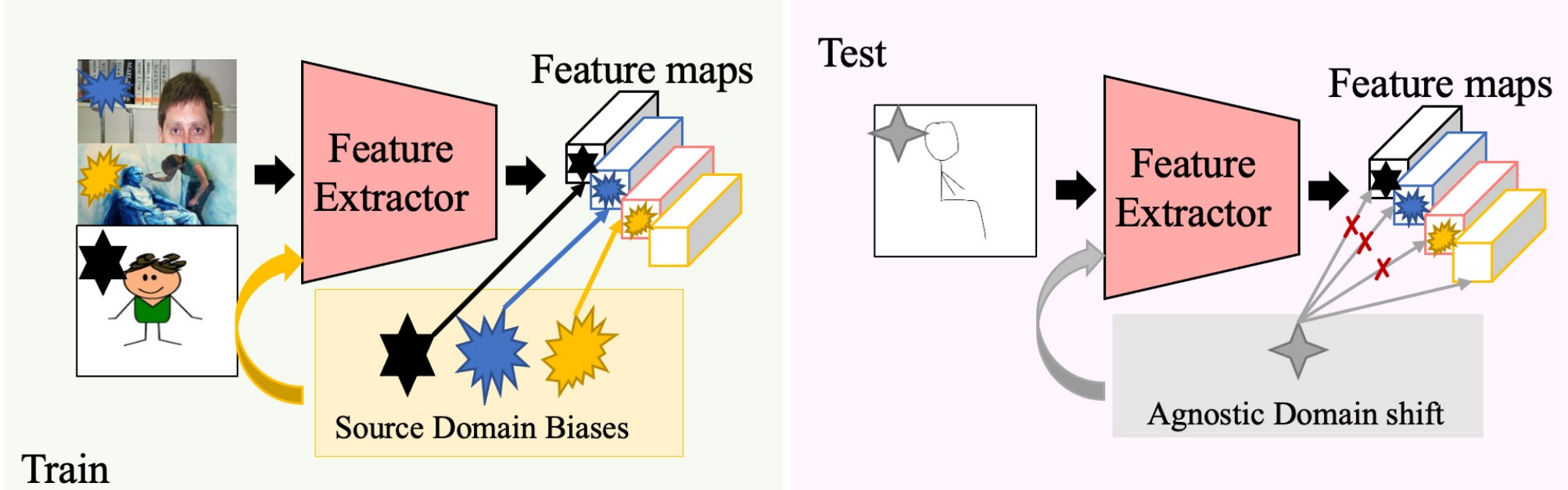




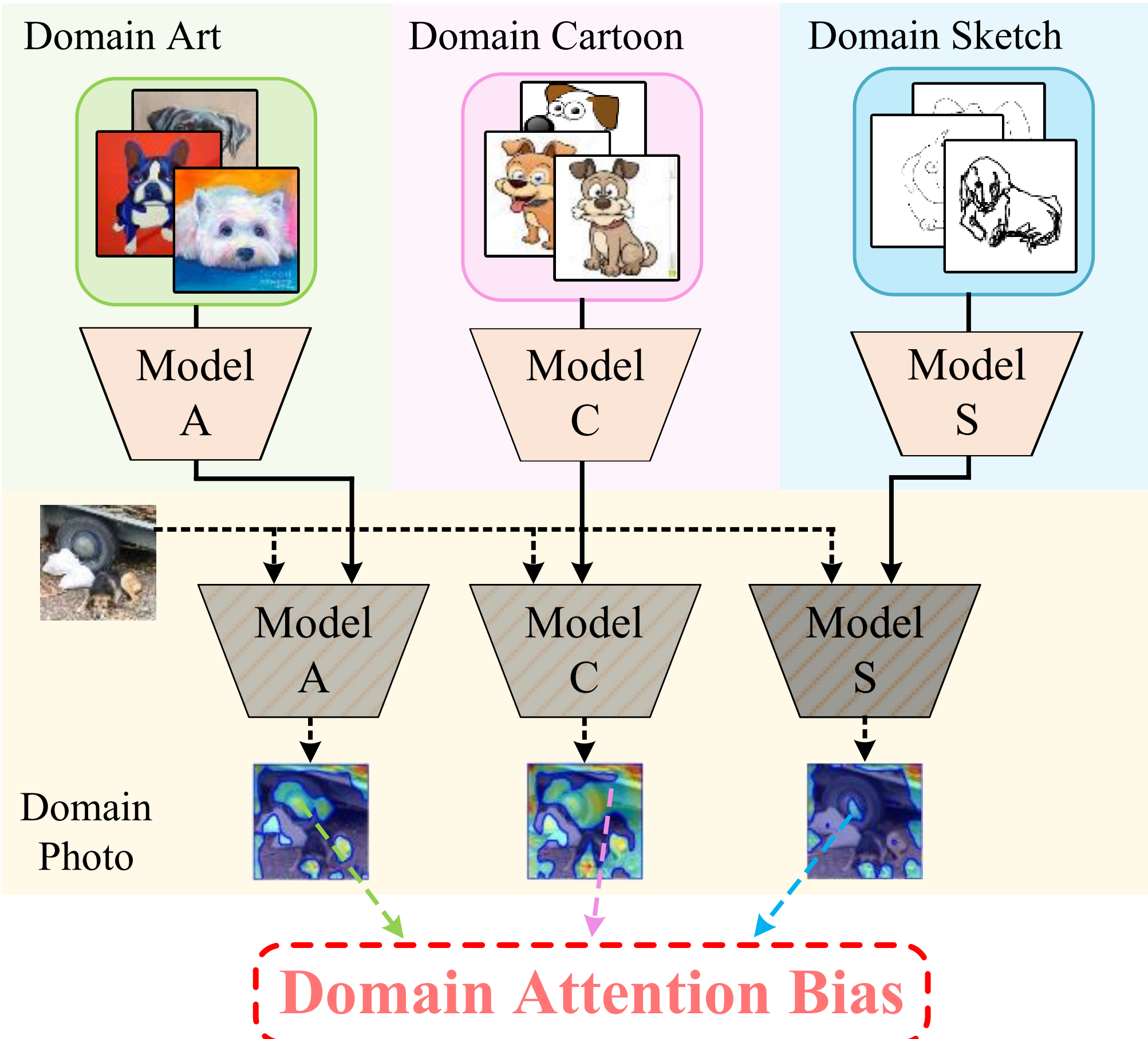
Motivation

Background

Deep models trained on seen domains perform poorly on unseen domains for the seen/unseen domain shifts.



