# Multi-adversarial Discriminative Deep Domain Generalization for Face Presentation Attack Detection


香港浸會大學
HONG KONG BAPTIST UNIVERSITY

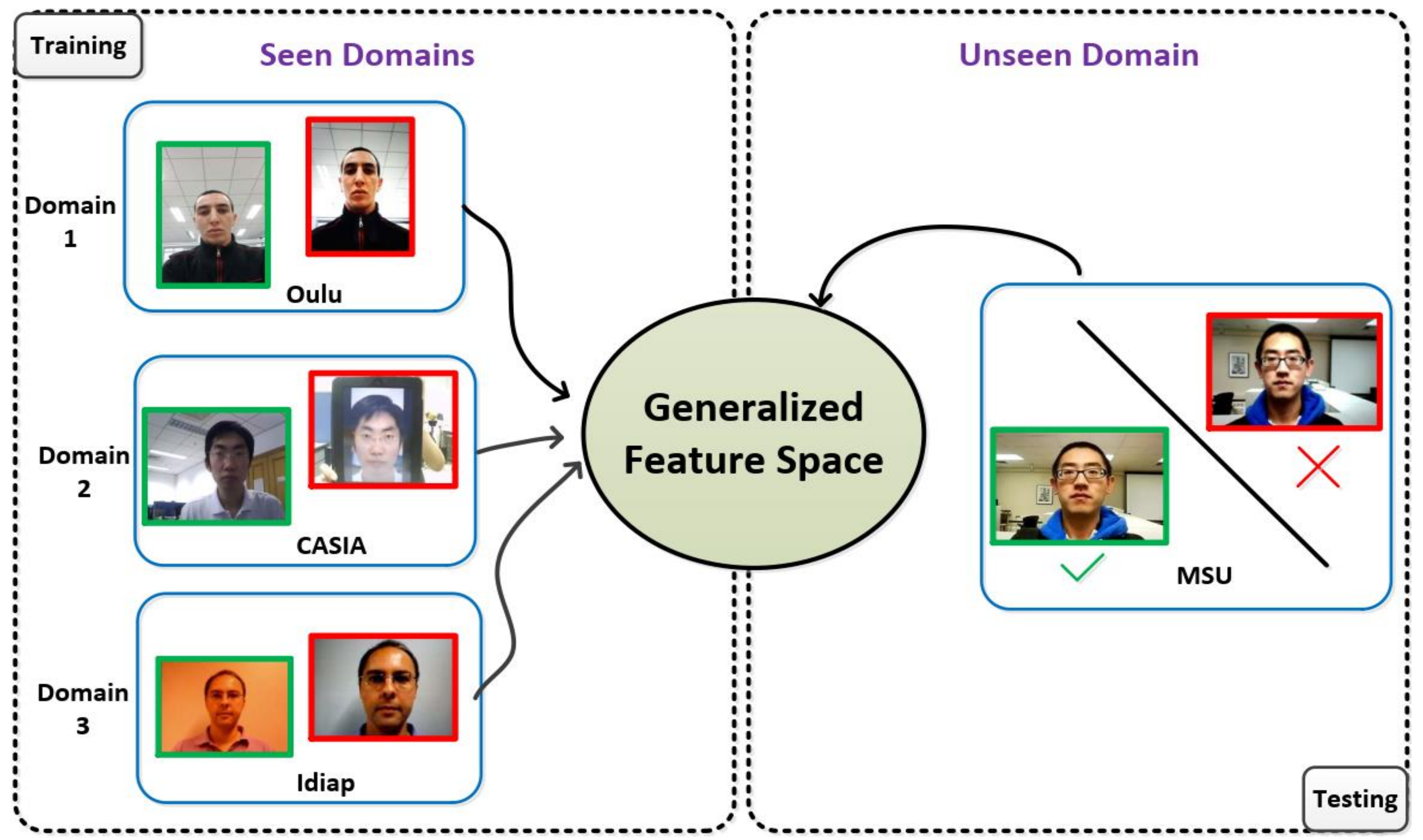
Rui Shao, Xiangyuan Lan, Jiawei Li, Pong C. Yuen

