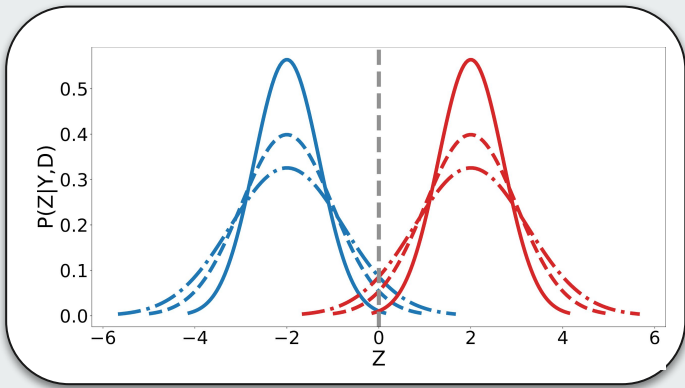
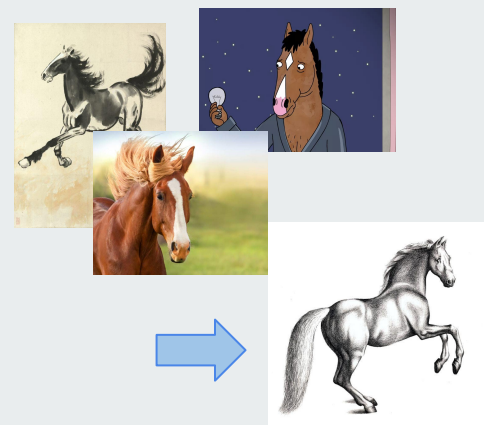


Respecting Domain Relations: Hypothesis Invariance for Domain Generalization

Ziqi Wang, Marco Loog, Jan van Gemert

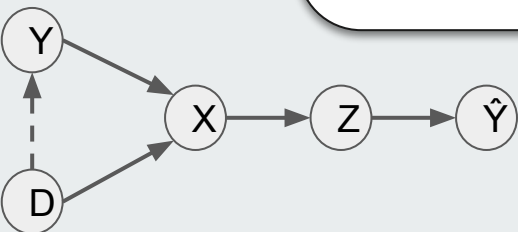


Domain Generalization



- Supervised learning
- Multiple non-i.i.d. Domains
- Target domain is unavailable

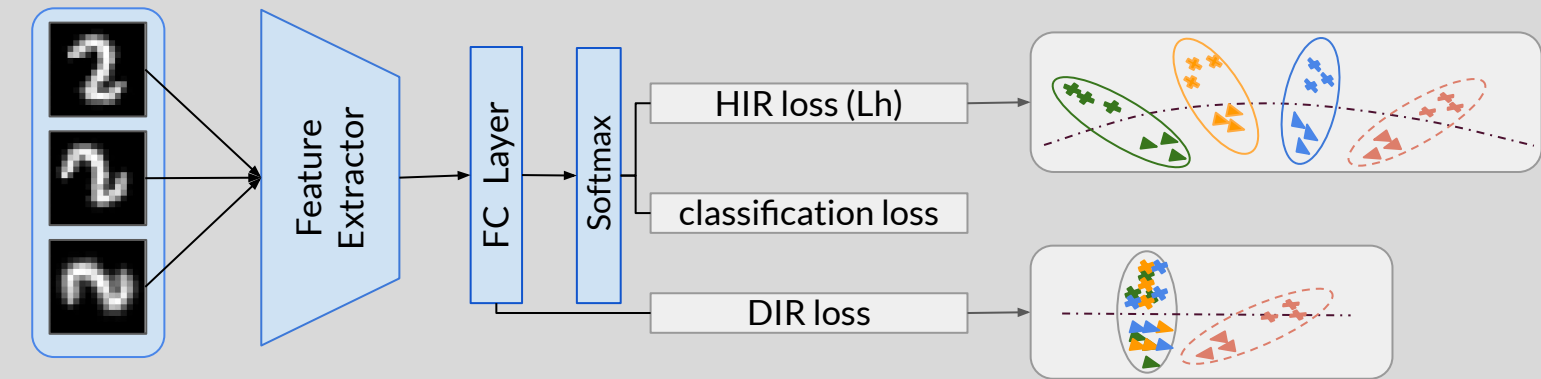
DIR & HIR



$$P_{Z|D} = P_Z, \quad P_{Z|Y,D} = P_{Z|Y}, \quad P_{Y|Z,D} = \frac{P_{Z|Y,D}P_{Y|D}}{P_{Z|D}},$$
$$\forall D \in \mathcal{D}. \quad \forall D \in \mathcal{D}.$$

$$\arg \max_Y P_{Y|Z,D} = \arg \max_Y P_{Y|Z},$$
$$\forall D \in \mathcal{D}.$$

Experiments & Results



	Methods	\mathcal{M}_0°	\mathcal{M}_{15°	\mathcal{M}_{30°	\mathcal{M}_{45°	\mathcal{M}_{60°	\mathcal{M}_{75°	Avg.
prior knowledge	LG	89.7	97.8	98.0	97.1	96.6	92.1	95.3
	DIVA	93.5	99.3	99.1	99.2	99.3	93.0	97.2
no prior knowledge	D-MTAE	82.5	96.3	93.4	78.6	94.2	80.5	87.5
	CCSA	84.6	95.6	94.6	82.9	94.8	82.1	89.1
	MMD-AAE	83.7	96.9	95.7	85.2	95.9	81.2	89.8
	DA	86.7	98.0	97.8	97.4	96.9	89.1	94.3
	HEX	90.1	98.9	98.9	98.8	98.3	90.0	95.8
Ours	AGG	89.87	99.41	98.98	95.14	98.63	91.13	95.53
	HIR	90.34 \pm 0.88	99.75 \pm 0.18	99.40 \pm 0.21	96.17 \pm 0.71	99.25 \pm 0.26	91.26 \pm 0.66	96.03

$$L_h = \sum_{i=1}^n \sum_{j=i+1}^n \sum_{c=1}^m P(\hat{Y}_i | Z_i, Y_i = c) \cdot \log \left(\frac{P(\hat{Y}_i | Z_i, Y_i = c)}{P(\hat{Y}_j | Z_j, Y_j = c)} \right)$$

Domains	D-MTAE	CIDDG	DBADG	MMD-AAE	MLDG	Epi-FCR	CCSA	AGG	HIR
V	63.90	64.38	69.99	67.70	67.7	67.1	67.10	65.4	69.10 \pm 1.8
L	60.13	63.06	63.49	62.60	61.3	64.3	62.10	60.6	62.22 \pm 1.7
C	89.05	88.83	93.63	94.40	94.4	94.1	92.30	93.1	95.39 \pm 0.9
S	61.33	62.10	61.32	64.40	65.9	65.9	59.10	65.8	65.71 \pm 1.6
Avg	68.60	69.59	72.11	72.28	72.3	72.9	70.15	71.2	73.10

Conclusions and take away

- Learning domain invariant representations is unnecessarily strict
- Invariance learning may lead to overfit or discarding too much information
- Domain adaptation approaches cannot be applied on domain generalization directly