Our Finding of Domain Attention Bias



We find that models trained on different domains have different attention bias

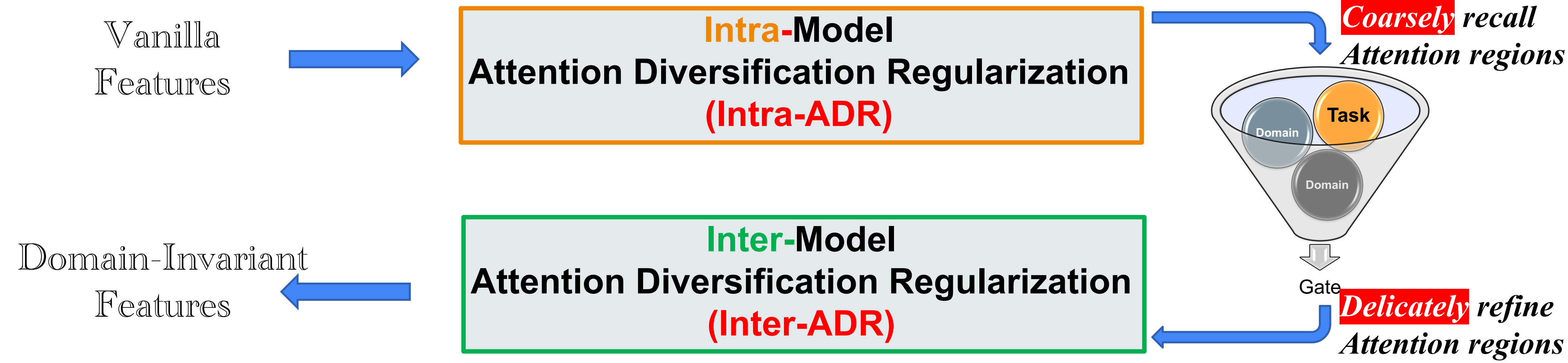
Our Revisiting of DG from Maximum Entropy Principle

