

Xiaosen Wang¹, Xuanran He², Jingdong Wang³, Kun He¹ ¹School of Computer Science and Technology, Huazhong University of Science and Technology ²Wee Kim Wee School of Communication and Information, Nanyang Technological University ³Microsoft Research Asia xiaosen@hust.edu.cn, xhe015@e.ntu.edu.sg, jingdw@microsoft.com, brooklet60@hust.edu.cn

Introduction

Adversarial Examples are imperceptible from legitimate ones by adding tiny perturbations, but lead to **incorrect model prediction**.

Transferability: adversarial examples generated for one model can still fool other models, that enables black-box attacks in the realworld applications without any knowledge of target model.

Background: existing attacks (e.g. PGD, CW, etc.) have exhibited great effectiveness, but with **low transferability**.

Methodology

Existing input transformations are effective to improve the transferability, but they are all applied on single input image. Could we further improve the transferability by **incorporating the informa**tion from other categories?

Mixup improves the model generalization by interpolating two randomly sampled samples (x, y) and (x', y') with $\lambda \in [0, 1]$ as follows: $\tilde{x} = \lambda \cdot x + (1 - \lambda) \cdot x', \quad \tilde{y} = \lambda \cdot y + (1 - \lambda) \cdot y'.$

MI-FGSM 100.0 43.6 42.4 35.7 13.1 12.8 Mixup 71.8 44.2 41.1 39.0 13.5 13.4	6
	6
Attack Inc-v3 Inc-v4 IncRes-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4}	IncRe

Table 1: Evaluations on MI-FGSM and *Mixup* transformation.

However, directly applying *mixup* for gradient calculation improves the transferability of crafted adversaries slightly but **degrades the** attack performance significantly under white-box setting. To utilize the information of images from other category without harming the white-box attack performance, we propose *admix* operation that admixes two images in a master and slave manner. $\tilde{x} = \gamma \cdot x + \eta' \cdot x' = \gamma \cdot (x + \eta \cdot x').$

We further propose an *Admix* attack method to improve the attack transferability, which calculates the average gradient on a set of **admixed images** $\{\tilde{x}\}$ of the input x by changing the value of γ or picking the add-in image x' from different categories in Eq. (2).

$$\bar{g}_{t+1} = \frac{1}{m_1 \cdot m_2} \sum_{x' \in X'} \sum_{i=0}^{m_1 - 1} \nabla_{x_t^{adv}} J(\gamma_i \cdot (x_t^{adv} + \eta \cdot x'), y; \theta), \quad (3)$$

Both *admix* and *mixup* generate a mixed image from an image pair, x and x'. Here we summarize their differences as follows:

- **Different goal**: *Mixup* aims to improve the model generalization while *admix* aims to generate more transferable adversaries.
- **Different Strategy**: *Mixup* treats *x* and *x'* equally and also mixes the label while *admix* treats *x* as the primary component and combines a small portion of x', and maintains the label of x.
- **Different interpolated image**: *Mixup* linearly interpolates x and x' while *admix* does not have such constraint, leading to more diversed transformed images.

Admix: Enhancing the Transferability of Adversarial Attacks

(1)

(2)

Algorithm

Algorithm 1 The Admix Attack Algorithm

- **Input:** A classifier f with loss function J and a benign example x with ground-truth label y
- **Input:** The maximum perturbation ϵ , number of iterations T and decay factor μ
- **Input:** The number of admixed copies m_1 and sampled images m_2 , and the strength of sampled image η

Output: An adversarial example
$$x^{adv} \in \mathcal{B}_{\epsilon}(x)$$

1.
$$\alpha = e/T$$
, $g_0 = 0$, $g_0 = 0$,
2. for $t = 0 \rightarrow T - 1$ do:

- Randomly sample a set X' of m_2 images from another category
- Calculate the average gradient \bar{g}_{t+1} by Eq. (3) 4:
- Update the enhanced momentum g_t : 5:

$$g_{t+1} = \mu \cdot g_t + \frac{\bar{g}}{\|\bar{g}_t\|}$$

Update x_{t+1}^{adv} by applying the gradient sign: 6:

$$x_{t+1}^{adv} = x_t^{adv} + \alpha \cdot \operatorname{sig}$$

7: end for

8: return
$$x^{adv} = x_T^{adv}$$
.

Model	Attack	Inc-v3	Inc-v4	IncRes-v2	Res-101	Inc-v3 _{ens3}	Inc-v3 _{ens4}	IncRes-v2 _{ens}
	DIM	99.0*	64.3	60.9	53.2	19.9	18.3	9.3
	TIM	100.0*	48.8	43.6	39.5	24.8	21.3	13.2
Inc-v3	SIM	100.0*	69.4	67.3	62.7	32.5	30.7	17.3
	Admix	100.0*	82.6	80.9	75.2	39.0	39.2	19.2
	DIM	72.9	97.4*	65.1	56.5	20.2	21.1	11.6
	TIM	58.6	99.6 *	46.5	42.3	26.2	23.4	17.2
Inc-v4	SIM	80.6	99.6 *	74.2	68.8	47.8	44.8	29.1
	Admix	87.8	99.4*	83.2	78.0	55.9	50.4	33.7
	DIM	70.1	63.4	93.5*	58.7	30.9	23.9	17.7
	TIM	62.2	55.4	97.4*	50.5	32.8	27.6	23.3
IncRes-v2	2 SIM	84.7	81.1	99.0*	76.4	56.3	48.3	42.8
	Admix	89.9	87.5	99.1*	81.9	64.2	56.7	50.0
	DIM	75.8	69.5	70.0	98.0*	35.7	31.6	19.9
	TIM	59.3	52.1	51.8	99.3*	35.4	31.3	23.1
Res-101	SIM	75.2	68.9	69.0	99.7 *	43.7	38.5	26.3
	Admix	85.4	80.8	79.6	99.7*	51.0	45.3	30.9
2: Eva	aluatior	ns on y	variou	ıs single	e input	: transfo	rmation	based atta
odel	Attack	Inc-v3	3 Inc-v4	IncRes-v2	2 Res-101	Inc-v 3_{ens3}	Inc-v 3_{ens4}	IncRes-v 2_{ens}
c-v3	SI-DIM	98.9*	85.0	81.3	76.3	48.0	45.1	24.9
	Admix-DIM	[99.8 *		87.7	83.5	52.2	49.9	28.6
c-v4	SI-DIM	89.3	98.8*	85.6	79.9	58.4	55.2	39.3
	Admix-DIM		99.2*	89.7	85.2	62.4	60.3	39.7
cRes-v2	SI-DIM	87.9	85.1	97.5*	82.9	66.0	59.3	52.2
	Admix-DIM		88.4	98.0*	85.8	70.5	63.7	55.3
es-101	SI-DIM	87.9	83.4	84.0	98.6*		57.5	42.0
	Admix-DIM	[91.9	89.0	89.6	99.8 *	69.7	62.3	46.6
Tał	ole 3: E	valua	tions of	on the a	ttacks	integrat	ed with	DIM.
odel	Attack	Inc-v3	Inc-v4	IncRes-v2	Res-101	Inc-v3 _{ens3}	Inc-v3 _{ens4}	IncRes-v2 _{ens}
c-v3	SI-TIM	100.0*	71.8	68.6	62.2	48.2	47.4	31.3
C-V.J	Admix-TIM	100.0*	83.9	80.4	74.4	59.1	57.9	39.2
c-v4	SI-TIM	78.2	99.6*	71.9	66.1	58.6	55.4	45.1
	Admix-TIM	87.4	99.7 *	82.3	77.0	68.1	65.3	53.1
cRes-v2	SI-TIM	84.5	82.2	98.8*	77.4	71.6	64.7	61.0
	Admix-TIM	90.2	88.2	98.6*	83.9	78.4	73.6	70.0
es-101	SI-TIM	74.2	69.9	70.2	99.8 *	59.5	54.5	42.8
5-101	Admix-TIM	83.2	78.9	80.7	99.7*	67.0	62.5	52.8
Table 4: Evaluations on the attacks integrated with TIM .								