Department of Computer Science, Hong Kong Baptist University



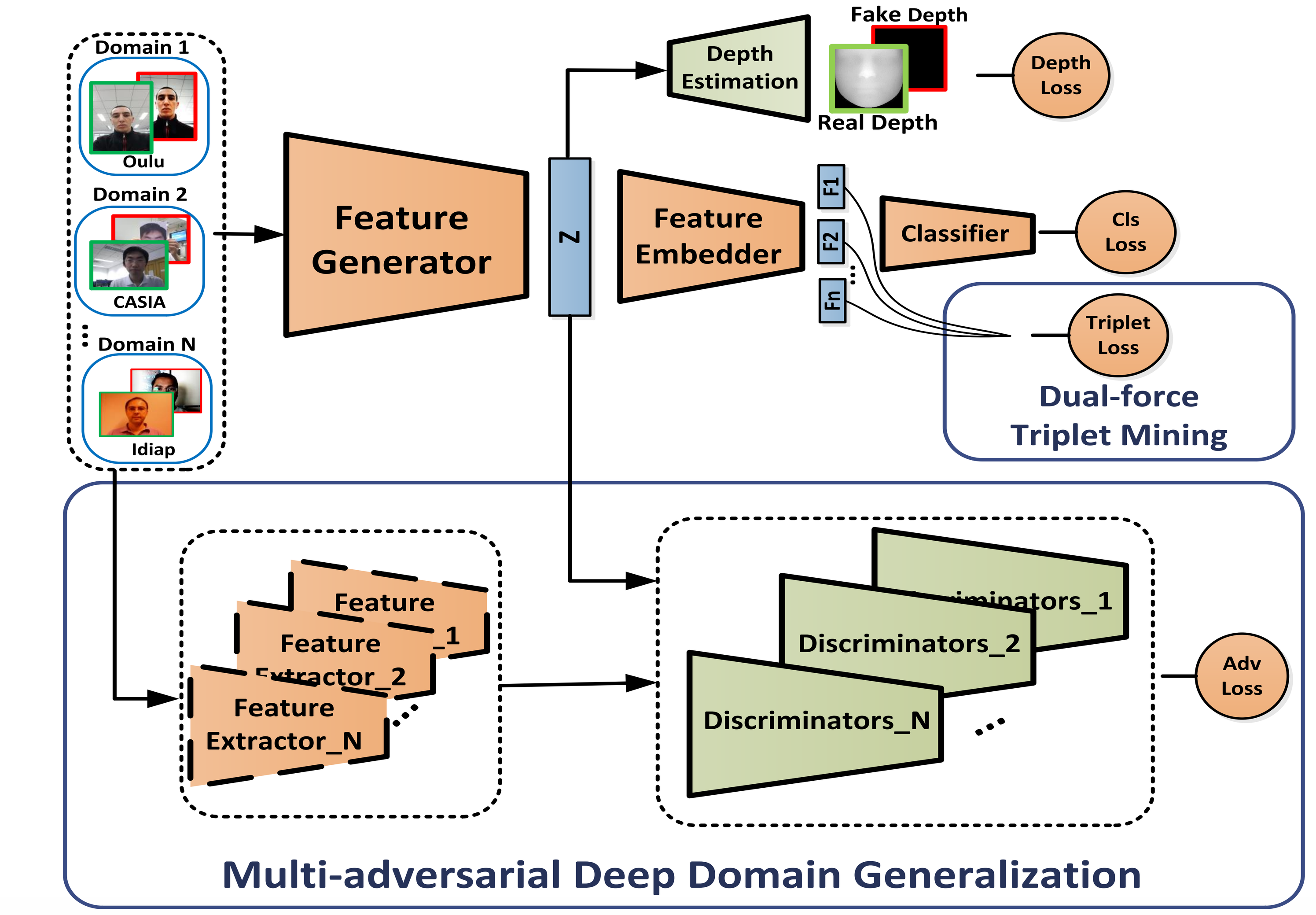
CVPR LONG BEACH CALIFORNIA June 16-20, 2019


## Introduction




- Improving the generalization ability of face anti-spoofing methods from the perspective of the domain generalization.
- Learning a generalized feature space that is shared and discriminative.

