

# Tensor Methods & Deep Learning

A modern love story



@JeanKossaifi  
jean.kossaifi@gmail.com

# Traditional approaches

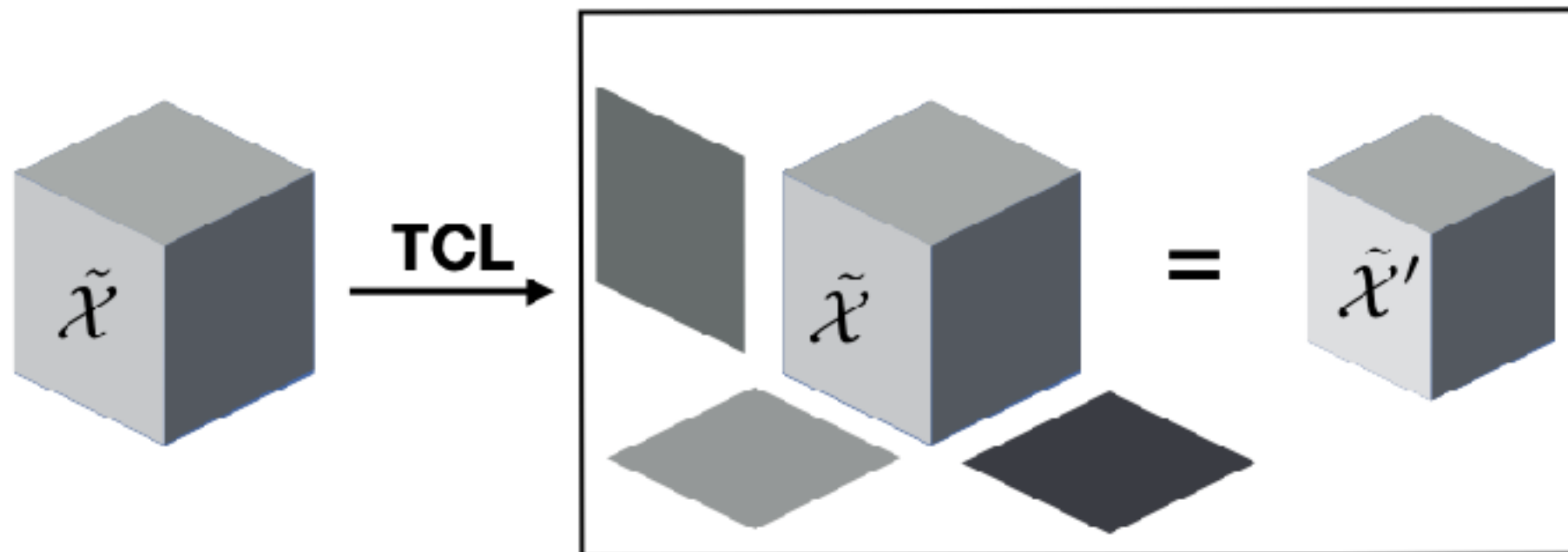
- DATA => CONV => RELU => POOL => Activation tensor
- Flattening loses information
- Can we leverage directly the activation tensor before the flattening?
- Potential space savings
- Performance improvement

