On genuine invariance learning without weight-tying

Artem Moskalev, Anna Sepliarskaia, Erik J. Bekkers, Arnold Smeulders

UvA - BOSCH DELTA LAB

VIS LAB VIDEO & IMAGE SENSE LAB

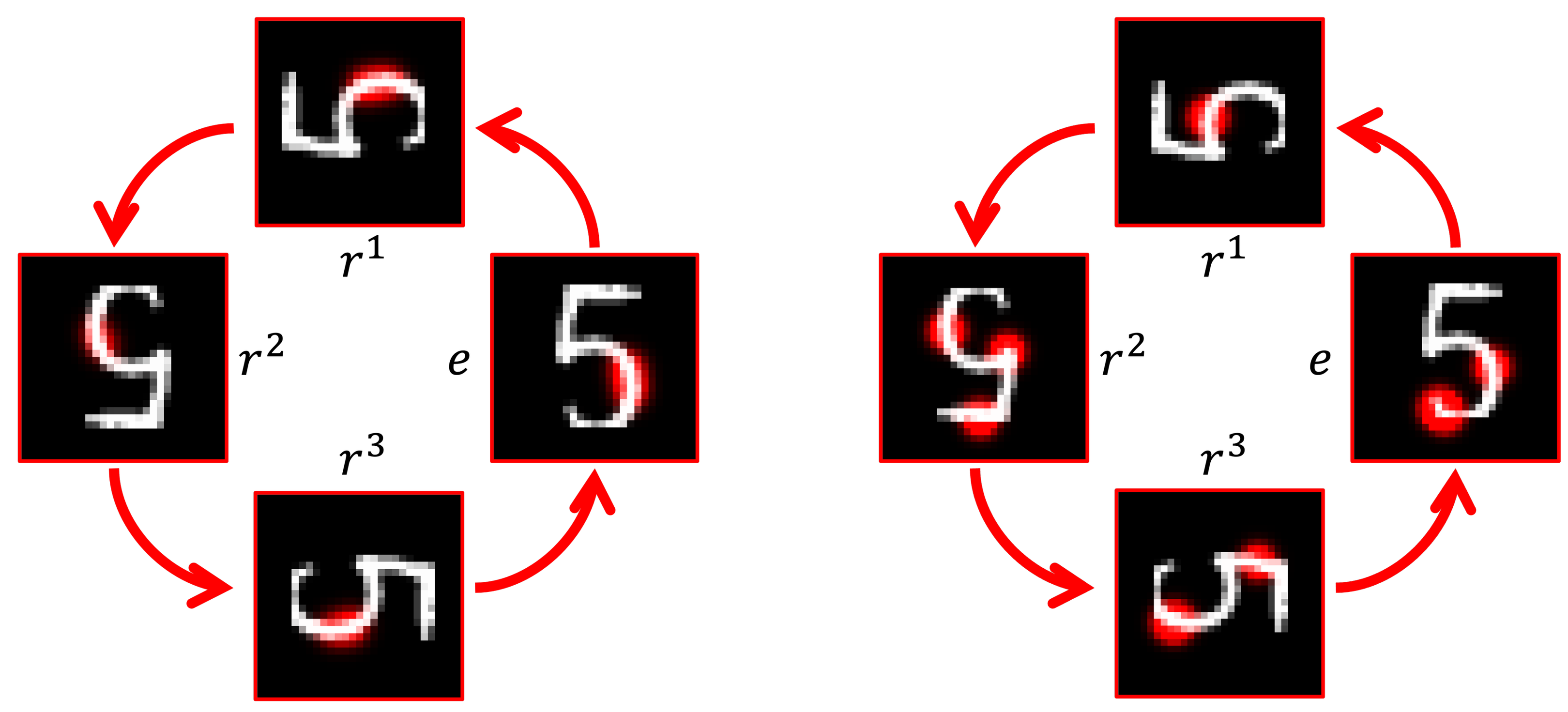
There exist **two methods** to achieve **invariance**:

1) Invariant weight-tying *(equivariant networks)*
2) Train invariance from data *(data augmentation)*

What are the **properties and limitations** on data-driven invariance learning?

