Gradient Alignment for Cross-domain Face Anti-Spoofing

Binh M. Le and Simon S. Woo*

Sungkyunkwan University, Suwon, South Korea

The IEEE/CVF Conference on Computer Vision and Pattern Recognition 2024

email: bmle[AT]g.skku.edu

Outline

1 – Introduction

2 – Methods

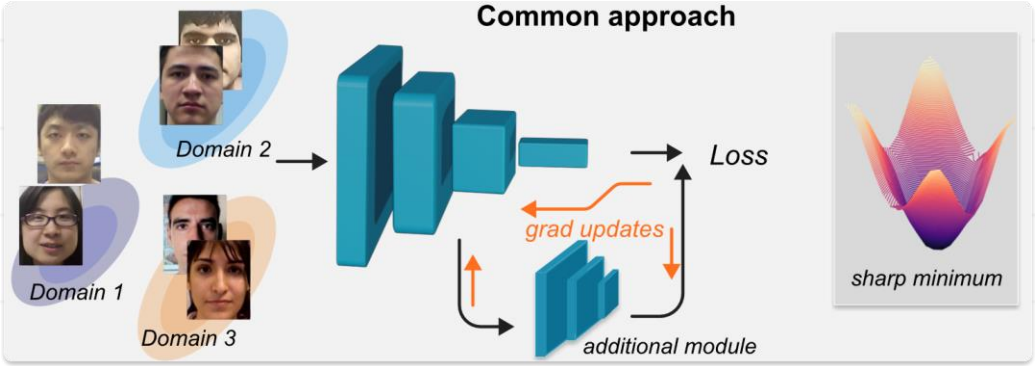
3 – Experimental Results

4 – Conclusion

Introduction

Motivation

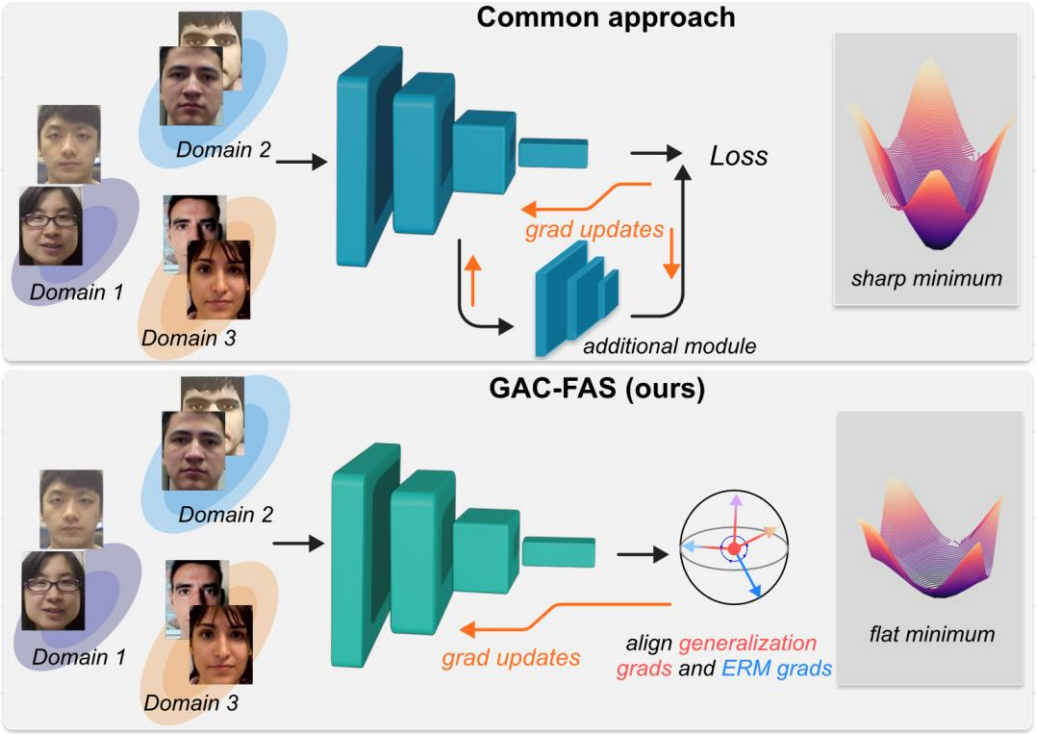
- Face recognition systems are vulnerable to presentation attacks.
- Recent studies have focused on improving DG for FAS by using auxiliary modules to learn domain-invariant features.
- Yet, they do not ensure convergence to a local flat minimum



Introduction

Motivation

- We develop a novel objective function that encourage the model to converge towards an optimal flat minimum without additional learning modules.



