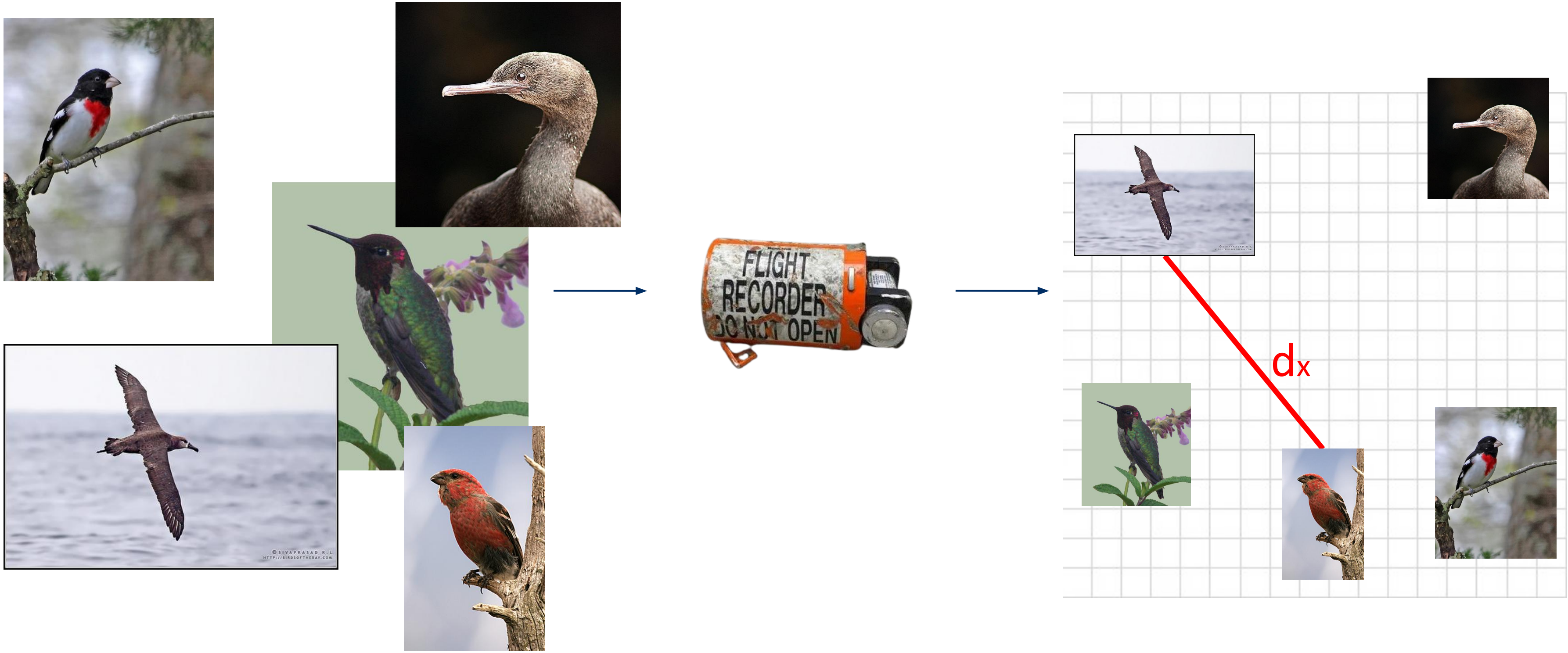
University of Antwerp
Faculty of Science