Both networks learn to correctly classify rotated digits

`Weight-Tying [WT]` `Data Aug. [DA]`



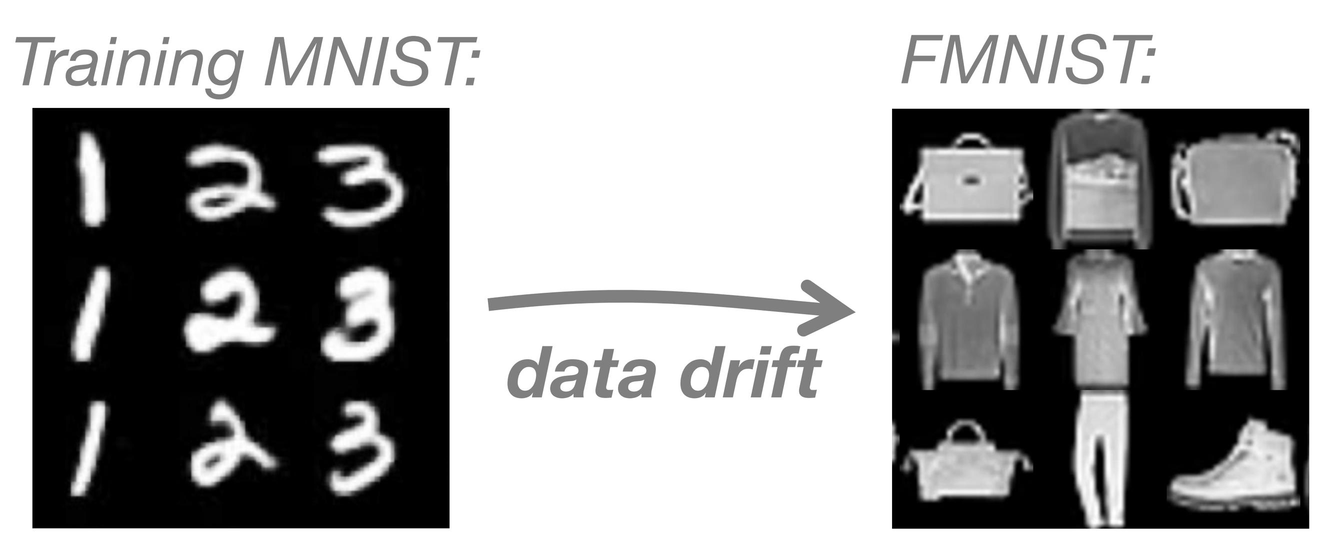
Check our paper:

Normalized saliency maps:

