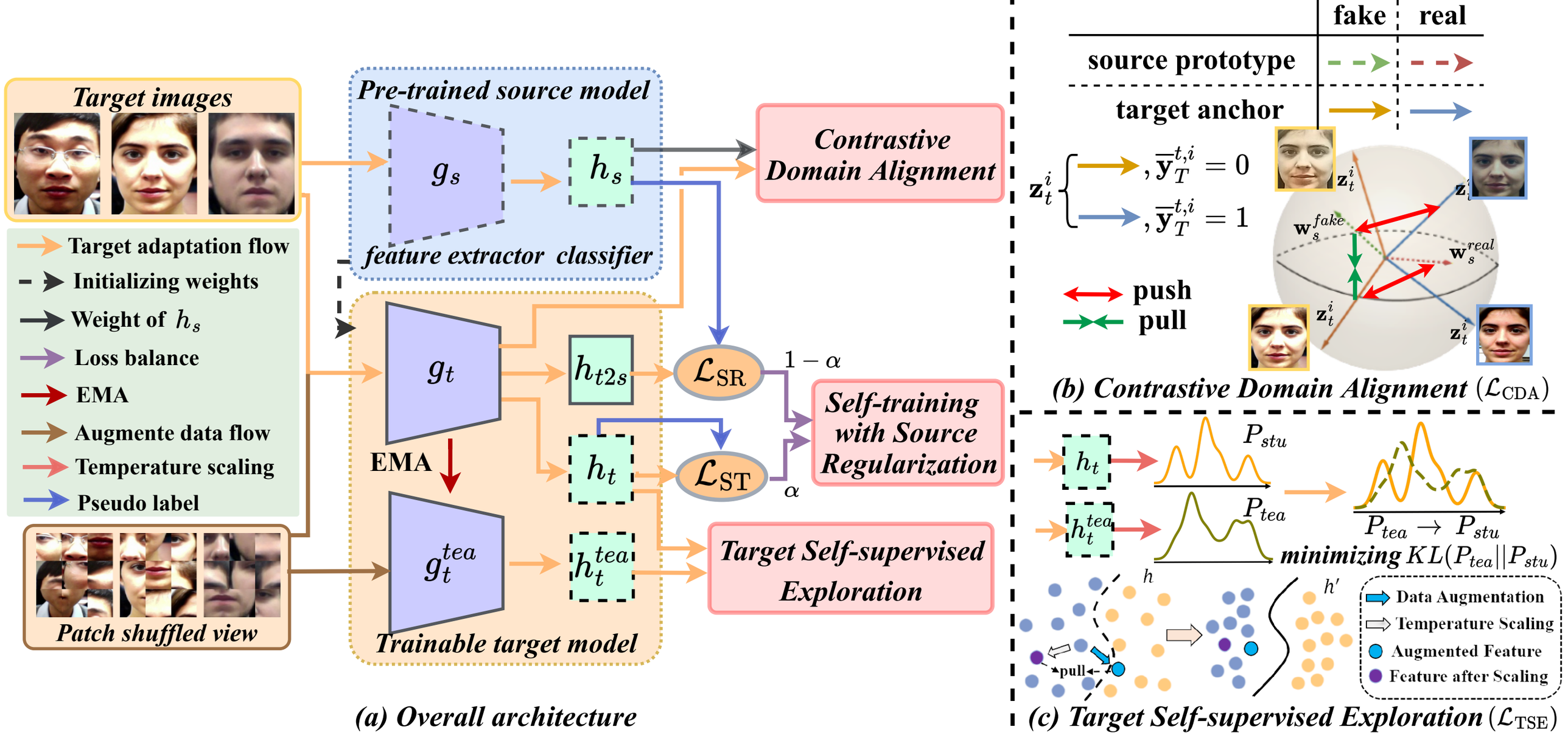
Target images
Pre-trained source model
g_s
h_s
feature extractor
classifier
Contrastive Domain Alignment
Target adaptation flow
Initializing weights
Weight of h_s
Loss balance
EMA
Augmente data flow
Temperature scaling
Pseudo label
g_t
h_{t2s}
h_t
EMA
$\mathcal{L}_{\text{SR}}$
$\mathcal{L}_{\text{ST}}$
$1-\alpha$
α
Self-training with Source Regularization
g_t^{tea}
h_t^{tea}
Target Self-supervised Exploration
Patch shuffled view
Trainable target model
(a) Overall architecture
fake
real
source prototype
target anchor
$\mathbf{z}_t^i$
$\overline{\mathbf{y}}_T^{t,i} = 0$
$\overline{\mathbf{y}}_T^{t,i} = 1$
push
pull
$\mathbf{w}_s^{fake}$
$\mathbf{w}_s^{real}$
(b) Contrastive Domain Alignment ($\mathcal{L}_{\text{CDA}}$)
h_t
h_t^{tea}
P_{stu}
P_{tea}
$P_{tea} \rightarrow P_{stu}$
minimizing $KL(P_{tea}||P_{stu})$
h
h'
pull
Data Augmentation
Temperature Scaling
Augmented Feature
Feature after Scaling
(c) Target Self-supervised Exploration ($\mathcal{L}_{\text{TSE}}$)