Random-Set Neural Networks

Shireen Kudukkil Manchingal[1*], Muhammad Mubashar[1] , Kaizheng Wang[2,4], Keivan Shariatmadar[3,4], Fabio Cuzzolin[1]

[1]School of Engineering, Computing and Mathematics, Oxford Brookes University, UK
[2]M-Group and DistriNet Division, Department of Computer Science, KU Leuven, Belgium
[3]LMSD Division, Mechanical Engineering, KU Leuven,
[4]Flanders Make@KU Leuven

ICLR

OXFORD BROOKES UNIVERSITY

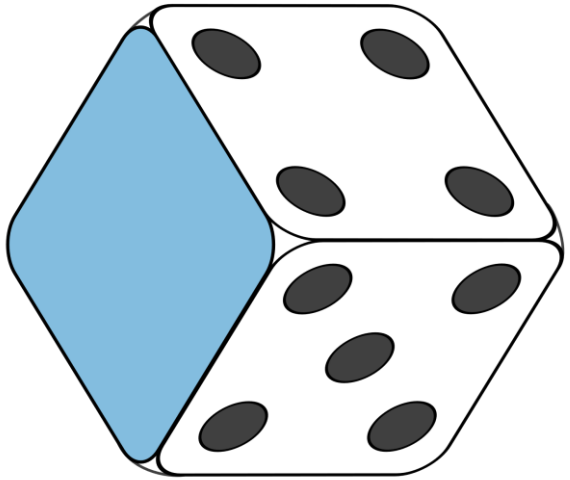
KU LEUVEN

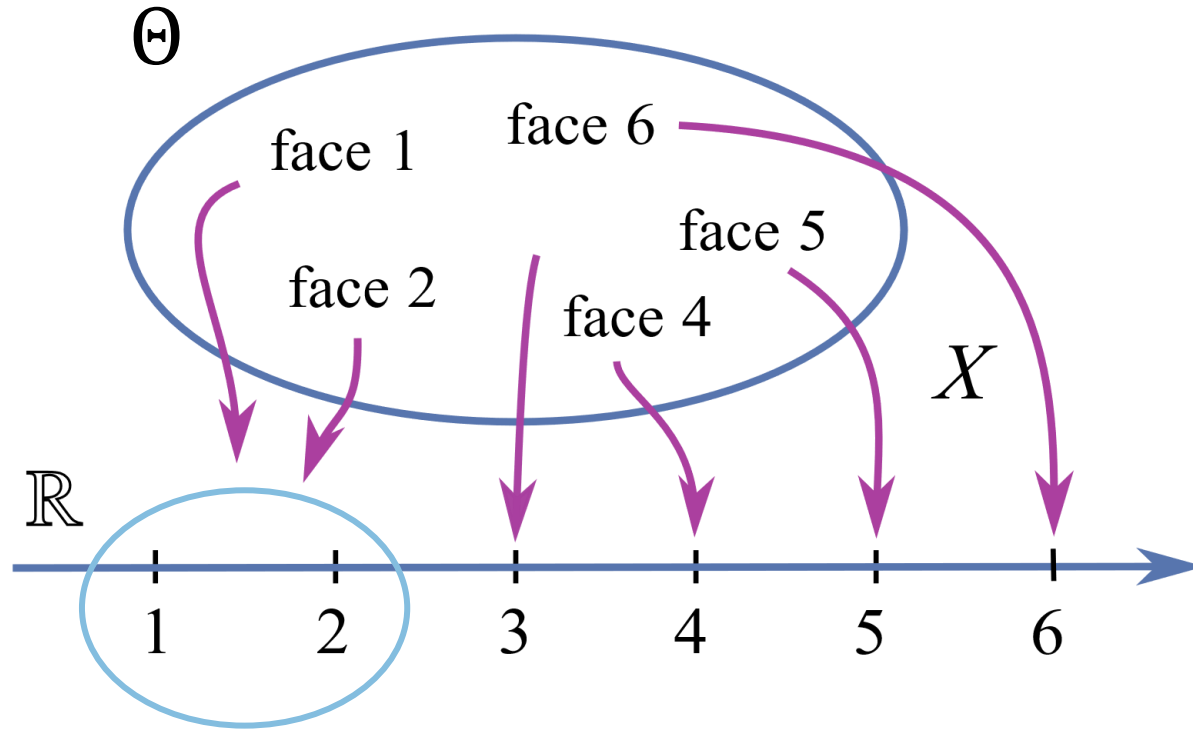
*Corresponding author: 19185895@brookes.ac.uk

