MaLP: Manipulation Localization Using a Proactive Scheme

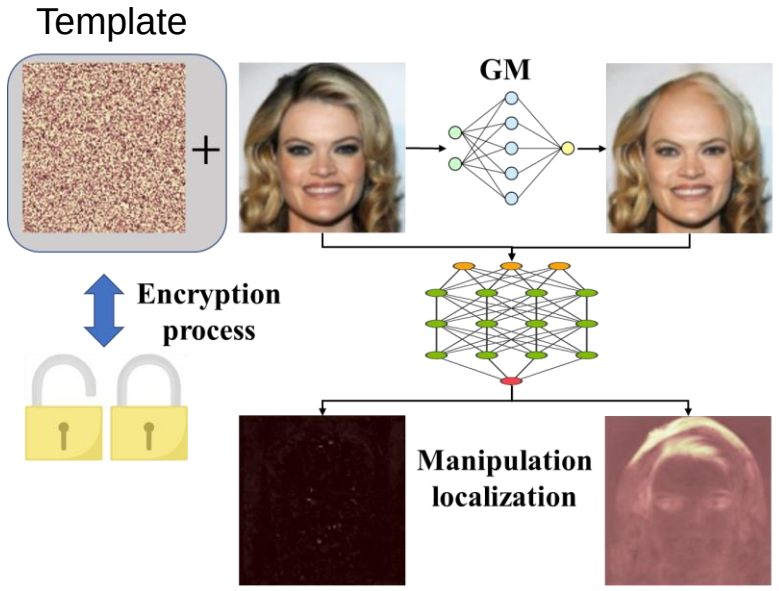
By:

Vishal Asnani[1] , Xi Yin[2] , Tal Hassner[2] , Xiaoming Liu[1]
[1]Michigan State University, [2]Meta AI

