Learning Domain-Invariant Relationship with Instrumental Variable for Domain Generalization

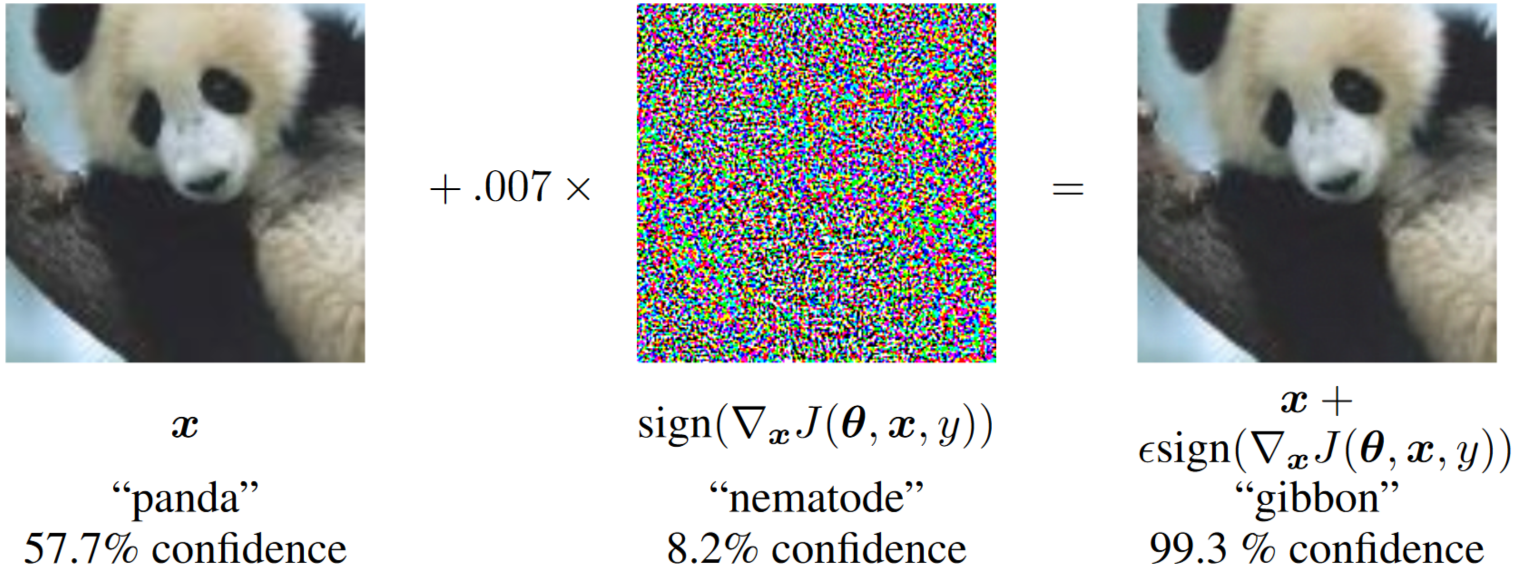
Junkun Yuan[1], Xu Ma[1] , Kun Kuang[1] , Ruoxuan Xiong[2] , Mingming Gong[3] , and Lanfen Lin[1]

[1] Zhejiang University

[2] Emory University

[3] The University of Melbourne

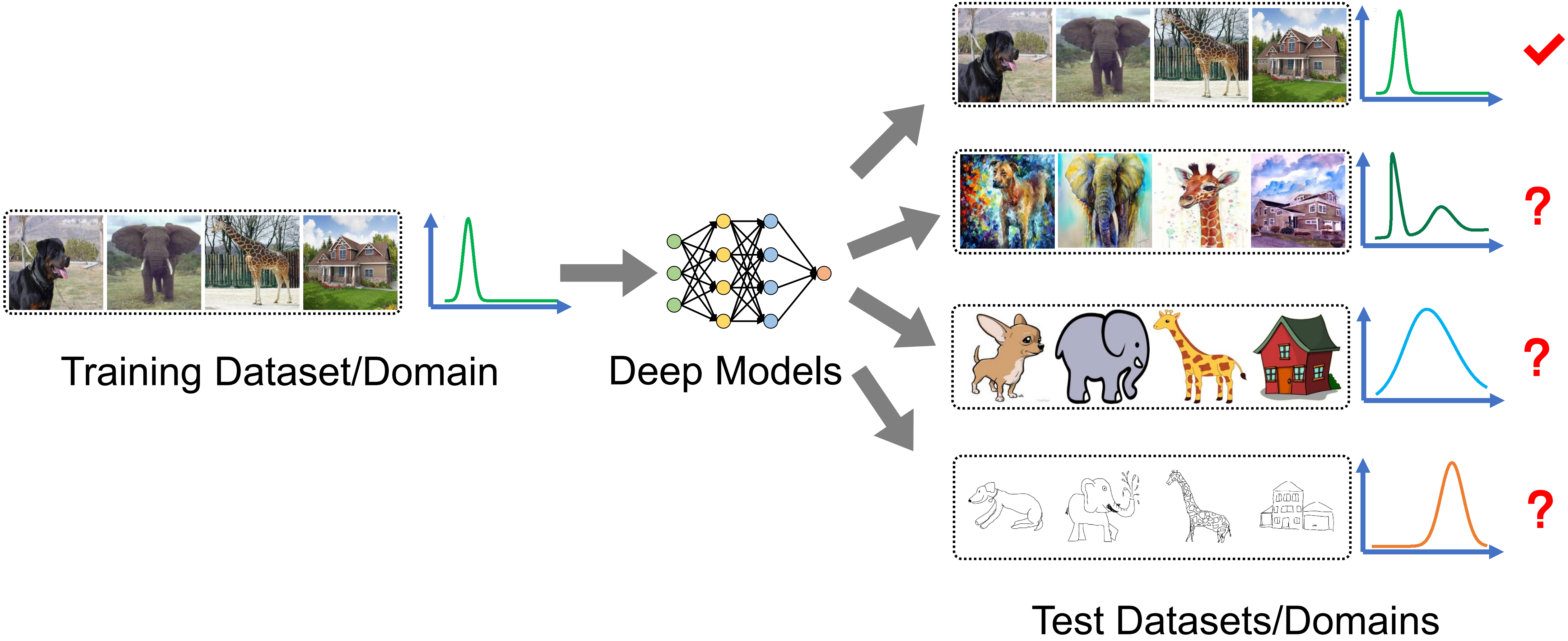
Deep Models Lack Robustness and Generalization