Maximum Entropy Principle: *when estimating the probability distribution, we should select that distribution with the largest uncertainty under given constraints.*

Our Insight: *when testing the unseen domains, each task-related attention is equally-useful (the maximum entropy)*

Methodology

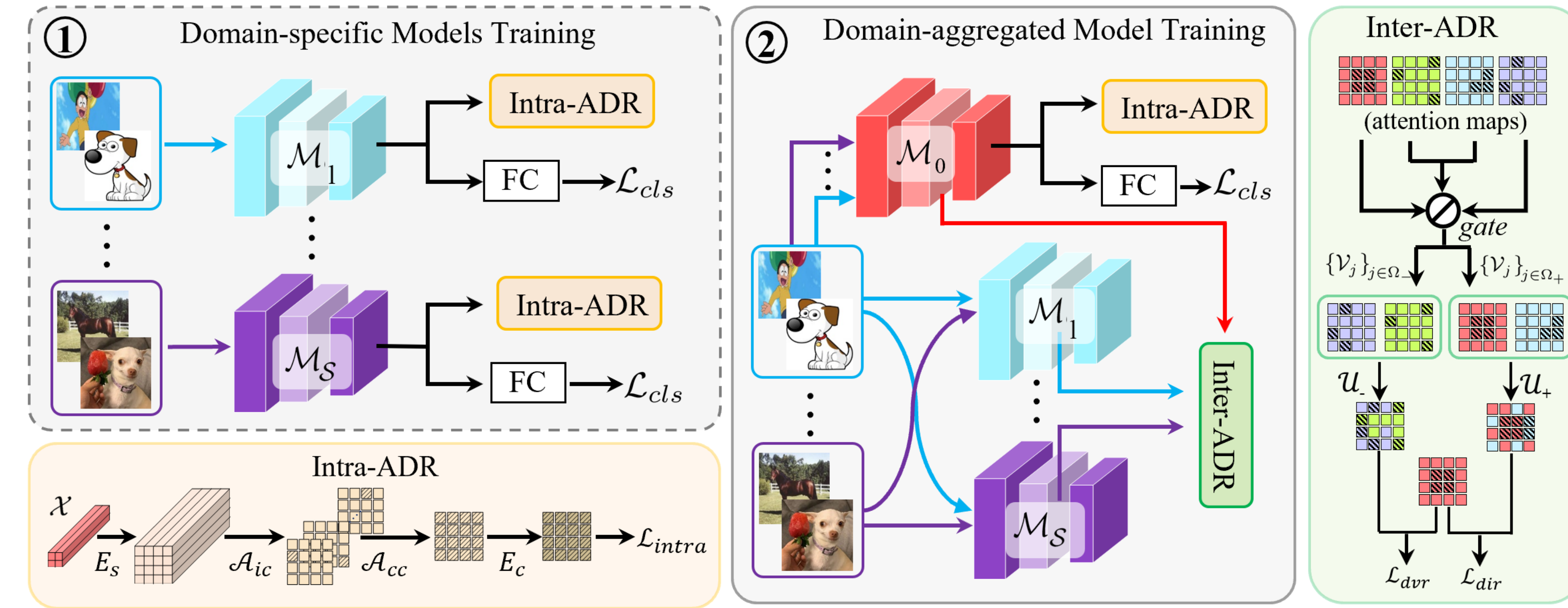
Workflow of Our Attention Diversification Framework



Training Scheme of Our Attention Diversification Framework

Intra-ADR forces different channels to concern on different regions and activates all the regions.

