# End-to-end Training for Whole Image Breast Cancer Diagnosis using An All Convolutional Design

Li Shen, li.shen@mssm.edu
Icahn School of Medicine at Mount Sinai, New York, New York

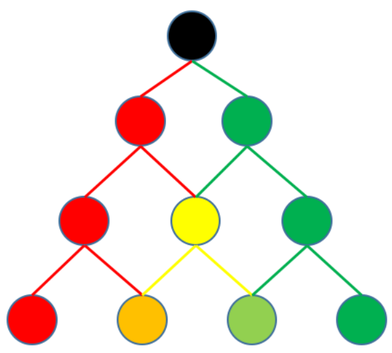
## Introduction

- Mammography is the most commonly used technique for breast cancer diagnosis.
- Harnessing machine learning for breast cancer diagnosis based on mammography is a very hot topic now.
- Large mammograms vs. small lesion sizes.
- Many mammography databases lack lesion annotation.

## Highlights

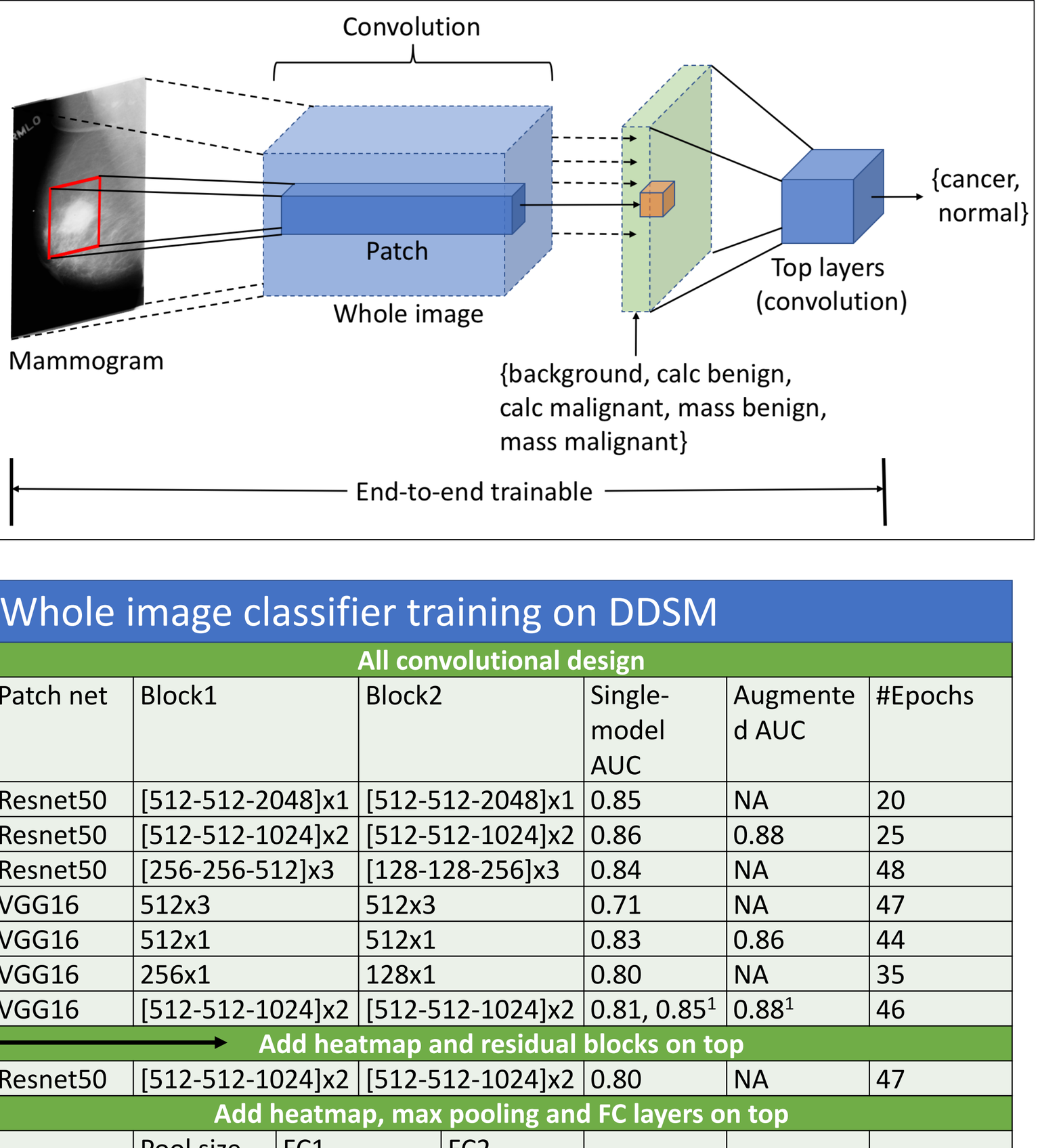
- Our algorithm requires lesion annotation only at the first stage of training.
- A model can be efficiently transferred to another database without lesion annotation.
- On DDSM, we achieve single model AUC score of 0.88; 3-model average AUC score of 0.91.
- On INbreast, we achieve single model AUC score of 0.96.

