Encouraging Intra-Class Diversity Through a Reverse Contrastive Loss for Single-Source Domain Generalization

ÉCOLE CENTRALE LYON

2021 ICCV OCTOBER 11-17 VIRTUAL

Thomas Duboudin (ECL), Emmanuel Dellandréa (ECL), Corentin Abgrall (Thales), Gilles Hénaff (Thales), Liming Chen (ECL) - corresponding author: Thomas Duboudin (thomas.duboudin@ec-lyon.fr)

Motivation

Making deep networks generalize to unseen distributions is a complicated task: see **DomainBed** https://github.com/facebookresearch/DomainBed

Conclusion: no published multi-source domain generalization method yields consistently significant improvements over the standard training procedure

Research focus: simpler situation !

A class correlated pattern in training

is missing at test-time

- The network only learns the most efficient patterns
- At test-time, the network fails to make a proper decision

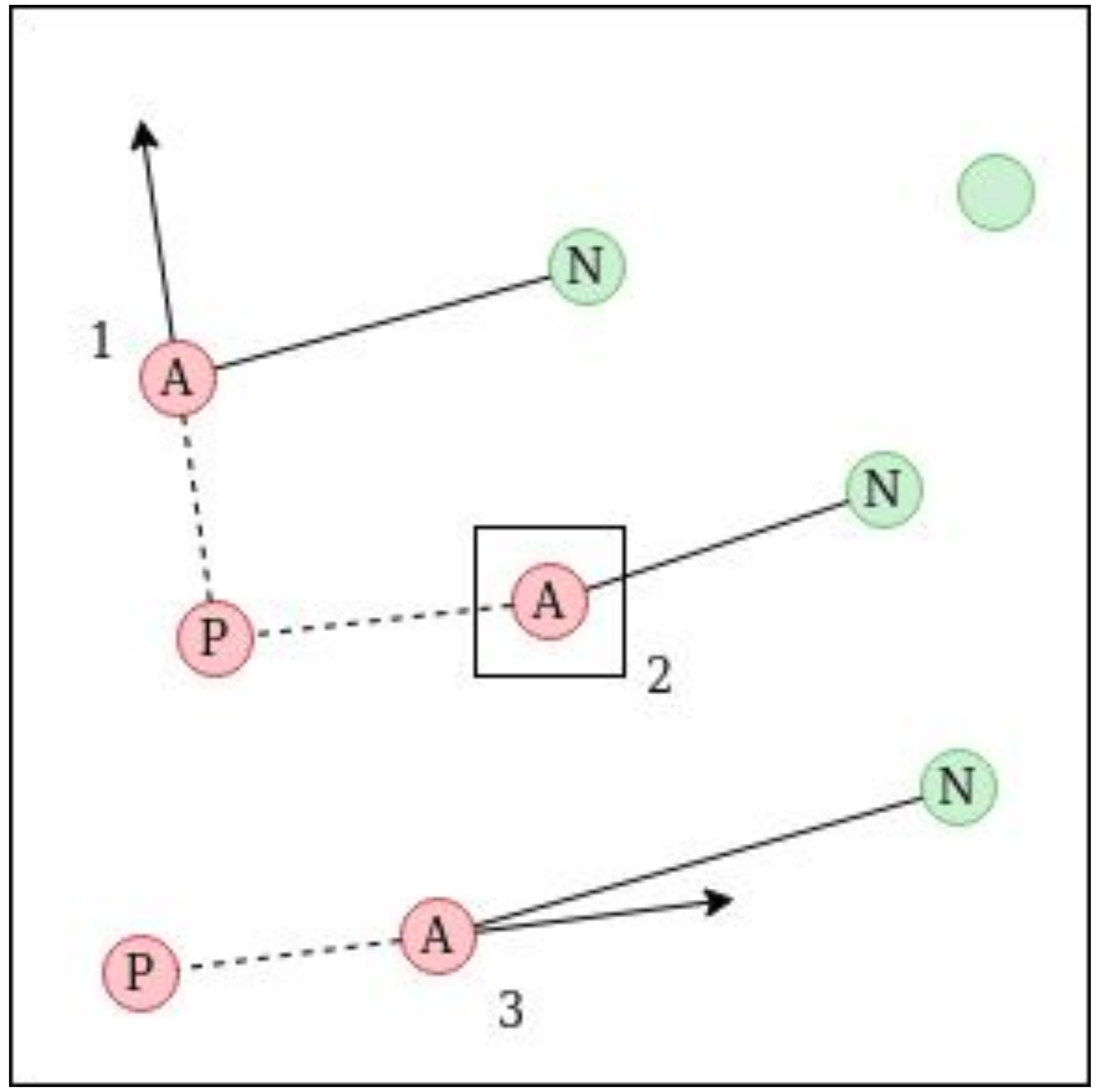
⇒ Networks focusing only on the most efficient patterns is one of the reasons why they fail to generalize outside of their training distribution

Solution

Use several semantically-different classification strategies to make a decision and not only the most obvious one

- No counter-examples and only one source domain

Our method : a reverse contrastive approach



- **By increasing the space between points of the same class in the latent space, the network is forced to learn different patterns for a single class**
- **We use the following reversed contrastive loss, along with the standard cross-entropy loss for classification (with m the multiplicative margin)**

$$\mathcal{L}_{RC} = \begin{cases} -d(f_a, f_p) \text{ } if \text{ } d(f_a, f_p) < m \times d(f_a, f_n) \\ 0 \text{ } otherwise \end{cases}$$

Results on our benchmark ↓

Method	Validation Accuracy	Test Accuracy
Standard Training Procedure	99.8 (± 0.006)	26.0 (± 4.7)
Dropout	99.8 (± 0.01)	31.9 (± 3.1)
Dropout & Orthogonality	99.8 (± 0.008)	42.7 (± 2.6)
Dropout & Covariance	99.8 (± 0.002)	42.6 (± 2.0)
Jigsaw Puzzle	99.8 (± 0.01)	43.0 (± 3.0)
(with early stopping Val>95%)	98.5 (± 0.3)	59.9 (± 4.5)
Auto-Encoder Reconstruction	99.8 (± 0.007)	28.5 (± 4.8)
Spectral Decoupling	99.8 (± 0.004)	47.7 (± 2.9)
(with early stopping Val>95%)	99.4 (± 0.09)	49.8 (± 1.5)
Representation Self-Challenging	99.2 (± 0.2)	43.5 (± 5.0)
RCL (ours - $m = \infty$)	**96.0 (± 0.3)**	**89.9 (± 0.9)**
Standard Training Procedure on the original MNIST	97.8 (± 0.12)	97.8 (± 0.12)

Conclusion

- Aiming for a tight-clusters large-margins latent space might be detrimental to out-of-domain generalization
- Existing methods only mitigate the damages, even in a very simple case
- Future work will be dedicated to more real-life like situations