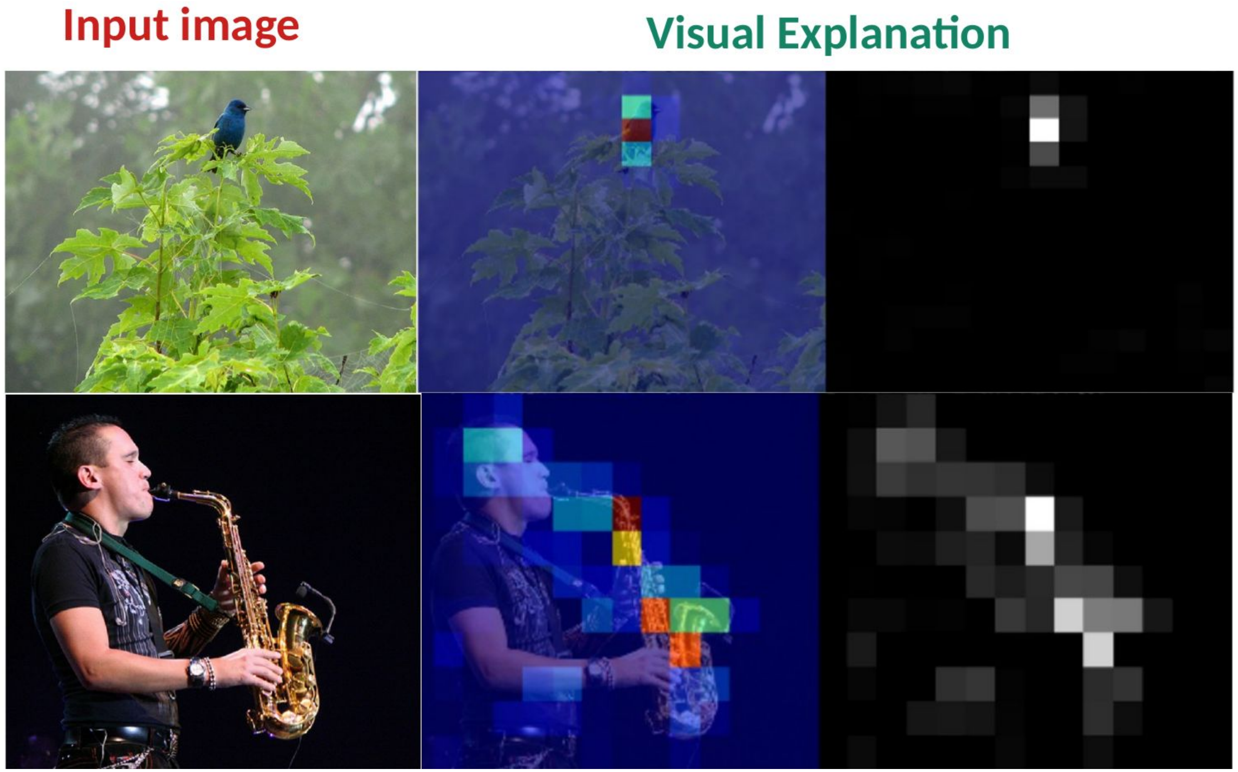
Neural Prototype Trees for Interpretable Fine-grained Image Recognition