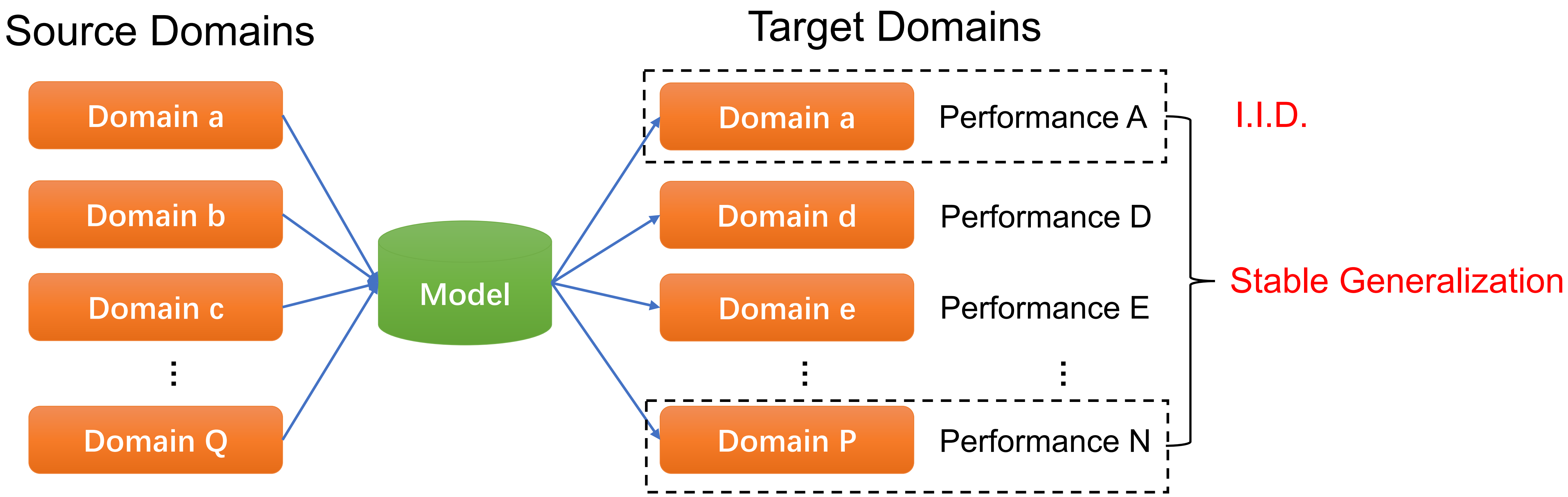
Fast Gradient Sign Method (FGSM) [1]

[1] Goodfellow I J, Shlens J, Szegedy C. Explaining and harnessing adversarial examples[J]. arXiv preprint arXiv:1412.6572, 2014.

Deep Learning Algorithms Mostly Rely on I.I.D. Assumption



Domain Generalization (DG): Learning Invariant Knowledge from Multiple Source Domains to Unknown Target Domains

