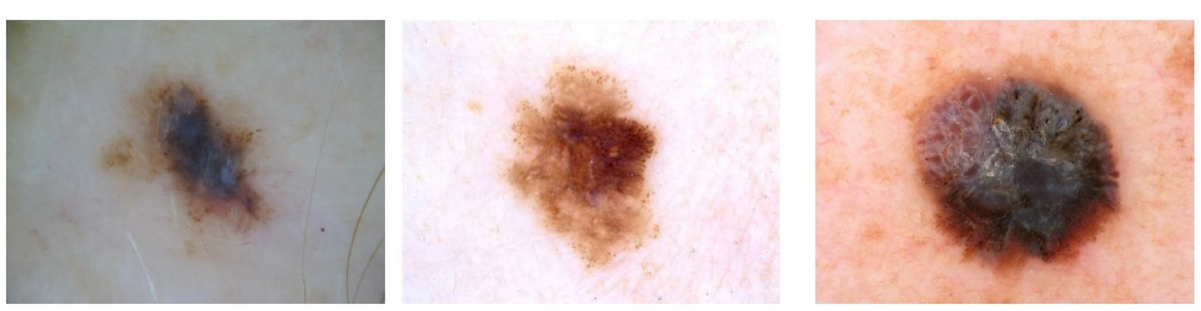
SkinCON: Towards consensus for the uncertainty of skin cancer sub-typing through distribution regularized adaptive predictive sets (DRAPS)

Zhihang Ren, Yunqi Li, Xinyu Li, Xinrong Xie, Erik P. Duhaime, Kathy Fang,
Tapabrata Chakraborti, Yunhui Guo, Stella X. Yu, David Whitney

Argmax Prediction is not Reliable for Diagnosis



Prediction with **uncertainty estimation** provides more diagnostic indication!

Contribution

- We collect and curate the **first** multi-label skin cancer dataset **SkinCon** that reveals **instance-level empirical response distribution**.
- We propose a novel method **Distribution Regularized Adaptive Prediction Sets, DRAPS**, satisfying the coverage rate and achieving state-of-the-art conformal set size.

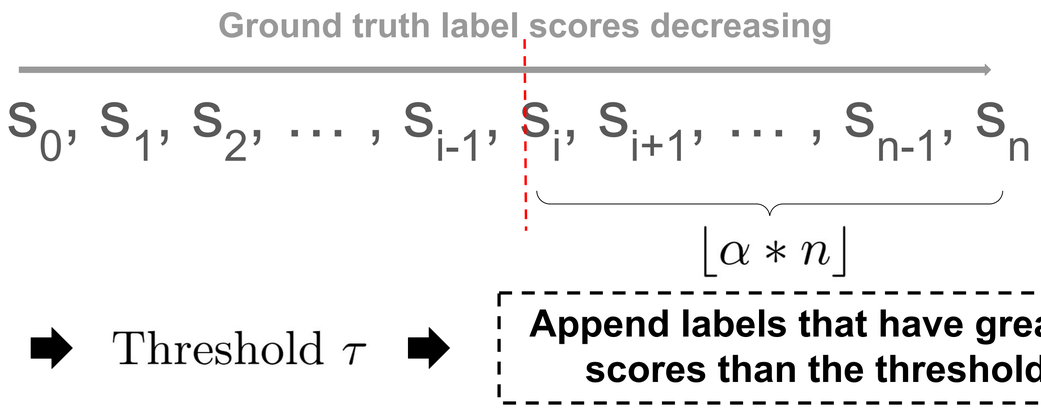
Prediction Set for Uncertainty Estimation

Imagine we have *n* data samples $\{(X_i, Y_i)\}_{i=1}^{n}$, and a discrete label Y_i. Given such data and a desired coverage level $1-\alpha \in (0,1)$,we seek to construct a prediction set $\hat{C}_{n,\alpha}$ for the unseen label of a new data point (X_{n+1}, Y_{n+1}) achieving marginal coverage; that is, obeying

$$\mathbb{P}[Y_{n+1} \in \hat{C}_{n,\alpha}(X_{n+1})] \geq 1-\alpha$$

Distribution Regularized Adaptive Predictive Sets (DRAPS)

Base model is trained with additional KL divergence loss via SkinCON empirical response distribution data.