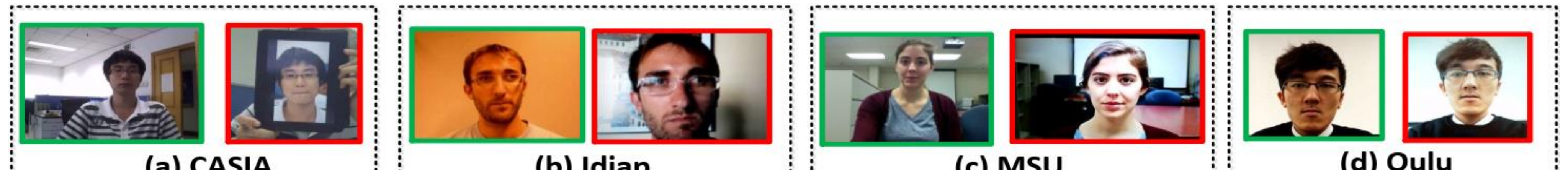
## Framework




$$\min_{G,E,C,Dep}\ \max_{D_1,D_2,\ldots,D_N}\ \mathcal{L}_{MADDG} = \mathcal{L}_{DG} + \mathcal{L}_{Trip} + \mathcal{L}_{Dep} + \mathcal{L}_{Cls}$$

## Experimental Results

### Datasets



(a) CASIA (b) Idiap (c) MSU (d) Oulu

### Comparison Results

| Methods | O&C&I to M | | O&M&I to C | | O&C&M to I | | I&C&M to O | |
|---|---|---|---|---|---|---|---|---|
| | HETR | AUC | HETR | AUC | HETR | AUC | HETR | AUC |
| MS_LBP | 29.76 | 78.50 | 54.28 | 44.98 | 50.30 | 51.64 | 50.29 | 49.31 |
| B_CNN | 29.25 | 82.87 | 34.88 | 71.94 | 34.47 | 65.88 | 29.61 | 77.54 |
| IDA | 66.67 | 27.86 | 55.17 | 39.05 | 28.35 | 78.25 | 54.20 | 44.59 |
| CT | 28.09 | 78.47 | 30.58 | 76.89 | 40.40 | 62.78 | 63.59 | 32.71 |
| LBPTOP | 36.90 | 70.80 | 42.60 | 61.05 | 49.45 | 49.54 | 53.15 | 44.09 |
| Auxiliary(Depth) | 22.72 | 85.88 | 33.52 | 73.15 | 29.14 | 71.69 | 30.17 | 77.61 |
| MMD_AAE | 27.08 | 83.19 | 44.59 | 58.29 | 31.58 | 75.18 | 40.98 | 63.08 |
| **Ours** | **17.69** | **88.06** | **24.5** | **84.51** | **22.19** | **84.99** | **27.98** | **80.02** |