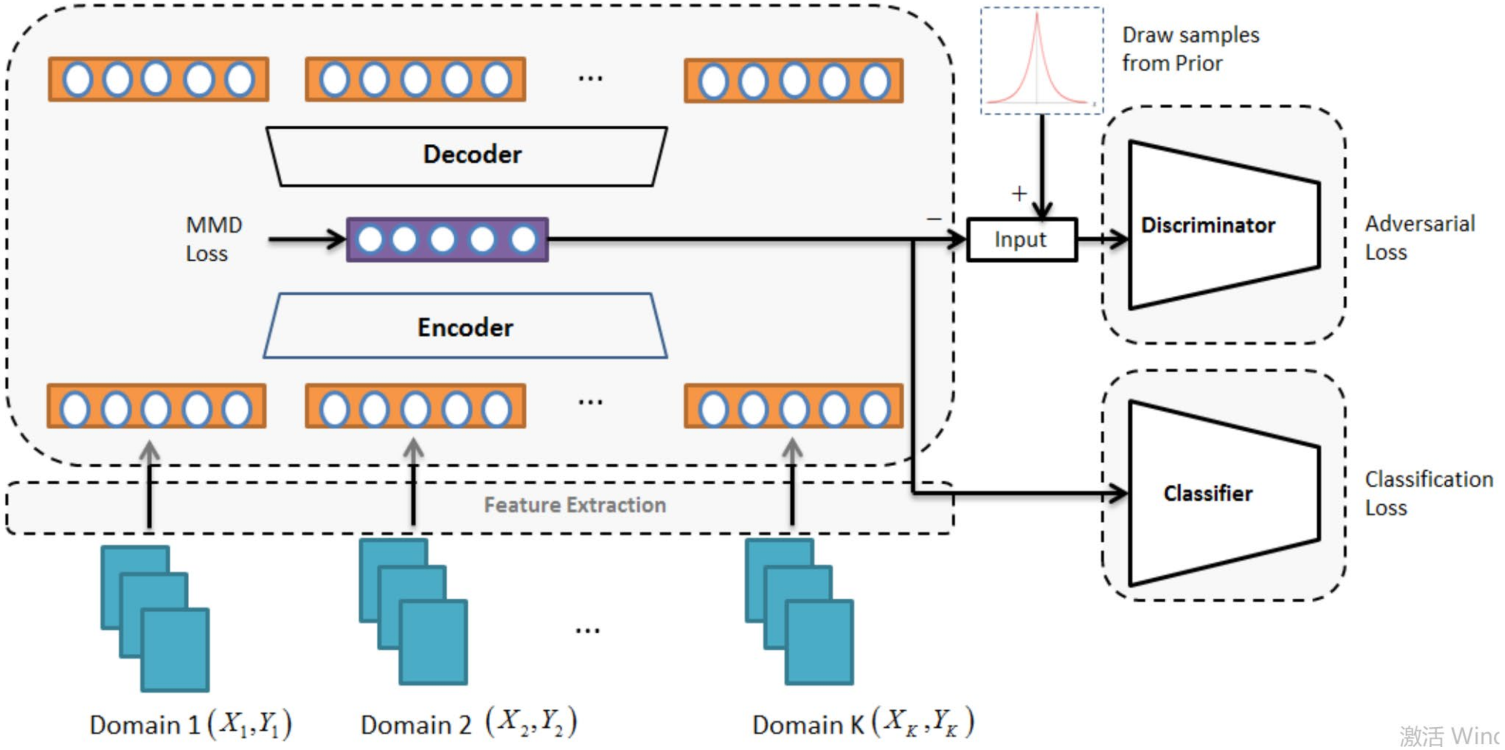


Problem Setup of Domain Generalization (DG)

- Q source datasets $\mathcal{D}^1, \mathcal{D}^2, \ldots, \mathcal{D}^Q$
- An unseen target dataset $\mathcal{D}^{Q+1}$
- N^q points are sampled for each dataset $\mathcal{D}^q$, i.e., $\mathcal{D}^q = \{\mathbf{x}_n^q, y_n^q\}_{n=1}^{N^q}$, $q = 1, \ldots, Q$
- Each dataset $\mathcal{D}^q$ is sampled from distribution P^q, $q = 1, \ldots, Q+1$, $P^q \neq P^p$ if $p \neq q$
- DG: Leverage the source datasets $\mathcal{D}^1, \mathcal{D}^2, \ldots, \mathcal{D}^Q$ to train a model and make it perform well on $\mathcal{D}^{Q+1}$.

Representative Works for DG

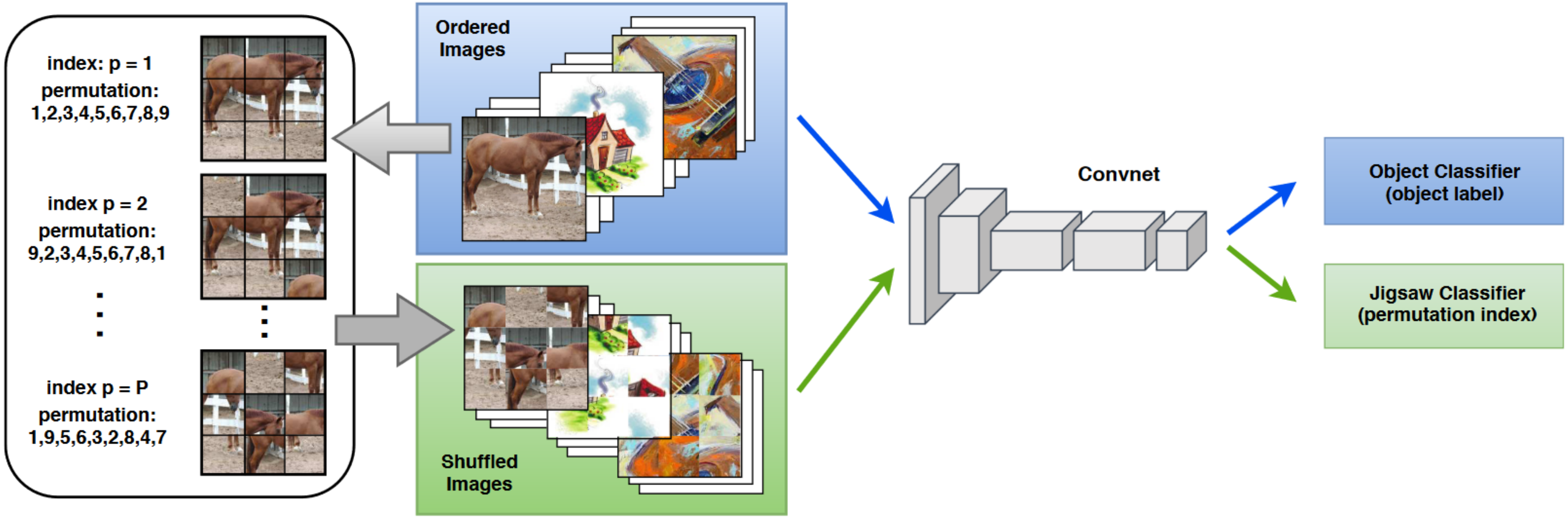
Domain invariant representation learning. [2]



[2] Li H, Pan S J, Wang S, et al. Domain generalization with adversarial feature learning[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 5400-5409.

