GlimpseNet: Deep Attentional Methods for Full-Image Mammogram Diagnosis

James Liu, William Hang
zliu19, willhang

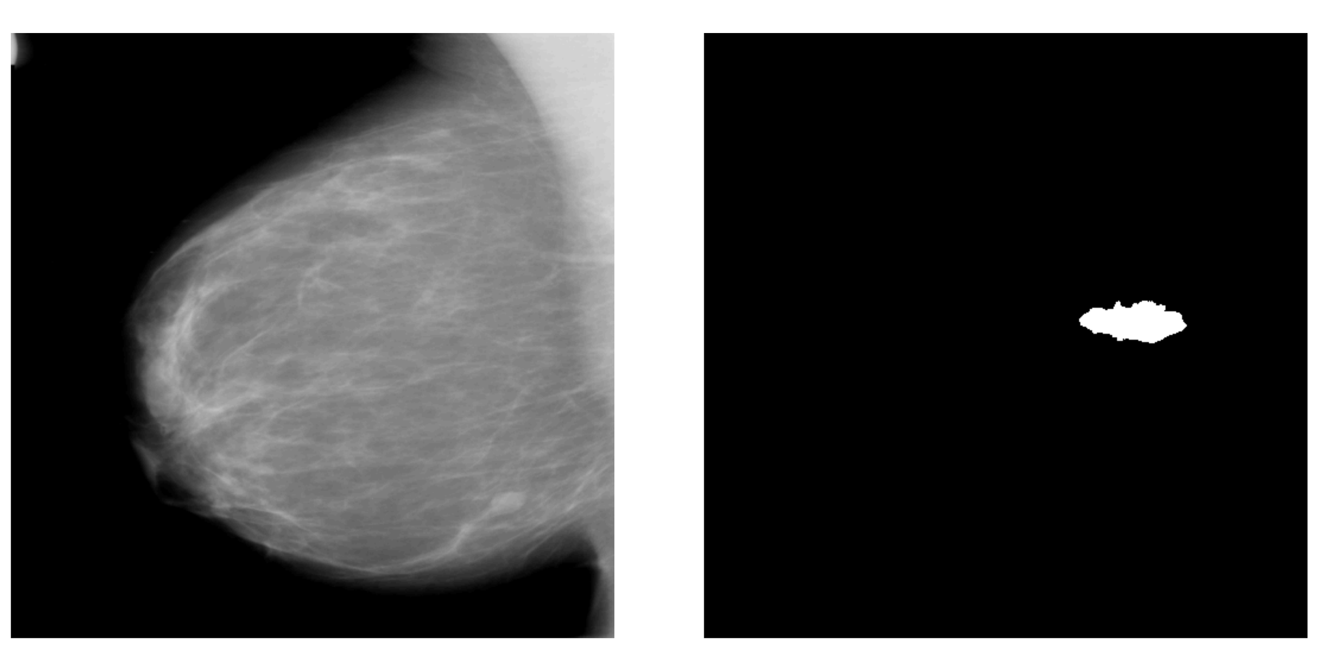
Motivation

1. Breast cancer accounts for 25% of cancers in all women, and relying on doctor mammogram diagnosis may miss 16-30% of cancers
2. Top performing papers on computational mammography focus on the radiologist pre-segmented tumor instead of on the full image
3. Full-mammogram diagnosis is important because it requires no human intervention
4. Fine-grained image recognition is an open problem in computer vision

We propose **GlimpseNet**, a fine-grained image recognition architecture that extracts salient regions in full mammogram images for diagnosis with multi-instance learning

