Appendix

We provide further evaluation on 1,000 images for TA and Surfree, and parameter studies on the dimension of directional line d and the bound τ for angle α.

Evaluation on more images. Due to the high computation cost, most attacks adopt hundreds of images for evaluation (*e.g.*, HSJA [8] (100), QEBA [29] (50), GeoDA [38] (350), Surfree [35] (200)). We follow Surfree [35] with 200 images in experiments. To better validate the effectiveness of TA, we further compare TA with Surfree [35] on VGG-16 [44] using 1,000 images. As shown in Table 4, TA consistently outperforms Surfree [35] under three RMSE thresholds, showing the superiority of TA.

Table 4: Attack success rate (%) on VGG-16 with 1000 images under different RMSE thresholds. The maximum number of queries is set to 1,000

RMSE	0.1	0.05	0.01
Surfree	98.4	90.2	36.5
TA (**Ours**)	**99.6**	**93.9**	**39.7**

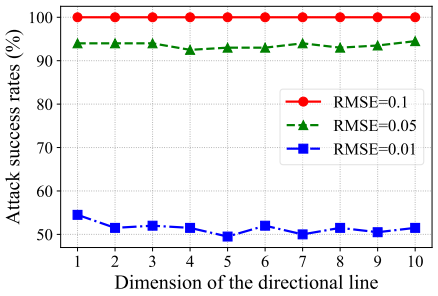

Fig. 10: Attack success rate (%) of TA on ResNet-18 within 1,000 queries with various dimensions of the directional line under three $RMSE$ thresholds

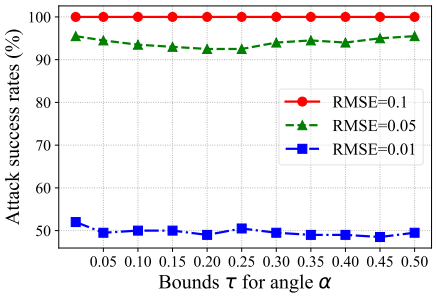

Fig. 11: Attack success rate (%) of TA on ResNet-18 within 1,000 queries with various bounds τ for angle α under three $RMSE$ thresholds

Parameter study on the dimension of directional line d. A small dimension of line d helps us sample diverse low-dimensional space in each iteration to boost the attack performance. To determine a good value for d, we conduct parameter studies by varying d from 1 to 10. As shown in Fig. 10, with the increment of d, the attack success rate continues to decrease, which is most obvious under the setting of $RMSE = 0.01$. Hence, we adopt $d = 3$ in experiments.

Parameter study on the bound τ **for angle** α. A small bound τ for α makes the learning strategy ineffective while a large bound might result in inaccurate estimation, which degrades the performance. We also conduct parameter studies for τ. As shown in Fig. 11, a larger τ will lead to lower attack success rate, which is also more obvious when $RMSE = 0.01$. Hence, we adopt $\tau = 0.1$ in experiments.