Representative Works for DG

Domain augmentation. [3]



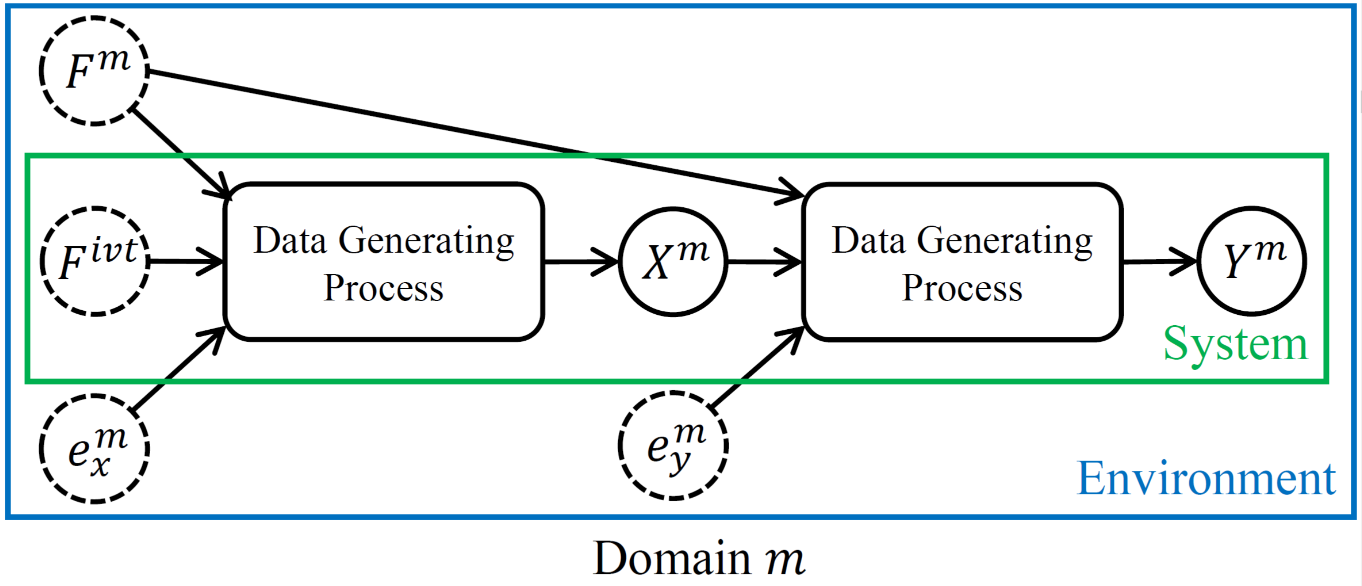
[3] Carlucci F M, D'Innocente A, Bucci S, et al. Domain generalization by solving jigsaw puzzles[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 2229-2238.

Problems of Existing DG Works

- Inefficient model training process.
- Learning domain-invariant marginal distribution $P(X)$, rather than invariant relationship between X and Y.
- Not explainable.

➔ Causality !

Data Generating Process (DGP) Assumption



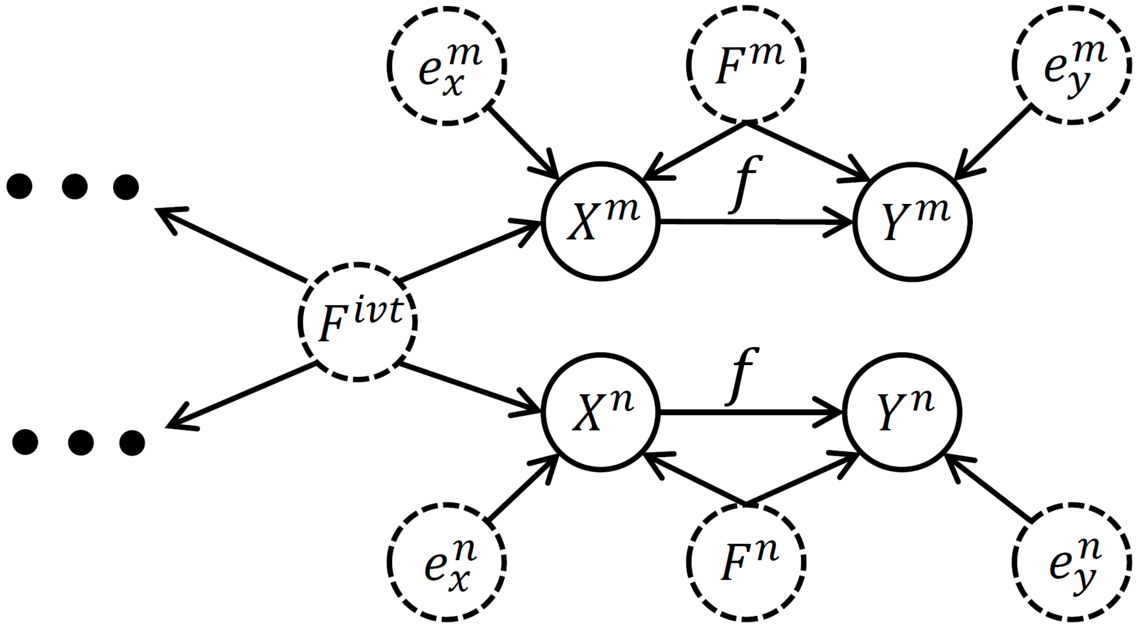
X^m & Y^m: Data (images) & label (categories).

F^{ivt}: Invariant factor (object shape).

F^m: Domain-specific factor (background, light condition).

e_x^m, e_y^m: error term.