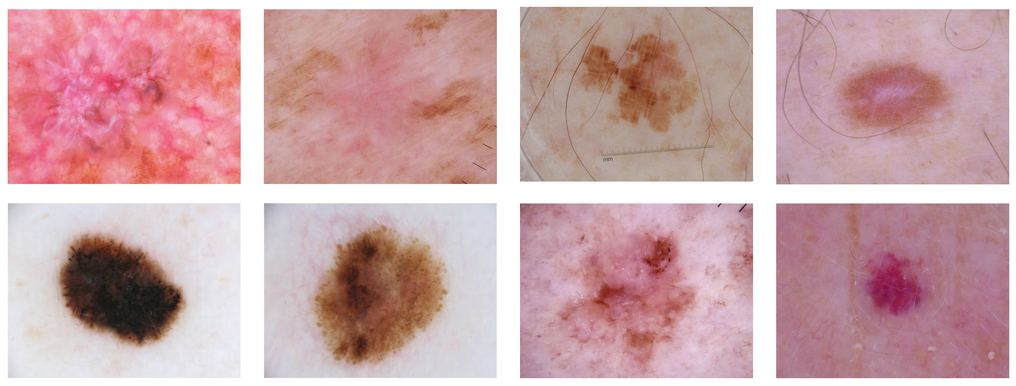


DRAPS Performance

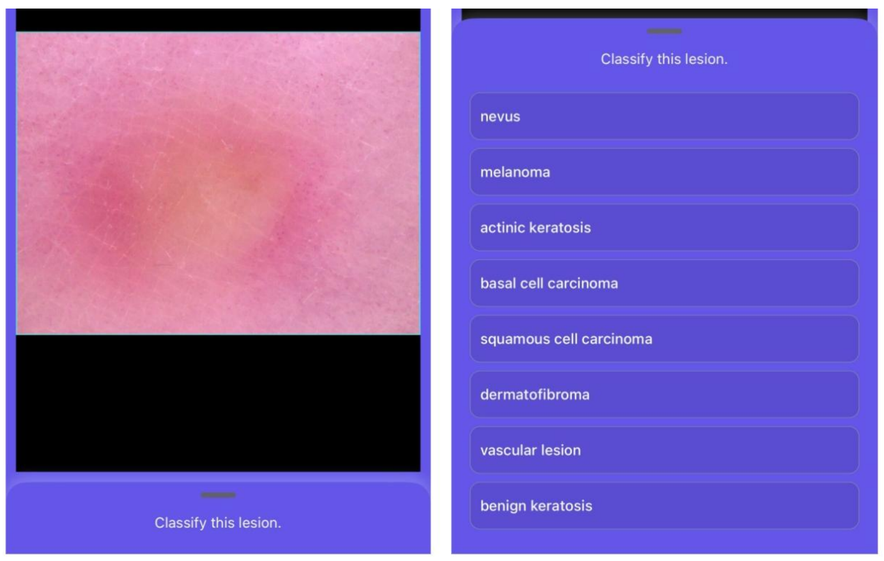
Model	Accuracy		Coverage			Size		
	Top-1	Top-5	Naive	RAPS	Ours	Naive	RAPS	Ours
ResNet18	91.62	99.94	0.928	0.964	0.935	1.278	1.183	**1.054**
ResNet50	92.07	100.0	0.935	0.967	0.941	1.165	1.192	**1.104**
ResNet101	93.02	100.0	0.947	0.961	0.931	1.367	1.163	**1.047**
ResNet152	91.50	99.94	0.934	0.961	0.936	1.241	1.167	**1.021**
ResNeXt101	92.77	99.87	0.938	0.972	0.941	1.050	1.179	**1.006**
VGG16	91.52	99.87	0.927	0.961	0.924	1.056	1.213	**1.008**
ShuffeNet	89.85	99.56	0.923	0.968	0.926	1.241	1.414	**1.154**
DenseNet161	92.83	99.94	0.941	0.971	0.936	1.141	1.146	**1.034**

SkinCON Dataset

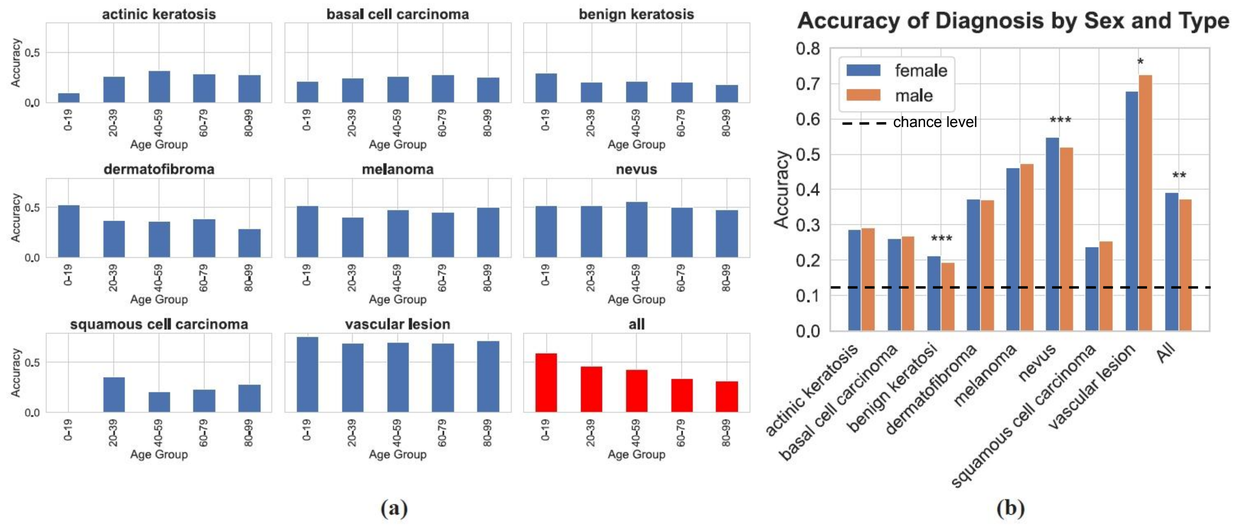
- **25,330** skin cancer images from the ISIC 2019 challenge dataset
- **937,167** diagnostic trials from **10,509** proficient participants
- **8** skin cancer types include: actinic keratosis, basal cell carcinoma, benign keratosis, dermatofibroma, melanoma, nevus, squamous cell carcinoma, and vascular lesion.
- An instance-level empirical response distribution dataset



Diagnosis Annotation UI



Diagnostic Biases over Age, Sex, and Lesion Type



Diverse Response Distributions in SkinCON Dataset