# Deep neural nets architectures

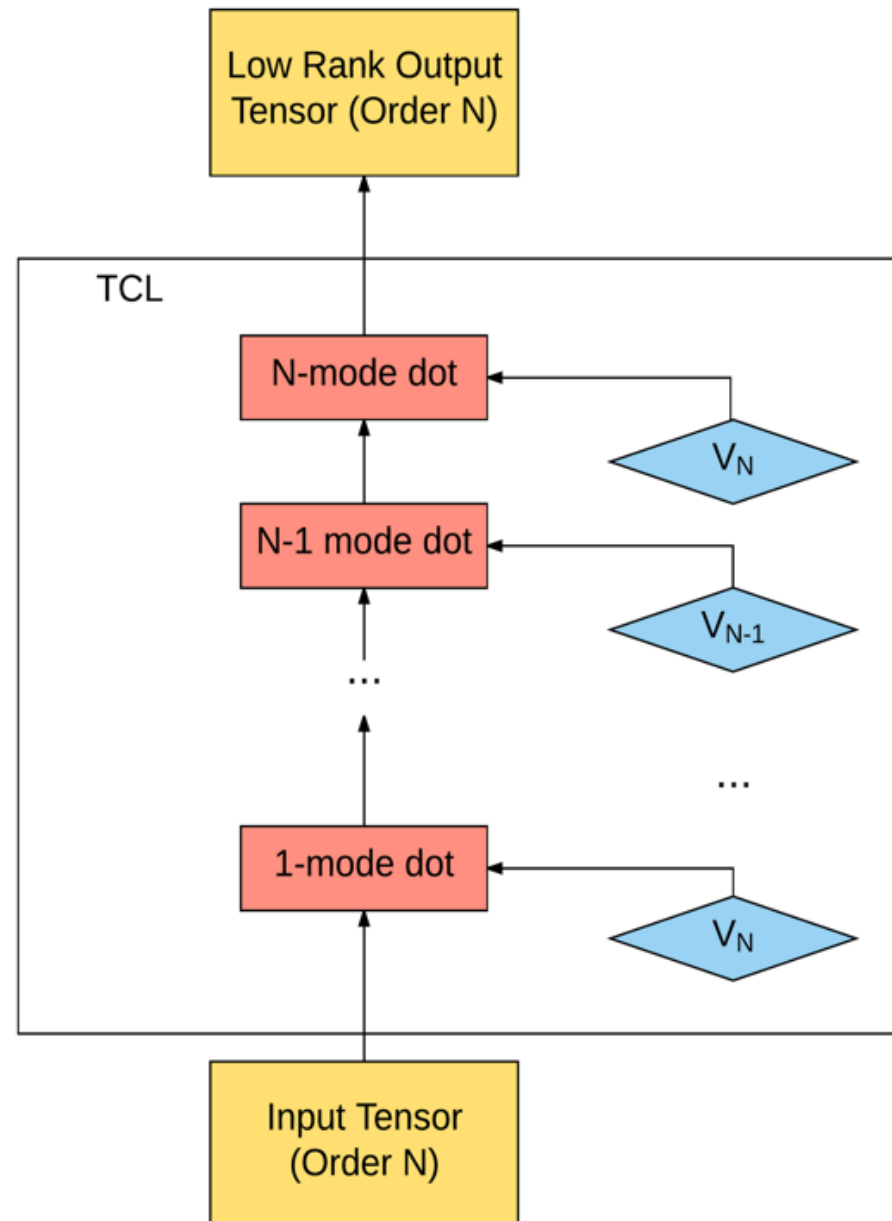
	Total params	Params in FC	%params in FC
AlexNet	61,100,840	58,631,144	96%
VGG-19	143,667,240	123,642,856	86%
ResNet-50	25,557,032	2,049,000	8%
RestNet-101	44,549,160	2,049,000	4.6%

# Tensor contraction

- Apply tensor contraction along each mode to obtain a low-rank tensor



# Tensor contraction



- Take activation tensor as input
- Feed it through a TCL
- Output a low-rank activation tensor

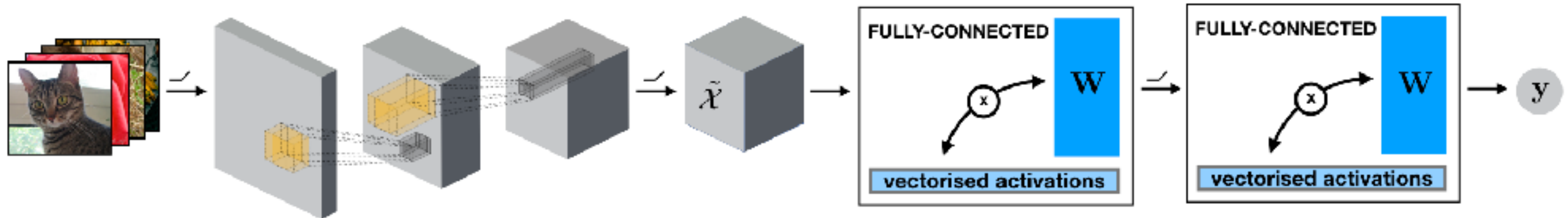
# Tensor contraction layers

- Trained end-to-end
- On ImageNet with VGG:
  - 65.9% space savings
  - performance drop of 0.6% only
- On ImageNet with AlexNet:
  - 56.6% space savings
  - Performance improvement of 0.5%