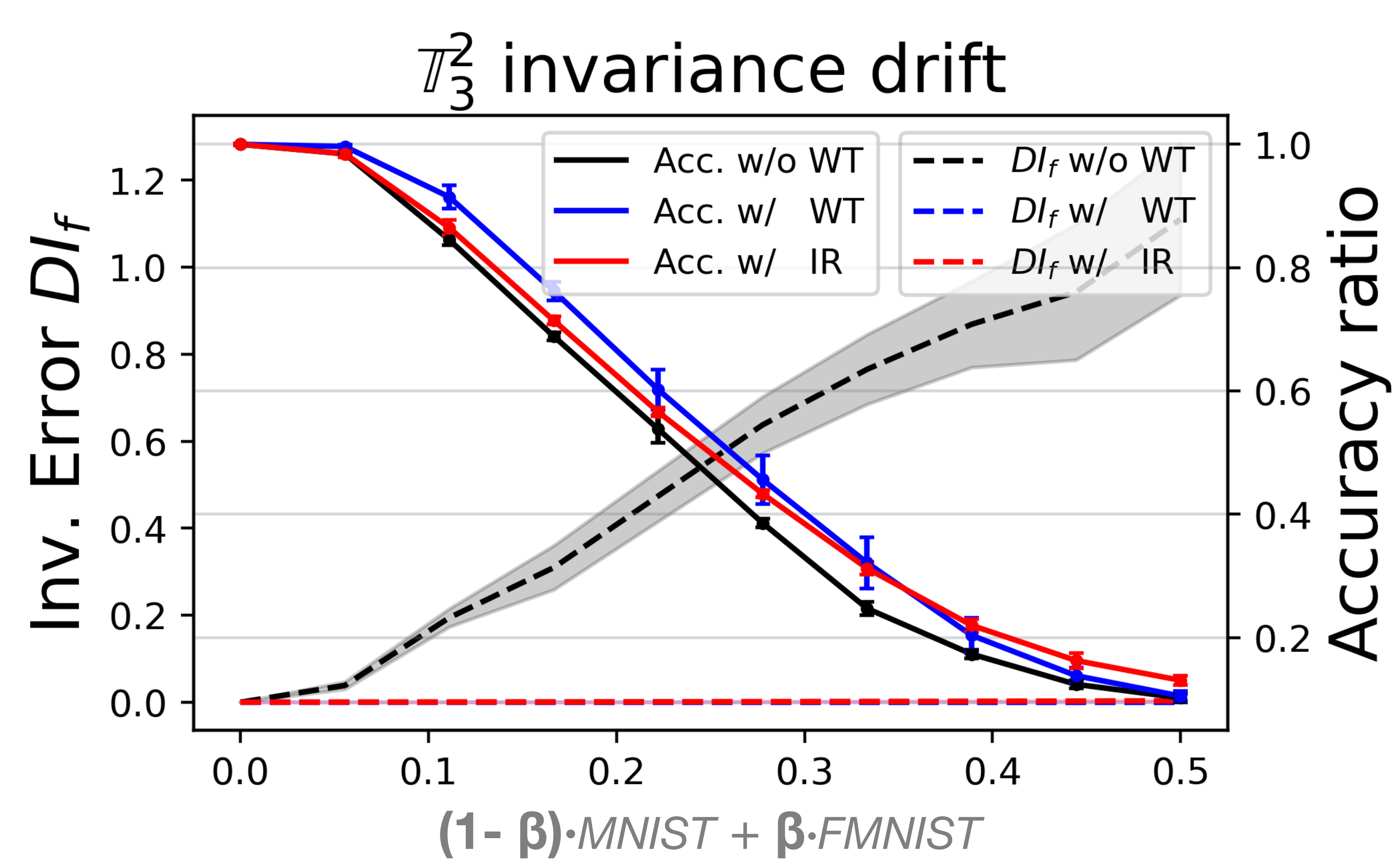
[WT] decision-making process is **genuinely invariant**

[DA] decision-making process **does not attain invariance**

Reliability of learned invariance:



Learned invariance is **strongly conditioned** on the input data



[IR] Invariance **Regularization**:

task loss

$$\min_\theta \mathcal{L}_f(\mathcal{D}) + \nu I_f(\mathcal{D},\mathcal{G})$$

invariance error

Significantly **improves invariance** but **decreases accuracy**

$\mathcal{G}$	Model	Acc. (%)	$LI \Downarrow$	$DI \Downarrow$	$SI \Uparrow$
$\mathbb{R}_4^2$	WT	94.6 ±0.1	0.00 ±0.0	0.0 ±0.0	1.00 ±0.00
	DA	94.0 ±0.5	98.6 ±5.4	0.3 ±0.1	0.17 ±0.04
	IR	87.9 ±0.8	0.02 ±0.0	0.0 ±0.0	0.95 ±0.03
$\mathbb{T}_3^2$	WT	96.6 ±0.1	0.0 ±0.0	0.0 ±0.0	1.00 ±0.0
	DA	96.2 ±0.2	50.8 ±6.8	0.1 ±0.0	0.57 ±0.08
	IR	93.1 ±0.1	0.01 ±0.0	0.0 ±0.0	0.95 ±0.07

3 metrics to measure **invariance error**

Invariance-induced **spectral decay**:

Proposition 3.1 (*Invariance-induced spectral decay*). *Logit invariance error minimization implies $\sigma_{max}(W(t)) \leq \sigma_{max}(W(0))$ when $t \rightarrow \infty$.*

Network routes for group invariance by reducing the sensitivity to **any input perturbation**