Comparatives, Quantifiers, Proportions: A Multi-Task Model for the Learning of Quantities from Vision

Sandro Pezzelle[1], Ionut-Teodor Sorodoc[2], Raffaella Bernardi[3]

[1,2,3]CIMeC - Center for Mind/Brain Sciences, [3]DISI, University of Trento
[2]Universitat Pompeu Fabra, Barcelona
sandro.pezzelle@unitn.it | sandropezzelle.github.io | quantit-clic.github.io

Motivation

Quantification: Operation of expressing *quantitative* information

- 'There are ***more*** cars than parking lots': **comparatives**
- '***Most of*** the supporters wear blue t-shirts': **quantifiers**
- '***20% of*** the trees have been planted last year': **proportions**
- '***Seven*** students passed the exam': **numbers**

Comparatives, *Quantifiers*, *Proportions* express a comparison or relation ***between sets***; *Numbers* denote cardinality of ***one set***

Different age of acquisition [1,2,3], no need of counting for using comparatives and quantifiers in *grounded* contexts [4]

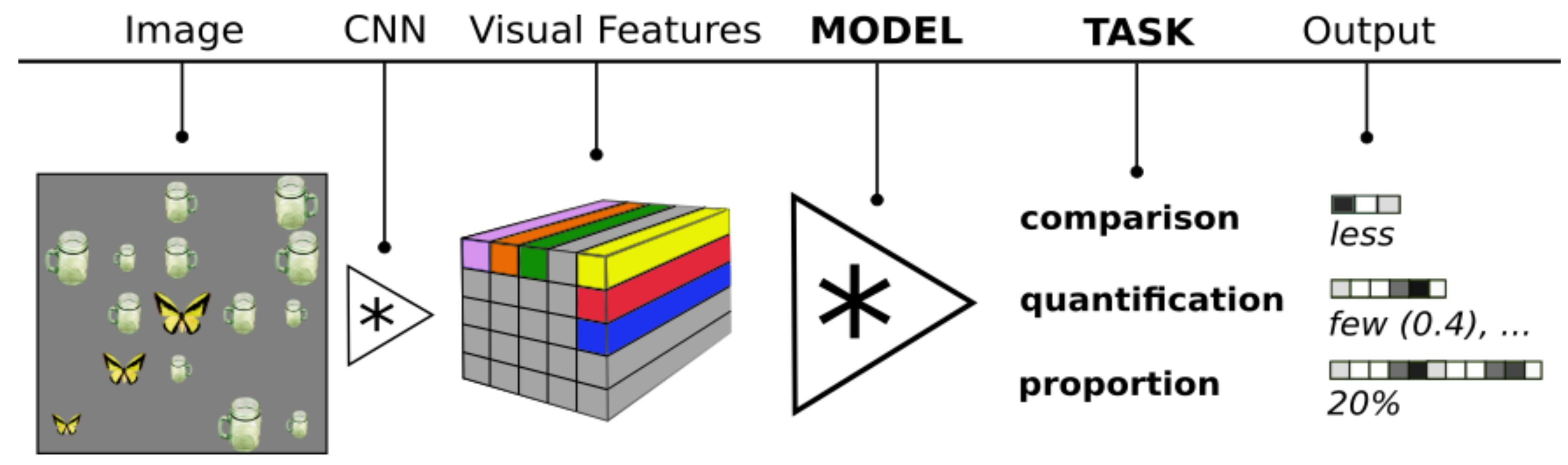
Hypothesis

Cs, *Qs*, *Ps* express increasingly-complex steps of **same** ratio-based **mechanism**; *Ns* require different, possibly interfering operation [5]