Konstantina Ellina
Pablo de Vicente Abad

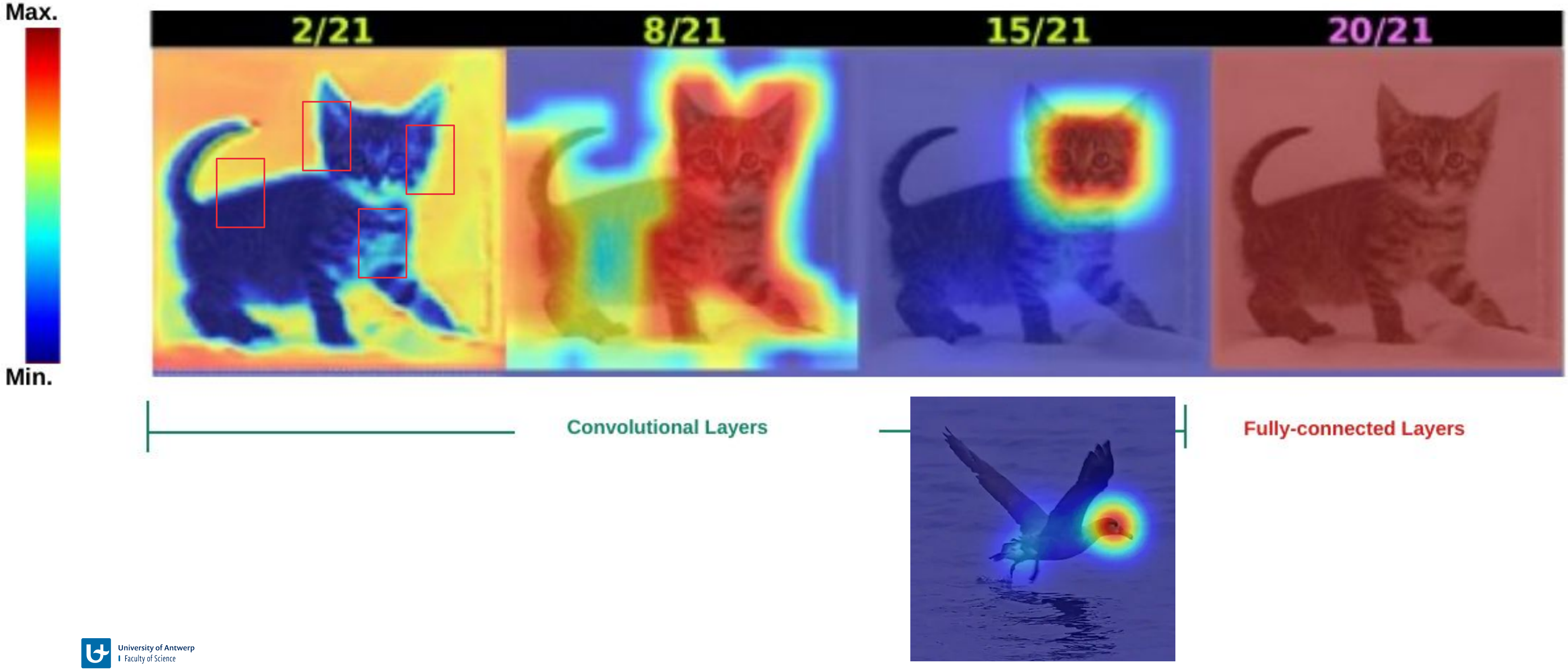
Traditional DNN’s



Model Explanation



Model Interpretation



Decision Trees

Small birds

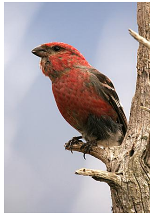
Big birds

Green birds

Red birds

Local birds

Migratory birds

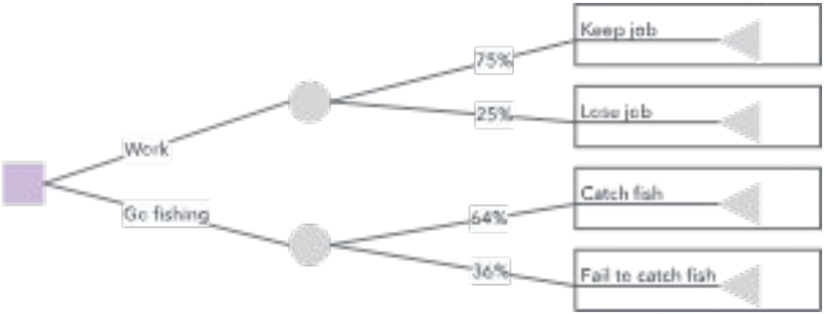


Neural Soft Decision Trees (SDTs)