This work has received funding from the European Union's Horizon 2020 Research and Innovation program under Grant Agreement No. 964505 (E-pi).

Random Sets

- Die: Θ = {face1, face2, ..., face6} → {1, 2, ..., 6}.
- Cloaked Faces: Faces 1 & 2 → mapped to set{1, 2}.
- **Random Set:** Set-valued random variable.

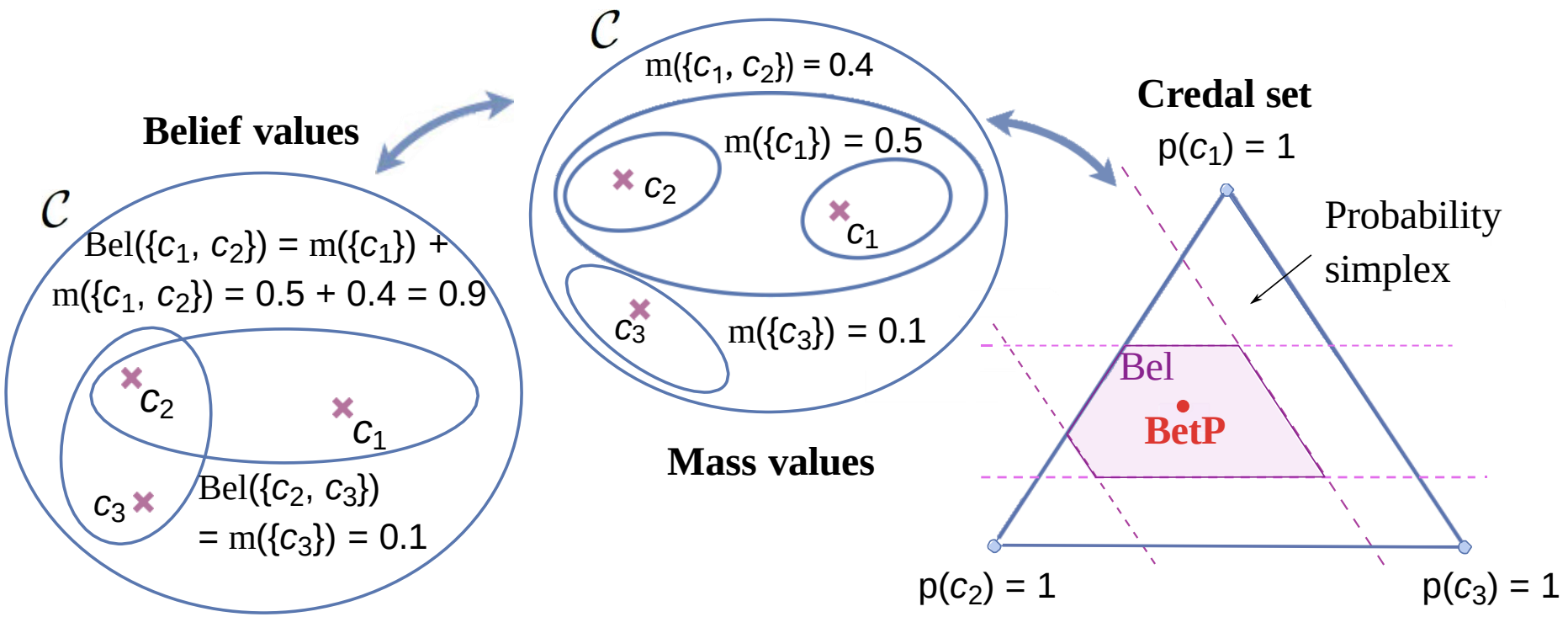




Belief Functions

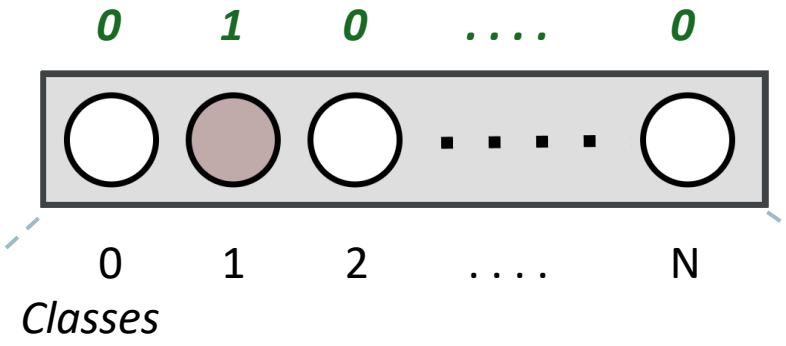
Definition:

- Assigns normalized, non-negative mass values to subsets $A \subseteq \Theta$.
- $m(A) \geq 0$; $\sum m(A) = 1$ for all $A \in P(\Theta)$