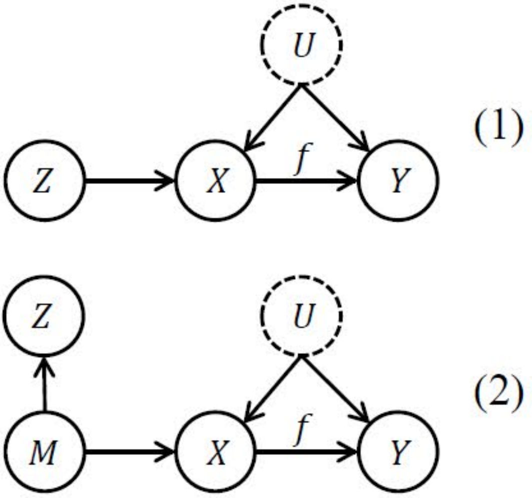
Robust Perception of Human & Invariant Relationship Assumption



f: Invariant relationship between data and label

Assumption 1. Data distributions of different domains satisfy the data generating process and causal graph, where only the factor F^{ivt} and relationship f are invariant.

Preliminary of Instrumental Variable (IV)

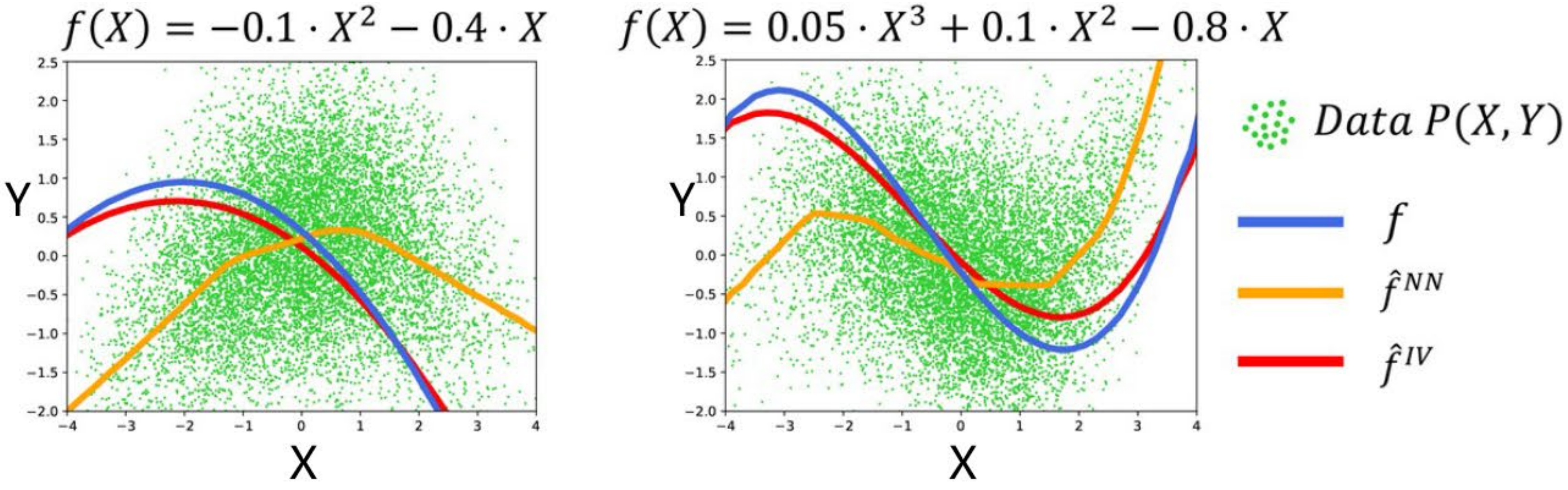


$Z \sim \mathcal{N}(0,1)$

$U \sim \mathcal{N}(0,1)$

$X = Z + U$

$Y = f(X) + U$

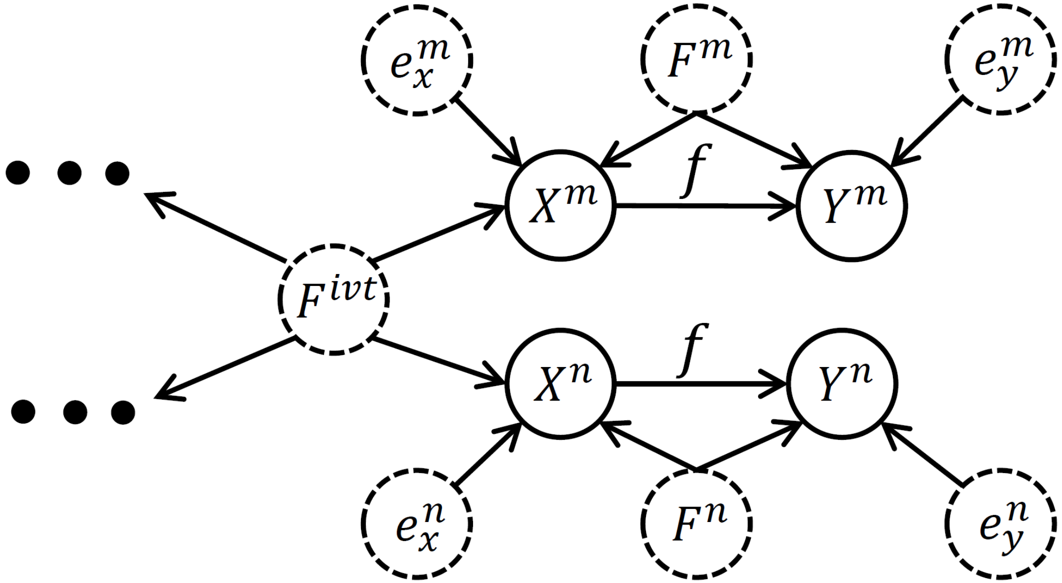


A valid IV should satisfy[4, 5]:

- **Relevance.** Z and X should be relevant, i.e., $P(X|Z) \neq P(X)$.
- **Exclusion.** Z is correlated to Y only through X, U, i.e., $Z \perp Y|(X, U)$.
- **Unconfounded instrument.** Z is independent of U, i.e., $Z \perp U$.

