KØBENHAVNS UNIVERSITET
DEPARTMENT OF COMPUTER SCIENCE

Reducing Annotation Need in Self-Explanatory Models for Lung Nodule Diagnosis (cRedAnno)

J. Lu[1,2], C. Yin[1,3], O. Krause[1], K. Erleben[1], M. B. Nielsen[2], S. Darkner[1]   lu@di.ku.dk

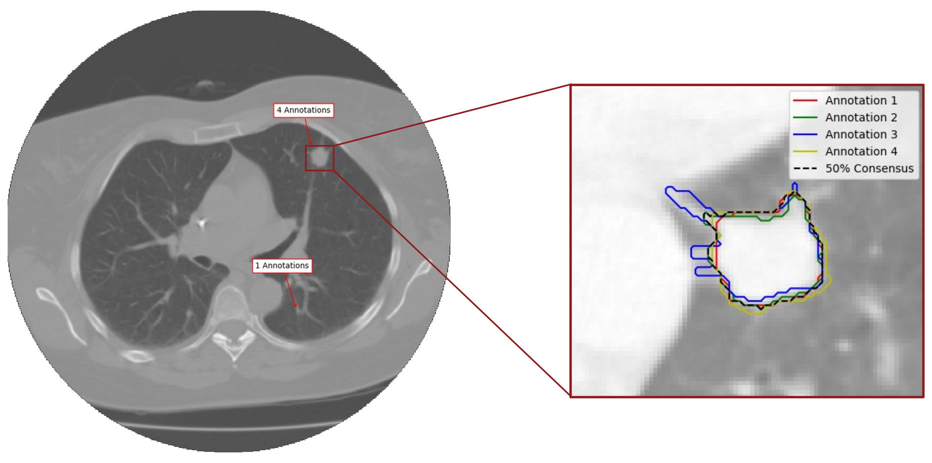
[1] Department of Computer Science, University of Copenhagen, Denmark
[2] Department of Diagnostic Radiology, Rigshospitalet, Copenhagen University Hospital, Denmark
[3] Department of Computer Science, Hong Kong Baptist University, China

github.com/diku-dk/credanno

Background

Lung nodule diagnosis

In clinical practice, accurate characterisation of pulmonary nodules in CT images is an essential step for effective lung cancer screening. Modern deep-learning-based "black box" algorithms, although achieving accurate classification performance, are hardly acceptable in high-stakes medical diagnosis.