Inter-ADR uses a paradigm of “simulate, divide and assemble”.



① Domain-specific Models Training → Warmup step equipped with Intra-ADR and for “simulate” part in Inter-ADR.

② Domain-aggregated Model Training → Main step equipped with Intra- and Inter-ADR for coarse-to-fine Attention Diversification.

Experiments

Results

PACS

| Methods | References | Art | Cartoon | Photo | Sketch | Avg. | Art | Cartoon | Photo | Sketch | Avg. |
|--------------------------------|------------|------|---------|-------|--------|------|------|---------|-------|--------|------|
| <i>ResNet-18</i> | | | | | | | | | | | |
| Baseline | - | 79.0 | 74.3 | 94.9 | 71.4 | 79.9 | 86.2 | 78.7 | 97.6 | 70.6 | 83.2 |
| MetaReg [1] | NeurIPS'18 | 83.7 | 77.2 | 95.5 | 70.3 | 81.7 | 87.2 | 79.2 | 97.6 | 70.3 | 83.6 |
| MASF [12] | NeurIPS'19 | 80.2 | 77.1 | 94.9 | 71.6 | 81.0 | 82.8 | 80.4 | 95.0 | 72.2 | 82.6 |
| Epi-FCR [26] | ICCV'19 | 82.1 | 77.0 | 93.9 | 73.0 | 81.5 | - | - | - | - | - |
| JiGen [4] | CVPR'19 | 79.4 | 75.2 | 96.0 | 71.3 | 80.5 | - | - | - | - | - |
| DMG [5] | ECCV'20 | 76.9 | 80.4 | 93.4 | 75.2 | 81.5 | 82.6 | 78.1 | 94.5 | 78.3 | 83.4 |
| RSC [21] | ECCV'20 | 84.4 | 80.3 | 95.9 | 80.8 | 85.1 | 87.8 | 82.1 | 97.9 | 83.3 | 87.9 |
| MixStyle [78] | ICLR'21 | 84.1 | 78.8 | 96.1 | 75.9 | 83.7 | - | - | - | - | - |
| SelfReg [24] | ICCV'21 | 82.3 | 78.4 | 96.2 | 77.5 | 83.6 | 87.9 | 79.4 | 96.8 | 78.3 | 85.6 |
| DAML [50] | CVPR'21 | 83.0 | 74.1 | 95.6 | 78.1 | 82.7 | - | - | - | - | - |
| SagNet [42] | CVPR'21 | 83.6 | 77.7 | 95.5 | 76.3 | 83.3 | 81.1 | 75.4 | 95.7 | 77.2 | 82.3 |
| FACT [65] | CVPR'21 | 85.4 | 78.4 | 95.2 | 79.2 | 84.5 | 89.6 | 81.7 | 96.8 | 84.4 | 88.1 |
| Intra-ADR | Ours | 82.4 | 79.4 | 95.3 | 82.3 | 84.9 | 87.7 | 81.2 | 97.1 | 83.8 | 87.5 |