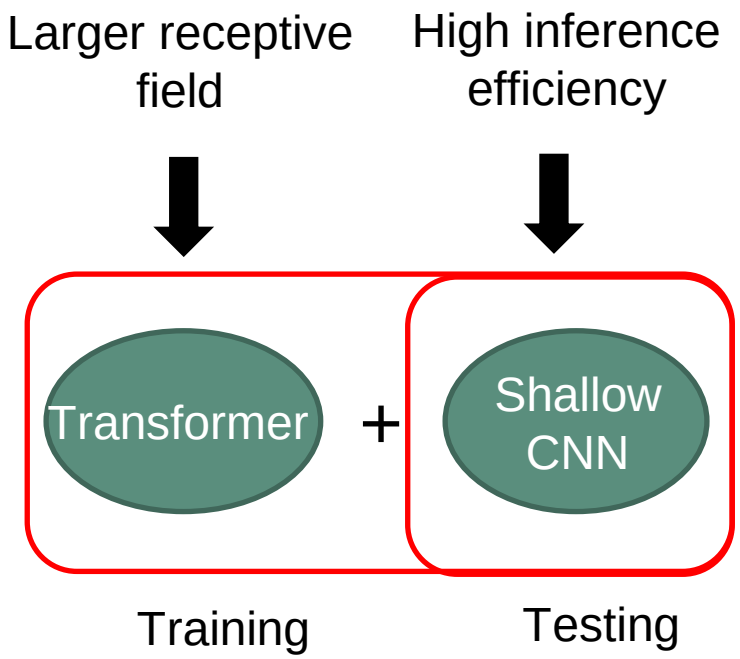
WED-AM-393

High-Level Summary of MaLP

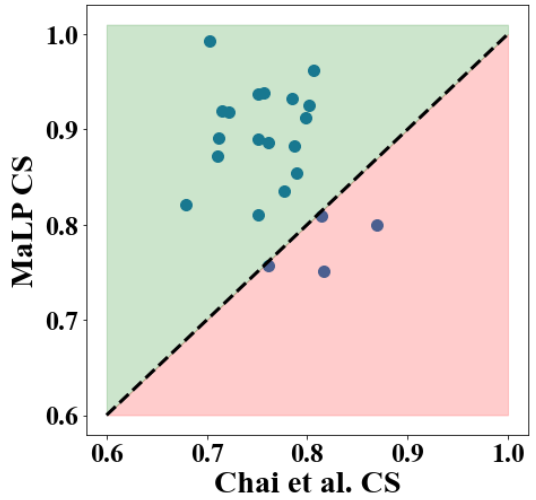
Overview



Two-branch architecture



Generalization across unseen GMS

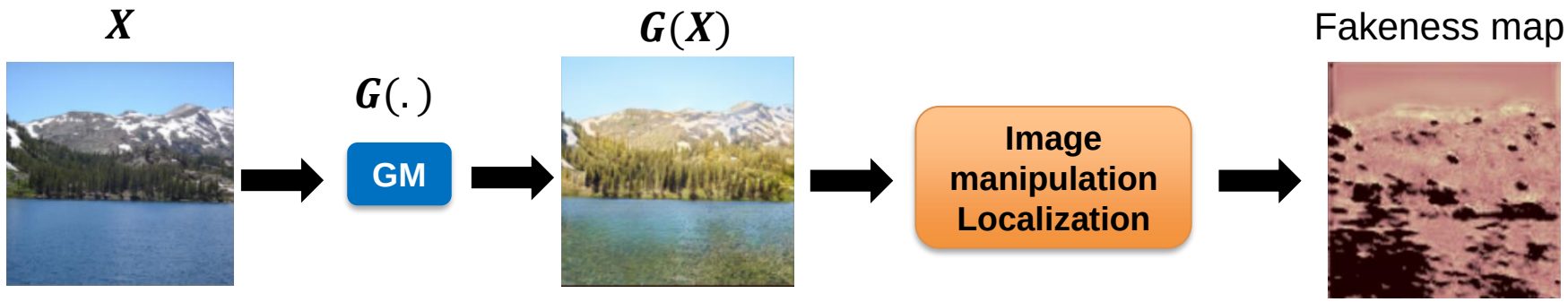


Contributions:

1. A novel proactive approach for manipulation localization.
2. Learning template set unsupervisedly to perform encryption.
3. A plug-and-play discriminator module to improve image generation quality for GMs.
4. A generalization benchmark for evaluating manipulation localization.

Chai et al. What makes fake images detectable? understanding properties that generalize. In *ECCV 2020:*

