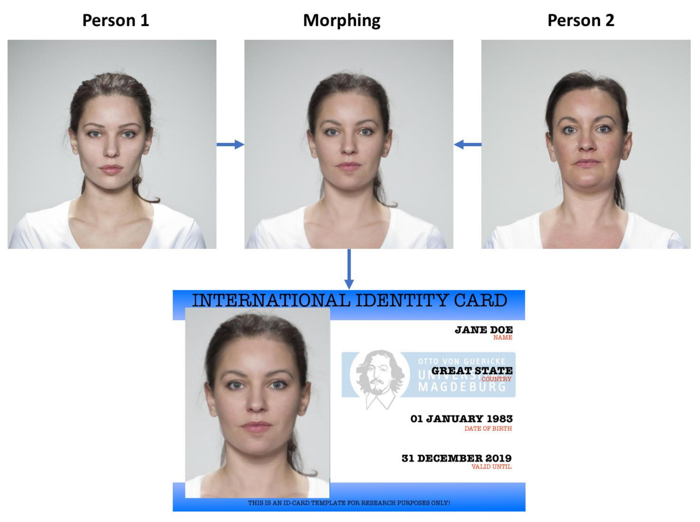
Towards Generalizable Morph Attack Detection with Consistency Regularization

Hossein Kashiani, Niloufar Alipour Talemi, Mohammad Saeed Ebrahimi Saadabadi, Nasser M. Nasrabadi
Lane Department of Computer Science and Electrical Engineering, WVU

Problem Statement

Face morphing is an image manipulation where two faces are blended together. At the time of passport enrollment, the passport photo can be easily manipulated with a morphing attack without the requirement of advanced forgery.



Motivation

- We can enhance the generalization capability of morph attack detection from the perspective of consistency regularization.
- Consistency regularization operates under the premise that generalizable morph attack detection should output consistent predictions irrespective of the possible variations that may occur in the input space.

Contribution

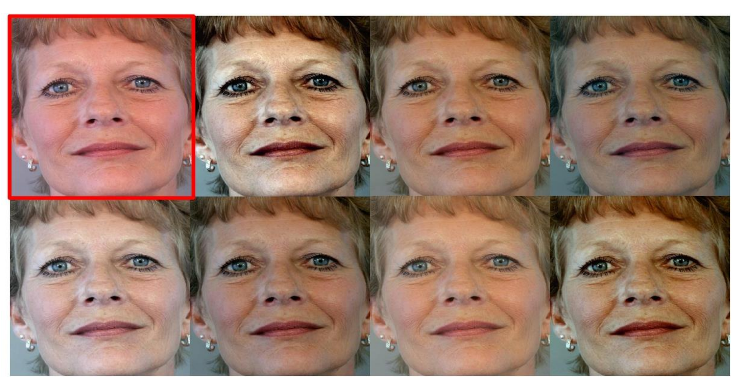
- We regularize morph attack detection model to predict consistent results regardless of potential variations caused by diverse morph attacks, image quality, and environmental situations.
- We propose two morph-wise augmentations to explore a wide space of realistic morph attack transformations in our consistency regularization.
- We carry out extensive evaluations on several datasets to validate the generalization capability of our morph attack detection.

Challenges

- There exist large domain shifts between different morph attacks.