Methods

Notations

- Learning network f parameterized by θ
- Training source domains: $\boldsymbol{\mathcal{S}} = \{\mathcal{S}_i\}_{i=1}^{k}$
- ERM objective: $\min_{\theta} \mathcal{L}(\theta; \boldsymbol{\mathcal{S}}) = \min_{\theta} \mathbb{E}_{\mathcal{S}_i \sim \boldsymbol{\mathcal{S}}} \mathcal{L}(\theta; \mathcal{S}_i)$

Methods

Sharpness-Aware Minimization (SAM)

- SAM minimizes perturbation loss:

$$\mathcal{L}_p(\theta;\mathcal{D}) = \mathcal{L}(\theta + \hat{\epsilon};\mathcal{D}), \qquad where\ \hat{\epsilon} = \frac{\nabla\mathcal{L}(\theta;\mathcal{D})}{\|\nabla\mathcal{L}(\theta;\mathcal{D})\|}$$

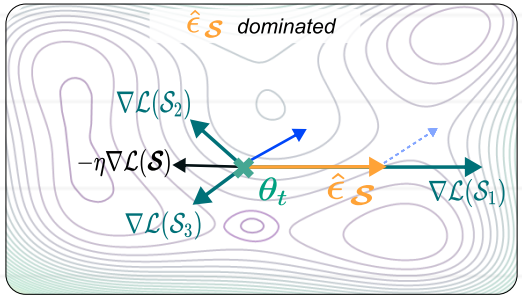
- We denote $\theta + \hat{\epsilon}$ as ascending point and $\hat{\epsilon}$ as ascending vector.

Methods

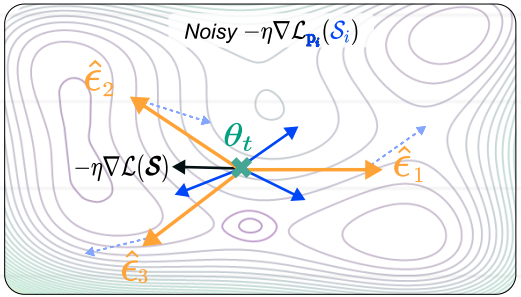
SAM-based approach's limitations

- Dominated $\hat{\epsilon}_{\mathcal{S}}$ in all-domain application ($\mathcal{D} = \mathcal{S}$),
- Gradient conflicts in domain-wise application ($\mathcal{D} = \mathcal{S}_i$).

See dashed blue arrows in the figures



(a) SAM on whole data



(b) SAM on each domain

Current model weight (current point) | Gradient direction | Ascending vector | ERM gradient update at current point | Gradient update at ascending point | Gradient update at ascending point translated to current point

Methods

GAC-FAS: Our refined objective

- Achieving optimal minimum: sufficient low and reside on a flat loss surface,
- Aligning cross-domain gradients: generalization gradient updates align with ERM gradient.

$$\mathcal{L}(\theta; \boldsymbol{S}) + \mathbb{E}_{\mathcal{S}_i \sim \boldsymbol{S}}[\mathcal{L}(\theta + \widehat{\epsilon_i} - \gamma \nabla \mathcal{L}(\theta; \boldsymbol{S}); \boldsymbol{S})] + \mathcal{R}(\theta; \boldsymbol{S})$$

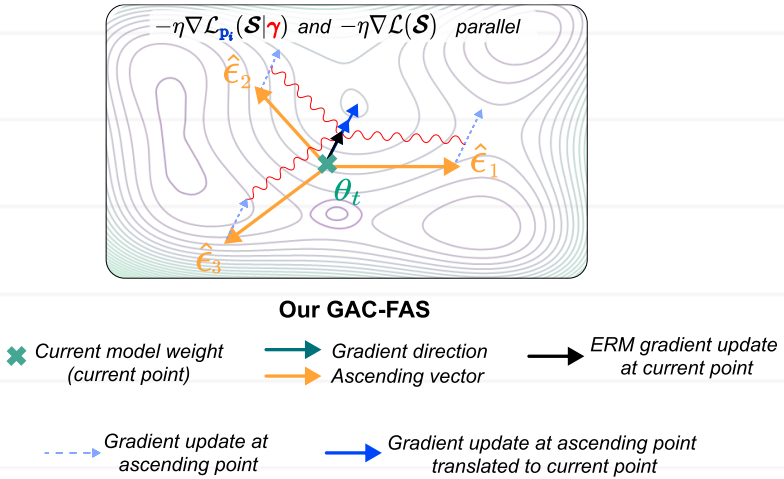
ERM | Generalization | Regularization

Methods

GAC-FAS: Benefits for DG

- Matching generalization gradient updates and ERM gradient updates.