Passive Manipulation Localization



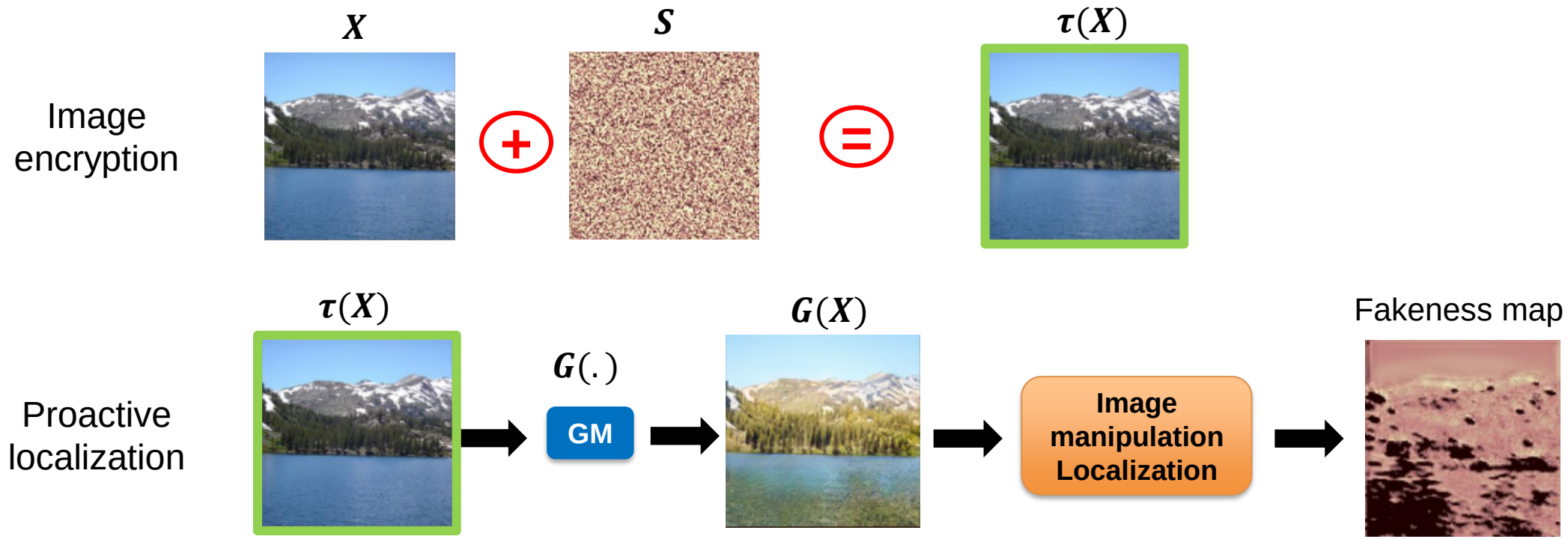
Weaknesses

Poor generalization across unseen **GMs**

Low-resolution fakeness map

Fail on unseen image **attribute modification**

Proactive Manipulation Localization



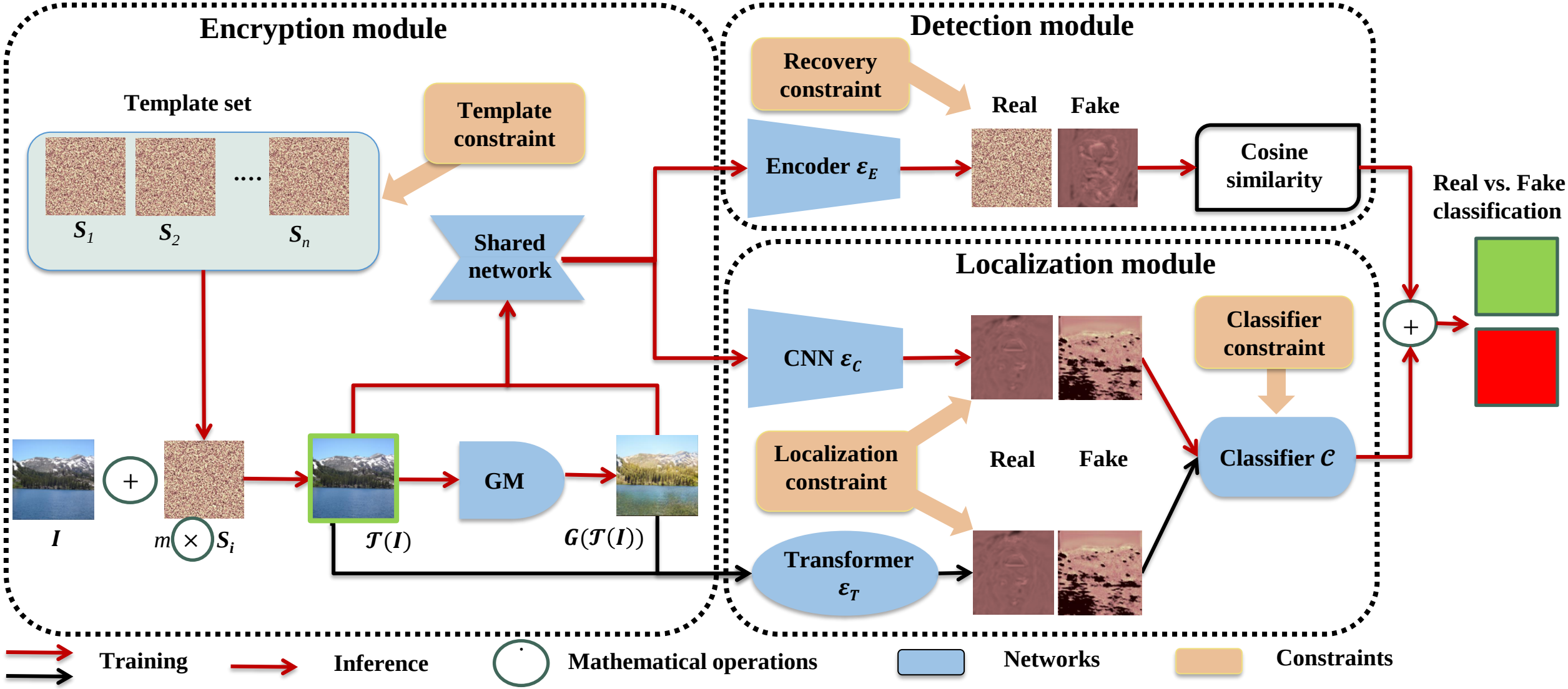
Challenges

- ❑ Formulate constraints for the template.
- ❑ Estimate a high-resolution fakeness map.
- ❑ Generalizable to unseen GMs and attributes manipulation.

Innovations

- ❑ Proactive scheme.
- ❑ Train on 1 GM and test on 22 GMs.
- ❑ Highly efficient.
- ❑ A convenient plug-and-play discriminator.
- ❑ Exhaustive evaluation benchmark.

Framework


Encryption module
Template set
S_1
S_2
....
S_n
Template constraint
Shared network
I
+
$m \times S_i$
$\mathcal{T}(I)$
GM
$G(\mathcal{T}(I))$
Detection module
Recovery constraint
Real
Fake
Encoder ε_E
Cosine similarity