$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$	eniin	ents							
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$	Model	Attack	Inc-v3	Inc-v4	IncRes-v2	Res-101	Inc-v3 _{ens3}	Inc-v3 _{ens4}	IncRes-v2 _{ens}
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$		DIM	99.0*	64.3	60.9	53.2	19.9	18.3	9.3
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$		TIM	100.0*	48.8	43.6	39.5	24.8	21.3	13.2
$\begin{array}{ c c c c c c c c c c c c c c c c c c c$	Inc-v3	SIM	100.0*	69.4	67.3	62.7	32.5	30.7	17.3
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$		Admix	100.0*	82.6	80.9	75.2	39.0	39.2	19.2
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $		DIM	72.9	97.4*	65.1	56.5	20.2	21.1	11.6
Admix87.899.4*83.278.055.950.433.7IncDIM70.163.493.5*58.730.923.917.7IncRes-V2SIM84.757.497.4*50.532.827.623.3IncRes-V2SIM84.781.199.0*76.456.348.342.8Admix89.987.599.1*81.964.256.750.0Admix89.987.599.1*81.964.256.750.0Res-101SIM75.869.570.098.0*35.731.619.9TIM59.352.151.899.3*35.431.323.1Res-101SIM75.268.969.099.7*43.738.526.3Admix85.480.879.699.7*51.045.330.9CSI-DIM85.480.879.699.7*51.045.330.9CAdmixInc-v3Inc-v3Inc-v3Inc-v230.9Admix99.8*85.679.958.455.239.3acev4SI-DIM89.398.8*85.679.958.455.239.3acesv2SI-DIM89.989.699.8*66.557.542.0acev4SI-DIM89.989.699.8*65.357.542.0acev4SI-DIM87.988.486.663.557.542.0		TIM	58.6	99.6 *	46.5	42.3	26.2	23.4	17.2
$\begin{array}{ c c c c c c c c c c c c c c c c c c c$	Inc-v4	SIM	80.6	99.6 *	74.2	68.8	47.8	44.8	29.1
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$		Admix	87.8	99.4*	83.2	78.0	55.9	50.4	33.7
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $		DIM	70.1	63.4	93.5*	58.7	30.9	23.9	17.7
Admix 89.9 87.5 99.1* 81.9 64.2 56.7 50.0 Res-101 DIM 75.8 69.5 70.0 98.0* 35.7 31.6 19.9 Res-101 DIM 75.8 69.5 70.0 98.0* 35.7 31.6 19.9 Res-101 TIM 59.3 52.1 51.8 99.3* 35.4 31.3 23.1 Admix SIM 75.2 68.9 69.0 99.7* 51.0 45.3 30.9 2 Carcolanian Nattice Nattice		TIM	62.2	55.4	97.4*	50.5	32.8	27.6	23.3
$\begin{array}{ c c c c c c c c c c c c c c c c c c c$	IncRes-v	2 SIM	84.7	81.1	99.0*	76.4	56.3	48.3	42.8
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$		Admix	89.9	87.5	99.1 *	81.9	64.2	56.7	50.0
Res-101SIM Admix75.2 85.468.9 80.869.0 79.699.7*43.7 91.038.5 45.326.3 30.9c2: Evaluationson varioussingleinputtransformationbased a lnc-v3 enstlodelAttackInc-v3Inc-v4IncRes-v2Res-101Inc-v3 enstInc-v3 enstInc-v3 enstInc-v3 enstc <v3< td="">SI-DIM Admix-DIM98.9*85.081.3 87.776.348.045.124.9 28.6c-v4SI-DIM Admix-DIM99.8* 90.590.587.7 85.783.552.249.9 28.628.6c-v4SI-DIM Admix-DIM93.0 90.299.2*89.785.262.460.339.7cRes-v2SI-DIM Admix-DIM87.9 91.085.1 91.097.5*82.966.059.352.2es-101SI-DIM Admix-DIM87.9 91.983.484.0 98.698.6*63.557.542.0dec-v3SI-DIM Admix-TIM87.9 91.989.699.8*69.762.346.6Ce-v3SI-DIM Admix-TIM87.9 91.989.699.8*69.762.346.6Ce-v3SI-TIM Admix-TIM100.0*71.868.662.248.247.431.3ce-v4SI-TIM Admix-TIM71.868.662.248.247.431.3ce-v4SI-TIM Admix-TIM71.868.662.2<!--</td--><td></td><td>DIM</td><td>75.8</td><td>69.5</td><td>70.0</td><td>98.0*</td><td>35.7</td><td>31.6</td><td>19.9</td></v3<>		DIM	75.8	69.5	70.0	98.0*	35.7	31.6	19.9
Admix 85.4 80.8 79.6 99.7* 51.0 45.3 30.9 2 2: Evaluations on various single input transformation based a odel Attack Inc-v3 Inc-v4 IncRes-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _e c-v3 SI-DIM 98.9* 85.0 81.3 76.3 48.0 45.1 24.9 c-v4 Admix-DIM 99.8* 90.5 87.7 83.5 52.2 49.9 28.6 c-v4 Admix-DIM 93.0 99.2* 89.7 85.2 62.4 60.3 39.7 cRes-v2 SI-DIM 87.9 85.1 97.5* 82.9 66.0 59.3 52.2 cRes-v2 SI-DIM 87.9 83.4 84.0 98.6* 63.5 57.5 42.0 cRes-v1 Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 cev4 Attack Inc-v3		TIM	59.3	52.1	51.8	99.3*	35.4	31.3	23.1
2: Evaluations on various single input transformation based a odelAttackInc-v3Inc-v4IncRes-v2Res-101Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2,c-v3SI-DIM98.9*85.081.376.348.045.124.9c-v4Admix-DIM99.8*90.587.783.552.249.928.6c-v4SI-DIM89.398.8*85.679.958.455.239.3c-v4Admix-DIM93.099.2*89.785.262.460.339.7cc-v4SI-DIM87.985.197.5*82.966.059.352.2admix-DIM90.288.498.0*85.870.563.755.3cc-v1SI-DIM87.983.484.098.6*63.557.542.0admix-DIM91.989.089.699.8*69.762.346.6Table 3: Evaluations on the attacks integrated with DIM.Ice-v3SI-TIM100.0*71.868.662.248.247.431.3ac-v4Admix-TIM100.0*83.980.474.459.157.939.2ac-v3SI-TIM74.299.7*82.377.068.165.353.1ac-v4Admix-TIM87.499.7*82.377.068.165.353.1ac-v4SI-TIM74.269.970.299.8*59.554.542.8ac-v4 <td>Res-101</td> <td>SIM</td> <td>75.2</td> <td>68.9</td> <td>69.0</td> <td>99.7*</td> <td>43.7</td> <td>38.5</td> <td>26.3</td>	Res-101	SIM	75.2	68.9	69.0	99.7 *	43.7	38.5	26.3