## A minimal example of multi-layer 1-D convolution



A CNN is recursively defined

## Converting from patch to whole image classifier




## Patch classifier training on DDSM

| Model | Accuracy | #Epochs |
|---|---|---|
| Resnet50 | 0.89 | 39 |
| VGG16 | 0.84 | 25 |

- 10 patches are generated from each ROI (overlap=0.9).
- 10 background patches from the same mammogram.
- Models are pretrained on the ImageNet.

**The heatmap layer may become an information barrier. Use Resnet50 as an example (dim of feature maps):**

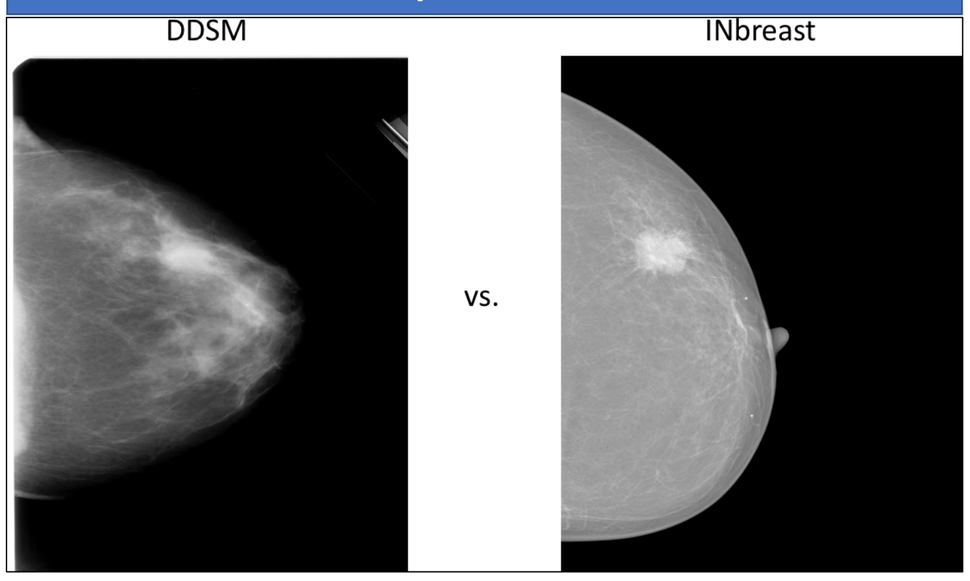
| With heatmap | No heatmap |
|---|---|
| 2048-5-512 | 2048-512 |

## Whole image classifier training on DDSM

| All convolutional design | | | | | |
|---|---|---|---|---|---|
| Patch net | Block1 | Block2 | Single-model AUC | Augmented AUC | #Epochs |
| Resnet50 | [512-512-2048]x1 | [512-512-2048]x1 | 0.85 | NA | 20 |
| Resnet50 | [512-512-1024]x2 | [512-512-1024]x2 | 0.86 | 0.88 | 25 |
| Resnet50 | [256-256-512]x3 | [128-128-256]x3 | 0.84 | NA | 48 |
| VGG16 | 512x3 | 512x3 | 0.71 | NA | 47 |
| VGG16 | 512x1 | 512x1 | 0.83 | 0.86 | 44 |
| VGG16 | 256x1 | 128x1 | 0.80 | NA | 35 |
| VGG16 | [512-512-1024]x2 | [512-512-1024]x2 | 0.81, 0.85[1] | 0.88[1] | 46 |
| Add heatmap and residual blocks on top | | | | | |
| Resnet50 | [512-512-1024]x2 | [512-512-1024]x2 | 0.80 | NA | 47 |
| Add heatmap, max pooling and FC layers on top | | | | | |
| | Pool size | FC1 | FC2 | | |
| Resnet50 | 5x5 | 64 | 32 | 0.73 | NA | 28 |
| VGG16 | 5x5 | 64 | 32 | 0.71 | NA | 26 |

[1]Result obtained from extended model training

## Different color profiles




## Transfer to INbreast database

| #Pat | #Img | Resnet50 | VGG16 | Hybrid |
|---|---|---|---|---|
| 20 | 79 | 0.78 | 0.87 | 0.89 |
| 30 | 117 | 0.78 | 0.90 | 0.90 |
| 40 | 159 | 0.82 | 0.90 | 0.93 |
| 50 | 199 | 0.80 | 0.93 | 0.93 |
| 60 | 239 | 0.84 | 0.95 | 0.91 |

The INbreast data are digital mammograms, the DDSM are scanned films.

VGG16 beats Resnet50 on INbreast. The reason: bottom layers!

**Computational setup**: a single NVIDIA Quadro M4000 GPU with 8GB memory. The deep learning framework is Keras 2 with Tensorflow as the backend.