| I ² -ADR | Ours | 82.9 | 80.8 | 95.3 | 82.3 | 85.6 | 88.5 | 83.2 | 95.2 | 85.8 | 88.2 |
| MixStyle + Intra-ADR | Ours | 86.0 | 80.3 | 96.0 | 84.4 | 86.7 | 88.6 | 83.2 | 98.0 | 85.2 | 88.7 |
| MixStyle + I ² -ADR | Ours | 85.3 | 81.2 | 95.4 | 86.1 | 87.0 | 87.7 | 84.5 | 98.2 | 85.6 | 89.2 |
| <i>ResNet-50</i> | | | | | | | | | | | |
| Baseline | - | 79.0 | 74.3 | 94.9 | 71.4 | 79.9 | 86.2 | 78.7 | 97.6 | 70.6 | 83.2 |
| MetaReg [1] | NeurIPS'18 | 83.7 | 77.2 | 95.5 | 70.3 | 81.7 | 87.2 | 79.2 | 97.6 | 70.3 | 83.6 |
| MASF [12] | NeurIPS'19 | 80.2 | 77.1 | 94.9 | 71.6 | 81.0 | 82.8 | 80.4 | 95.0 | 72.2 | 82.6 |
| Epi-FCR [26] | ICCV'19 | 82.1 | 77.0 | 93.9 | 73.0 | 81.5 | - | - | - | - | - |
| JiGen [4] | CVPR'19 | 79.4 | 75.2 | 96.0 | 71.3 | 80.5 | - | - | - | - | - |
| DMG [5] | ECCV'20 | 76.9 | 80.4 | 93.4 | 75.2 | 81.5 | 82.6 | 78.1 | 94.5 | 78.3 | 83.4 |
| RSC [21] | ECCV'20 | 84.4 | 80.3 | 95.9 | 80.8 | 85.1 | 87.8 | 82.1 | 97.9 | 83.3 | 87.9 |
| MixStyle [78] | ICLR'21 | 84.1 | 78.8 | 96.1 | 75.9 | 83.7 | - | - | - | - | - |
| SelfReg [24] | ICCV'21 | 82.3 | 78.4 | 96.2 | 77.5 | 83.6 | 87.9 | 79.4 | 96.8 | 78.3 | 85.6 |
| DAML [50] | CVPR'21 | 83.0 | 74.1 | 95.6 | 78.1 | 82.7 | - | - | - | - | - |
| SagNet [42] | CVPR'21 | 83.6 | 77.7 | 95.5 | 76.3 | 83.3 | 81.1 | 75.4 | 95.7 | 77.2 | 82.3 |
| FACT [65] | CVPR'21 | 85.4 | 78.4 | 95.2 | 79.2 | 84.5 | 89.6 | 81.7 | 96.8 | 84.4 | 88.1 |
| Intra-ADR | Ours | 82.4 | 79.4 | 95.3 | 82.3 | 84.9 | 87.7 | 81.2 | 97.1 | 83.8 | 87.5 |
| I ² -ADR | Ours | 82.9 | 80.8 | 95.3 | 82.3 | 85.6 | 88.5 | 83.2 | 95.2 | 85.8 | 88.2 |
| MixStyle + Intra-ADR | Ours | 86.0 | 80.3 | 96.0 | 84.4 | 86.7 | 88.6 | 83.2 | 98.0 | 85.2 | 88.7 |
| MixStyle + I ² -ADR | Ours | 85.3 | 81.2 | 95.4 | 86.1 | 87.0 | 87.7 | 84.5 | 98.2 | 85.6 | 89.2 |