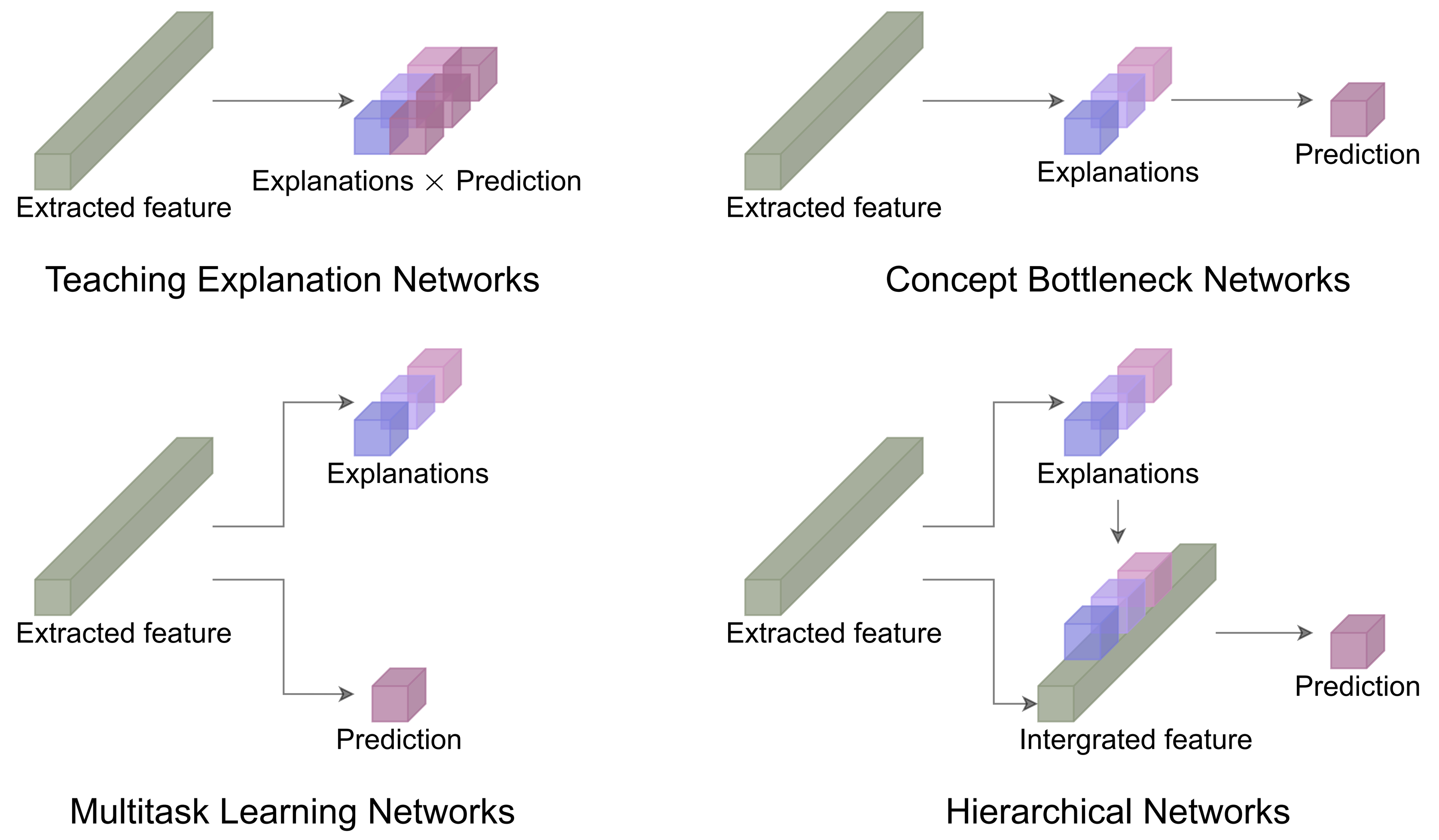
CT scan   Nodule segmentation

Nodule attributes	Subtlety	Obvious
	InternalStructure	Soft Tissue
	Calcification	Non-central
	Sphericity	Ovoid
	Margin	Sharp
	Lobulation	Nearly No Lobulation
	Spiculation	Medium Spiculation
	Texture	Solid
	Malignancy	Moderately Suspicious

