Breast Cancer Classification with Capsules and Convolutional Neural Nets

Marcell Veiner
University of Aberdeen: Department of Computing Science

UNIVERSITY OF ABERDEEN

Introduction

- Breast cancer is one of the most common type of cancer in the UK.
- 1 out of every 8 females are diagnosed with it at some point.
- 4th most common cause of cancer death in the UK.

Figure 1: Breast Cancer Statistics in the UK.

- Diagnosis often set up with the help of computer aided techniques.
- Current approaches (CNNs), discard valuable low-level in-formation.
- New approach: Capsule Networks.

Capsule Networks

- Novel deep learning approach proposed to address the flaws of CNNs.
- Its first version appeared in [6], which has been refined several times since.
- Learning process similar to inverse graphics.
- Recognise not only objects but their attributes (orientation, size, skew…)

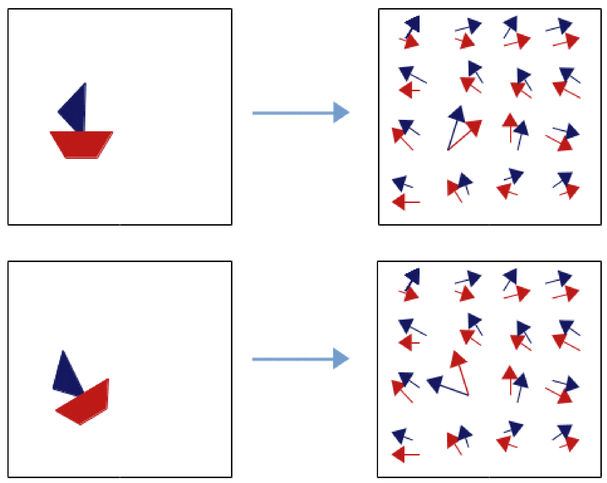

Figure 2: Capsules Finding Parts of Objects [3].

- Capsules are more robust to adversarial attacks [4].
- Viewpoint changes have linear effects on part-whole relationships [4].
- Capsules are equivariant to translation and affine transformations [6].
- Capsules do not discard information about the position of an entity [6].

Datasets

- 3 datasets, patched to 512 x 512 using 2 different strides.

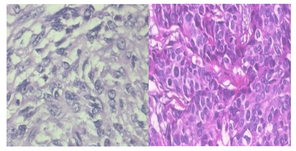

Figure 3: Examples of the BreakHis dataset (Benign & Malignant) [8].

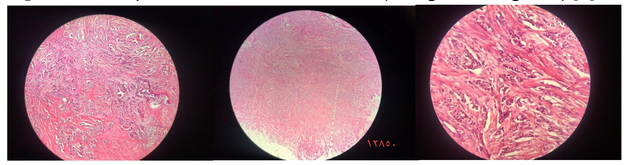

Figure 4: Examples of the Databiox dataset (Grade I – III) [2].

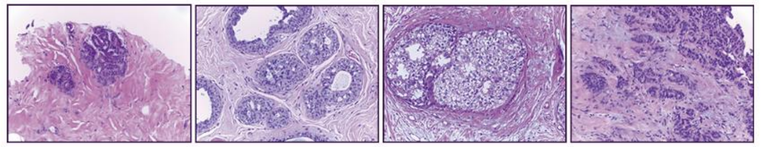

Figure 5: Examples of the BACH dataset [1].

Architecture

- Two stage architecture as in [5].
- First model (Patch-wise network) downscales patches.
- Second model (Image-wise network) uses patch-voting to establish label.

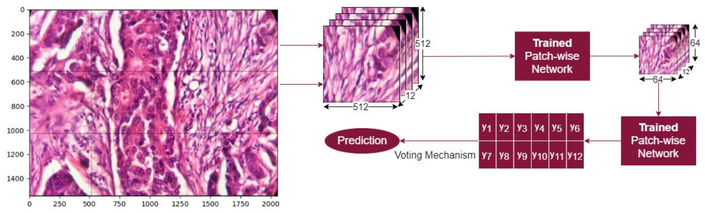

Figure 6: Image-wise Phase of Training.