Real vs. Fake classification
Localization module
CNN ε_C
Classifier constraint
Localization constraint
Real
Fake
Classifier C
Transformer ε_T
+
Training
Inference
Mathematical operations
Networks
Constraints

Results

Manipulation localization comparison with prior works.

Method	Localization			Detection		
	CS	PSNR	SSIM	Accuracy	EER	AUC
Dang et al.	0.6230	6.214	0.2178	0.9975	**0.0050**	0.9975
Huang el al.	0.8831	22.890	**0.7876**	0.9945	0.0077	0.9998
MaLP	**0.9394**	**23.020**	0.7312	**0.9991**	0.0072	**1.0**

Comparison of localization performance across unseen GMs and attribute modifications.

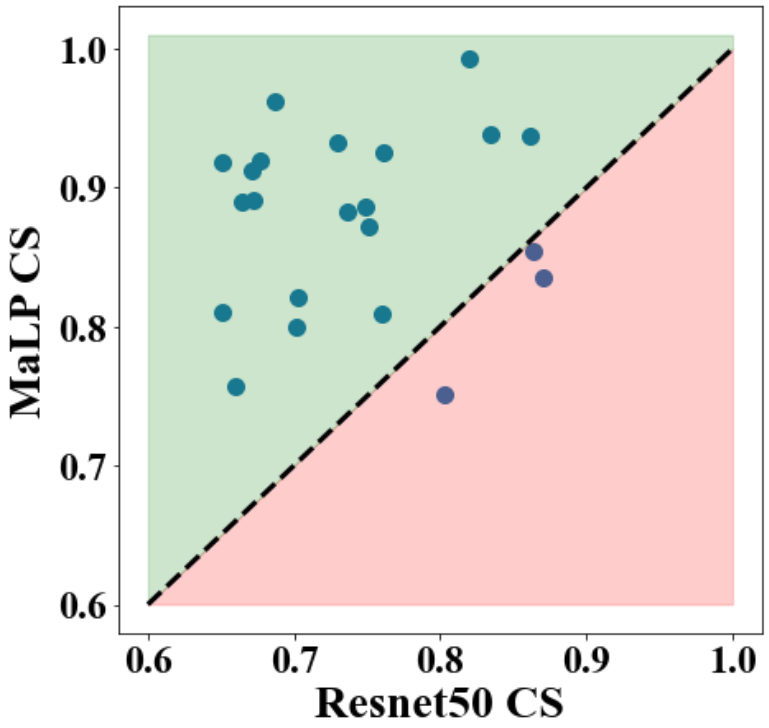
Method	Cosine similarity (AttGAN)			Cosine similarity (StyleGAN)		
	Bald	Black hair	Eyeglasses	Smile	Age	Gender
Huang el al.	0.8141	0.6932	0.6950	0.6176	0.3141	0.6470
MaLP	**0.8201**	**0.7940**	**0.8557**	**0.8159**	**0.8255**	**0.8016**