Annotation info

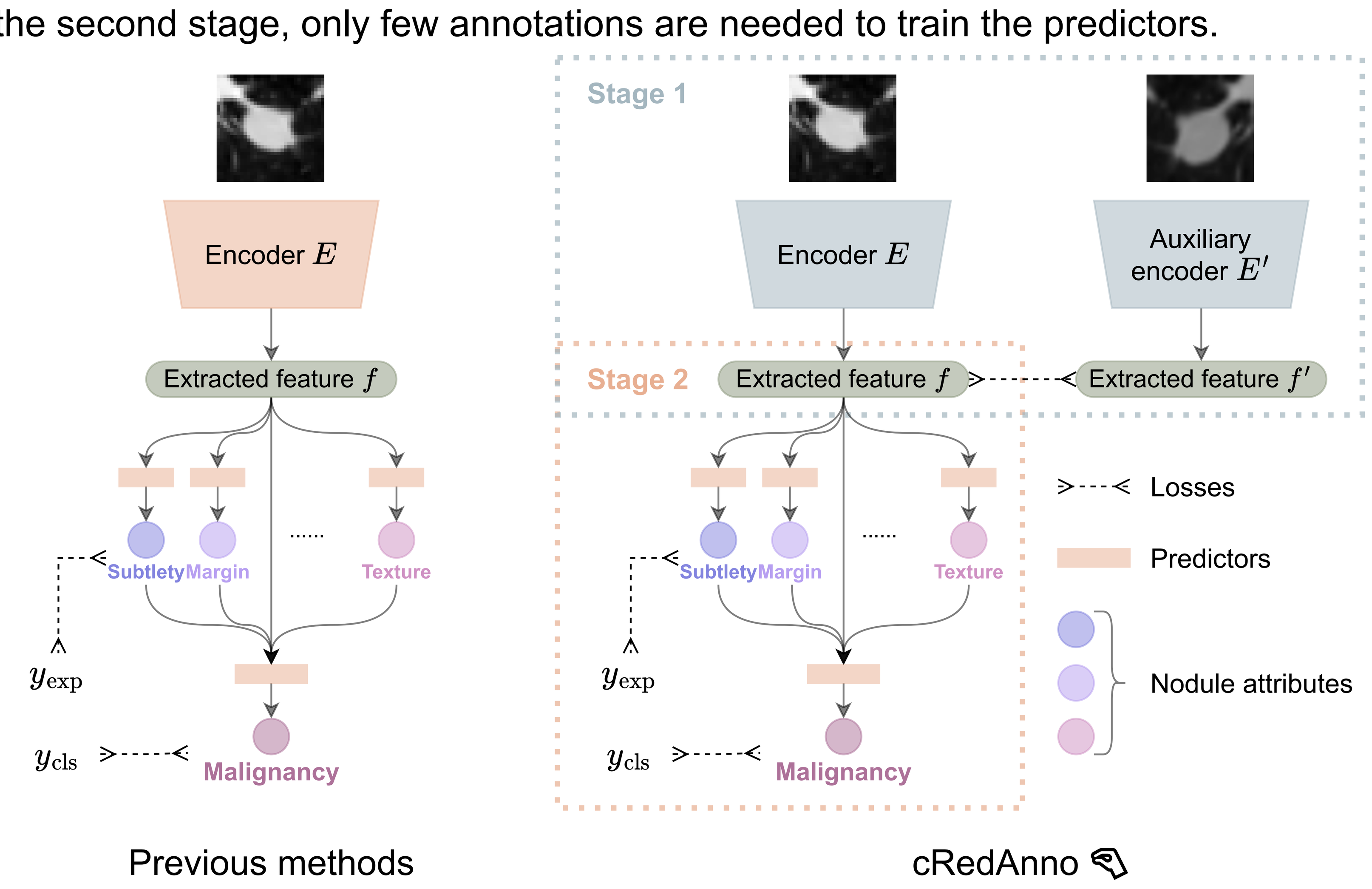
Self-explanatory models

Feature-based self-explanatory methods explain their classification in terms of human-understandable features. In medical applications, this semantic matching of clinical knowledge adds significantly to the trustworthiness of the AI. However, the cost of additional annotation of features remains a pressing issue.



Method

(a) Previous works are trained end-to-end, where all parameters are learned from the annotations. (b) Our proposed cRedAnno uses two-stage training, where most of the parameters are learned during the first stage in a self-supervised manner. Therefore, in the second stage, only few annotations are needed to train the predictors.



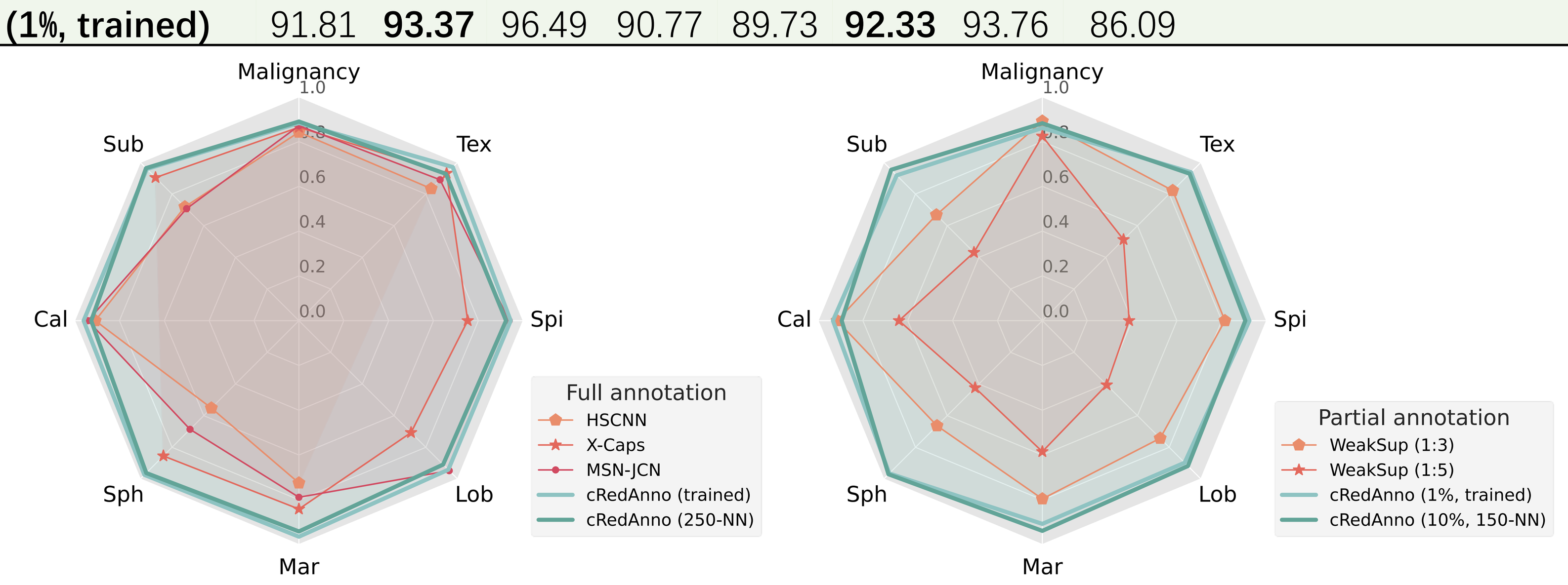
Stage 1: unsupervised feature extraction. The majority of parameters are trained using self-supervised contrastive learning as an encoder to map the input images to a latent space that complies with radiologists' reasoning for nodule malignancy.