New Augmentation

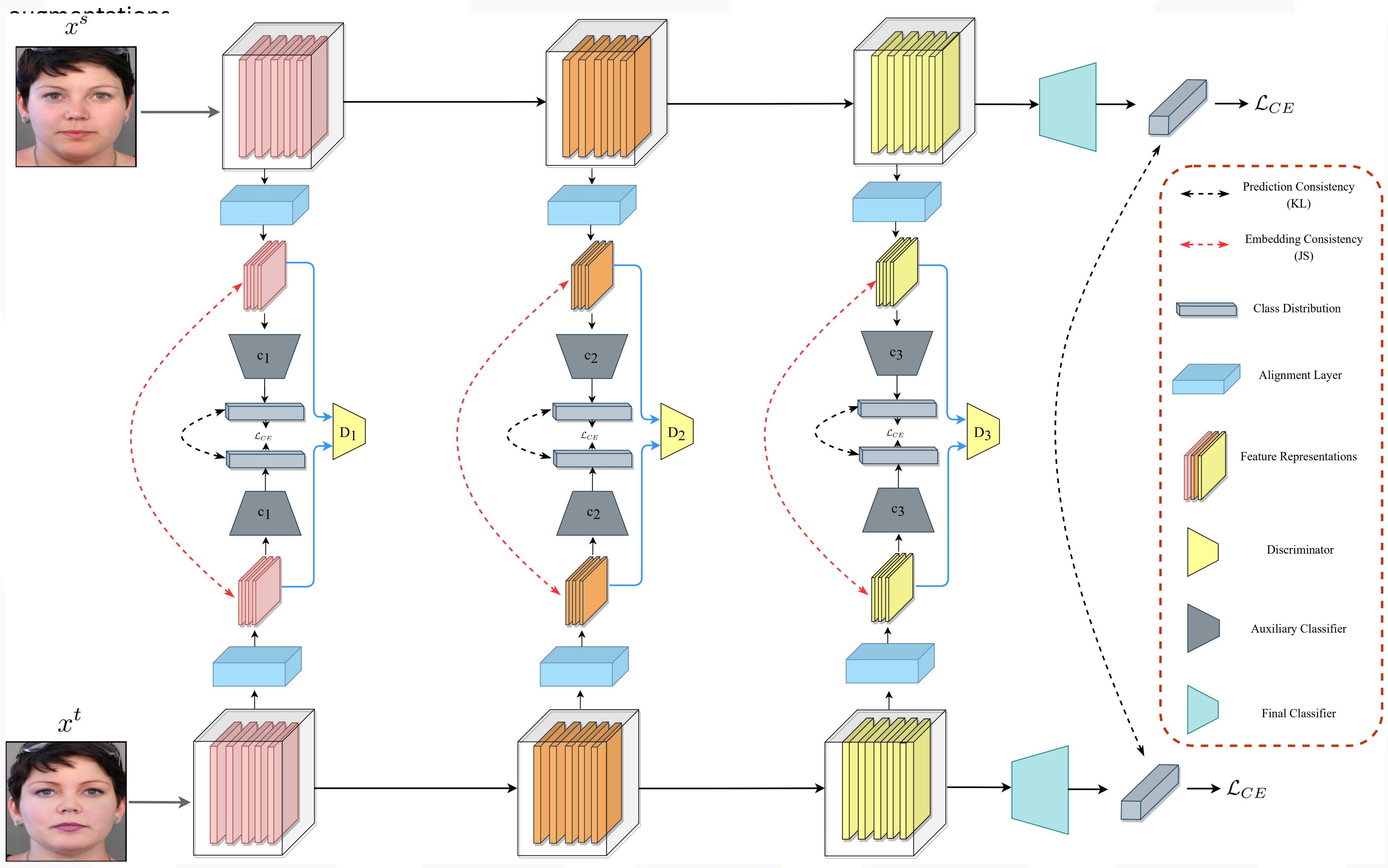
- We propose the Inter-domain Style Mixup (ISM) augmentation, which employs the photo-realistic style transfer to synthesizes unseen morph attacks with new styles, while keeping the content of the input morph images unchanged.



- In addition, we propose the Self-morphing (SM) augmentation to synthesize morph attacks with minimal visual artifacts using several instances of the same identity.

Proposed Method

- To improve a generalized morph attack detection, we impose consistency regularization at both the logit and feature representation levels using Kullback–Leibler (KL) divergence minimization. For this objective, several regularization branches are first integrated into the intermediate layers of our model and the embedding levels are computed at these branches.
- To learn the domain-shared feature representation, we employ adversarial feature learning at different feature representation levels. A feature extractor competes with a domain discriminator to learn a domain-shared feature representation and the domain discriminator determines whether the input images come from the intact morph images or the augmented ones.
- To explore a wide space of realistic morph transformations in our consistency regularization, we propose the ISM and the SM augmentations.



Evaluations

- In the cross-morph evaluations, we use the FRGC FaceMorpher the training data and the test sets belongs to the FRGC morph faces with StyleGAN2, MIPGAN, and OpenCV morphing attacks. Our method is denoted by GRL. As reported in Table 1, the proposed GRL outperforms its competitors in all comparisons.

Table 1. Cross-morph evaluations of the proposed method with the state-of-the-art studies on FRGC datasets. The results are in terms of APCER1 (@BPCER=1%), APCER5 (@BPCER=5%), APCER (@BPCER10=10%), EER, and AUC metrics.