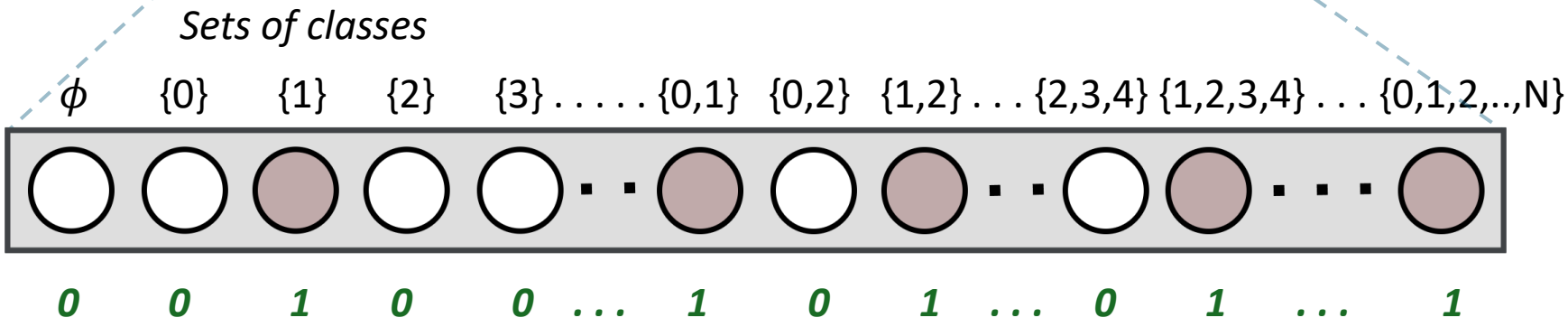


Belief function encoding of ground-truth

Original one-hot encoded ground truth

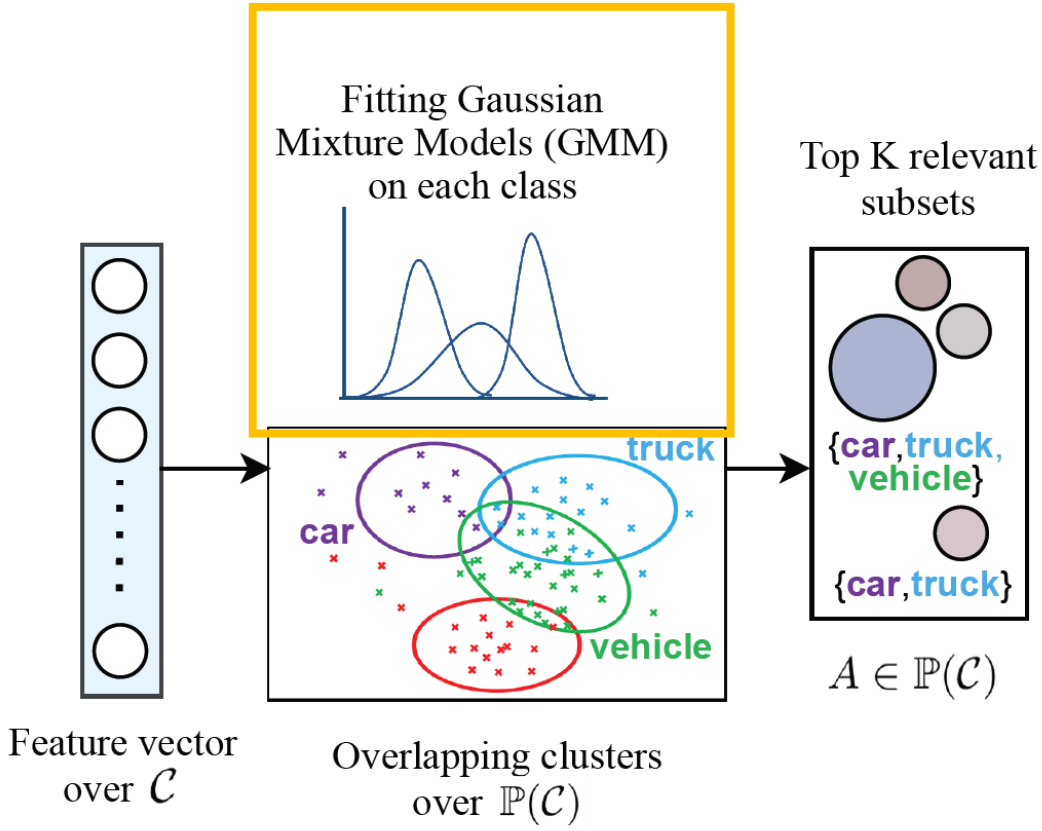


Belief function encoded ground truth



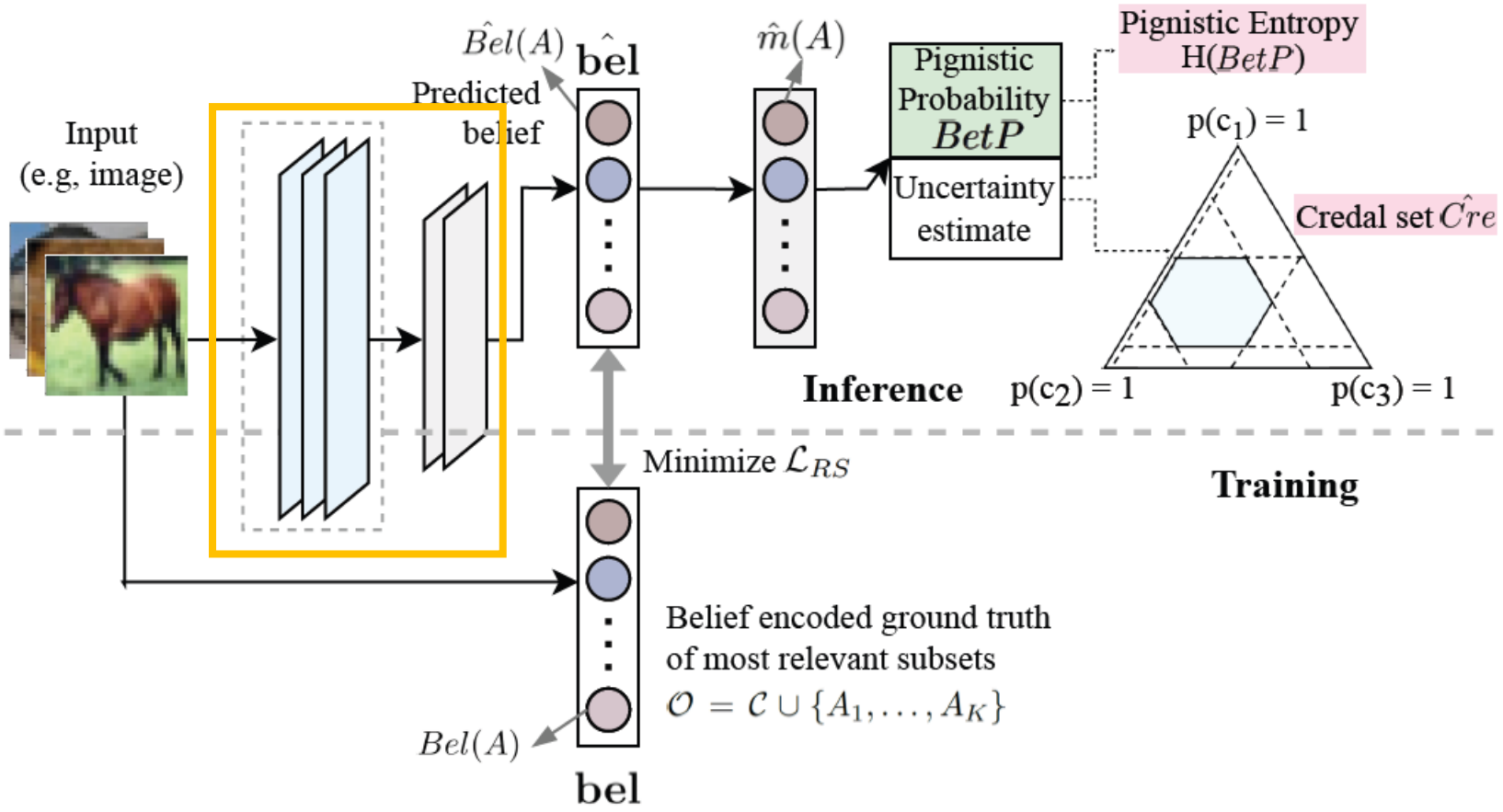
RS-NN Architecture

- Step 1: Selection of an optimal budget of such sets from the training set



RS-NN Architecture

- Step 2: New output layers and loss that predict random sets for relevant sets of classes



