Teaching AI to Teach: Leveraging Limited Human Salience Data Into Unlimited Saliency-Based Training (Supplementary Materials)

Colton R. Crum
ccrum@nd.edu
Aidan Boyd
aboyd3@nd.edu
Kevin Bowyer
kwb@nd.edu
Adam Czajka
aczajka@nd.edu

University of Notre Dame
Notre Dame, IN 46556, USA

These supplementary materials illustrate selected aspects of the presented work.

Figures S1 and S2 show authentic and synthetically-generated (or faked in other ways) face and iris examples, respectively, taken from datasets used in this work.

Table S1 mimics Table 1 in the main paper, except that we present results for the task of iris presentation attack detection.

Table S2 shows the performance impact of AI Students after reducing the number of human annotations used to train AI Teachers for iris PAD.

Figure S3 mimics Figure 3 in the main paper, except that we present results for the task of iris presentation attack detection.

Finally, Figures S4 and S5 show examples of AI Teachers-generated saliency for models trained in various ways (with and without human saliency), for two types of saliency generation (CAM and RISE) and across various architectures.

Table S1: Same as in Table 1 in the main paper, except that results for iris PAD are shown.

How the AI Teachers were trained on TAIT data	Mean AUC on the EAIS data			
	DenseNet	ResNet	Xception	Inception
Without human salience	0.944±0.009	0.908±0.042	**0.952±0.003**	0.939±0.015
With human salience	**0.950±0.013**	**0.915±0.018**	0.950±0.003	**0.947±0.010**

© 2022. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

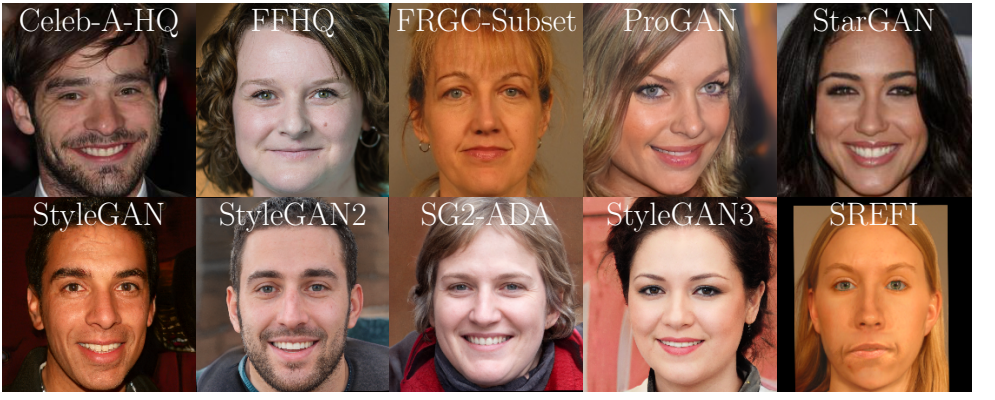


Figure S1: Examples of the **authentic** (Celeb-A-HQ, FFHQ, FRGC-Subset) and **synthetic** (ProGAN, StarGAN, StyleGAN, StyleGAN2, StyleGAN2-ADA, StyleGAN3, SREFI) samples from the dataset used to train the AI Students.

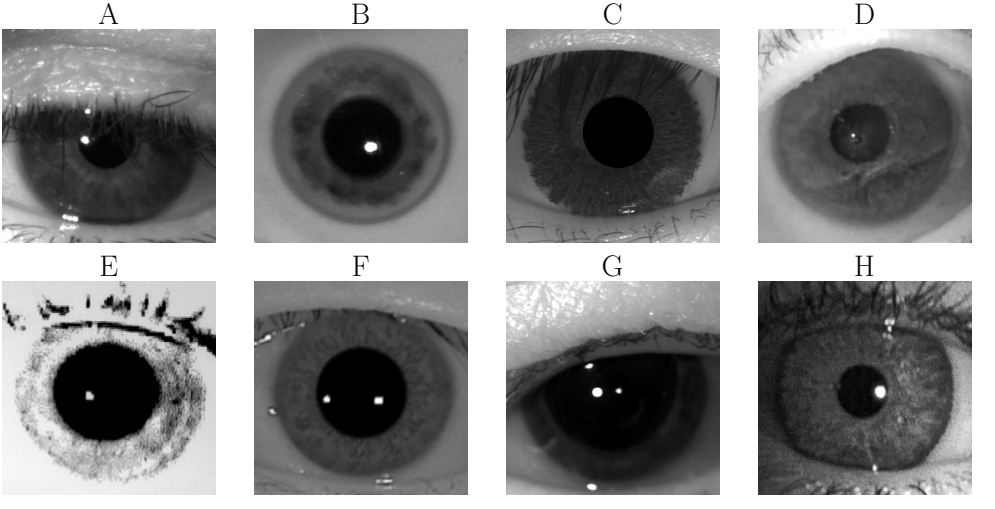
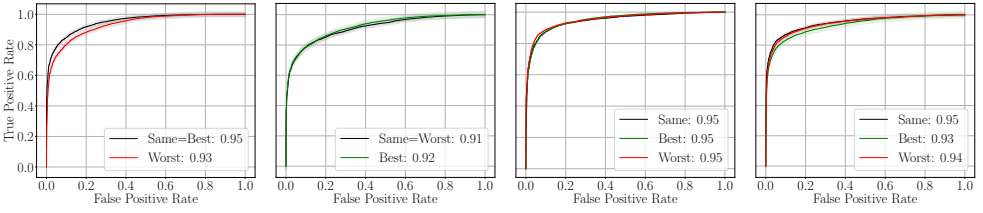


Figure S2: Examples of **live sample** (A) and **spoof samples:** aratifical (B), textured contact lenses (C), post-mortem (D), paper printouts (E), synthetically-generated (F), diseased (G), and wearing contacts then printed (H). Samples taken from the dataset used to train the AI Students.

Table S2: AUC for iris PAD using DenseNet as the backbone for Teachers and Students. Same training configurations as the main paper, except different amounts of data were used to train Teachers. The number of samples are split evenly between both classes. The table shows a single train-test experiment.

# of human annotated samples used to train Teacher models	**Student performance**
100 (13% of human annotated set)	0.905
382 (50% of human annotated set)	0.940



(a) DenseNet (b) ResNet (c) Xception (d) Inception

Figure S3: Same as in Figure 3 in the main paper, except that results for iris PAD are shown.

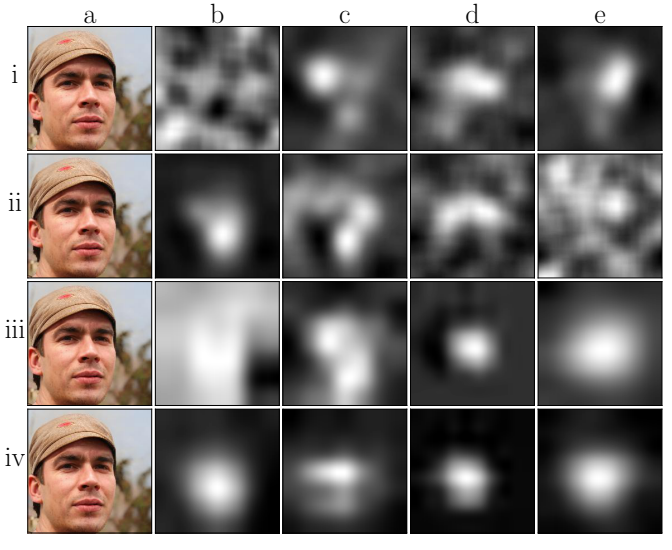


Figure S4: Illustration of AI Teacher model's salience maps on the TAIT training data for **synthetic face detection**: (a) input sample, (b) DenseNet, (c) ResNet, (d) Xception, (e) Inception; (i) RISE-based salience + cross-entropy training, (ii) RISE-based salience + human-guided training, (iii) CAM-based salience + cross-entropy training, (iv) CAM-based salience + human-guided training.

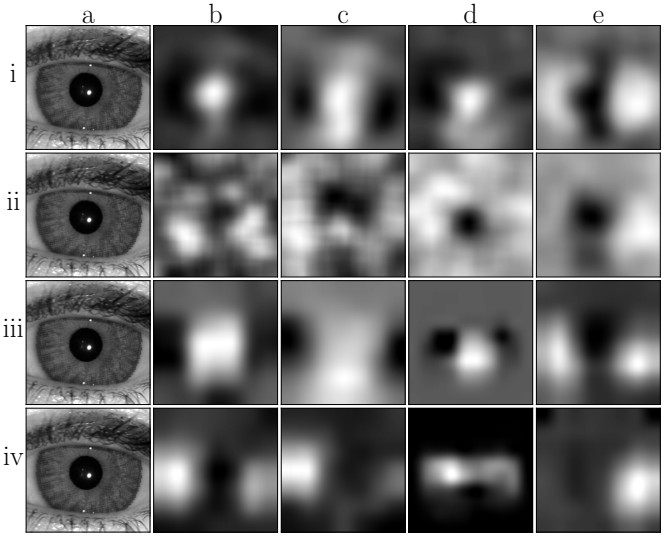


Figure S5: Illustration of AI Teacher model's salience maps on the TAIT training data for **iris PAD detection**: (a) input sample, (b) DenseNet, (c) ResNet, (d) Xception, (e) Inception; (i) RISE-based salience + cross-entropy training, (ii) RISE-based salience + human-guided training, (iii) CAM-based salience + cross-entropy training, (iv) CAM-based salience + human-guided training.