Office Home

| Methods | Ar | Cl | Pr | Rw | Avg. | Methods | Ar | Cl | Pr | Rw | Avg. |
|--------------------------------|------|------|------|------|------|--------------------------------|------|------|------|------|------|
| <i>ResNet-18</i> | | | | | | | | | | | |
| Baseline | 57.8 | 52.7 | 73.5 | 74.8 | 64.7 | Baseline | 61.3 | 52.4 | 75.8 | 76.6 | 66.5 |
| RSC [21] | 58.4 | 47.9 | 71.6 | 74.5 | 63.1 | MLDG [28] | 61.5 | 53.2 | 75.0 | 77.5 | 66.8 |
| MixStyle [78] | 58.7 | 53.4 | 74.2 | 75.9 | 65.5 | RSC [21] | 50.7 | 51.4 | 74.8 | 75.1 | 65.5 |
| SagNet [42] | 60.2 | 45.4 | 70.4 | 73.4 | 62.3 | SelfReg [24] | 63.6 | 53.1 | 76.9 | 78.1 | 67.9 |
| FACT [65] | 60.3 | 54.9 | 74.5 | 76.6 | 66.6 | SagNet [42] | 63.4 | 54.8 | 75.8 | 78.3 | 68.1 |
| Intra-ADR | 64.5 | 54.0 | 73.9 | 74.7 | 66.8 | Intra-ADR | 67.3 | 54.1 | 78.8 | 78.8 | 69.8 |
| I ² -ADR | 66.4 | 53.3 | 74.9 | 75.3 | 67.5 | I ² -ADR | 70.3 | 55.1 | 80.7 | 79.2 | 71.4 |
| MixStyle + Intra-ADR | 65.9 | 55.3 | 74.3 | 75.1 | 67.7 | MixStyle + Intra-ADR | 69.5 | 55.9 | 80.6 | 80.4 | 71.4 |
| MixStyle + I ² -ADR | 66.8 | 56.8 | 75.3 | 75.7 | 68.7 | MixStyle + I ² -ADR | 71.1 | 56.9 | 81.8 | 80.5 | 72.5 |
| <i>ResNet-50</i> | | | | | | | | | | | |
| Baseline | 57.8 | 52.7 | 73.5 | 74.8 | 64.7 | Baseline | 61.3 | 52.4 | 75.8 | 76.6 | 66.5 |
| RSC [21] | 58.4 | 47.9 | 71.6 | 74.5 | 63.1 | MLDG [28] | 61.5 | 53.2 | 75.0 | 77.5 | 66.8 |
| MixStyle [78] | 58.7 | 53.4 | 74.2 | 75.9 | 65.5 | RSC [21] | 50.7 | 51.4 | 74.8 | 75.1 | 65.5 |
| SagNet [42] | 60.2 | 45.4 | 70.4 | 73.4 | 62.3 | SelfReg [24] | 63.6 | 53.1 | 76.9 | 78.1 | 67.9 |
| FACT [65] | 60.3 | 54.9 | 74.5 | 76.6 | 66.6 | SagNet [42] | 63.4 | 54.8 | 75.8 | 78.3 | 68.1 |
| Intra-ADR | 64.5 | 54.0 | 73.9 | 74.7 | 66.8 | Intra-ADR | 67.3 | 54.1 | 78.8 | 78.8 | 69.8 |
| I ² -ADR | 66.4 | 53.3 | 74.9 | 75.3 | 67.5 | I ² -ADR | 70.3 | 55.1 | 80.7 | 79.2 | 71.4 |
| MixStyle + Intra-ADR | 65.9 | 55.3 | 74.3 | 75.1 | 67.7 | MixStyle + Intra-ADR | 69.5 | 55.9 | 80.6 | 80.4 | 71.4 |
| MixStyle + I ² -ADR | 66.8 | 56.8 | 75.3 | 75.7 | 68.7 | MixStyle + I ² -ADR | 71.1 | 56.9 | 81.8 | 80.5 | 72.5 |