Task & Dataset

Research Question

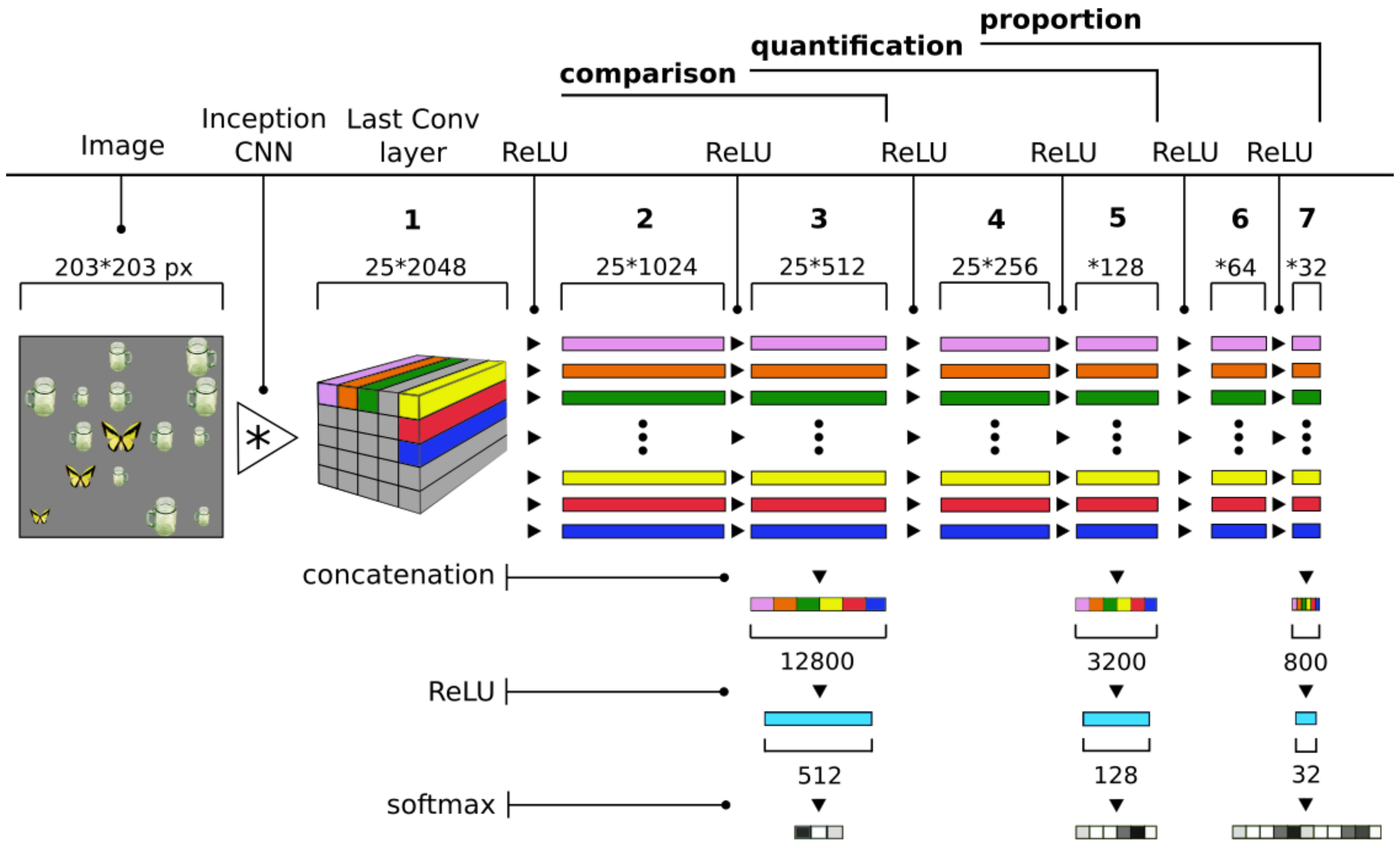
Can 3 ratio-based quantification tasks be modeled by a single, **Multi-Task Learning** (MTL) neural network **from Vision**?



Dataset

17K (70% train, 10% val, 20% test) synthetic scenes depicting 17 ***ratios*** targets (animals):non-targets (artifacts) from [6]

Multi-Task Learning Model



Predictions

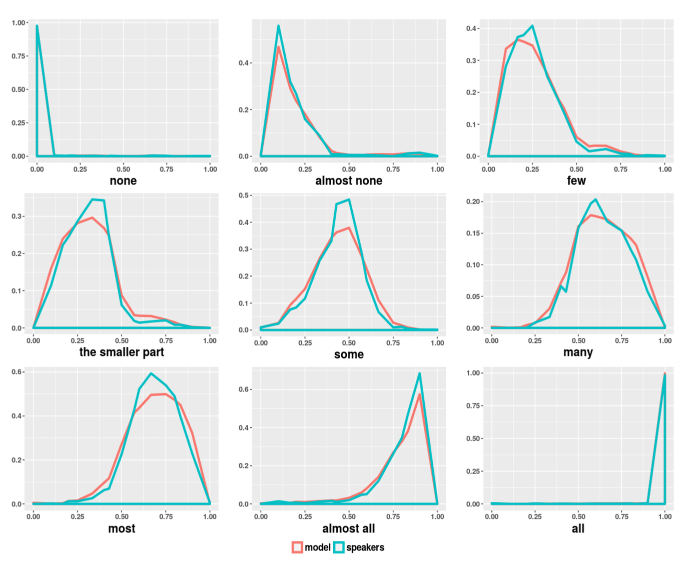
3	9	17
'less'	'few (0.4) ...'	'20%'
class	**probabilities** [7]	**class**