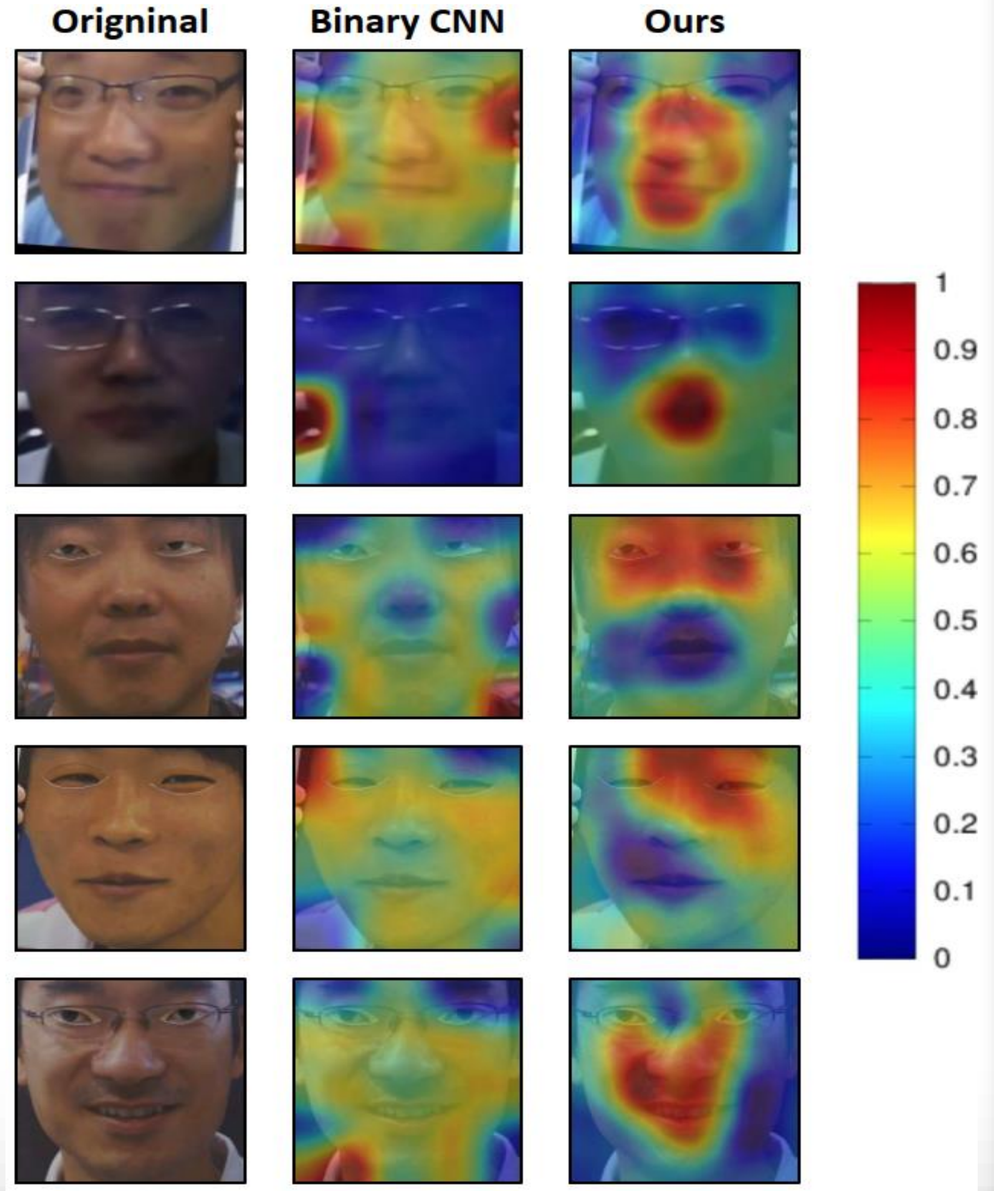
### Ablation Study

| Methods | O&C&M to I | |
|---|---|---|
| | HETR | AUC |
| Ours_wo/mgan | 36.50 | 63.15 |
| Ours_wo/trip | 34.99 | 71.37 |
| Ours_wo/dep | 37.44 | 62.82 |
| **Ours** | **22.19** | **84.99** |