[4] Pearl J. Causality[M]. Cambridge university press, 2009.
[5] Hartford J, Lewis G, Leyton-Brown K, et al. Deep IV: A flexible approach for counterfactual prediction[C]//International Conference on Machine Learning. PMLR, 2017: 1414-1423.

Instrumental Variable for Domain Generalization Problem



Proposition 1. For any two domains m and n, if $m \neq n$, then the following conditions hold: (1) $P(X^m|X^n) \neq P(X^m)$; (2) $X^n \perp Y^m|(X^n, F^m)$; (3) $X^n \perp F^m$; (4) $X^n \perp e_y^m$.

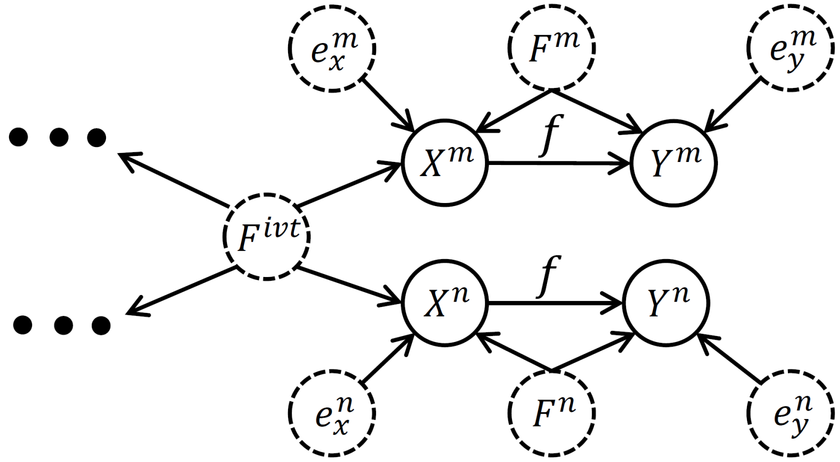
Theorem 1. For any two domains m and n, if $m \neq n$, then X^n is a valid instrumental variable of domain m.

Two-stage Method to Learn Domain-Invariant Relationship

By following [3, 4, 5], we assume DGP as:

$$Y^m = f(X^m) + F^m + e_y^m,$$

where $\mathbb{E}[e_y^m] = \mathbb{E}[F^m] = 0.$



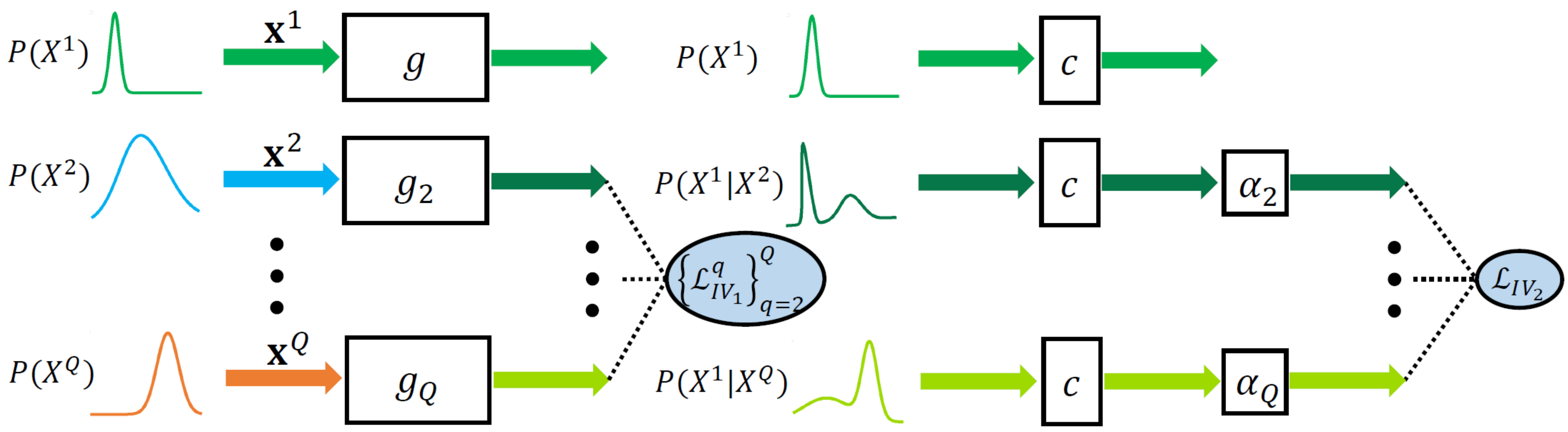
We have:

$$\mathbb{E}[Y^m|X^n] = \mathbb{E}[f(X^m)|X^n] + \mathbb{E}[F^m|X^n] + \mathbb{E}[e_y^m|X^n]$$

$$= \int f(X^m) dP(X^m|X^n)$$

[5] Hartford J, Lewis G, Leyton-Brown K, et al. Deep IV: A flexible approach for counterfactual prediction[C]//International Conference on Machine Learning. PMLR, 2017: 1414-1423.
[6] R. Singh, M. Sahani, and A. Gretton. Kernel instrumental variable regression. In Advances in Neural Information Processing Systems (NeurIPS), pages 4593–4605, 2019.
[7] A. Bennett, N. Kallus, and T. Schnabel. Deep generalized method of moments for instrumental variable analysis. In Advances in Neural Information Processing Systems (NeurIPS), pages 3564–3574, 2019.