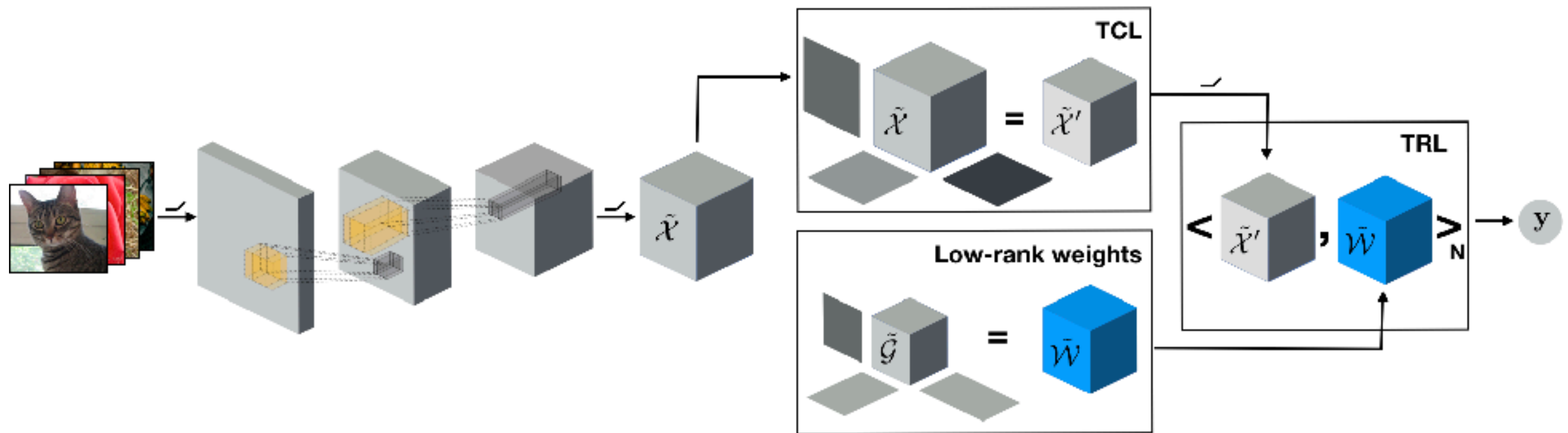
SPACE savings :