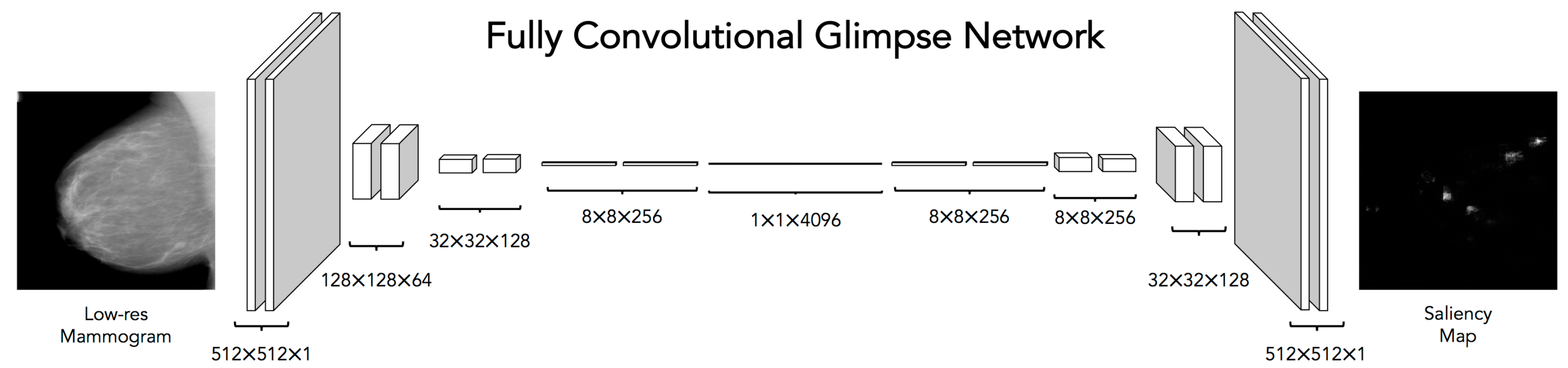
Dataset

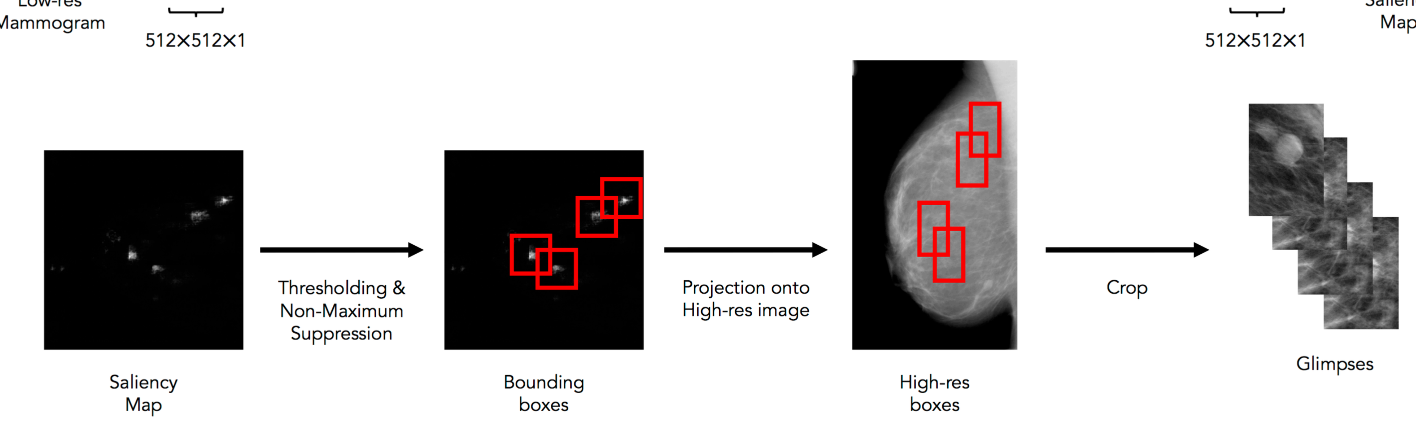
- Digital Database for Screening Mammography (DDSM) provided by the University of South Florida and The Cancer Imaging Archive.
- 1318 mammograms from 691 patients
- Groups of mammograms that belong to the same patient were taken at different views
- Each patient is given one of three diagnoses: BENIGN WITHOUT CALLBACK, BENIGN, and MALIGNANT
- Each mammogram image comes with a segmentation mask of the tumor, where the pixels that belong to the tumor are set to white. We use these masks to train our model to focus on the most salient regions