RS-NN Performance

- Compared RS-NN with all the most recent baselines
- **Better accuracy**, same **inference time** as standard NN with no uncertainty estimation

Datasets	MNIST	CIFAR-10	Intel Image	CIFAR-100	ImageNet (Top-1)	ImageNet (Top-5)	Inference time (ms)
RS-NN (ours)	$\mathbf{99.71 \pm 0.03}$	$\mathbf{93.53 \pm 0.09}$	$\mathbf{94.22 \pm 0.03}$	$\mathbf{71.61 \pm 0.07}$	**79.92**	**94.47**	$\mathbf{1.91 \pm 0.02}$
LB-BNN (Hobbhahn et al., 2022)	99.58 ± 0.04	89.95 ± 0.81	90.49 ± 0.42	59.89 ± 1.96	72.48	90.85	7.11 ± 0.89
FSVI (Rudner et al., 2022)	99.18 ± 0.03	80.29 ± 0.05	88.92 ± 0.13	53.34 ± 0.09	62.56	84.69	340.25 ± 0.76
DE (Lakshminarayanan et al., 2017)	99.25 ± 0.01	92.73 ± 0.04	91.98 ± 0.11	70.53 ± 0.07	78.77	94.37	13163.50 ± 3.37
ENN (Osband et al., 2024)	99.07 ± 0.11	91.55 ± 0.60	91.49 ± 0.19	68.02 ± 0.26	71.82	89.48	3.10 ± 0.03
CNN	99.12 ± 0.04	92.08 ± 0.42	90.89 ± 0.10	65.50 ± 0.08	78.56	94.34	$\mathbf{1.91 \pm 0.03}$