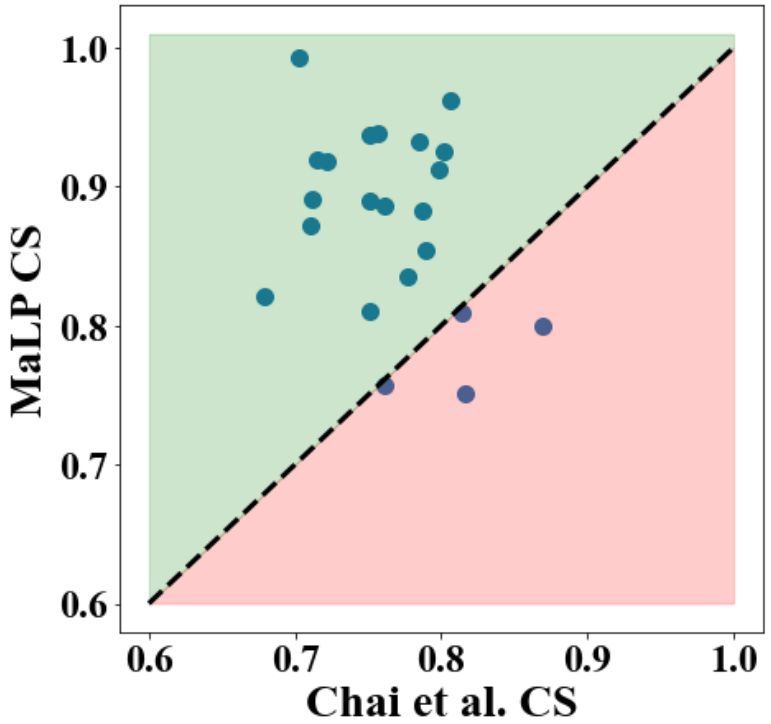
MaLP **outperforms prior works** for localization and detection

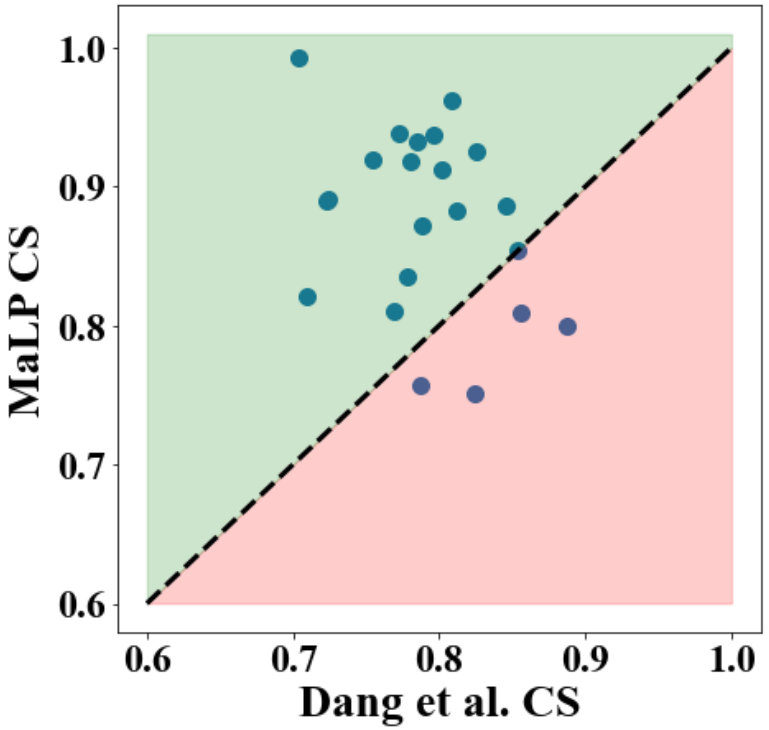
MaLP has **better generalization** across different facial attributes modifications

