Low Resource Breast Cancer Detection with Mammograms

Sara Ebrahim

Agenda

- Motivation.
- Previous work.
- Dataset.
- Approach.
- Results.
- Conclusions.

Motivation

AI advanced in breast cancer detection recently.

How we use these advancement in Africa.

Africa has many hospitals with few mammograms dataset; some may not keep all performed mammograms stored locally.

Motivation

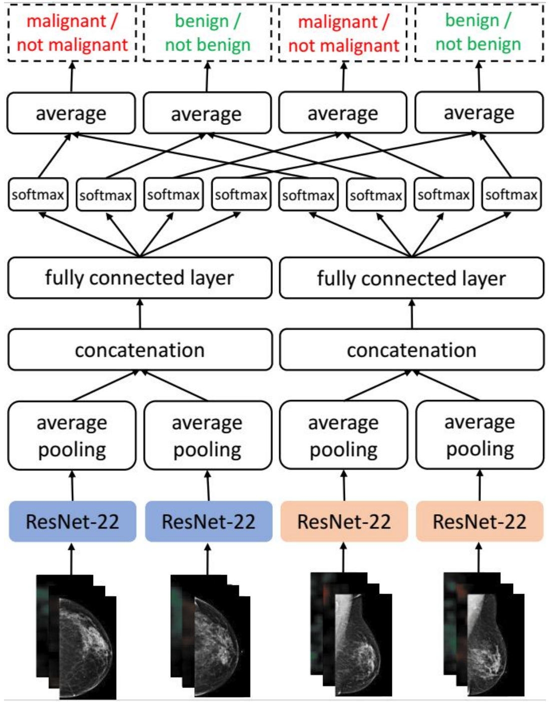
We decided to start with NYU recent research that reached 89% AUC score, that was trained on a large dataset (~1m mammogram).

What can we get if we used such released model trained on a private huge dataset of mammograms?

Previous work: NYU v1.0

NYU trained a deep learning (DL) model on ~1m images from the NYU v1.0 private dataset.

Multiview DL model + heatmaps as channels from an auxiliary patch-based DL model.



Wu, Nan, et al. "Deep neural networks improve radiologists' performance in breast cancer screening." *IEEE transactions on medical imaging* (2019).

Dataset

INBreast dataset: 115 cases (410 images).

Similar to NYU v1.0 dataset: in terms of resolution.

Moreira, Inês C., et al. "Inbreast: toward a full-field digital mammographic database." *Academic radiology* 19.2 (2012): 236-248.

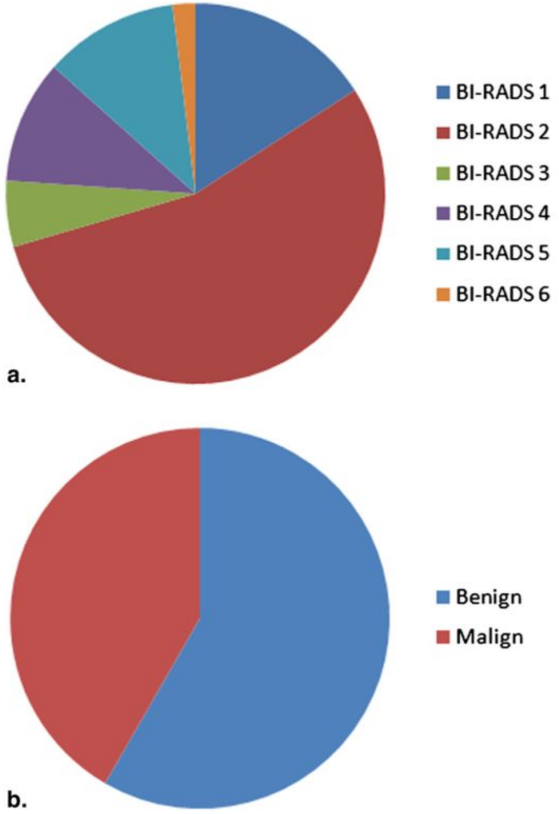


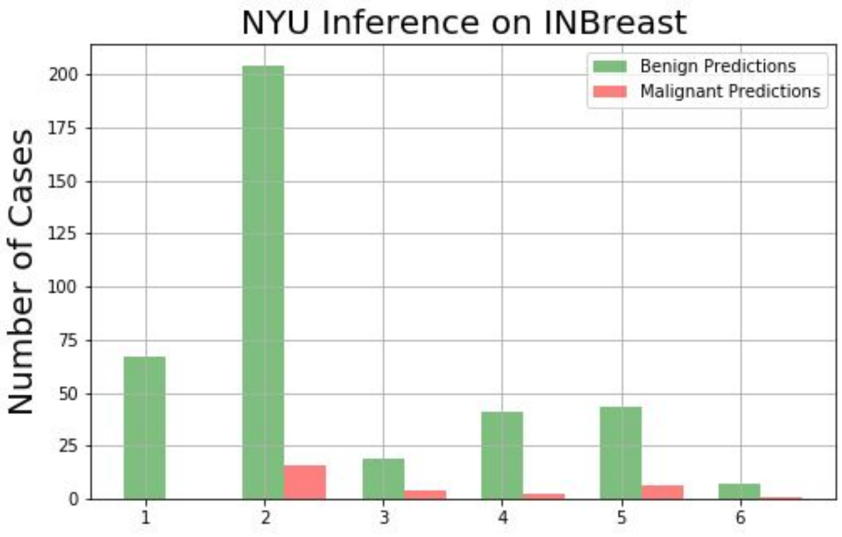
Figure 8. Charts of **(a)** the Breast Imaging Reporting and Data System images distribution **(b)** benign/malignant cases distribution.