odelAttackInc-v3Inc-v4IncRes-v2Res-101Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _e uc-v3SI-DIM98.9*85.081.376.348.045.124.9uc-v4Admix-DIM99.8*90.587.783.552.249.928.6uc-v4SI-DIM89.398.8*85.679.958.455.239.3uc-v4Admix-DIM93.099.2*89.785.262.460.339.7ucRes-v2SI-DIM87.985.197.5*82.966.059.352.2Admix-DIM90.288.498.0*85.870.563.755.3es-101SI-DIM87.983.484.098.6*63.557.542.0Admix-DIM91.989.089.699.8*69.762.346.6Table 3: Evaluations on the attacks integrated with DIM.IodelAttackInc-v3Inc-v4IncRes-v2Res-101Inc-v3 _{ens4} IncRes-v2_euc-v3SI-TIM100.0*71.868.662.248.247.431.3uc-v4Admix-TIM78.299.6*71.966.158.655.445.1uc-v4SI-TIM78.299.6*71.966.158.655.445.1uc-v4SI-TIM87.499.7*82.377.068.165.353.1uc-v4Admix-TIM87.499.7*82.377.0 <td< td=""><td></td><td>Admix</td><td>85.4</td><td>80.8</td><td>79.6</td><td>99.7*</td><td>51.0</td><td>45.3</td><td>30.9</td></td<>		Admix	85.4	80.8	79.6	99.7 *	51.0	45.3	30.9
Ic-v3SI-DIM Admix-DIM98.9* 99.8*85.081.376.348.045.124.9Ic-v4Admix-DIM Admix-DIM99.8*90.587.783.552.249.928.6Ic-v4SI-DIM Admix-DIM89.398.8*85.679.958.455.239.3Ic-v4Admix-DIM Admix-DIM93.099.2*89.785.262.460.339.7IcRes-v2SI-DIM Admix-DIM87.985.197.5*82.966.059.352.2es-101SI-DIM Admix-DIM91.983.484.098.6*63.557.542.0es-101SI-DIM Admix-DIM91.989.089.699.8*69.762.346.6Table 3: Evaluations on the attacks integrated with DIM.Ice-v3SI-TIM Admix-TIM100.0*71.868.662.248.247.431.3Ice-v4SI-TIM Admix-TIM78.299.6*71.966.158.655.445.1Ice-v4SI-TIM Admix-TIM87.499.7*82.377.068.165.353.1IncRes-v2SI-TIM Admix-TIM84.582.298.6*83.978.473.670.0es-101SI-TIM Admix-TIM74.269.970.299.8*59.554.542.8es-101SI-TIM Admix-TIM83.278.980.799.7*67.062.552.8	e 2: Ev	aluatior	ns on	variou	as single	e inpu ⁻	t transfo	rmation	based att
Admix-DIM 99.8* 90.5 87.7 83.5 52.2 49.9 28.6 .c-v4 SI-DIM 89.3 98.8* 85.6 79.9 58.4 55.2 39.3 .cc-v4 Admix-DIM 93.0 99.2* 89.7 85.2 62.4 60.3 39.7 .ccRes-v2 SI-DIM 87.9 85.1 97.5* 82.9 66.0 59.3 52.2 Admix-DIM 90.2 88.4 98.0* 85.8 70.5 63.7 55.3 es-101 SI-DIM 87.9 83.4 84.0 98.6* 69.7 62.3 46.6 admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 acte-101 Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 acte-v2 Inc-v3 Inc-v4 IncRes-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _e ac-v3 SI-TIM 100.0*	lodel	Attack	Inc-v3	3 Inc-v4	IncRes-v2	2 Res-10	1 Inc-v 3_{ens}	$_{3}$ Inc-v 3_{ens}	IncRes-v2 $_{ens}$
Admix-DIM 99.8* 90.5 87.7 83.5 52.2 49.9 28.6 c-v4 SI-DIM 89.3 98.8* 85.6 79.9 58.4 55.2 39.3 c-v4 Admix-DIM 93.0 99.2* 89.7 85.2 62.4 60.3 39.7 cRes-v2 SI-DIM 87.9 85.1 97.5* 82.9 66.0 59.3 52.2 Admix-DIM 90.2 88.4 98.0* 85.8 70.5 63.7 55.3 es-101 SI-DIM 87.9 83.4 84.0 98.6* 63.5 57.5 42.0 des-101 Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 Table 3: Evaluations on the attacks integrated with DIM. ice-v3 SI-TIM 100.0* 71.8 68.6 62.2 48.2 47.4 31.3 ice-v4 Admix-TIM 100.0* 83.9 80.4 74.4 59.1 57.9	c w3	SI-DIM	98.9*	* 85.0	81.3	76.3	48.0	45.1	24.9
Admix-DIM 93.0 99.2* 89.7 85.2 62.4 60.3 39.7 cRes-v2 SI-DIM 87.9 85.1 97.5* 82.9 66.0 59.3 52.2 Admix-DIM 90.2 88.4 98.0* 85.8 70.5 63.7 55.3 cs-101 SI-DIM 87.9 83.4 84.0 98.6* 63.5 57.5 42.0 dss-101 SI-DIM 87.9 83.4 84.0 98.6* 63.5 57.5 42.0 dss-101 Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 Table 3: Evaluations on the attacks integrated with DIM. Inc-v3 Inc-v4 Inc-v3 Inc-v4 Inc-v3 Inc-v4 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2e nc-v3 SI-TIM 100.0* 71.8 68.6 62.2 48.2 47.4 31.3 nc-v4 Admix-TIM 78.2 99.6* 71.9 66.1 58.6	C- VJ	Admix-DIM	99.8 [°]	* 90.5	87.7	83.5	52.2	49.9	28.6
Admix-DIM 93.0 99.2* 89.7 85.2 62.4 60.3 39.7 ccRes-v2 SI-DIM 87.9 85.1 97.5* 82.9 66.0 59.3 52.2 Admix-DIM 90.2 88.4 98.0* 85.8 70.5 63.7 55.3 es-101 SI-DIM 87.9 83.4 84.0 98.6* 63.5 57.5 42.0 Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 Table 3: Evaluations on the attacks integrated with DIM. Iodel Attack Inc-v3 Inc-v4 IncRes-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _e ic-v3 SI-TIM 100.0* 71.8 68.6 62.2 48.2 47.4 31.3 ic-v4 Admix-TIM 100.0* 83.9 80.4 74.4 59.1 57.9 39.2 ic-v4 Admix-TIM 78.2 99.6* 71.9 66.1 58.6 55.4 45.1 ic-v4 Admix-TIM 87.4 99.7* 82.3<		SI-DIM	89.3	98.8*	85.6	79.9	58.4	55.2	39.3
Admix-DIM 90.2 88.4 98.0* 85.8 70.5 63.7 55.3 es-101 SI-DIM Admix-DIM 87.9 83.4 84.0 98.6* 63.5 57.5 42.0 Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 Table 3: Evaluations on the attacks integrated with DIM. Inc-v3 Inc-v4 Inc-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _e Ic-v3 SI-TIM 100.0* 71.8 68.6 62.2 48.2 47.4 31.3 Admix-TIM 100.0* 83.9 80.4 74.4 59.1 57.9 39.2 nc-v4 SI-TIM 100.0* 83.9 80.4 74.4 59.1 57.9 39.2 nc-v4 SI-TIM 78.2 99.6* 71.9 66.1 58.6 55.4 45.1 nc-v4 Admix-TIM 87.4 99.7* 82.3 77.0 68.1 65.3 53.1 ncRes-v2		Admix-DIM	93.0	99.2 *	89.7	85.2	62.4	60.3	39.7