Stage 2: supervised prediction. A small random portion of labelled samples is used to train a simple predictor for each nodule attribute. Then the predicted human-interpretable nodule attributes are used jointly with the extracted features to make the final classification.

Results

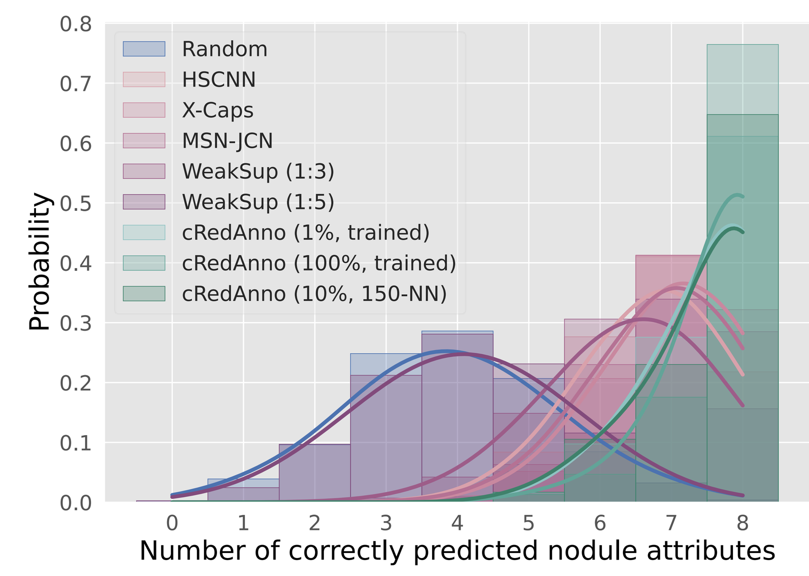
Prediction performance

	Nodule attributes							Malig-nancy	#nodules	No additional info
	Sub	Cal	Sph	Mar	Lob	Spi	Tex			
Full annotation										
HSCNN[1]	71.90	90.80	55.20	72.50	-	-	83.40	84.20	4252	✗
X-Caps[2]	90.39	-	85.44	84.14	70.69	75.23	93.10	86.39	1149	✓
MSN-JCN[3]	70.77	94.07	68.63	78.88	**94.75**	93.75	89.00	87.07	2616	✗
MTMR[4]	-	-	-	-	-	-	-	**93.50**	1422	✗
(50-NN)	94.93	92.72	95.58	93.76	91.29	92.72	94.67	87.52		
(250-NN)	**96.36**	92.59	96.23	94.15	90.90	92.33	92.72	88.95	**730**	✓
*** (trained)**	95.84	**95.97**	**97.40**	**96.49**	94.15	**94.41**	**97.01**	88.30		
Partial annotation										
WeakSup[5] (1:5)	43.10	63.90	42.40	58.50	40.60	38.70	51.20	82.40	2558	✗
WeakSup[5] (1:3)	66.80	91.50	66.40	79.60	74.30	81.40	82.20	**89.10**		
(10%, 50-NN)	94.93	92.07	**96.75**	**94.28**	**92.59**	91.16	**94.15**	87.13		
*** (10%, 150-NN)**	**95.32**	89.47	97.01	93.89	91.81	90.51	92.85	88.17	**730**	✓
*** (1%, trained)**	91.81	**93.37**	96.49	90.77	89.73	**92.33**	93.76	86.09		