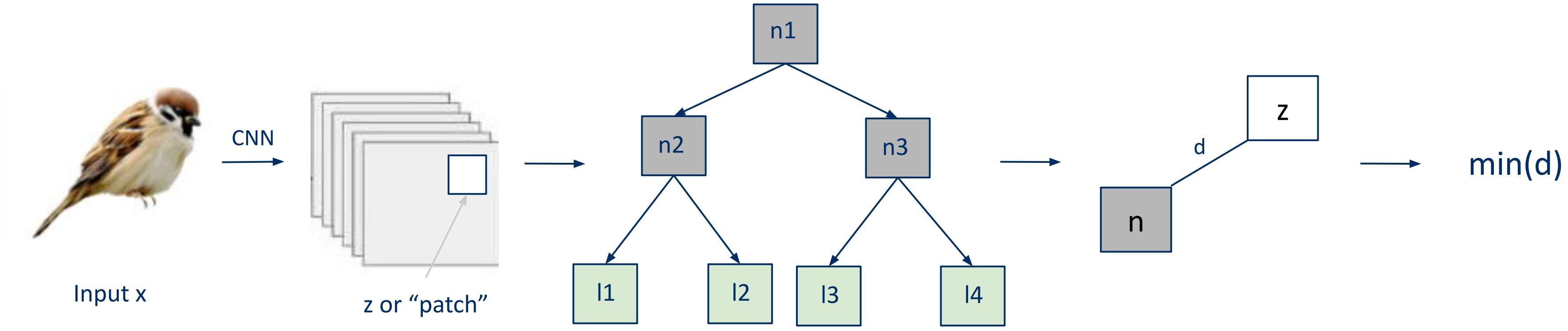
Hard decision tree

Soft decision tree

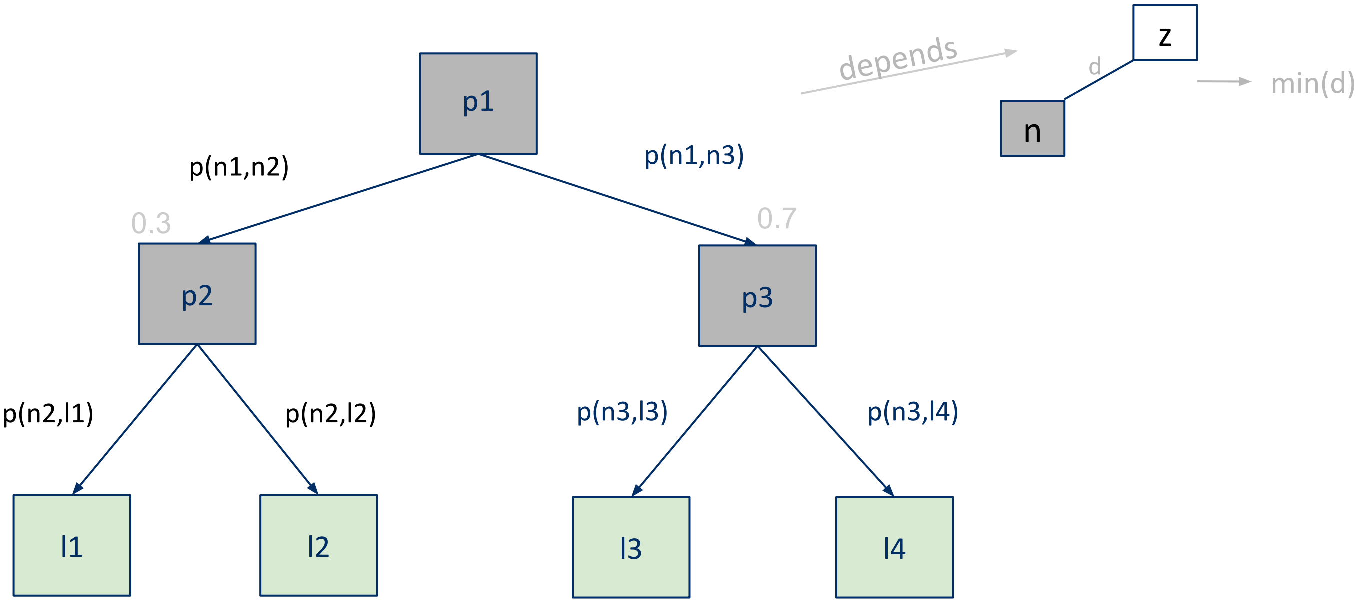
- More flexible
- Complex relationships



Model architecture



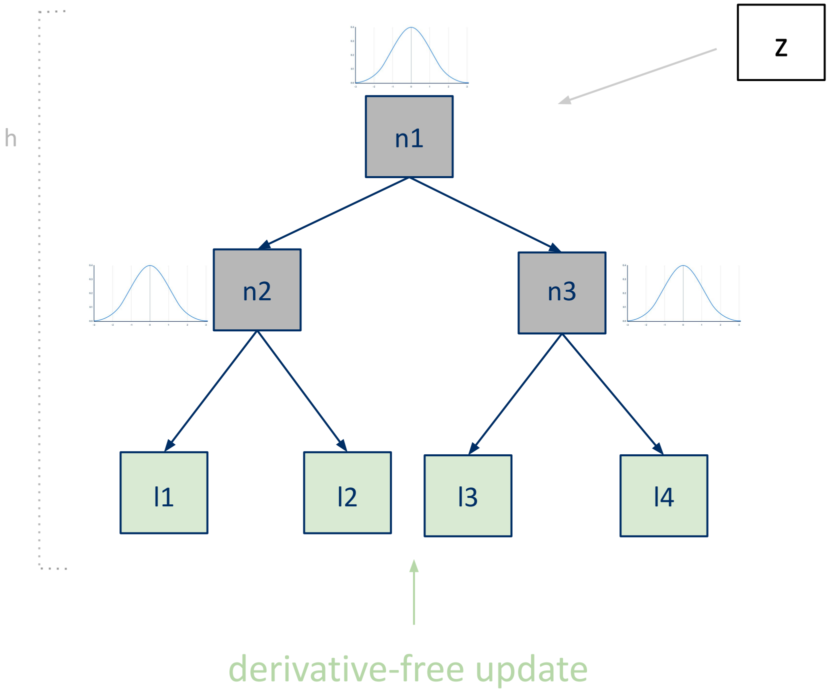
Model architecture



$\pi(l1) * \sigma(c1) + \pi(l2) * \sigma(c2) + \pi(l3) * \sigma(c3) + \pi(l4) * \sigma(c4)$ = final prediction

Training

Pretrained CNN



Backpropagation in mini batches

Deterministic reasoning

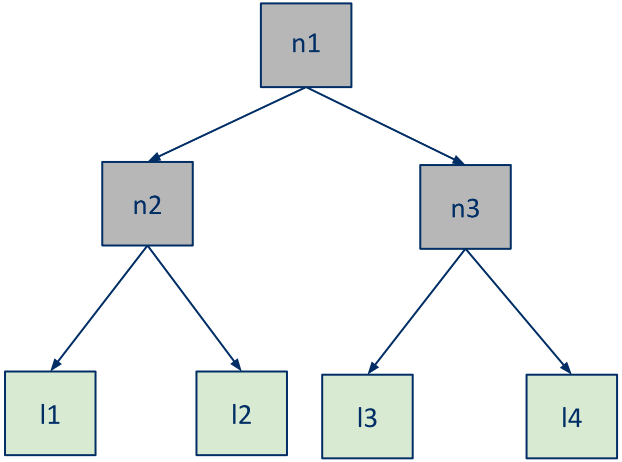
- From soft decision tree to hard decision tree at test time
- Each input gets the same prediction every time
- Same accuracy