Admix-DIM 90.2 88.4 98.0* 85.8 70.5 63.7 55.3 es-101 SI-DIM 87.9 83.4 84.0 98.6* 63.5 57.5 42.0 Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 Table 3: Evaluations on the attacks integrated with DIM. Incrv3 Inc-v4 IncRes-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _e hc-v3 SI-TIM 100.0* 71.8 68.6 62.2 48.2 47.4 31.3 hc-v3 SI-TIM 100.0* 83.9 80.4 74.4 59.1 57.9 39.2 hc-v4 Admix-TIM 78.2 99.6* 71.9 66.1 58.6 55.4 45.1 hc-v4 SI-TIM 78.2 99.6* 71.9 66.1 58.6 55.4 45.1 hc-v4 Admix-TIM 87.4 99.7* 82.3 77.0 68.1 65.3 53.1 hc-	R_{0}	SI-DIM	87.9	85.1	97.5*	82.9	66.0	59.3	52.2
Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 Table 3: Evaluations on the attacks integrated with DIM. Iodel Attack Inc-v3 Inc-v4 IncRes-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _e Iodel Attack Inc-v3 Inc-v4 IncRes-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _e Inc-v3 SI-TIM 100.0* 71.8 68.6 62.2 48.2 47.4 31.3 Inc-v3 Admix-TIM 100.0* 83.9 80.4 74.4 59.1 57.9 39.2 Inc-v4 SI-TIM 78.2 99.6* 71.9 66.1 58.6 55.4 45.1 Inc-v4 Admix-TIM 87.4 99.7* 82.3 77.0 68.1 65.3 53.1 IncRes-v2 SI-TIM 84.5 82.2 98.6* 83.9 78.4 73.6 70.0 IncRes-v2 SI-TIM 74.2 69.9	ICINES-VZ	Admix-DIM	90.2	88.4	98.0*	85.8	70.5	63.7	55.3
Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 Table 3: Evaluations on the attacks integrated with DIM. Iodel Attack Inc-v3 Inc-v4 IncRes-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _e nc-v3 SI-TIM 100.0* 71.8 68.6 62.2 48.2 47.4 31.3 nc-v3 SI-TIM 100.0* 83.9 80.4 74.4 59.1 57.9 39.2 nc-v4 SI-TIM 78.2 99.6* 71.9 66.1 58.6 55.4 45.1 nc-v4 SI-TIM 87.4 99.7* 82.3 77.0 68.1 65.3 53.1 ncRes-v2 SI-TIM 84.5 82.2 98.8* 77.4 71.6 64.7 61.0 ncRes-v2 SI-TIM 90.2 88.2 98.6* 83.9 78.4 73.6 70.0 es-101 SI-TIM 74.2 69.9 70.2 99.8* 59.5 54.5 42.8 Admix-TIM 83.2 78.9	os_101	SI-DIM	87.9	83.4	84.0	98.6*	63.5	57.5	42.0
IodelAttackInc-v3Inc-v4IncRes-v2Res-101Inc-v3_{ens3}Inc-v3_{ens4}IncRes-v2_enc-v3SI-TIM100.0*71.868.662.248.247.431.3 $Admix$ -TIM100.0*83.980.474.459.157.939.2nc-v4SI-TIM78.299.6*71.966.158.655.445.1 $Admix$ -TIM87.499.7*82.377.068.165.353.1 $nc-v4$ SI-TIM84.582.298.8*77.471.664.761.0 $ncRes-v2$ SI-TIM90.288.298.6*83.978.473.670.0 $es-101$ SI-TIM74.269.970.299.8*59.554.542.8 $Admix$ -TIM83.278.980.799.7*67.062.552.8		Admix-DIM	91.9	89.0	89.6	99.8*	69.7	62.3	46.6
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	Ta	ble 3: Ev	valua	tions	on the a	ttacks	integrat	ed with	DIM.
Admix-TIM 100.0* 83.9 80.4 74.4 59.1 57.9 39.2 nc-v4 SI-TIM 78.2 99.6* 71.9 66.1 58.6 55.4 45.1 Admix-TIM 87.4 99.7* 82.3 77.0 68.1 65.3 53.1 ncRes-v2 SI-TIM 84.5 82.2 98.8* 77.4 71.6 64.7 61.0 ncRes-v2 SI-TIM 84.5 82.2 98.6* 83.9 78.4 73.6 70.0 es-101 SI-TIM 74.2 69.9 70.2 99.8* 59.5 54.5 42.8 Admix-TIM 83.2 78.9 80.7 99.7* 67.0 62.5 52.8	lodel	Attack	Inc-v3	3 Inc-v4	IncRes-v2	2 Res-102	1 Inc-v 3_{ens3}	Inc-v 3_{ens4}	IncRes-v2 _{ens}
Admix-IIM 100.0* 83.9 80.4 74.4 59.1 57.9 39.2 nc-v4 SI-TIM 78.2 99.6* 71.9 66.1 58.6 55.4 45.1 Admix-TIM 87.4 99.7* 82.3 77.0 68.1 65.3 53.1 nc-v4 Admix-TIM 84.5 82.2 98.8* 77.4 71.6 64.7 61.0 ncRes-v2 SI-TIM 84.5 82.2 98.6* 83.9 78.4 73.6 70.0 es-101 SI-TIM 74.2 69.9 70.2 99.8* 59.5 54.5 42.8 Admix-TIM 83.2 78.9 80.7 99.7* 67.0 62.5 52.8	0.0-372	SI-TIM	100.0*	71.8	68.6	62.2	48.2	47.4	31.3
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $		Admix-TIM	100.0*	83.9	80.4	74.4	59.1	57.9	39.2
Admix-TIM 87.4 99.7* 82.3 77.0 68.1 65.3 53.1 ncRes-v2 SI-TIM 84.5 82.2 98.8* 77.4 71.6 64.7 61.0 Admix-TIM 90.2 88.2 98.6* 83.9 78.4 73.6 70.0 es-101 SI-TIM 74.2 69.9 70.2 99.8* 59.5 54.5 42.8 Admix-TIM 83.2 78.9 80.7 99.7* 67.0 62.5 52.8		SI-TIM	78.2	99.6*	71.9	66.1	58.6	55.4	45.1
Admix-TIM90.288.298.6*83.978.473.670.0es-101SI-TIM74.269.970.299.8*59.554.542.8Admix-TIM83.278.980.799.7*67.062.552.8	IC-V4	<i>Admix-</i> TIM	87.4	99.7 *	82.3	77.0	68.1	65.3	53.1
Admix-TIM 90.2 88.2 98.6* 83.9 78.4 73.6 70.0 es-101 SI-TIM 74.2 69.9 70.2 99.8* 59.5 54.5 42.8 Admix-TIM 83.2 78.9 80.7 99.7* 67.0 62.5 52.8	R_{0} π^{0}	SI-TIM	84.5	82.2	98.8*	77.4	71.6	64.7	61.0
es-101 <i>Admix</i> -TIM 83.2 78.9 80.7 99.7* 67.0 62.5 52.8		Admix-TIM	90.2	88.2	98.6*	83.9	78.4	73.6	70.0
Admix-TIM 83.2 78.9 80.7 99.7* 67.0 62.5 52.8	00 101	SI-TIM	74.2	69.9	70.2	99.8*	59.5	54.5	42.8
Table 4: Evaluations on the attacks integrated with TIM.	es-101	Admix-TIM	83.2	78.9	80.7	99.7*	67.0	62.5	52.8
\sim									