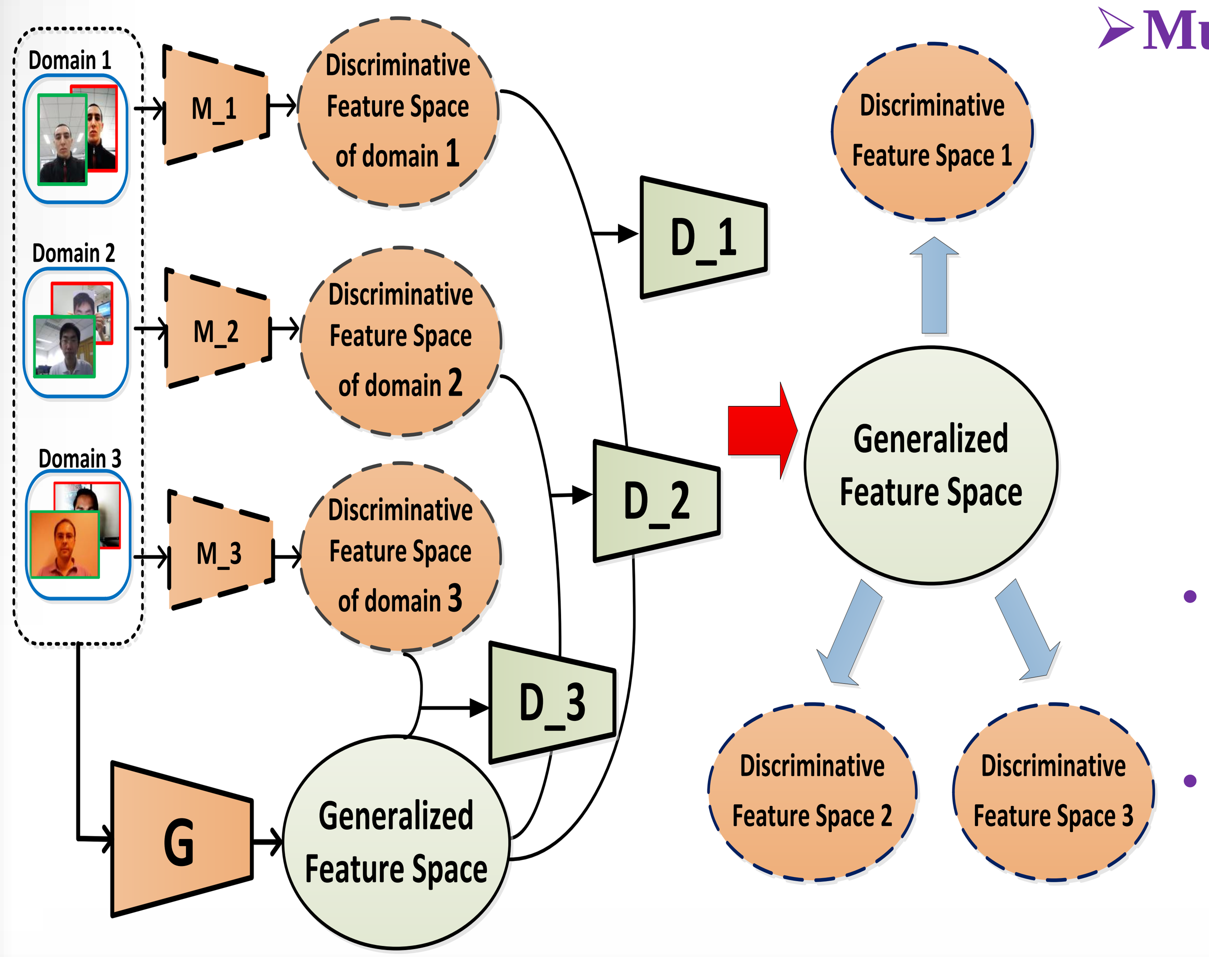
### Limited source domains

| Methods | M&I to C | | M&I to O | |
|---|---|---|---|---|
| | HETR | AUC | HETR | AUC |
| MS_LBP | 51.16 | 52.09 | 43.63 | 58.07 |
| IDA | 45.16 | 58.8 | 54.52 | 42.17 |
| LBPTOP | 45.27 | 54.88 | 47.26 | 50.21 |
| **Ours** | **41.02** | **64.33** | **39.35** | **65.10** |