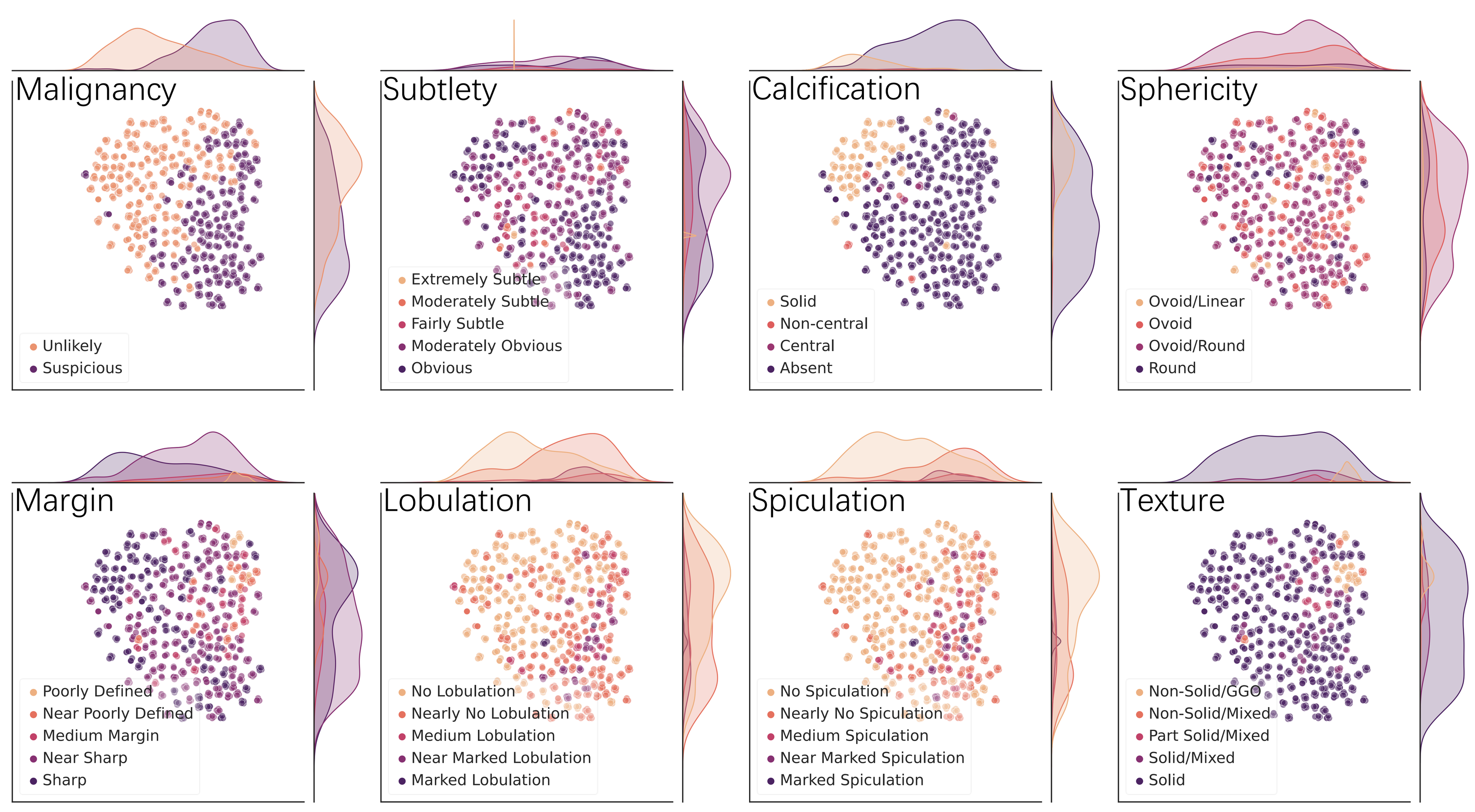


cRedAnno in almost all cases outperforms other methods in nodule attributes significantly, and shows robustness w.r.t. configurations, meanwhile using the fewest nodules and no additional information.

cRedAnno shows a significantly larger probability of simultaneously predicting all 8 nodule attributes correctly. Approximately 90% nodules have at least 7 attributes correctly predicted.



Analysis of extracted features



Malignancy shows highly separable in the learned space and correlates with the clustering in each nodule attribute.

Conclusion

- We propose cRedAnno, a data-/annotation-efficient self-explanatory approach for lung nodule diagnosis.
- When training with hundreds of nodule samples and only 1% of their annotations, cRedAnno significantly outperforms previous works in predicting nodule attributes, meanwhile achieving competitive accuracy in predicting malignancy.
- Visualisation of the learned space further indicates that the correlation between the clustering of malignancy and nodule attributes coincides with clinical knowledge.

References
[1] S. Shen et al., "An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification," Expert Systems with Applications, 2019.
[2] R. LaLonde et al., "Encoding Visual Attributes in Capsules for Explainable Medical Diagnoses," MICCAI, 2020.
[3] W. Chen et al., "End-to-End Multi-Task Learning for Lung Nodule Segmentation and Diagnosis," ICPR, 2021.
[4] L. Liu et al., "Multi-Task Deep Model With Margin Ranking Loss for Lung Nodule Analysis," IEEE Trans. Med. Imaging, 2020.
[5] A. Joshi et al., "Lung nodule malignancy classification with weakly supervised explanation generation," J. Med. Imag., 2021.