$$\mathcal{L}(\theta;\boldsymbol{S}) + \mathbb{E}_{\mathcal{S}_i\sim\boldsymbol{S}}\left[\mathcal{L}_{p_i}(\theta;\boldsymbol{S}) - \gamma\left\langle\nabla\mathcal{L}_{p_i}(\theta;\boldsymbol{S}), \nabla\mathcal{L}(\theta;\boldsymbol{S})\right\rangle\right] + \mathcal{R}(\theta;\boldsymbol{S})$$

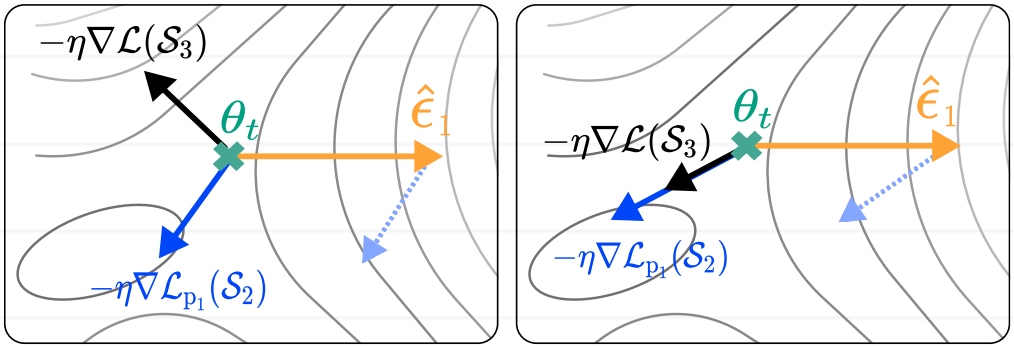


GAC-FAS: Benefits for DG

- Matching generalization gradient updates and ERM gradient updates.

$$\mathcal{L}(\theta;\boldsymbol{S}) + \mathbb{E}_{\mathcal{S}_i\sim\boldsymbol{S}}\left[\mathcal{L}_{p_i}(\theta;\boldsymbol{S}) - \gamma\left\langle\nabla\mathcal{L}_{p_i}(\theta;\boldsymbol{S}), \nabla\mathcal{L}(\theta;\boldsymbol{S})\right\rangle\right] + \mathcal{R}(\theta;\boldsymbol{S})$$

$$\sum_{m=1}^{k}\sum_{n=1}^{k}\left\langle\nabla\mathcal{L}_{p_i}(\theta;\mathcal{S}_m), \nabla\mathcal{L}(\theta;\mathcal{S}_n)\right\rangle$$



w/o our regularization

with our regularization

Experimental results

Setting up

Datasets (4): `Idiap Replay Attack` (**I**), `OULU-NPU` (**O**), `CASIA-MFSD` (**C**), `MSU-MFSD` (**M**).

Backbone (1): ResNet18.

Metric (2) : HTER↓ and AUC↑.

Experimental results

Results

Leave-one-out

Methods	ICM→O		OCM→I		OCI→M		OMI→C	
	HTER↓	AUC↑	HTER↓	AUC↑	HTER↓	AUC↑	HTER↓	AUC↑
MMD-AAE	40.98	63.08	31.58	75.18	27.08	83.19	44.59	58.29
MADDG	27.98	80.02	22.19	84.99	17.69	88.06	24.50	84.51
RFM	16.45	91.16	17.30	90.48	13.89	93.98	20.27	88.16
SSDG-M	25.17	81.83	18.21	94.61	16.67	90.47	23.11	85.45
SSDG-R	15.61	91.54	11.71	96.59	7.38	97.17	10.44	95.94
D2AM	15.27	90.87	15.43	91.22	12.70	95.66	20.98	85.58
SDA	23.10	84.30	15.60	90.10	15.40	91.80	24.50	84.40
DRDG	15.63	91.75	15.56	91.79	12.43	95.81	19.05	88.79
ANRL	15.67	91.90	16.03	91.04	10.83	96.75	17.85	89.26
SSAN	13.72	93.63	8.88	96.79	6.67	98.75	10.00	96.67
AMEL	11.31	93.96	18.60	88.79	10.23	96.62	11.88	94.39
EBDG	15.66	92.02	18.69	92.28	9.56	97.17	18.34	90.01
PathNet	11.82	95.07	13.40	95.67	7.10	98.46	11.33	94.58
IADG	8.86	97.14	10.62	94.50	5.41	98.19	8.70	96.40
SA-FAS	10.00	96.23	6.58	97.54	5.95	96.55	8.78	95.37
UDG-FAS	10.97	95.36	5.86	98.62	5.95	**98.47**	9.82	**96.76**
GAC-FAS	**8.60**$^{0.28}$	**97.16**$^{.40}$	**4.29** $^{.83}$	**98.87** $^{.60}$	**5.00** $^{.00}$	97.56 $^{.06}$	**8.20**$^{0.43}$	95.16$^{.09}$