Dang et al, On the detection of digital face manipulation. In *CVPR 2020*
Huang et al, Fakelocator: Robust localization of GAN-based face manipulations. *IEEE Transactions on Information Forensics and Security*, 2022.

Generalization Across Multiple GMs







Better generalization performance across **unseen GMs** compared to prior passive methods

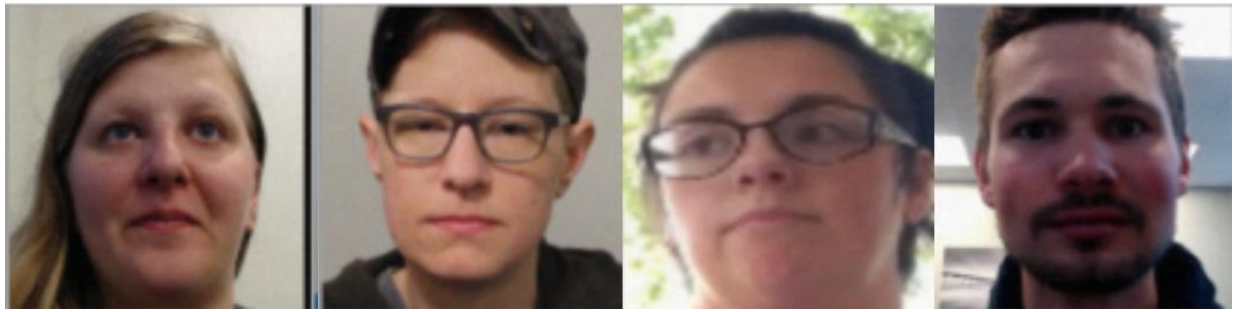
Dang et al, On the detection of digital face manipulation. In *CVPR 2020*
Chai et al. What makes fake images detectable? understanding properties that generalize. In *ECCV 2020:*

Manipulation Localization Visualization

MaLP is able to **localize** manipulated regions across **different facial attributes and non-face images**

MaLP as Discriminator

Real image



0.58 0.62 0.61 0.59

Manipulation before fine-tuning

0.42 0.45 0.40 0.44

Manipulation after fine-tuning

FID score comparison for the application of MaLP as a discriminator

State	Fine-tune	StarGAN FID
Before	-	60.49
After	G	51.91
	G + MaLP	52.07

Fine-tuning GM with MaLP as discriminator produces **better quality** generated images

Conclusions

- **Proactive scheme** for image manipulation localization is proposed.
- A **template set** is estimated using defined constraints to benefit localization.
- The fakeness map is estimated via a **two-branch architecture** using local and global-level features.
- **A generalization benchmark** for manipulation localization consisting of evaluation on 22 GMs is proposed.
- **Better localization performance** than prior SoTA works.
- A **plug-and-play discriminator module** for improving generation quality of GMs.

Thank you for listening!

GitHub

Paper

MICHIGAN STATE UNIVERSITY

Meta

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA