Detecting Anomalous Faces with "No Peeking" Autoencoders

Anand Bhattad Jason Rock David Forsyth

ILLINOIS

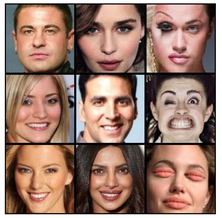
2018 CVPR SALT LAKE CITY • JUNE 18-22

1. Motivation: Detect Sleepy Drivers, Heart Attacks

Task: Given a pool of faces images, find the most anomalous

Applications

- Incapacitated machine operators
- X-Rays; Fabrics



CelebA Typical Anomaly

Problems in Previous work

- One class from MNIST as inlier and others are outliers [1]
- Training with Anomaly examples (supervised) [2]
- Impractical for real life applications

Contributions

- Inpainting Autoencoders as a feature extractor
- A dataset of real anomalous faces and real typical faces
- A complete unsupervised framework for *True* anomaly detection

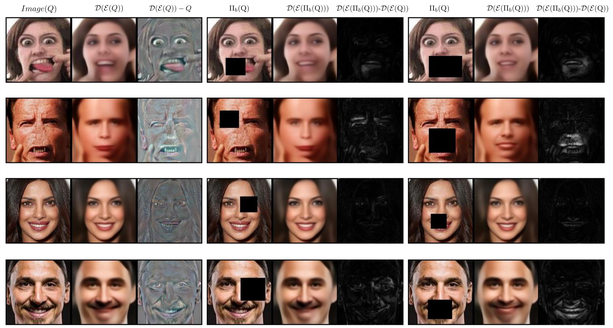
2. Beating Peeking with Inpainting Autoencoders

Problems with AE

- AE reconstruction based on pixel values
- AE residual is a poor anomaly signal ('peeking')

Solution

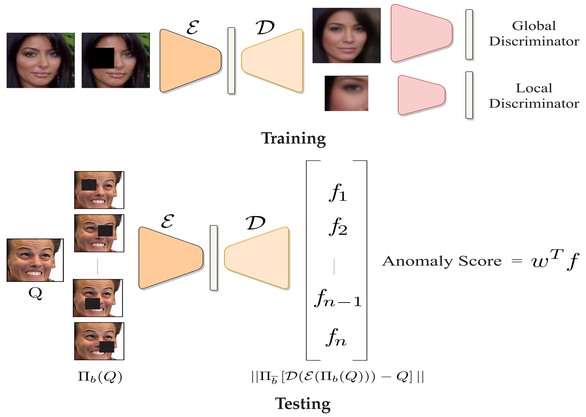
- Mask large regions



Top two rows: Examples from Anomaly Set
Bottom two rows: Examples from Typical Set

- Inpainting residuals are strong cues for anomaly detection
- Each mask is a local anomaly detector

3. Inpainting Residual Features (IRF)



4. Eyeglasses Experiment

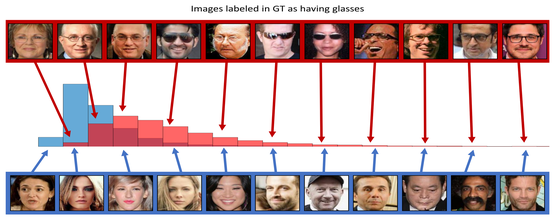
- Viola Jones
- 128 x 128 images
- 32x32 boxes
- 7700 G; 7000 NG

	L_1 Regularized Logistic Regressor		
Feature	Resnet50	AE Code	Res Patch
Accuracy	**99.4**	85.6	93.5

	1-Class SVM		
Feature	Resnet50	AE Code	Res Patches
AUC	52.3	51.8	53.8

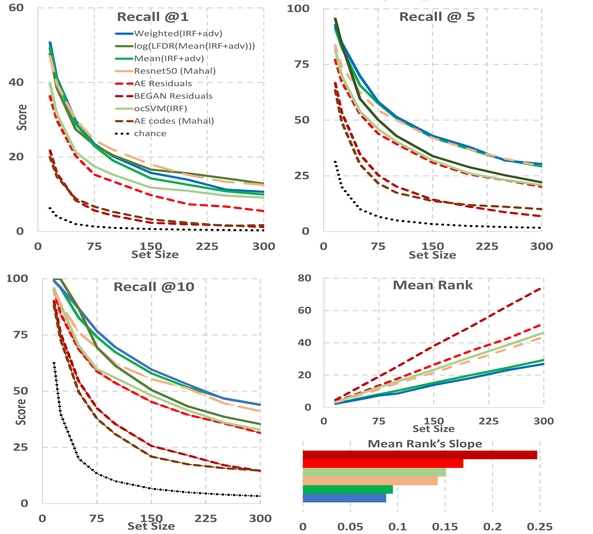
	Mahalanobis Distance		
Feature	Resnet50	AE Code	Res Patches
AUC	52.8	50.8	**92.5**

	L_∞ norm		
Feature	Resnet50	AE Code	Res Patches
AUC	NA	NA	**81.5**



5. True Anomaly Experiment

Dataset: 316 Anomalous & 100 Typical Images (passes Viola Jones)



Solid lines are for IRF and dashed are baselines. For recalls, higher the better and for mean rank, lower the better

6. Qualitative Results

Anomaly set: Median image from each decile

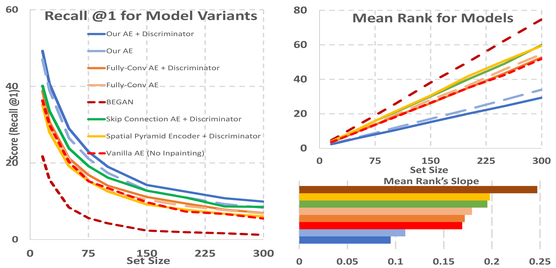
Typical set: Median image from each decile

celebA: Median image from each decile

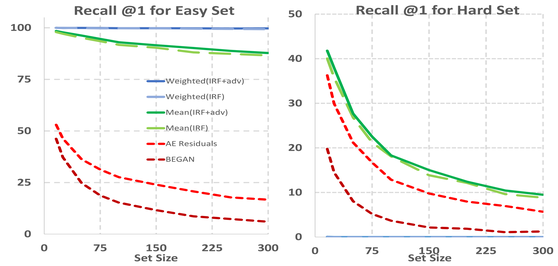
Bar charts: Recall@1 (rosy brown), @5 (slate blue), & @10 (gray) for 16, 64, 128 & 256 sets. All images are arranged by their decreasing anomaly scores.

7. Ablation Experiments

Experiments with different architectures used for image inpainting



Affect of dataset bias and models performance on Easy vs Hard data



8. Conclusions

- Unsupervised framework for evaluating future anomaly detection, enabled through two small datasets
- IRF combats the over-generalization of compression losses
- IRF works well for both supervised and unsupervised tasks

Limitation

- Detects concentrated anomalies well as compared to the diffused ones
- **Solution:** Multiple inpainting boxes

References

1. Zhai, Shuangfei, et al. "Deep structured energy based models for anomaly detection." ICML 2016.
2. Zaheer, Manzil, et al. "Deep sets." NIPS 2017.

Acknowledgement

We thank Karan Ganju and Unnat Jain for insightful discussions