Domain Invariant Relationship with Instrumental Variable



Stage 1: $\mathcal{L}_{IV_1}^q = \mathbb{I}(y^q = y^1)d_k^2(g_q(x^q), g(x^1))$, d_k^2: distance metric like MMD

Stage 2: $\mathcal{L}_{IV_2} = \frac{1}{Q-1}\sum_{q=2}^{Q} \alpha_q \mathbb{E}_{(x^q,y^q),(x^1,y^1)}[\mathbb{I}(y^q = y^1)\ell(c \circ g(x^q), y^1)]$, ℓ: classification loss

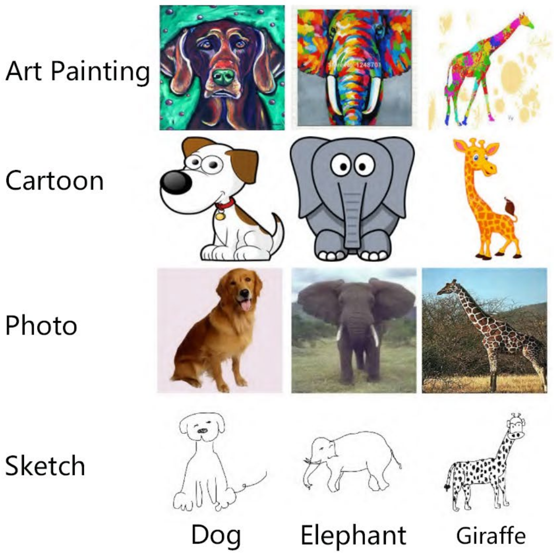
Results on Real-World Datasets

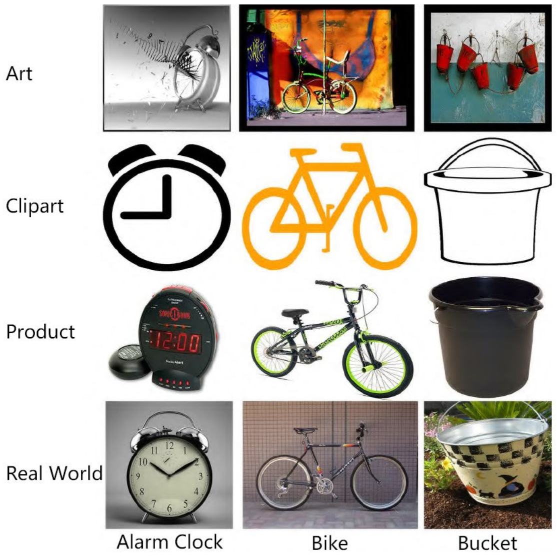
RESULTS (%) FOR DOMAIN GENERALIZATION ON PACS DATASET.

Methods	Art	Cartoon	Photo	Sketch	Average
DeepAll [8]	78.96	72.93	96.28	70.59	79.94
JiGen [8]	79.42	75.25	96.03	71.35	80.51
MASF [10]	80.29	77.17	94.99	71.69	81.04
DGER [61]	80.70	76.40	96.65	71.77	81.38
Epi-FCR [21]	82.1	77.0	93.9	73.0	81.5
MMLD [36]	81.28	77.16	96.09	72.29	81.83
EISNet [49]	81.89	76.44	95.93	74.33	82.15
L2A-OT [63]	83.3	78.2	96.2	73.6	82.8
DDAIG [62]	**84.2**	78.1	95.3	74.7	83.1
DRIVE w/o IV	79.40 ± 0.10	76.93 ± 0.09	95.75 ± 0.10	74.44 ± 0.07	81.63 ± 0.03
DRIVE w/o pre	81.95 ± 0.25	77.55 ± 0.31	96.64 ± 0.34	75.65 ± 0.10	82.95 ± 0.14
DRIVE	83.36 ± 0.70	**78.76 ± 0.08**	**96.87 ± 0.18**	**78.68 ± 0.96**	**84.42 ± 0.11**

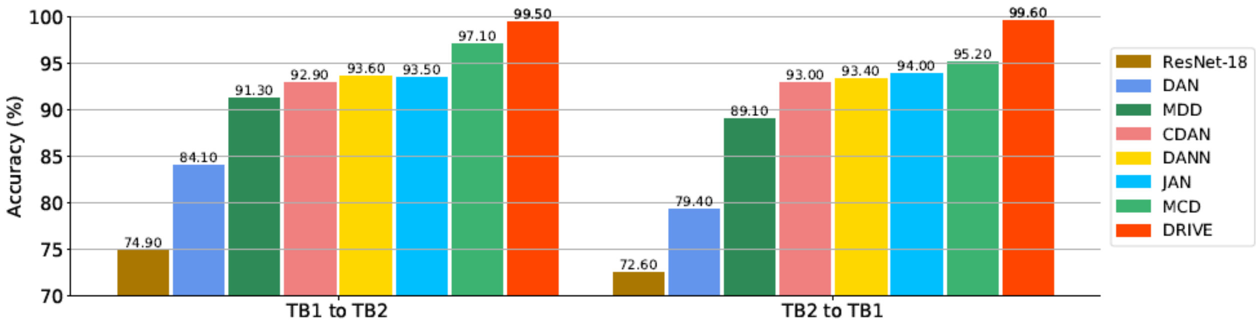
RESULTS (%) FOR DOMAIN GENERALIZATION ON OFFICE-HOME DATASET.