$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$	erim	ents							
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$	Model	Attack	Inc-v3	Inc-v4	IncRes-v2	Res-101	Inc-v3 _{ens3}	Inc-v3 _{ens4}	IncRes-v2 _{ens}
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$		DIM	99.0*	64.3	60.9	53.2	19.9	18.3	9.3
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$		TIM	100.0*	48.8	43.6	39.5	24.8	21.3	13.2
$\begin{array}{ c c c c c c c c c c c c c c c c c c c$	Inc-v3	SIM	100.0*	69.4	67.3	62.7	32.5	30.7	17.3
$\begin{array}{ c c c c c c c c c c c c c c c c c c c$		Admix	100.0*	82.6	80.9	75.2	39.0	39.2	19.2
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$		DIM	72.9	97.4*	65.1	56.5	20.2	21.1	11.6
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$		TIM	58.6	99.6 *	46.5	42.3	26.2	23.4	17.2
$\begin{array}{ c c c c c c c c c c c c c c c c c c c$	Inc-v4	SIM	80.6	99.6 *	74.2	68.8	47.8	44.8	29.1
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$		Admix	87.8	99.4*	83.2	78.0	55.9	50.4	33.7
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$		DIM	70.1	63.4	93.5*	58.7	30.9	23.9	17.7
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$		TIM	62.2	55.4	97.4*	50.5	32.8	27.6	23.3
$\begin{array}{ c c c c c c c c c c c c c c c c c c c$	IncRes-v	2 SIM	84.7	81.1	99.0*	76.4	56.3	48.3	42.8
$\begin{array}{ c c c c c c c c c c c c c c c c c c c$		Admix	89.9	87.5	99.1 *	81.9	64.2	56.7	50.0
Res-101SIM Admix75.268.969.099.7*43.738.526.3admix85.480.879.699.7*51.045.330.9e 2: Evaluations on various single input transformation based attModelAttackInc-v3Inc-v4IncRes-v2Res-101Inc-v3_{ens3}Inc-v3_{ens4}IncRes-v2_{ens}nc-v3Admix-DIM98.9*85.081.376.348.045.124.9admix-DIM99.8*90.587.783.552.249.928.6nc-v4SI-DIM89.398.8*85.679.958.455.239.3admix-DIM93.099.2*89.785.262.460.339.7ncRes-v2SI-DIM87.985.197.5*82.966.059.352.2Admix-DIM90.288.498.0*85.870.563.755.3Res-101SI-DIM87.983.484.098.6*63.557.542.0Admix-DIM91.989.089.699.8*69.762.346.6Table 3: Evaluations on the attacks integrated with DIM.ModelAttackInc-v3Inc-v4IncRes-v2Res-101Inc-v3_{ens3}Inc-v3_{ens4}IncRes-v2_{ens}nc-v3SI-TIM100.0*71.868.662.248.247.431.3nc-v4Admix-TIM90.288.980.474.459.157.939.2		DIM	75.8	69.5	70.0	98.0*	35.7	31.6	19.9
Admix 85.4 80.8 79.6 99.7* 51.0 45.3 30.9 e 2: Evaluations on various single input transformation based att Model Attack Inc-v3 Inc-v4 IncRes-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _{ens} nc-v3 SI-DIM 98.9* 85.0 81.3 76.3 48.0 45.1 24.9 Admix-DIM 98.9* 85.0 81.3 76.3 48.0 45.1 24.9 nc-v3 SI-DIM 98.9* 90.5 87.7 83.5 52.2 49.9 28.6 nc-v4 Admix-DIM 93.9 98.8* 85.6 79.9 58.4 55.2 39.3 ncRes-v2 SI-DIM 87.9 85.1 97.5* 82.9 66.0 59.3 52.2 Admix-DIM 90.2 88.4 98.0* 85.8 70.5 63.7 55.3 Res-101 SI-DIM 87.9 83.4 84.0 98.6* 63.5 57.5		TIM	59.3	52.1	51.8	99.3*	35.4	31.3	23.1
e 2: Evaluations on various single input transformation based attAddelAttackInc-v3Inc-v4IncRes-v2Res-101Inc-v3Inc-v3Inc-v3IncRes-v2Inc-v3nc-v3SI-DIM98.9*85.081.376.348.045.124.9Admix-DIM99.8*90.587.783.552.249.928.6nc-v4SI-DIM89.398.8*85.679.958.455.239.3nc-v4SI-DIM89.398.8*85.679.958.460.339.7nc-v4Admix-DIM93.099.2*89.785.262.460.339.7ncRes-v2Admix-DIM90.288.498.0*85.870.563.755.3Res-101SI-DIM87.983.484.098.6*63.557.542.0Admix-DIM91.989.089.699.8*69.762.346.6Table 3: Evaluations on the attacks integrated with DIM.ModelAttackInc-v3Inc-v4IncRes-v248.247.431.3nc-v3SI-TIM100.0*71.868.662.248.247.431.3nc-v4Admix-TIM100.0*83.980.474.459.157.939.2nc-v3SI-TIM78.299.6*71.966.158.655.445.1nc-v4Admix-TIM84.582.298.8*77.471.664.761.0 <td>Res-101</td> <td>SIM</td> <td>75.2</td> <td>68.9</td> <td>69.0</td> <td>99.7*</td> <td>43.7</td> <td>38.5</td> <td>26.3</td>	Res-101	SIM	75.2	68.9	69.0	99.7*	43.7	38.5	26.3
AodelAttackInc-v3Inc-v4IncRes-v2Res-101Inc-v3_{ens3}Inc-v3_{ens4}IncRes-v2_{ens} $nc-v3$ SI-DIM98.9*85.081.376.348.045.124.9 $nc-v3$ Admix-DIM99.8*90.587.783.552.249.928.6 $nc-v4$ SI-DIM89.398.8*85.679.958.455.239.3 $nc-v4$ Admix-DIM93.099.2*89.785.262.460.339.7 $ncRes-v2$ SI-DIM87.985.197.5*82.966.059.352.2 $ncRes-v2$ SI-DIM87.983.498.0*85.870.563.755.3 $admix-DIM$ 90.288.498.0*85.870.563.755.3 $acs-101$ SI-DIM87.983.484.098.6*63.557.542.0 $Admix-DIM$ 91.989.089.699.8*69.762.346.6Table 3: Evaluations on the attacks integrated with DIM.ModelAttackInc-v3Inc-v4IncRes-v248.247.431.3 $nc-v3$ SI-TIM100.0*71.868.662.248.247.431.3 $nc-v4$ Admix-TIM78.299.6*71.966.158.655.445.1 $nc-v4$ SI-TIM78.299.7*82.377.068.165.353.1 $nc-v4$ SI-TIM87.499.7*82.3<		Admix	85.4	80.8	79.6	99.7 *	51.0	45.3	30.9
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	e 2: Ev	aluation	ns on T	variou	ıs single	e input	t transfo	rmation	based att
Admix-DIM 99.8* 90.5 87.7 83.5 52.2 49.9 28.6 nc-v4 SI-DIM 89.3 98.8* 85.6 79.9 58.4 55.2 39.3 nc-v4 Admix-DIM 93.0 99.2* 89.7 85.2 62.4 60.3 39.7 ncRes-v2 SI-DIM 87.9 85.1 97.5* 82.9 66.0 59.3 52.2 Admix-DIM 90.2 88.4 98.0* 85.8 70.5 63.7 55.3 Admix-DIM 90.2 88.4 98.0* 98.6* 63.5 57.5 42.0 Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 Table 3: Evaluations on the attacks integrated with DIM. Model Attack Inc-v3 Inc-v4 IncRes-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _{ens} nc-v3 SI-TIM 100.0* 71.8 68.6 62.2 48.2 47.4 31.3<	Aodel	Attack	Inc-v3	3 Inc-v4	IncRes-v2	2 Res-101	Inc-v3 _{ens}	$_{3}$ Inc-v3 _{ens4}	IncRes-v2 _{ens}
Admix-DIM 99.8* 90.5 87.7 83.5 52.2 49.9 28.6 nc-v4 SI-DIM 89.3 98.8* 85.6 79.9 58.4 55.2 39.3 Admix-DIM 93.0 99.2* 89.7 85.2 62.4 60.3 39.7 ncRes-v2 SI-DIM 87.9 85.1 97.5* 82.9 66.0 59.3 52.2 Admix-DIM 90.2 88.4 98.0* 85.8 70.5 63.7 55.3 ces-101 SI-DIM 87.9 83.4 84.0 98.6* 63.5 57.5 42.0 Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 Table 3: Evaluations on the attacks integrated with DIM. Admix-TIM 100.0* 71.8 68.6 62.2 48.2 47.4 31.3 nc-v3 SI-TIM 100.0* 83.9 80.4 74.4 59.1 57.9 39.2 nc-v4 <td< td=""><td>n_{C-V}</td><td>SI-DIM</td><td>98.9*</td><td>85.0</td><td>81.3</td><td>76.3</td><td>48.0</td><td>45.1</td><td>24.9</td></td<>	n_{C-V}	SI-DIM	98.9*	85.0	81.3	76.3	48.0	45.1	24.9
Admix-DIM 93.0 99.2* 89.7 85.2 62.4 60.3 39.7 ncRes-v2 SI-DIM 87.9 85.1 97.5* 82.9 66.0 59.3 52.2 Admix-DIM 90.2 88.4 98.0* 85.8 70.5 63.7 55.3 Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 Case-101 Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 Model Attack Inc-v3 Inc-v4 IncRes-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _{ens} nc-v3 SI-TIM 100.0* 71.8 68.6 62.2 48.2 47.4 31.3 nc-v4 SI-TIM 78.2 99.6* 71.9 66.1 58.6 55.4 45.1 nc-v4 SI-TIM 77.4 <t< td=""><td></td><td>Admix-DIM</td><td>1 99.8*</td><td>90.5</td><td>87.7</td><td>83.5</td><td>52.2</td><td>49.9</td><td>28.6</td></t<>		Admix-DIM	1 99.8 *	90.5	87.7	83.5	52.2	49.9	28.6