DomainNet

| Methods | References | Clipart | Infograph | Painting | Quickdraw | Real | Sketch | Avg. |
|--------------------------------|------------|----------|-----------|----------|-----------|----------|----------|------|
| <i>ResNet-18</i> | | | | | | | | |
| Baseline | - | 57.1 | 17.6 | 43.2 | 13.8 | 54.9 | 39.4 | 37.6 |
| MetaReg [1] | NeurIPS'18 | 53.7 | 21.1 | 45.3 | 10.6 | 58.5 | 42.3 | 38.6 |
| DMG [5] | ECCV'20 | 60.1 | 18.8 | 44.5 | 14.2 | 54.7 | 41.7 | 39.0 |
| Intra-ADR | Ours | 57.3±0.1 | 14.9±0.3 | 42.8±0.2 | 12.2±0.4 | 52.9±0.5 | 46.0±0.2 | 37.7 |
| I ² -ADR | Ours | 57.3±0.3 | 15.2±0.3 | 44.1±0.1 | 12.1±0.4 | 53.9±0.6 | 46.7±0.2 | 38.2 |
| MixStyle + Intra-ADR | Ours | 57.4±0.2 | 15.3±0.1 | 43.3±0.2 | 12.3±0.4 | 53.5±0.3 | 46.5±0.2 | 38.1 |
| MixStyle + I ² -ADR | Ours | 57.4±0.4 | 15.7±0.2 | 44.7±0.1 | 12.3±0.4 | 54.4±0.2 | 47.4±0.1 | 38.7 |
| <i>ResNet-50</i> | | | | | | | | |
| Baseline | - | 62.2 | 19.9 | 45.5 | 13.8 | 57.5 | 44.4 | 40.5 |
| MetaReg [1] | NeurIPS'18 | 59.8 | 25.6 | 50.2 | 11.5 | 64.6 | 50.1 | 43.6 |
| MLDG [28] | AAAI'18 | 59.1±0.2 | 19.1±0.3 | 45.8±0.7 | 13.4±0.3 | 59.6±0.2 | 50.2±0.4 | 41.2 |
| C-DANN [31] | ECCV'18 | 54.6±0.4 | 17.3±0.1 | 43.7±0.9 | 12.1±0.7 | 56.2±0.4 | 45.9±0.5 | 38.3 |
| RSC [21] | ECCV'20 | 55.0±1.2 | 18.3±0.5 | 44.4±0.6 | 12.2±0.2 | 55.7±0.7 | 47.8±0.9 | 38.9 |
| DMG [5] | ECCV'20 | 65.2 | 22.2 | 50.0 | 15.7 | 59.6 | 49.0 | 43.6 |
| SagNet [42] | CVPR'21 | 57.7±0.3 | 19.0±0.2 | 45.3±0.3 | 12.7±0.5 | 58.1±0.5 | 48.8±0.2 | 40.3 |
| SelfReg [24] | ICCV'21 | 60.7±0.1 | 21.6±0.1 | 49.4±0.2 | 12.7±0.1 | 60.7±0.1 | 51.7±0.1 | 42.8 |
| Intra-ADR | Ours | 63.6±0.1 | 20.0±0.1 | 49.4±0.1 | 14.8±0.3 | 60.0±0.4 | 54.4±0.1 | 43.7 |
| I ² -ADR | Ours | 64.4±0.2 | 20.2±0.6 | 49.2±0.5 | 15.0±0.2 | 61.6±0.4 | 53.3±0.1 | 44.0 |
| MixStyle + Intra-ADR | Ours | 63.9±0.1 | 20.1±0.5 | 49.4±0.2 | 15.0±0.4 | 60.4±0.3 | 54.4±0.1 | 43.9 |
| MixStyle + I ² -ADR | Ours | 64.1±0.1 | 20.4±0.2 | 49.2±0.4 | 15.1±0.2 | 61.3±0.4 | 54.3±0.4 | 44.1 |

Ablation Study

| Method | \mathcal{L}_{intra} | \mathcal{L}_{div} | \mathcal{L}_{dir} | Art | Cartoon | Photo | Sketch | Avg. |
|---------------------|-----------------------|---------------------|---------------------|------|---------|-------|--------|------|
| I ² -ADR | ✓ | ✓ | ✓ | 82.4 | 79.4 | 95.3 | 82.3 | 84.9 |
| | ✓ | ✓ | ✗ | 82.3 | 80.0 | 95.1 | 82.6 | 85.0 |
| | ✓ | ✗ | ✓ | 82.7 | 80.5 | 95.0 | 83.2 | 85.4 |
| | ✓ | ✗ | ✗ | 82.5 | 80.2 | 95.1 | 82.9 | 85.2 |
| Intra-ADR | ✓ | ✓ | ✓ | 81.3 | 77.3 | 94.7 | 78.8 | 83.0 |
| | ✓ | ✓ | ✗ | 80.0 | 77.2 | 96.0 | 80.9 | 83.5 |
| | ✓ | ✗ | ✓ | 81.9 | 79.3 | 95.5 | 79.3 | 84.0 |
| | ✓ | ✗ | ✗ | 82.4 | 79.4 | 95.3 | 82.3 | 84.9 |

Attention Visualization