### Attention Map



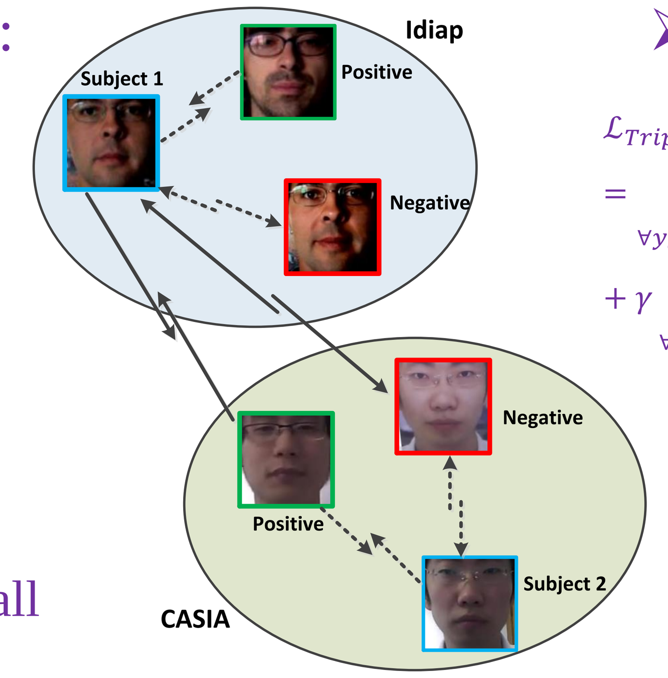

## Network Components

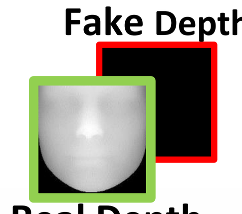



### Multi-adversarial Domain Generalization :

$$\mathcal{L}_{DG} = (\mathbf{X}, \mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_N; G, D_1, D_2, \ldots, D_N) = \sum_{i=1}^{N}\left( \begin{array}{c} \mathbb{E}_{x\sim\mathbf{X}}\left[\log\left(D_i(G(x))\right)\right] \\ +\mathbb{E}_{x_i\sim\mathbf{X}_i}\left[\log\left(1 - D_i(M_i(x_i))\right)\right] \end{array}\right)$$

- Train one feature generator to compete with all the N domain discriminators simultaneously.
- A shared feature space is learned after one feature generator fools all the N domain discriminators.




### Dual-force Triplet-mining Constraint

$$\mathcal{L}_{Trip} = (\mathbf{X}, \mathbf{Y}; G, E) = \sum_{\forall y_a = y_p, \forall y_a \neq y_n, i=j}\left[\left\|E(G(x_i^a)) - E(G(x_j^p))\right\|_2^2 - \left\|E(G(x_i^a)) - E\left(G(x_j^n)\right) + \alpha_1\right\|_2^2\right]_+ + \gamma \sum_{\forall y_a = y_p, \forall y_a \neq y_n, i\neq k}\left[\left\|E(G(x_i^a)) - E(G(x_k^p))\right\|_2^2 - \left\|E(G(x_i^a)) - E(G(x_k^n)) + \alpha_2\right\|_2^2\right]_+$$

- Fake face with the same identity has similar facial characteristics; real face with the different identity has different facial characteristics.
- Distance of each subject to its intra/cross-domain positive smaller than to its intra/cross-domain negative.

### Auxiliary Face Depth Information

$$\mathcal{L}_{Dep}(X; Dep) = \left\|Dep(G(X)) - I\right\|_2^2$$

- Feature space guided to exploit generalized differentiation cues related to the face depth in the learning process.