$$1 - \frac{n_{\text{compressed}}}{n_{\text{original}}}$$

# Traditional Architectures

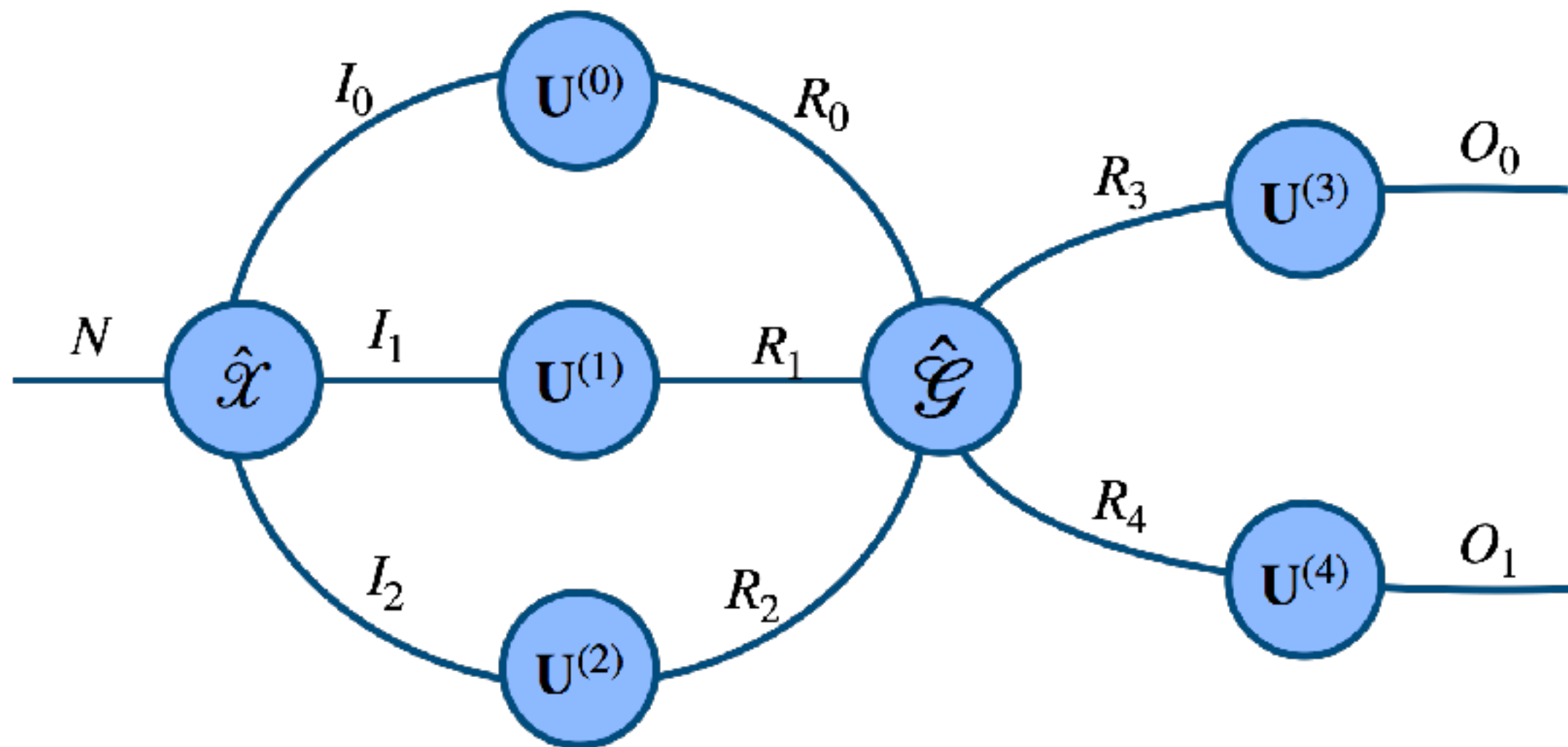


# Tensor Regression Networks





# Tensor Regression Networks



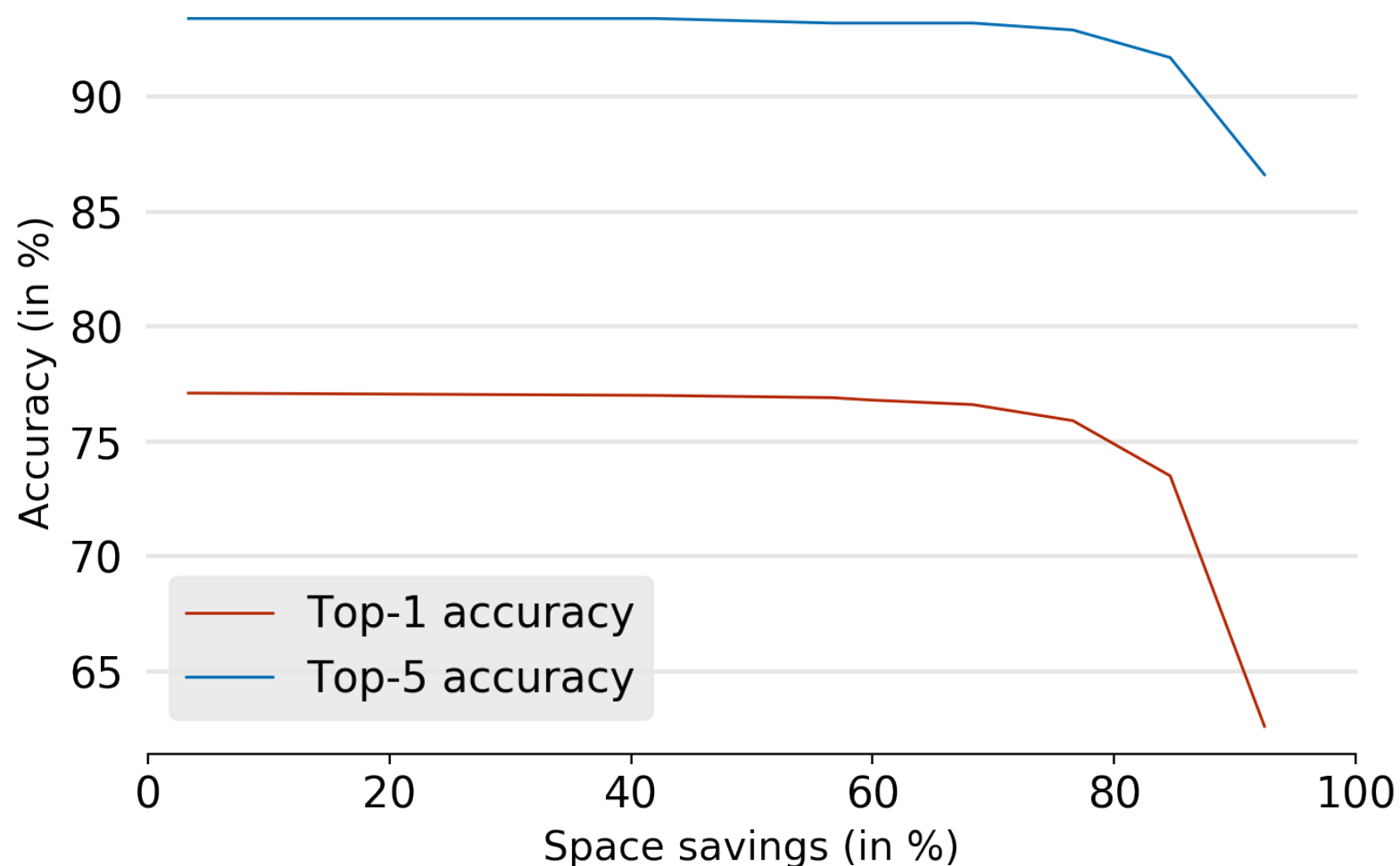
# Performance of the trl

TRL rank	Performance (in %)		
	Top-1	Top-5	Space savings
baseline	77.1	93.4	0
(200, 1, 1, 200)	77.1	93.2	68.2
(150, 1, 1, 150)	76	92.9	76.6
(100, 1, 1, 100)	74.6	91.7	84.6
(50, 1, 1, 50)	73.6	91	92.4

