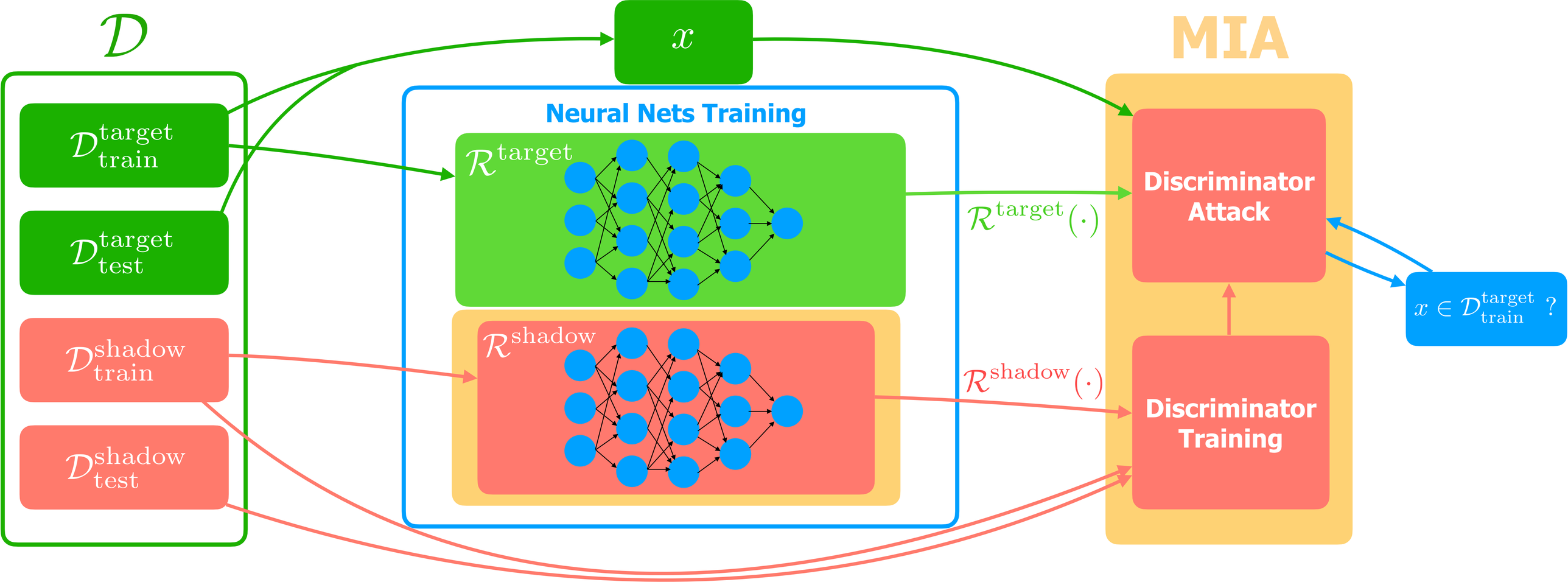
$\mathcal{D}$
$\mathcal{D}^{\text{target}}_{\text{train}}$
$\mathcal{D}^{\text{target}}_{\text{test}}$
$\mathcal{D}^{\text{shadow}}_{\text{train}}$
$\mathcal{D}^{\text{shadow}}_{\text{test}}$
x
Neural Nets Training
$\mathcal{R}^{\text{target}}$
$\mathcal{R}^{\text{shadow}}$
$\mathcal{R}^{\text{target}}(\cdot)$
$\mathcal{R}^{\text{shadow}}(\cdot)$
MIA
Discriminator Attack
Discriminator Training
$x \in \mathcal{D}^{\text{target}}_{\text{train}}$?