Results and fine tuning

Data set	Method	Inter-pret.	Top-1 Accuracy	#Proto types
CUB (224 × 224)	Triplet Model [34]	-	**87.5**	n.a.
	TranSlider [58]	-	85.8	n.a.
	TASN [57]	o	87.0	n.a.
	ProtoPNet [9]	+	79.2	2000
	ProtoTree h=9 (ours)	++	82.2±0.7	**202**
	ProtoPNet ens. (3) [9]	+	84.8	6000
	ProtoTree ens. (3)	+	86.6	605
	ProtoTree ens. (5)	+	**87.2**	1008
CARS (224 × 224)	RAU [36]	-	**93.8**	n.a.
	Triplet Model [34]	-	93.6	n.a.
	TASN [57]	o	93.8	n.a.
	ProtoPNet [9]	+	86.1	1960
	ProtoTree h=11 (ours)	++	86.6±0.2	**195**
	ProtoPNet ens. (3) [9]	+	91.4	5880
	ProtoTree ens. (3)	+	90.3	586
	ProtoTree ens. (5)	+	**91.5**	977

Pruning

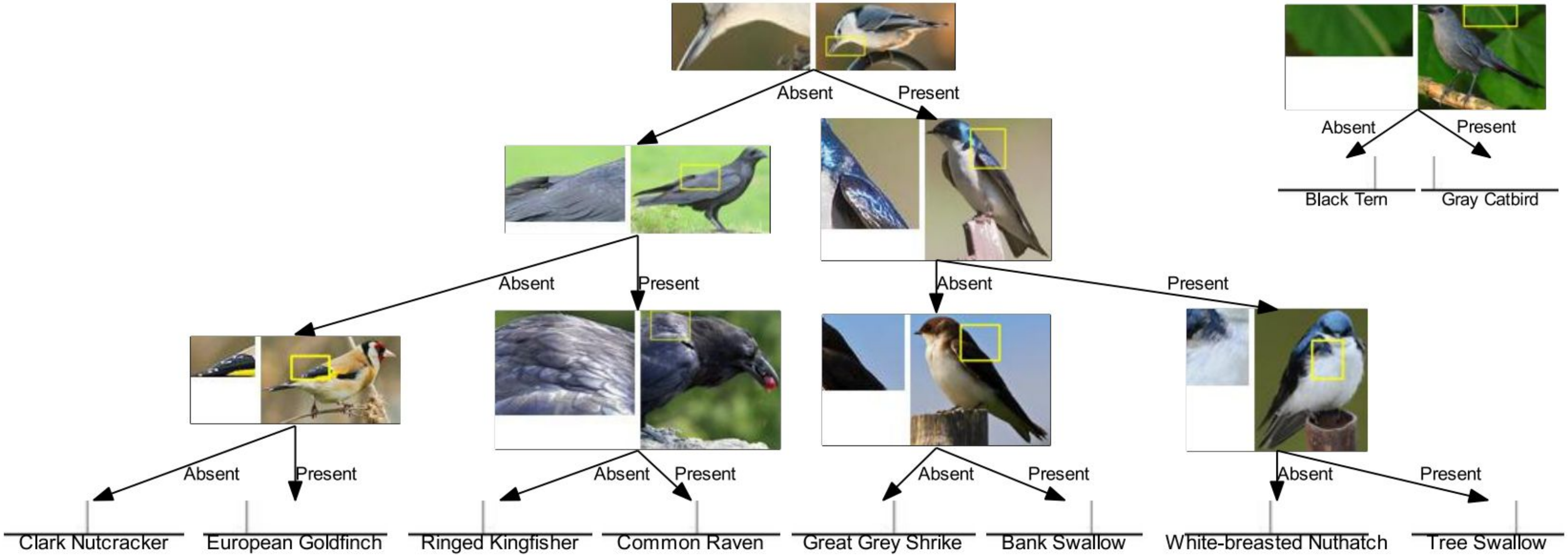
- Reduce complexity
- Enhance interpretability without losing significant accuracy



$\max(\ \sigma(c1)\quad \sigma(c2)\quad \sigma(c3)\quad \sigma(c4)\ ) <= \text{threshold } \tau$

Summarizing

- Need to create interpretable models by design
- Accuracy-Interpretability trade-off



University of Antwerp
Faculty of Science

Neural Prototype Trees for Interpretable Fine-grained Image Recognition

Konstantina Ellina
Pablo de Vicente Abad