Limited source domains

Methods	MI→C		MI→O	
	HTER↓	AUC↑	HTER↓	AUC↑
MSLBP	51.16	52.09	43.63	58.07
Color Texture	55.17	46.89	53.31	45.16
LBPTOP	45.27	54.88	47.26	50.21
MADDG	41.02	64.33	39.35	65.10
SSDG-M	31.89	71.29	36.01	66.88
D2AM	32.65	72.04	27.70	75.36
DRDG	31.28	71.50	33.35	69.14
ANRL	31.06	72.12	30.73	74.10
SSAN	30.00	76.20	29.44	76.62
EBDG	27.97	75.84	25.94	78.28
AMEL	24.52	82.12	19.68	87.01
IAGD	24.07	85.13	18.47	**90.49**
GAC-FAS	**16.91**$^{1.17}$	**88.12**$^{.58}$	**17.88**$^{.15}$	89.67$^{.39}$

Experimental results

Results

Unseen attacks

Methods	AUC↑
SVM1+IMQ	70.23 $^{12.69}$
CDCN	88.69 $^{10.56}$
CDCN++	87.53 $^{10.90}$
SSAN	88.01 $^{9.93}$
TTN-S	89.7 $^{19.17}$
UDG-FAS	92.43 $^{6.86}$
GAC-FAS	**93.39** $^{4.27}$

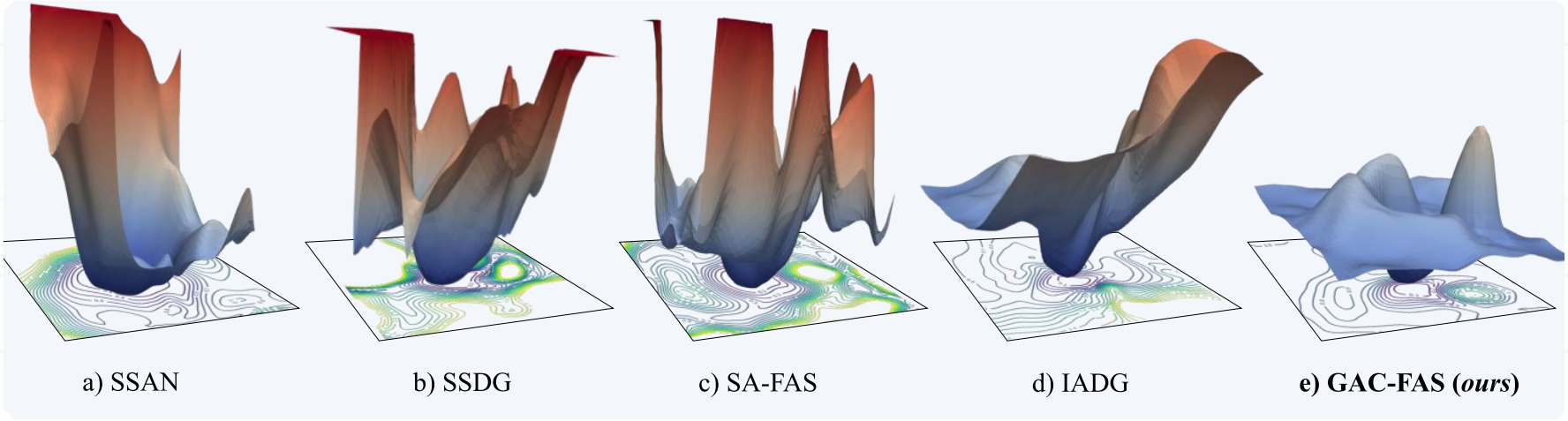
a) 2D attack

Methods	AUC↑
Saha *et al.*	79.20
Panwar *et al.*	80.00
SSDG-R	82.11
CIFAS	83.20
UDG-FAS	87.26
GAC-FAS	**89.27** $^{.58}$

b) 3D attack

Experimental results

Ablation Studies



a) SSAN b) SSDG c) SA-FAS d) IADG **e) GAC-FAS (*ours*)**

Conclusion

- Inspired by recent studies of SAM, we introduce a novel objective to optimize the minimum for DG in FAS.
- Our method regulates SAM generalization gradients of whole data at ascending points to be aligned with gradients derived from ERM.
- Comprehensive analysis demonstrates the generalization of GAC-FAS across evaluation settings.

Thank you!