Admix-DIM 93.0 99.2* 89.7 85.2 62.4 60.3 39.7 ncRes-v2 SI-DIM 87.9 85.1 97.5* 82.9 66.0 59.3 52.2 Admix-DIM 90.2 88.4 98.0* 85.8 70.5 63.7 55.3 Admix-DIM 90.2 88.4 98.0* 85.8 70.5 63.7 55.3 Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 Table 3: Evaluations on the attacks integrated with DIM. Model Attack Inc-v3 Inc-v4 IncRes-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _{ens} nc-v3 SI-TIM 100.0* 71.8 68.6 62.2 48.2 47.4 31.3 nc-v3 SI-TIM 100.0* 83.9 80.4 74.4 59.1 57.9 39.2 nc-v4 Admix-TIM 87.4 99.7* 82.3 77.0 68.1 65.3 53.1<	nc-v/	SI-DIM	89.3	98.8*	85.6	79.9	58.4	55.2	39.3
Admix-DIM 90.2 88.4 98.0* 85.8 70.5 63.7 55.3 des-101 SI-DIM 87.9 83.4 84.0 98.6* 63.5 57.5 42.0 Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 Table 3: Evaluations on the attacks integrated with DIM. Model Attack Inc-v3 Inc-v4 IncRes-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _{ens} nc-v3 SI-TIM 100.0* 71.8 68.6 62.2 48.2 47.4 31.3 nc-v3 SI-TIM 100.0* 83.9 80.4 74.4 59.1 57.9 39.2 nc-v4 Admix-TIM 87.4 99.7* 82.3 77.0 68.1 65.3 53.1 ncRes-v2 SI-TIM 84.5 82.2 98.8* 77.4 71.6 64.7 61.0 ncRes-v2 SI-TIM 90.2 88.2 98.6* 83.9 78.4		Admix-DIM	1 93.0	99.2 *	89.7	85.2	62.4	60.3	39.7
Admix-DIM 90.2 88.4 98.0* 85.8 70.5 63.7 55.3 Res-101 SI-DIM 87.9 83.4 84.0 98.6* 63.5 57.5 42.0 Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 Table 3: Evaluations on the attacks integrated with DIM. Incrv3 Inc-v3 Inc-v3 Inc-v3 Inc-v4 Incres-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _{ens} nc-v3 SI-TIM 100.0* 71.8 68.6 62.2 48.2 47.4 31.3 nc-v4 SI-TIM 100.0* 83.9 80.4 74.4 59.1 57.9 39.2 nc-v4 SI-TIM 78.2 99.6* 71.9 66.1 58.6 55.4 45.1 nc-v4 Admix-TIM 87.4 99.7* 82.3 77.0 68.1 65.3 53.1 ncRes-v2 Admix-TIM 90.2 88.2 98.6* 83.9 78.4	$n_{\rm CRos-y2}$	SI-DIM	87.9	85.1	97.5*	82.9	66.0	59.3	52.2
Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 Table 3: Evaluations on the attacks integrated with DIM. Model Attack Inc-v3 Inc-v4 IncRes-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _{ens} Model Attack Inc-v3 Inc-v4 IncRes-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _{ens} Model Attack Inc-v3 Inc-v4 IncRes-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _{ens} Model Attack Inc-v3 Inc-v4 IncRes-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _{ens} Model SI-TIM 100.0* 83.9 80.4 74.4 59.1 57.9 39.2 Mc-v4 SI-TIM 78.2 99.6* 71.9 66.1 58.6 55.4 45.1 Mc-v4 SI-TIM 87.4 99.7* 82.3 77.0 68.1 65.3 53.1 McRes-v2	IICINE5-VZ	Admix-DIM	1 90.2	88.4	98.0 *	85.8	70.5	63.7	55.3
Admix-DIM 91.9 89.0 89.6 99.8* 69.7 62.3 46.6 Table 3: Evaluations on the attacks integrated with DIM. Model Attack Inc-v3 Inc-v4 IncRes-v2 Res-101 Inc-v3 _{ens3} Inc-v3 _{ens4} IncRes-v2 _{ens} nc-v3 SI-TIM 100.0* 71.8 68.6 62.2 48.2 47.4 31.3 nc-v3 SI-TIM 100.0* 83.9 80.4 74.4 59.1 57.9 39.2 nc-v4 SI-TIM 78.2 99.6* 71.9 66.1 58.6 55.4 45.1 nc-v4 SI-TIM 87.4 99.7* 82.3 77.0 68.1 65.3 53.1 ncRes-v2 SI-TIM 84.5 82.2 98.8* 77.4 71.6 64.7 61.0 ncRes-v2 SI-TIM 84.5 82.2 98.6* 83.9 78.4 73.6 70.0 Res-101 SI-TIM 74.2 69.9 70.2 99.8* 59.5 54.5 42.8 Admix-TIM 83.2 <td>205-101</td> <td>SI-DIM</td> <td>87.9</td> <td>83.4</td> <td>84.0</td> <td>98.6*</td> <td>63.5</td> <td>57.5</td> <td>42.0</td>	205-101	SI-DIM	87.9	83.4	84.0	98.6*	63.5	57.5	42.0
ModelAttackInc-v3Inc-v4IncRes-v2Res-101Inc-v3 $_{ens3}$ Inc-v3 $_{ens4}$ IncRes-v2 $_{ens}$ nc-v3SI-TIM100.0*71.868.662.248.247.431.3Admix-TIM100.0*83.980.474.459.157.939.2nc-v4SI-TIM78.299.6*71.966.158.655.445.1Admix-TIM87.499.7*82.377.068.165.353.1nc-v4SI-TIM84.582.298.8*77.471.664.761.0ncRes-v2SI-TIM84.582.298.6*83.978.473.670.0Res-101SI-TIM74.269.970.299.8*59.554.542.8Admix-TIM83.278.980.799.7*67.062.552.8	C5 101	Admix-DIM	1 91.9	89.0	89.6	99.8*	69.7	62.3	46.6
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	Ta	ble 3: E	valua	tions	on the a	ttacks	integrat	ed with	DIM.
nc-v3 Admix-TIM 100.0* 83.9 80.4 74.4 59.1 57.9 39.2 nc-v4 SI-TIM 78.2 99.6* 71.9 66.1 58.6 55.4 45.1 Admix-TIM 87.4 99.7* 82.3 77.0 68.1 65.3 53.1 nc-v4 SI-TIM 84.5 82.2 98.8* 77.4 71.6 64.7 61.0 ncRes-v2 SI-TIM 84.5 82.2 98.6* 83.9 78.4 73.6 70.0 Res-101 SI-TIM 74.2 69.9 70.2 99.8* 59.5 54.5 42.8 Admix-TIM 83.2 78.9 80.7 99.7* 67.0 62.5 52.8	Model	Attack	Inc-v3	Inc-v4	IncRes-v2	2 Res-101	Inc-v 3_{ens}	Inc-v 3_{ens4}	IncRes-v2 _{ens}
Admix-TIM 100.0* 83.9 80.4 74.4 59.1 57.9 39.2 nc-v4 SI-TIM 78.2 99.6* 71.9 66.1 58.6 55.4 45.1 Admix-TIM 87.4 99.7* 82.3 77.0 68.1 65.3 53.1 ncRes-v2 SI-TIM 84.5 82.2 98.8* 77.4 71.6 64.7 61.0 Admix-TIM 90.2 88.2 98.6* 83.9 78.4 73.6 70.0 Res-101 SI-TIM 74.2 69.9 70.2 99.8* 59.5 54.5 42.8 Admix-TIM 83.2 78.9 80.7 99.7* 67.0 62.5 52.8	n_{c-w}	SI-TIM	100.0*	71.8	68.6	62.2	48.2	47.4	31.3
nc-v4Admix-TIM87.499.7*82.377.068.165.353.1ncRes-v2SI-TIM84.582.298.8*77.471.664.761.0Admix-TIM90.288.298.6*83.978.473.670.0Res-101SI-TIM74.269.970.299.8*59.554.542.8Admix-TIM83.278.980.799.7*67.062.552.8		Admix-TIM	[100.0*	83.9	80.4	74.4	59.1	57.9	39.2
Admix-IIM 87.4 99.7* 82.3 77.0 68.1 65.3 53.1 ncRes-v2 SI-TIM 84.5 82.2 98.8* 77.4 71.6 64.7 61.0 Admix-TIM 90.2 88.2 98.6* 83.9 78.4 73.6 70.0 Res-101 SI-TIM 74.2 69.9 70.2 99.8* 59.5 54.5 42.8 Admix-TIM 83.2 78.9 80.7 99.7* 67.0 62.5 52.8	nc-v/	SI-TIM	78.2	99.6*	71.9	66.1	58.6	55.4	45.1
Admix-TIM90.288.298.6*83.978.473.670.0Res-101SI-TIM74.269.970.299.8*59.554.542.8Admix-TIM83.278.980.799.7*67.062.552.8	LIC-V4	Admix-TIM	[87.4	99.7*	82.3	77.0	68.1	65.3	53.1
Admix-IIM 90.2 88.2 98.6* 83.9 78.4 73.6 70.0 Res-101 SI-TIM 74.2 69.9 70.2 99.8* 59.5 54.5 42.8 Admix-TIM 83.2 78.9 80.7 99.7* 67.0 62.5 52.8	$n_{\rm c}R_{\rm os}$ $v^{\rm 2}$	SI-TIM	84.5	82.2	98.8*	77.4	71.6	64.7	61.0
Admix-TIM 83.2 78.9 80.7 99.7* 67.0 62.5 52.8		Admix-TIM	[90.2	88.2	98.6*	83.9	78.4	73.6	70.0
Admix-11M 83.2 78.9 80.7 99.7* 67.0 62.5 52.8	P_{00-101}	SI-TIM	74.2	69.9	70.2	99.8*	59.5	54.5	42.8
	NC2-101	Admix-TIM	[83.2	78.9	80.7	99.7*	67.0	62.5	52.8
Table 4: Evaluations on the attacks integrated with TIM .	Ta	ble 4: E	valua	tions	on the a	ttacks	integra	ted with	TIM.