Approach

1. Direct inference with NYU model.
2. Transfer learning.
3. Feature extractor then SVM linear classifier.

Results

1. **Direct inference with NYU model.**
2. Transfer learning.
3. Feature extractor then SVM linear classifier.



Approach(es)

1. Direct inference with NYU model.
2. **Transfer learning.**
3. Feature extractor then SVM linear classifier.

Model	Pretrained on	Task	Accuracy	AUROC	Precision	Recall
resnet50	ImageNet	BIRADS	0.549	0.475	0.513	0.481
resnet18	ImageNet	BIRADS	0.573	0.390	0.686	0.524
squeeznet	ImageNet	BIRADS	0.427	0.459	0.270	0.210
resnet50	**ImageNet**	**Binary**	**0.866**	**0.9**	**0.737**	**0.7**
resnet18	ImageNet	Binray	0.854	0.872	0.75	0.6
squeeznet	ImageNet	Binary	0.732	0.695	0.437	0.35
NYU (L-MLO)	NYU v1.0	Binary	0.780	0.544	0.625	0.25
NYU (R-CC)	**NYU v1.0**	**Binary**	**0.793**	**0.581**	**0.667**	**0.3**

Table 2: Fine tuning both Resnet50 pretrained on ImageNet and the NYU network trained on the NYU v1.0 dataset. Common performance metrics are in the last four columns.

Results

1. Direct inference with NYU model.
2. Transfer learning.
3. **Feature extractor then SVM linear classifier.**

Model	balanced	Feature Extractor	# Features	Accuracy	AUC	mean	std
SVM+Linear	No	ImageNet Resnet50	2048	0.6544	0.562	0.58	0.09
SVM+Linear	No	ImageNet Resnet50	217	0.8676	0.8362	0.74	0.11
SVM+RBF	No	ImageNet Resnet50	217	0.75	0.5143	0.51	0.03
LR	No	ImageNet Resnet50	217	0.8676	0.8362	0.76	0.12
SVM+Linear	**SMOTE**	**ImageNet Resnet50**	217	**0.9317**	**0.9345**	**0.94**	**0.03**
LR	SMOTE	ImageNet Resnet50	217	0.9122	0.9157	0.95	0.04
SVM+Linear	No	NYU L-CC	4096	0.6912	0.5680	0.56	0.07
SVM+RBF	No	NYU L-CC	4096	0.7426	0.5	0.50	0.03
SVM+Linear	No	NYU L-CC	350	0.8750	0.8412	0.74	0.10
SVM+RBF	No	NYU L-CC	350	0.7426	0.5	0.50	0.03
LR	No	NYU L-CC	350	0.8456	0.8027	0.72	0.11
SVM+Linear	SMOTE	NYU L-CC	350	0.9610	0.9631	0.96	0.02
LR	SMOTE	NYU L-CC	350	0.9317	0.9361	0.94	0.04
SVM+Linear	No	NYU Avg	325	0.8162	0.7269	0.73	0.07
LR	No	NYU Avg	325	0.8162	0.7082	0.72	0.08
SVM+Linear	**SMOTE**	**NYU Avg**	325	**0.9756**	**0.9758**	**0.97**	**0.04**
SVM+RBF	SMOTE	NYU Avg	325	0.6683	0.6733	0.72	0.07
LR	SMOTE	NYU Avg	325	0.9610	0.9623	0.97	0.03

Table 3: Results of the binary task classification of the INBreast dataset (i.e. BIRADS1,2 and 3 are benign and BIRADS4,5 and 6 are malignant). The used models are SVMs with two different kernel functions (Linear and RBF) and Logistic Regression (LR). Lasso feature selection is used and the number of used features is stated in the 4th column. A 10-fold cross validation of the balanced accuracy is depicted in the last two columns (the mean and standard deviation of the 10 folds). SMOTE was used as an oversampling technique for the malignant class. We use both one feature extractor from NYU model and an average over the four extractors.

Conclusions

- ML + DL sometimes is better than focusing only on one field.
- We might think of a medical-ImageNet dataset for cancer diseases.
- There is an urge for medical centers to release a private-preserved medical datasets to do a wide test on different datasets.

Thank You