	Method	APCER1%	APCER5%	APCER10%	EER	AUC
MIPGAN	ConvNext	17.40	3.07	1.20	16.33	99.17
	Inception	61.98	36.68	23.82	17.26	91.12
	Residual	-	-	-	6.67	-
	GRL	00.00	00.00	00.00	4.28	99.99
StyleGAN	ConvNext	44.60	14.52	2.80	7.65	97.57
	Inception	50.60	32.39	25.56	17.26	94.89
	GRL	00.00	00.00	00.00	00.00	100.00
OpenCV	ConvNext	60.68	29.66	12.65	11.50	95.27
	Inception	00.00	00.00	00.00	00.00	100.00
	GRL	00.00	00.00	00.00	00.00	100.00

- In the cross-domain evaluations, we use Twins morph dataset as the training set and the test sets belongs to the FRGC, AMSL, FERET, VISAPP17, and FRLL datasets with landmark-based and GAN-based morphing attacks. These attacks consists of StyleGAN2, WebMorph, OpenCV, and FaceMorpher attacks. As reported in Table 2, the proposed GRL outperforms its competitors in all comparisons.

Table 2. Cross-domain comparison of the proposed GRL with the state-of-the-art studies. The evaluations are in terms of the EER metric.

	Methods	FRLL AMSL	FRLL Webmorpher	FRLL OpenCV	FRLL StyleGAN	FRLL FaceMorpher	FERET OpenCV	FERET StyleGAN	FERET FaceMorpher	FRGC OpenCV	FRGC StyleGAN	FRGC FaceMorpher	FRGC MIPGAN	VISAPP17	LMA-DRD
EER	SPL-MAD	12.09	15.72	5.78	12.92	4.67	30.21	28.95	25.76	19.54	15.57	18.42	-		29.54
	MixFacenet	15.18	12.35	4.39	8.99	3.87	-	-	-	-	-	-	-		23.72
	Inception	10.79	9.86	5.38	11.37	3.17	-	-	-	-	-	-	-		19.01
	PW-MAD	15.18	16.65	2.42	16.64	2.20	-	-	-	-	-	-	-		20.39
	Hamza	-	-	-	-	-	13.5	-	11.5	-	-	-	-	-	-
	Quality	7.91	7.13	5.41	7.04	3.60	12.29	13.99	10.80	24.48	14.32	24.17	-		25.09
	OrthoMAD	14.80	15.23	0.73	6.54	0.98	-	-	-	-	-	-	-	-	-
	Residuals (LMA)	-	-	-	-	-	-	-	-	-	0.17	-	13.92	-	-
	Mutual	3.11	-	-	-	-	-	-	-	-	-	-	-	4.69	-
	Scale-Space Gradients	-	-	-	-	-	-	-	0.98	-	-	-	6.67	-	

Conclusion

- Our paper introduces a morph attack detection system with strong generalization capabilities across various morph attacks, as demonstrated in experiments on multiple datasets..

References

1. C. Szegedy, et al. Inception-v4, Inception-ResNet and the impact of residual connections on learning. In AAAI, 2017.
2. P. Phillips, et al. The FERET database and evaluation procedure for face-recognition algorithms. Image and Vision Computing, 16:295–306, 1998.
3. E. Sarkar, et al. Vulnerability analysis of face morphing attacks from landmarks and generative adversarial networks. arXiv preprint, Oct. 2020.
4. P. J. Phillips, et al. Overview of the face recognition grand challenge. In CVPR, pages 947–954, 2005.
5. T. Neubert, et al. Extended StirTrace benchmarking of biometric and forensic qualities of morphed face images. IET Biometrics, 7(4):325–332, 2018.
6. H. Zhang, et al. MIPGAN—generating strong and high-quality morphing attacks using identity prior driven GAN. IEEE Transactions on Biometrics, Behavior, and Identity Science, 3(3):365–383, 2021.
7. Damer, C. A. F. López, M. Fang, N. Spiller, M. V. Pham, & F. Boutros. "Privacy-friendly synthetic data for face morphing attack detectors." CVPRW, 2022.
8. Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, S. Xie. "A ConvNet for the 2020s." CVPR, 2022, pp. 11976-11986.
9. Yoo, Y., Uh, Y., Chun, S., Kang, B., & Ha, J.-W. "Photorealistic style

Acknowledgment

This project is supported by the Center for Identification Technology Research (CITeR) and the National Science Foundation (NSF) under Grant #1650474