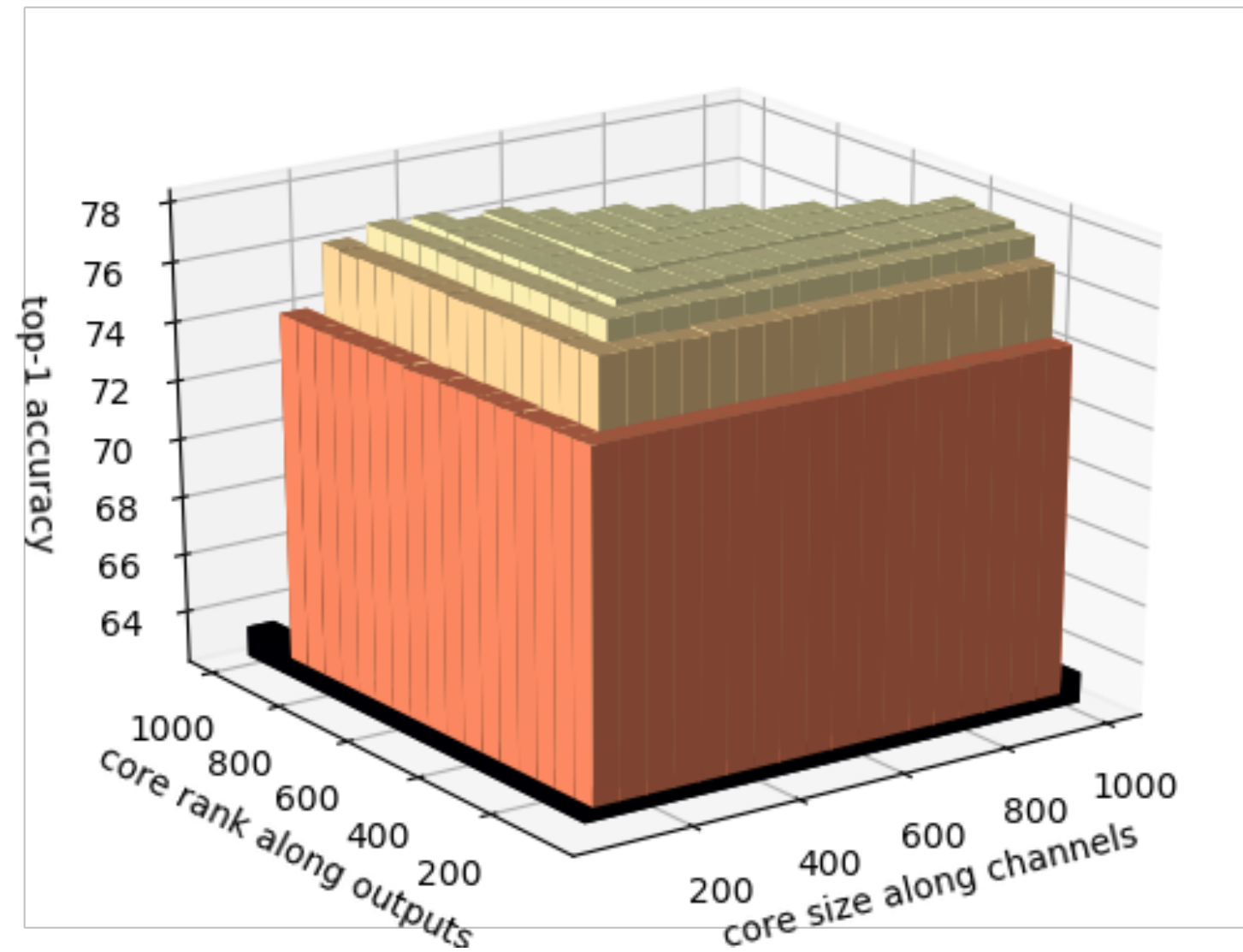
**ResNet-101**

# Performance of the TRL

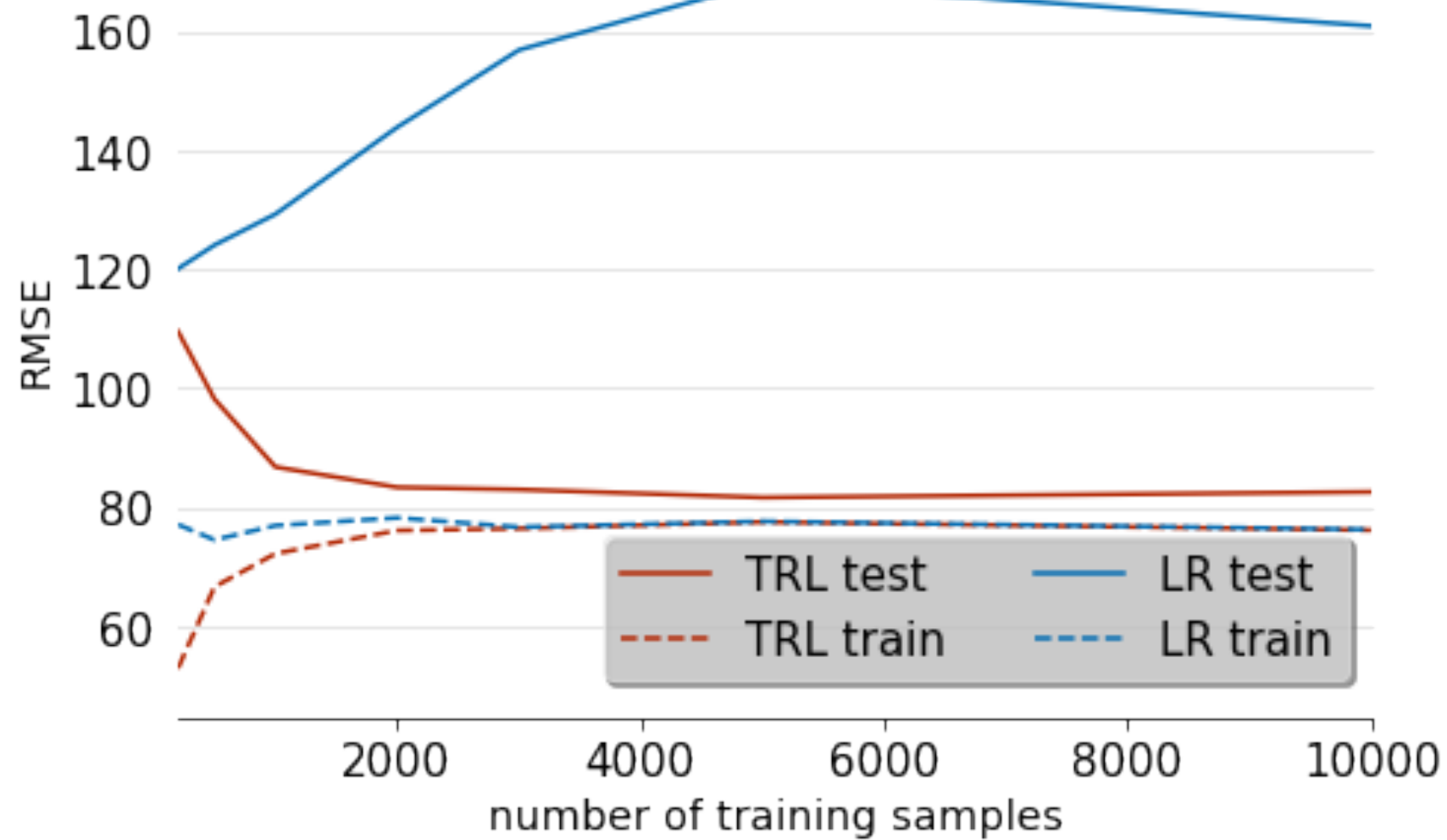
- 92.4% space savings, 4% decrease in Top-1 accuracy
- 68.2% space savings, no decrease in Top-1 accuracy