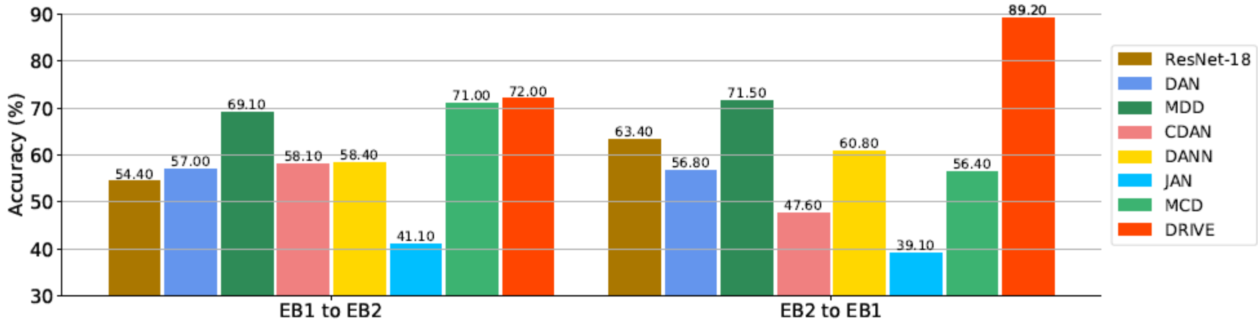
Methods	Art	Clipart	Product	Real-World	Average
DeepAll [8]	52.15	45.86	70.86	73.15	60.51
JiGen [8]	53.04	47.51	71.47	72.79	61.20
DSON [44]	59.37	44.70	71.84	74.68	62.90
RSC [16]	58.42	**47.90**	71.63	74.54	63.12
DRIVE w/o IV	55.53 ± 0.21	45.92 ± 0.50	71.64 ± 0.35	74.49 ± 0.05	61.90 ± 0.20
DRIVE w/o pre	59.30 ± 0.06	47.65 ± 0.30	72.03 ± 0.57	75.55 ± 0.24	63.63 ± 0.11
DRIVE	**60.40 ± 0.26**	47.73 ± 0.28	**72.63 ± 0.18**	**76.14 ± 0.10**	**64.23 ± 0.09**

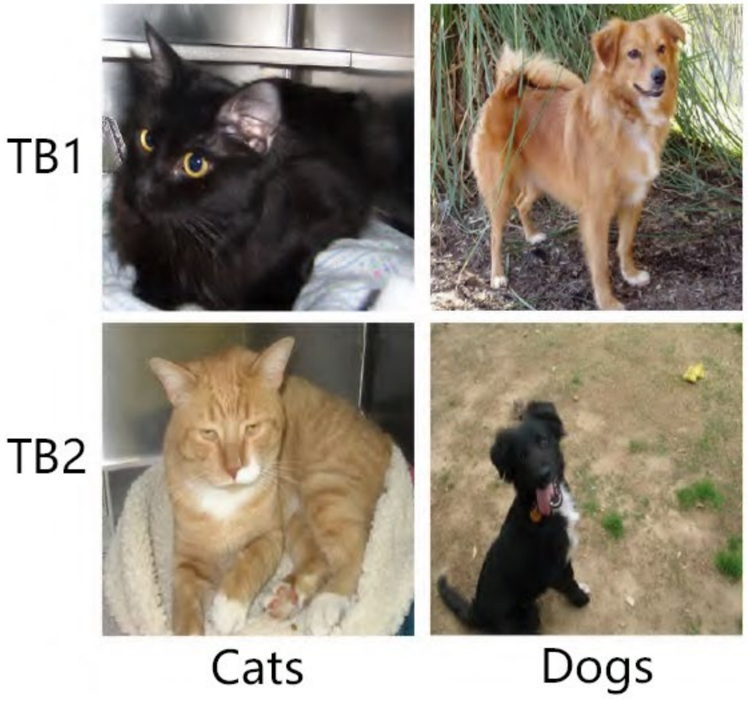




Results on Real-World Datasets







Conclusions

- Build generalizable models is important for AI applications.
- Domain generalization is one solution.
- Causality techniques has been demonstrated to be useful in exploring invariant causal relations.
- How to develop causality-inspired algorithms for enabling domain generalization and other machine learning problems.

References

[1] Goodfellow I J, Shlens J, Szegedy C. Explaining and harnessing adversarial examples[J]. arXiv preprint arXiv:1412.6572, 2014.

[2] Li H, Pan S J, Wang S, et al. Domain generalization with adversarial feature learning[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 5400-5409.

[3] Carlucci F M, D'Innocente A, Bucci S, et al. Domain generalization by solving jigsaw puzzles[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 2229-2238.

[4] Pearl J. Causality[M]. Cambridge university press, 2009.

[5] Hartford J, Lewis G, Leyton-Brown K, et al. Deep IV: A flexible approach for counterfactual prediction[C]//International Conference on Machine Learning. PMLR, 2017: 1414-1423.

[6] R. Singh, M. Sahani, and A. Gretton. Kernel instrumental variable regression. In Advances in Neural Information Processing Systems (NeurIPS), pages 4593–4605, 2019.

[5] A. Bennett, N. Kallus, and T. Schnabel. Deep generalized method of moments for instrumental variable analysis. In Advances in Neural Information Processing Systems (NeurIPS), pages 3564–3574, 2019.

[7] Yuan J, Ma X, Kuang K, et al. Learning Domain-Invariant Relationship with Instrumental Variable for Domain Generalization[J]. arXiv preprint arXiv:2110.01438, 2021.

Thank You!