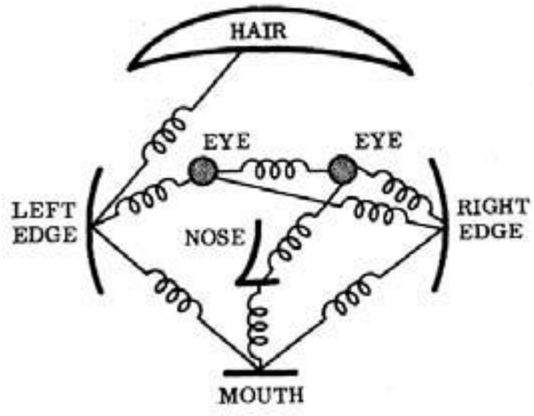
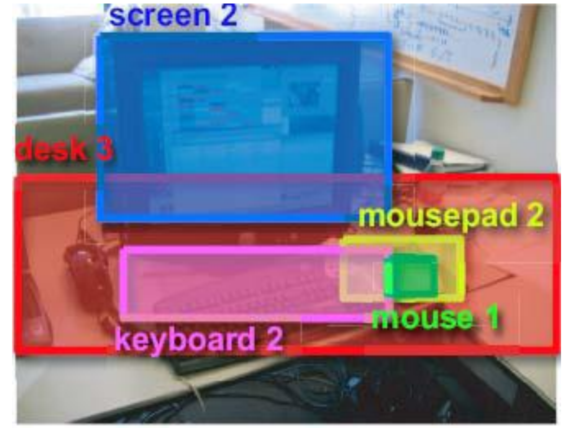


COMP 9517 Computer Vision

Applications

What we have learned

- Image processing
- Feature representation
- Machine learning
- Segmentation
- Deep learning
- Motion and tracking

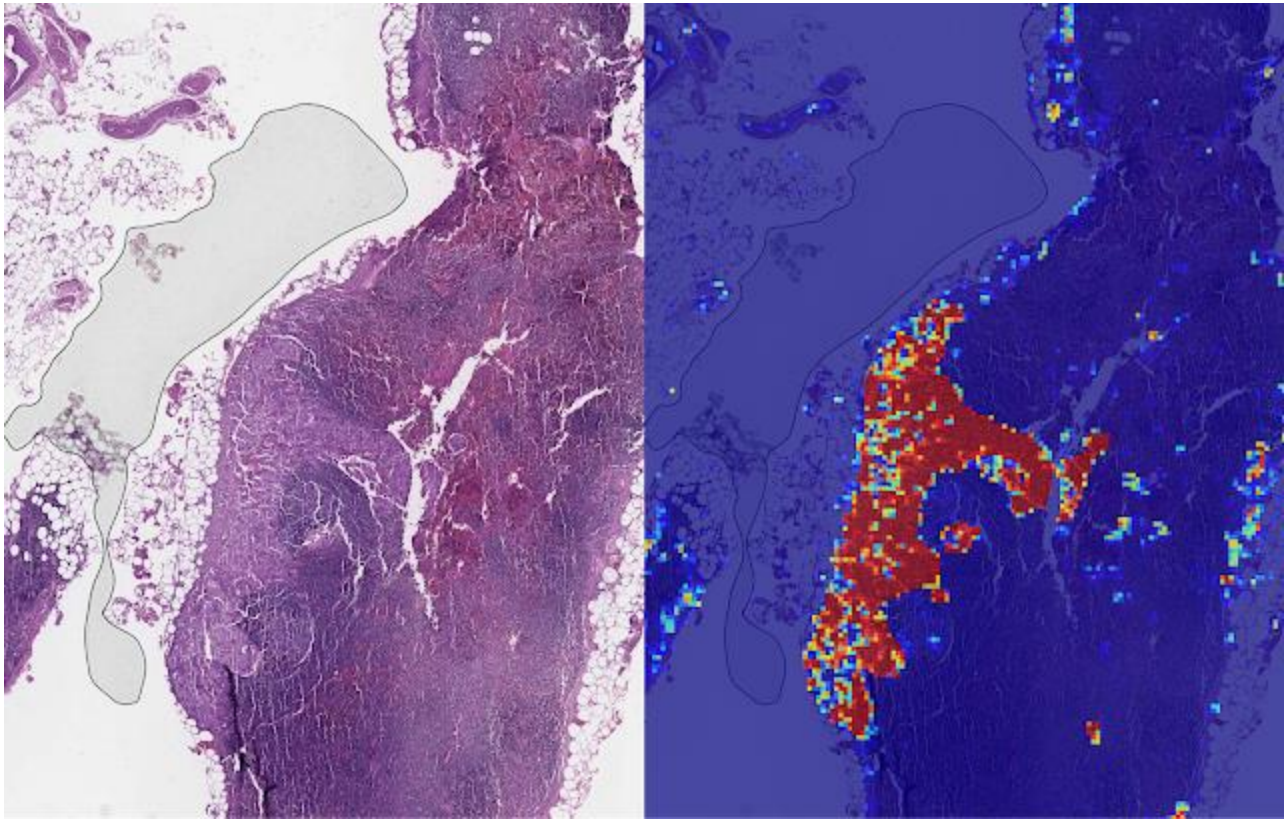
=> These are the main methodological components useful for building computer vision applications

Outline

- Case studies of computer vision applications:
  - WSI analysis
  - Mitosis detection
  - Interpreting deep learning