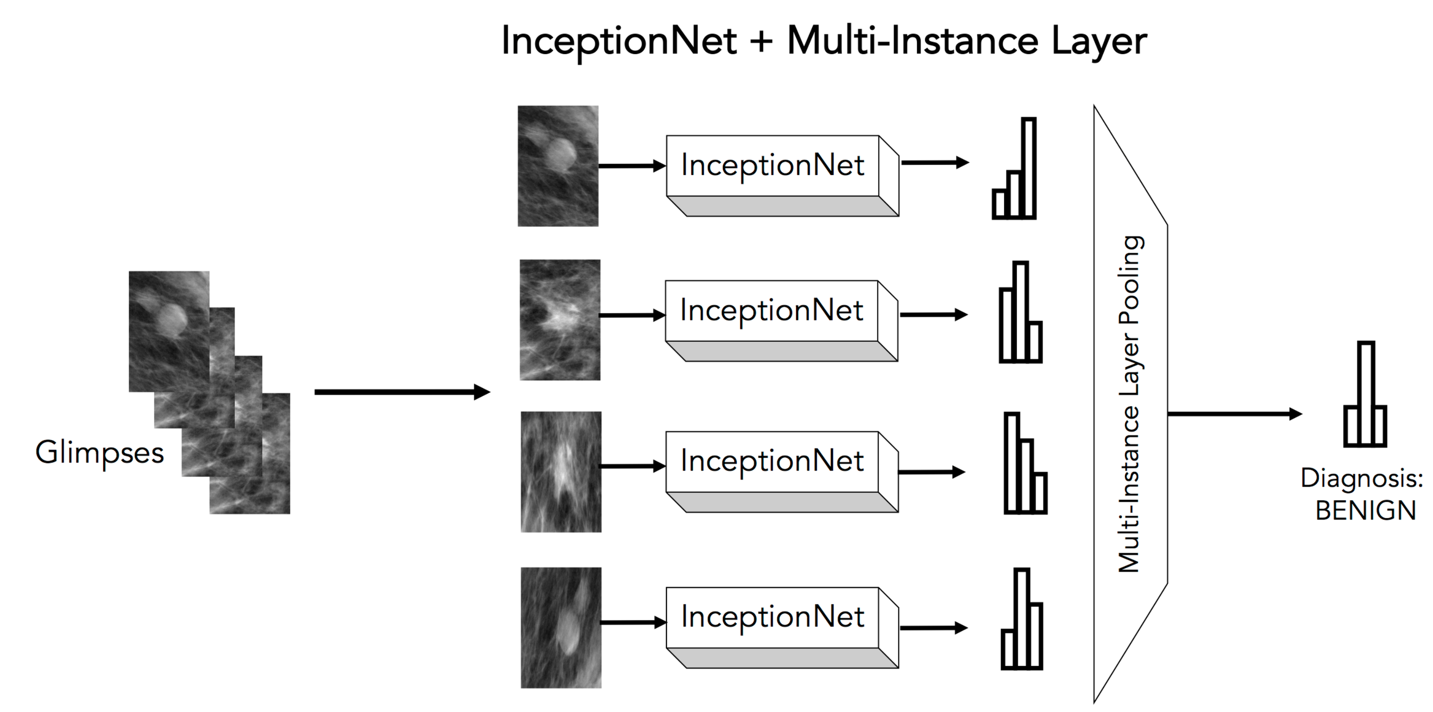
Model Architecture

Our model, **GlimpseNet**, consists of a **Fully Convolutional Glimpse Network (FCGN)**, **InceptionNet**, and a **Multi-Instance Layer (MIL)**