- Patch-wise network fixed.
- Image-wise network: BaseCNN, NazeriCNN, DynamicCapsules, SRCapsules, VariationalCapsules.
- Mixed Networks: EfficientNet, VariationalMixedCapsules.

Results

Dataset	Model	Training Accuracy	Validation Accuracy	Test Accuracy	F1	Maj. Vote	Sum Vote	Max Vote
BreakHis	BaseCNN	0.96	0.95	0.95	0.96	N/A		
	NazeriCNN	0.94	0.92	0.94	0.90			
	DynamicCaps	0.96	0.95	0.95	0.95			
	SRCaps	0.96	0.96	0.95	**0.97**			
	VarCaps	0.93	0.95	0.95	0.96			
	VarMixedCaps	0.89	0.92	0.93	0.95			
	EffNet	**1.00**	**0.98**	**0.98**	**0.97**			
Databiox	BaseCNN	0.69	0.56	0.57	0.57	0.63	0.66	0.66
	NazeriCNN	0.72	0.54	0.58	0.58	0.63	0.63	0.65
	DynamicCaps	0.73	0,53	**0.61**	0.59	0.64	0.63	**0.67**
	SRCaps	0.71	0.58	**0.61**	0.61	0.64	0,65	**0.67**
	VarCaps	0.58	0.50	0.55	0.55	0.58	0.60	0.58
	VarMixedCaps	0.54	0.49	0.52	0.52	0.57	0.55	0.54
	EffNet	**0.84**	**0.59**	**0.61**	**0.64**	**0.71**	**0.68**	0.64
BACH	BaseCNN	0.92	0.69	0.72	0.71	0.83	0.77	0.75
	NazeriCNN	0.80	0.63	0.61	0.61	0.73	0.68	0.58
	DynamicCaps	0.92	0.68	0.67	0.67	0,75	0.78	0.73
	SRCaps	0.86	0.68	0.68	0.68	0.77	0.75	0.70
	VarCaps	0.63	0.67	0.64	0.64	0.79	0.74	0.70
	VarMixedCaps	0.63	0.59	0.55	0.55	-	-	-
	EffNet	**0.98**	**0.75**	**0.81**	**0.80**	**0.90**	**0.88**	**0.87**

Table 1: Overall Performance.

Conclusion

- BreakHis: competing with published results.
- Databiox: Poor overall performance, models stuck.
- BACH: EffNet competing performance.

- Two-staged approach: Capsules outperforming baseline CNNs.
- Overall: Transfer Learning superior.

References

[1] Guilherme Aresta, et al. Bach: Grand challenge on breast cancer histology images. Medical image analysis, 56:122–139, 2019.
[2] Hamidreza Bolhasani, et al. Informaticsin Medicine Unlocked, 19:100341, 2020.
[3] Charlotte Burmeister. Capsule networks - better cnns?, Jan 2020.
[4] Geoffrey E Hinton, Sara Sabour, and Nicholas Frosst. Matrix capsules with em routing. ininternational conference on learning representations, 2018.
[5] Kamyar Nazeri, Azad Aminpour, and Mehran Ebrahimi. Two-stage convolutional neural network for breast cancer histology image classification. International Conference Image Analysis and Recognition, pages 717–726. Springer, 2018.
[6] Sara Sabour, Nicholas Frosst, and Geoffrey E Hinton. Dynamic routing between capsules. arXiv preprint arXiv:1710.09829, 2017.
[7] Sara Sabour, et al. Unsupervised part representation by flow capsules. arXiv preprint arXiv:2011.13920, 2020.
[8] Fabio A Spanhol, et al. A dataset for breast cancer histopathological image classification. Ieee transactions on biomedical engineering, 63(7):1455–1462, 2015

Contact Information

- Marcell Veiner
- m.veiner.17@adbn.ac.uk
- BSc. Computing Science & Mathematics