Model	Attack	Inc-v3	Inc-v4	IncRes-v2	Res-101	Inc-v3 _{ens3}	Inc-v3 _{ens4}	IncRes-v2 _{ens}
	DIM	99.0*	64.3	60.9	53.2	19.9	18.3	9.3
	TIM	100.0*	48.8	43.6	39.5	24.8	21.3	13.2
Inc-v3	SIM	100.0*	69.4	67.3	62.7	32.5	30.7	17.3
	Admix	100.0*	82.6	80.9	75.2	39.0	39.2	19.2
	DIM	72.9	97.4*	65.1	56.5	20.2	21.1	11.6
	TIM	58.6	99.6 *	46.5	42.3	26.2	23.4	17.2
Inc-v4	SIM	80.6	99.6 *	74.2	68.8	47.8	44.8	29.1
	Admix	87.8	99.4*	83.2	78.0	55.9	50.4	33.7
	DIM	70.1	63.4	93.5*	58.7	30.9	23.9	17.7
	TIM	62.2	55.4	97.4*	50.5	32.8	27.6	23.3
IncRes-v	2 SIM	84.7	81.1	99.0*	76.4	56.3	48.3	42.8
	Admix	89.9	87.5	99.1 *	81.9	64.2	56.7	50.0
	DIM	75.8	69.5	70.0	98.0*	35.7	31.6	19.9
	TIM	59.3	52.1	51.8	99.3*	35.4	31.3	23.1
Res-101	SIM	75.2	68.9	69.0	99.7 *	43.7	38.5	26.3
	Admix	85.4	80.8	79.6	99.7 *	51.0	45.3	30.9
e 2: Ev	aluation	ns on v	variou	ıs single	e input	t transfo	rmation	based att
lodel	Attack	Inc-v3	Inc-v4	IncRes-v2	2 Res-101	Inc-v3 _{ens}	Inc-v 3_{ens4}	IncRes-v2 _{ens}
nc-v3	SI-DIM	98.9*	85.0	81.3	76.3	48.0	45.1	24.9
	Admix-DIN	1 99.8 *	90.5	87.7	83.5	52.2	49.9	28.6
nc-v4	SI-DIM	89.3	98.8*	85.6	79.9	58.4	55.2	39.3
	<i>Admix-</i> DIN	4 93.0	99.2 *	89.7	85.2	62.4	60.3	39.7
			85.1	97.5*	020			
R_{0}	SI-DIM	87.9	00.1	97.5	82.9	66.0	59.3	52.2
ncRes-v2			88.4	98.0 *	82.9 85.8	66.0 70.5	59.3 63.7	52.2 55.3
	SI-DIM					70.5		
ncRes-v2 Res-101	SI-DIM Admix-DIN	4 90.2 87.9	88.4	98.0*	85.8	70.5 63.5	63.7	55.3
les-101	SI-DIM Admix-DIN SI-DIM Admix-DIN	4 90.2 87.9 87.9 91.9 91.9	88.4 83.4 89.0	98.0 * 84.0 89.6	85.8 98.6* 99.8 *	70.5 63.5	63.7 57.5 62.3	55.3 42.0 46.6
.es-101 Ta	SI-DIM Admix-DIN SI-DIM Admix-DIN	4 90.2 87.9 87.9 91.9 91.9	88.4 83.4 89.0 tions o	98.0* 84.0 89.6 on the a	85.8 98.6* 99.8* ttacks	70.5 63.5 69.7 integrat	63.7 57.5 62.3 ed with	55.3 42.0 46.6 DIM.
es-101 Ta Iodel	SI-DIM Admix-DIN SI-DIM Admix-DIN ble 3: E	1 90.2 87.9 1 91.9 valua	88.4 83.4 89.0 tions (Inc-v4	98.0* 84.0 89.6 on the a	85.8 98.6* 99.8* ttacks	70.5 63.5 69.7 integrat	63.7 57.5 62.3 ed with	55.3 42.0 46.6 DIM.
es-101 Ta Iodel	SI-DIM Admix-DIN SI-DIM Admix-DIN ble 3: E Attack	4 90.2 87.9 91.9 Valuat 100.0*	88.4 83.4 89.0 tions (Inc-v4 71.8	98.0* 84.0 89.6 On the a IncRes-v2	85.8 98.6* 99.8* ttacks Res-101	70.5 63.5 69.7 integrat Inc-v3 _{ens3}	63.7 57.5 62.3 ed with Inc-v3 _{ens4}	55.3 42.0 46.6 DIM. IncRes-v2 _{ens}
es-101 Ta Iodel nc-v3	SI-DIM Admix-DIN SI-DIM Admix-DIN ble 3: E Attack SI-TIM	4 90.2 87.9 91.9 Valuat 100.0*	88.4 83.4 89.0 tions (Inc-v4 71.8	98.0* 84.0 89.6 On the a IncRes-v2 68.6 80.4	85.8 98.6* 99.8* ttacks Res-101 62.2	70.5 63.5 69.7 integrat Inc-v3 _{ens3} 48.2	63.7 57.5 62.3 ed with Inc-v3 _{ens4} 47.4	55.3 42.0 46.6 DIM. IncRes-v2 _{ens} 31.3
es-101 Ta Iodel nc-v3	SI-DIM Admix-DIN SI-DIM Admix-DIN ble 3: E Attack SI-TIM Admix-TIN	4 90.2 87.9 4 91.9 Valuat Inc-v3 100.0* 100.0* 78.2	88.4 83.4 89.0 tions (Inc-v4 71.8 83.9	98.0* 84.0 89.6 0n the a IncRes-v2 68.6 80.4 71.9	85.8 98.6* 99.8* ttacks Res-101 62.2 74.4	70.5 63.5 69.7 integrat Inc-v3 _{ens3} 48.2 59.1	63.7 57.5 62.3 ed with Inc-v3 _{ens4} 47.4 57.9	55.3 42.0 46.6 DIM. IncRes-v2 _{ens} 31.3 39.2
es-101 Ta Iodel nc-v3 nc-v4	SI-DIM Admix-DIN SI-DIM Admix-DIN ble 3: E Attack SI-TIM Admix-TIN SI-TIM	4 90.2 87.9 4 91.9 Valuat Inc-v3 100.0* 100.0* 78.2	88.4 83.4 89.0 tions (Inc-v4 71.8 83.9 99.6*	98.0* 84.0 89.6 0n the a IncRes-v2 68.6 80.4 71.9	85.8 98.6* 99.8* ttacks Res-101 62.2 74.4 66.1	70.5 63.5 69.7 integrat Inc-v3 _{ens3} 48.2 59.1 58.6	63.7 57.5 62.3 ed with Inc-v3 _{ens4} 47.4 57.9 55.4	55.3 42.0 46.6 DIN. IncRes-v2 _{ens} 31.3 39.2 45.1
es-101 Ta fodel nc-v3 nc-v4	SI-DIM Admix-DIN SI-DIM Admix-DIN ble 3: E Attack SI-TIM Admix-TIN SI-TIM Admix-TIN	4 90.2 87.9 91.9 Valuat 100.0* 100.0* 100.0* 187.2 87.4 84.5 84.5	88.4 83.4 89.0 tions (Inc-v4 71.8 83.9 99.6* 99.7*	98.0* 84.0 89.6 0n the a IncRes-v2 68.6 80.4 71.9 82.3	85.8 98.6* 99.8* ttacks Res-101 62.2 74.4 66.1 77.0	70.5 63.5 69.7 integrat Inc-v3 _{ens3} 48.2 59.1 58.6 68.1	63.7 57.5 62.3 ed with Inc-v3 _{ens4} 47.4 57.9 55.4 65.3	55.3 42.0 46.6 DIM. IncRes-v2 _{ens} 31.3 39.2 45.1 53.1
es-101 Ta Iodel nc-v3 nc-v4 ncRes-v2	SI-DIM Admix-DIN SI-DIM Admix-DIN ble 3: E Attack SI-TIM Admix-TIN SI-TIM Admix-TIN SI-TIM	4 90.2 87.9 91.9 Valuat 100.0* 100.0* 78.2 78.2 87.4 84.5 84.5	88.4 83.4 89.0 tions (Inc-v4 71.8 83.9 99.6* 99.7* 82.2	98.0* 84.0 89.6 0n the a IncRes-v2 68.6 80.4 71.9 82.3 98.8*	85.8 98.6* 99.8* ttacks Res-101 62.2 74.4 66.1 77.0 77.4	70.5 63.5 69.7 integrat Inc-v3 _{ens3} 48.2 59.1 58.6 68.1 71.6	63.7 57.5 62.3 ed with Inc-v3 _{ens4} 47.4 57.9 55.4 65.3 64.7	55.3 42.0 46.6 DIM. IncRes-v2 _{ens} 31.3 39.2 45.1 53.1 61.0
les-101	SI-DIM Admix-DIN SI-DIM Admix-DIN ble 3: E Attack SI-TIM Admix-TIN SI-TIM Admix-TIN SI-TIM	4 90.2 87.9 91.9 Valuat 1.0 Inc-v3 100.0* 1 100.0* 1 87.4 84.5 90.2 74.2 74.2	88.4 83.4 89.0 tions (Inc-v4 71.8 83.9 99.6* 99.7* 82.2 88.2	98.0* 84.0 89.6 0n the a IncRes-v2 68.6 80.4 71.9 82.3 98.8* 98.6*	85.8 98.6* 99.8* ttacks Res-101 62.2 74.4 66.1 77.0 77.4 83.9	70.5 63.5 69.7 integrat Inc-v3 _{ens3} 48.2 59.1 58.6 68.1 71.6 78.4	63.7 57.5 62.3 ed with Inc-v3 _{ens4} 47.4 57.9 55.4 65.3 64.7 73.6	55.3 42.0 46.6 DIN. IncRes-v2 _{ens} 31.3 39.2 45.1 53.1 61.0 70.0