FCGN takes in a low-res image and generates a saliency map describing the importance of each pixel in the original image. With thresholding and NMS, we recover the important regions, which we call glimpses, from the high-resolution image



Glimpses are fed into InceptionNet, generating multiple class probability distributions, which are then pooled with MIL to generate a single distribution for final classification

Experiments

Procedures

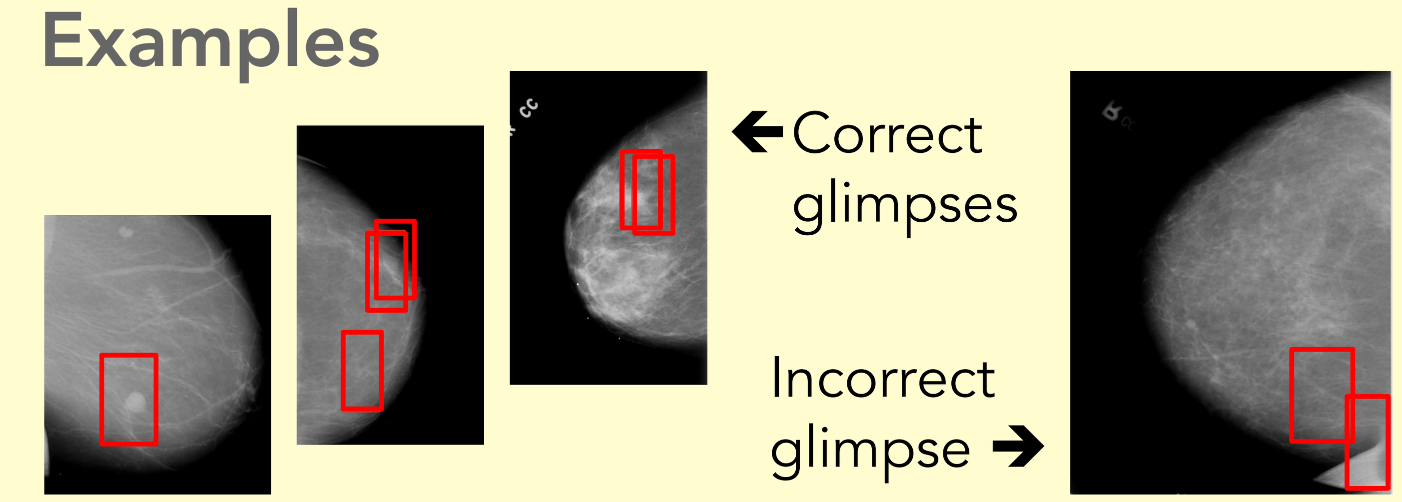
- 10-fold cross validation on our model. Training set ~ 1186 images, test set ~ 132 images
- Three class classification problem
- Model is trained for 50 epochs before validating the model on the test set
- FCGN loss function is MSRE between the segmentation mask and the generated attention image. InceptionNet + MIL loss function is softmax loss

Results

Method	3-class accuracy
Baseline – Inception v3	0.490
Multiview Model	0.620
Sanchez et al.	0.621
Attribute Model	0.645
Coattention Model	0.647
GlimpseNet NAND (a=10, b=0.2)	0.589
GlimpseNet topk (k=0.3)	0.571
GlimpseNet LSE (r=5)	**0.695**

- Our model outperforms the state of the art
- Train/test curves provided in the final report

Examples



Conclusions/Future Directions

- Our contribution is GlimpseNet, a fine-grained image recognition architecture capable of diagnosing full mammograms with hard attention on salient regions, outperforming the current state of the art by 7.4%
- Future steps include further validation on other mammogram datasets, hyperparameter tuning to achieve higher accuracies, augmenting the data to improve generalizability, and attempting to make GlimpseNet fully differentiable