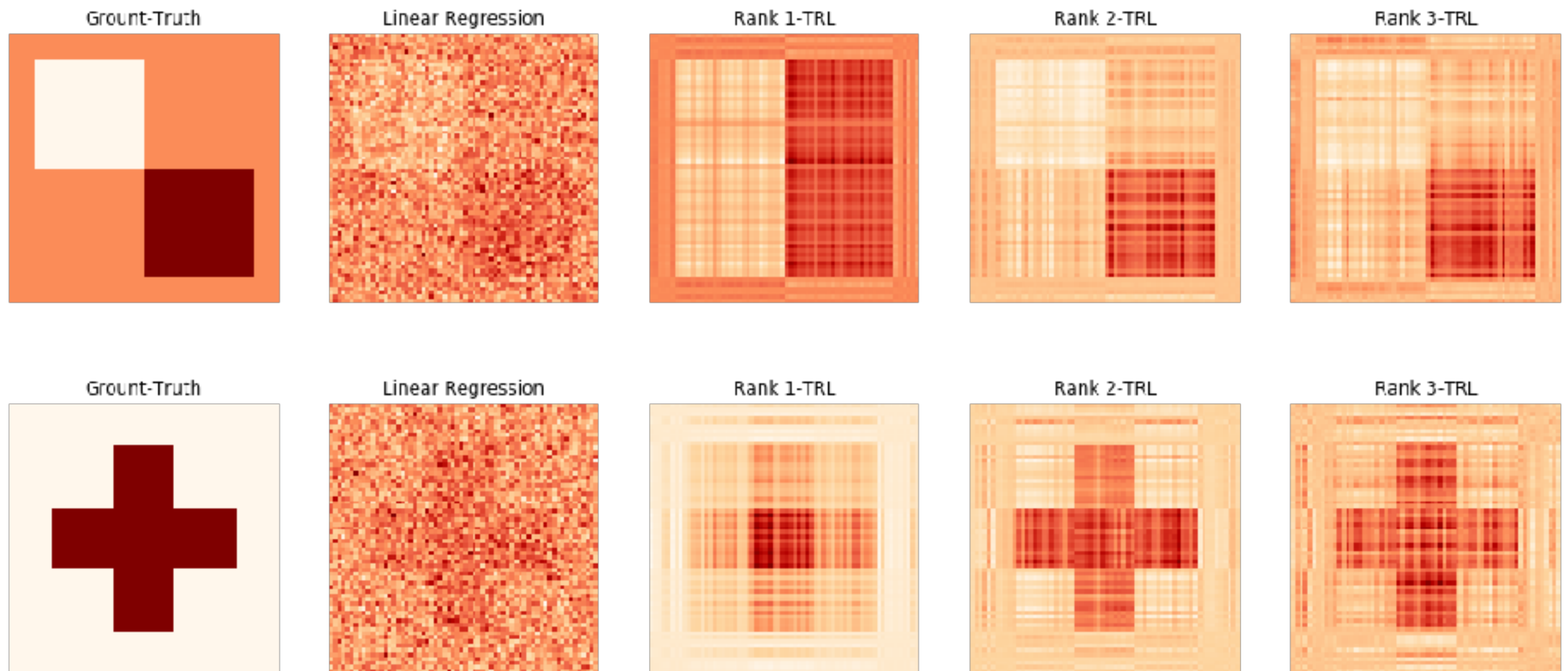
# Performance of the TRL



# TRL vs Linear



# Structure in the Regression Weights





Any questions?



@JeanKossaifi  
jean.kossaifi@gmail.com