RS-NN Performance

- But also: better ability to identify **out-of-distribution (OoD)** data (i.e., unusual or rare data)
- Better in-distribution vs out-of-distribution **uncertainty** estimation

In-distribution (iD)						Out-of-distribution (OoD)					
						SVHN			Intel Image		
Dataset	Model	Test accuracy (%) (↑)	Uncertainty measure	In-distribution Entropy (↓)	ECE (↓)	AUROC (↑)	AUPRC (↑)	Entropy (↑)	AUROC (↑)	AUPRC (↑)	Entropy (↑)
CIFAR-10	RS-NN	**93.53**	Pignistic entropy	**0.088 ± 0.308**	0.0484	**94.91**	**93.72**	1.132 ± 0.855	**97.39**	**90.27**	**1.517 ± 0.740**
	LB-BNN	89.95	Predictive Entropy	0.191 ± 0.412	0.0585	88.14	81.96	0.828 ± 0.243	82.21	55.17	0.763 ± 0.722
	FSVI	80.29	Predictive Entropy	0.118 ± 0.563	0.0521	80.59	80.84	0.413 ± 0.461	74.27	72.51	0.289 ± 0.670
	DE	92.73	Mean Entropy	0.154 ± 0.367	**0.0482**	93.84	91.88	0.939 ± 0.554	94.25	79.36	1.166 ± 0.552
	ENN	91.55	Mean Entropy	0.126 ± 0.323	0.0556	92.76	89.05	0.887 ± 0.514	85.67	58.09	0.600 ± 0.578
	CNN	92.08	Softmax Entropy	0.114 ± 0.304	0.0669	93.11	91.0	0.930 ± 0.610	87.75	65.54	0.719 ± 0.673
						F-MNIST			K-MNIST		
MNIST	RS-NN	**99.71**	Pignistic entropy	**0.010 ± 0.111**	**0.0029**	**93.89**	**93.98**	0.530 ± 0.770	**96.75**	**96.58**	**0.740 ± 0.917**
	LB-BNN	99.58	Predictive Entropy	**0.001 ± 0.085**	0.0032	89.65	90.36	0.287 ± 0.442	95.61	95.65	0.540 ± 0.621
	FSVI	99.18	Predictive Entropy	0.006 ± 0.265	0.0047	92.79	91.17	0.264 ± 0.289	91.65	95.75	0.313 ± 0.381
	DE	99.25	Mean Entropy	0.031 ± 0.155	0.0031	92.30	92.05	**0.584 ± 0.587**	95.81	94.71	0.564 ± 0.715
	ENN	99.07	Mean Entropy	0.022 ± 0.127	0.0039	81.79	82.92	0.313 ± 0.464	95.94	95.45	0.503 ± 0.672
	CNN	98.90	Softmax Entropy	0.023 ± 0.135	0.0052	83.77	84.14	0.278 ± 0.426	94.46	93.94	0.616 ± 0.688
						ImageNet-O					
						AUROC	AUPRC	Entropy			
ImageNet	RS-NN	**79.92**	Pignistic entropy	2.972 ± 2.108	0.1416	**60.38**	**55.16**	3.659 ± 3.771			
	LB-BNN	72.48	Predictive Entropy	2.471 ± 2.972	0.5812	41.08	30.99	1.383 ± 0.028			
	FSVI	62.56	Predictive Entropy	**1.328 ± 1.966**	0.3890	50.55	49.88	1.637 ± 1.328			
	DE	78.77	Mean Entropy	1.532 ± 1.325	0.1940	55.37	53.20	1.775 ± 1.343			
	ENN	71.82	Mean Entropy	1.395 ± 1.510	0.5961	54.67	43.73	1.617 ± 1.597			
	CNN	78.56	Softmax Entropy	6.386 ± 1.388	0.4004	54.28	48.73	**6.575 ± 1.512**			

RS-NN Performance

- Increased robustness to adversarial attacks
- Statistical guarantees, using conformal learning, on how often the prediction is correct!

Read the full paper here:

Acknowledgement

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 964505 (E-pi).

Thank you!