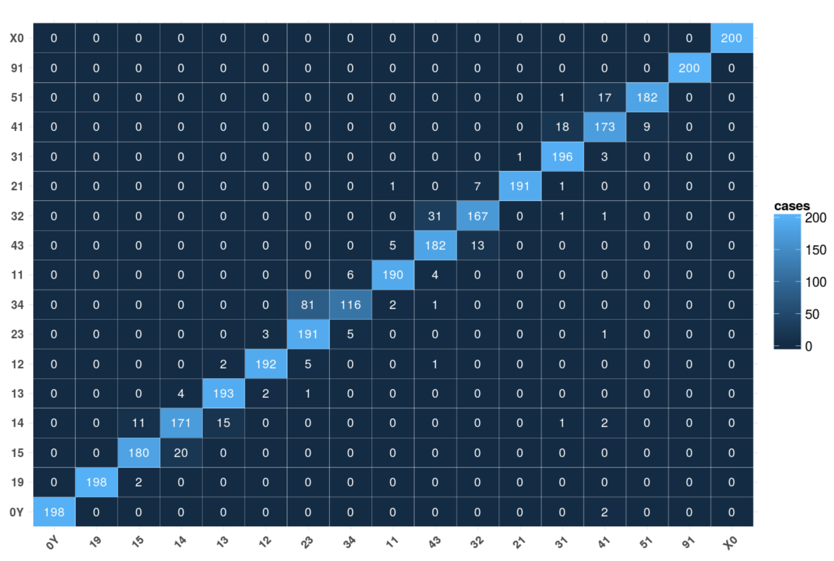
Results

MTL outperforms one-task models: sharing weights ***helps***!

MTL model approximates human data and makes 'plausible' errors

model	setComp	vagueQ	propTarg	nTarg
	accuracy	*Pearson r*	*accuracy*	*accuracy*
chance/majority	0.470	0.320	0.058	0.132
one-task-frozen	0.783	0.622	0.210	0.312
one-task-end2end	0.902	0.964	0.659	**0.966**
multi-task-prop	**0.995**	**0.982**	**0.918**	–
multi-task-number	0.854	0.807	–	0.478





In-Depth Evaluation

Numbers in the loop

Introducing number of targets in the pipeline ***hurts*** performance!

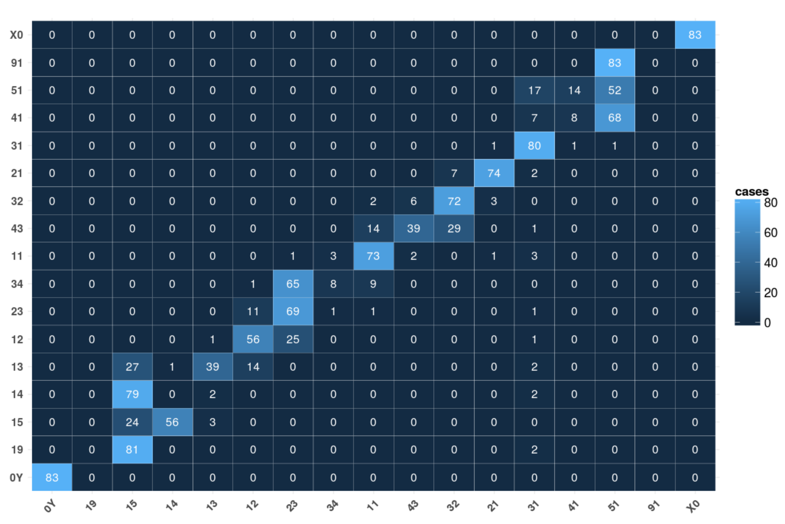
Reversing the architecture

Proportions (.08 acc) > Quantifiers (.32 *r*) > Comparatives (.65 acc)

Does MTL generalize?

Train w/ 80 combinations, test w/ 17 **unseen** combinations (1/*ratio*)

model	setComp	vagueQ	propTarg
	accuracy	*Pearson r*	*accuracy*
chance/majority	0.470	0.320	0.058
one-task-frozen	0.763	0.548	0.068
one-task-end2end	0.793	0.922	0.059
multi-task-prop	**0.943**	**0.960**	**0.539**



Discussion & References

Sharing a **common core** boosts performance in all tasks, proving their (a) interdependency and (b) increasing complexity

Are representations learned from one modality abstract enough to be transferable to different modalities, e.g. ***language, sounds***?

References

[1] Odic, D. et al. (2013). Young children's understanding of "more" and discrimination of number and surface area. Journal of Experimental Psychology:Learning, Memory, and Cognition 39(2):451.
[2] Hurewitz, F. et al. (2006). Asymmetries in the acquisition of numbers and quantifiers. Language learning and development 2(2):77–96.
[3] Le Corre, M. and Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. Cognition 105(2):395–438.
[4] Halberda,J. et al. (2008). The development of 'most' comprehension and its potential dependence on counting ability in preschoolers. Language Learning and Development 4(2):99–121
[5] Fabbri, S. et al. (2012). The role of numerosity in processing nonsymbolic proportions. The Quarterly Journal of Experimental Psychology 65(12):2435–2446.
[6] Kiani, R. et al. (2007). Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. Journal of neurophysiology 97(6):4296–4309.
[7] Pezzelle, S., Bernardi, R., Piazza, M. (*submit.*). Probing the mental scale of quantifiers. Cognition.