 \mathcal{J}_{t+1} \bar{y}_{t+1}

 $\operatorname{gn}(g_{t+1})$

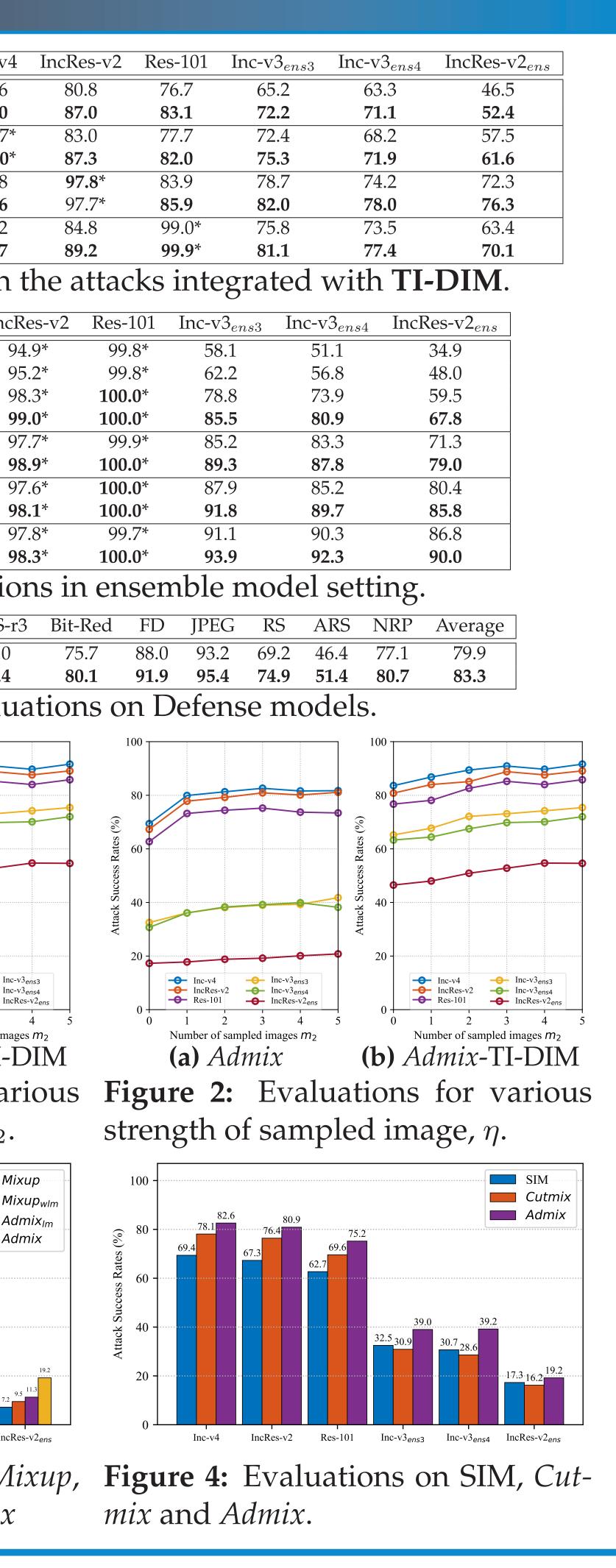
Experiments

	he		lents)		
	Mod	01	Attac		Inc-v3	Inc-v4
			SI-TI-E		99.1*	83.6
	Inc-v	3	Admix-TI		99.9 *	89.0
	Inc-v	г.Л	SI-TI-D	DIM	87.9	98.7*
		4	Admix-TI		90.4	99.0*
	IncR	es-v2	SI-TI-E Admix-TI		88.8	86.8
			SI-TI-E		90.1 84.7	89.6 82.2
	Res-1	101	Admix-TI		91.0	87.7
		Tab	ole 5: E	Evalı	atior	ns on
		A	Attack	Inc-v	3 Inc-	v4 Inck
			DIM	99.4	l* 97.4	L* 94
			TIM	99.8		
			SIM Admix	99.9 100.0		
			I-DIM	99.7		
			nix-DIM	99.7		
		_	I-TIM	99.7		
			<i>mix-</i> TIM	99.7		
			TI-DIM <i>ix</i> -TI-DIM	99.6 99.7		
		2 10////				luatio
	Г	A	ttack	HGD	R&P	NIPS-r
		SI-7	TI-DIM	91.4	88.0	90.0
		Admi	x-TI-DIM	93.7	90.3	92.4
				lab	le 7: 1	Evalu
10	00			100 -		
	30			9 80 9 3		.
cess Rates (%) 9	50			Attack Success Rates (%)		0
ess Ra				ess Re		
2 Succ	+0 -			$\frac{2}{9}$ $\frac{2}{3}$ $\frac{2}{3}$	6	
Attack Suc F	8			Attacl		
2	20		_ 	• 20 ·		
	-0-	IncRes-v2	Inc-v3 _{ens3}		Inc-v4	s-v2 Inc-v
	$0 \downarrow 0 1$	Res-101	IncRes-v2 _{ens}	5 0	0 1	2 3
		ber of samp) Ad1	bled images m_2	(b)		f sampled imag
Fi.			Evalu			
nu	imbe	er of	samp	lea 1	mage	m_{2} .
1	00	99.4100.0				····· Mi>
			82.6).9		Mix
(%)	80		74.6 72.1	75	.2	Adi
Success Rates (%)	60 -		67.6	63.8		
ccess				55.1		
	40	44.2	41.1	39.0	39.0	39.2
Attack					25.0	23.8
	20 -				19.4	3.4
	0					
	-	nc-v3	Inc-v4 IncRes-v2	2 Res-101	Inc-v3 _{ens3}	Inc-v3 _{ens4} IncRe
Fig	gure	3:	Evalı	Jatic	ons o	n Ma
M	ixup	wlm	, Admiz	x_{lm} a	and A	dmix
	1					

Conclusion

- We propose a novel input transformation based attack, *Admix*, utilizing the information of images
- gradient-based attacks.
- transferability of various attacks.

SCOCTOBER 11-17 VIRTUAL



from other category to enhance the transferability.
Our method is geneally applicable to other input transformations (*i.e.* DIM, TIM, SIM *etc.*) and gradient-based attacks



• Experiments show our method could significantly enhance the