WSI Analysis

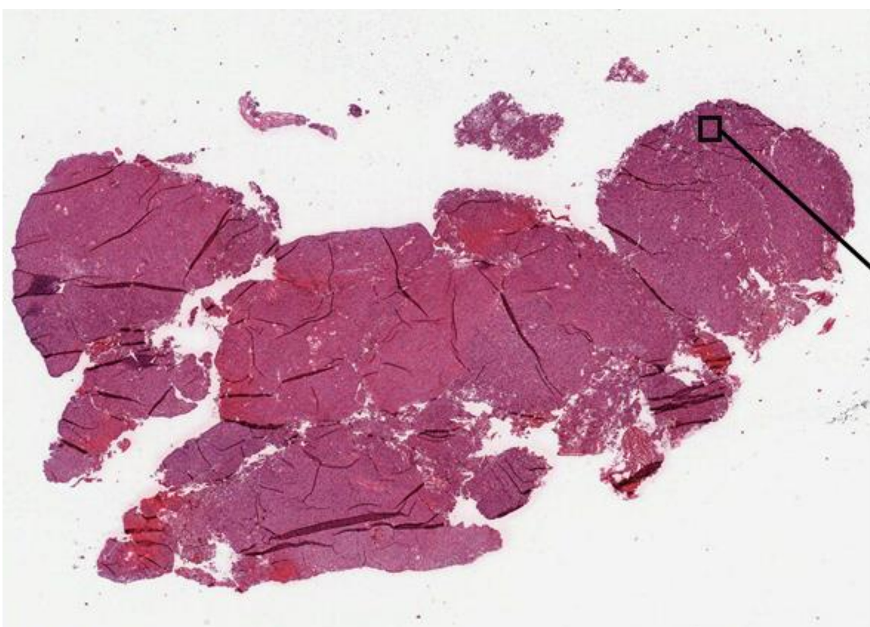
Cancer Detection

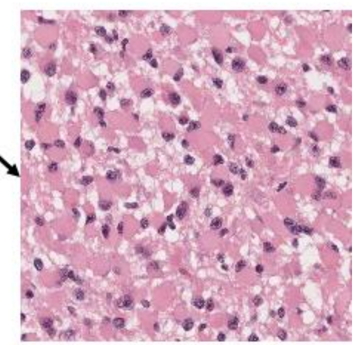


Source: https://ai.googleblog.com/2018/10/applying-deep-learning-to-metastatic.html

Cancer Detection

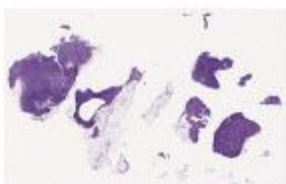
- Whole-slide image (WSI): very high resolution images
- A shift to fully digital environment for pathology





40x magnification

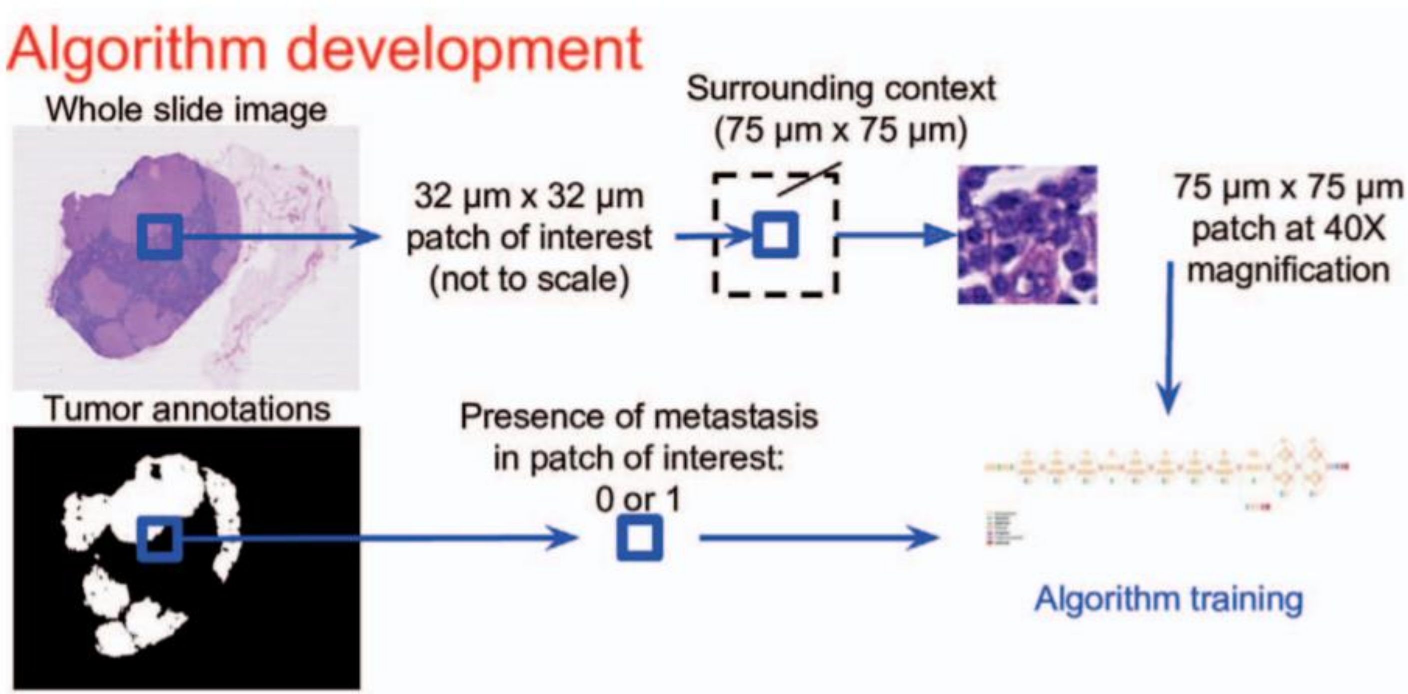
Cancer Detection



Source: https://ai.googleblog.com/2018/10/applying-deep-learning-to-metastatic.html

Cancer Detection

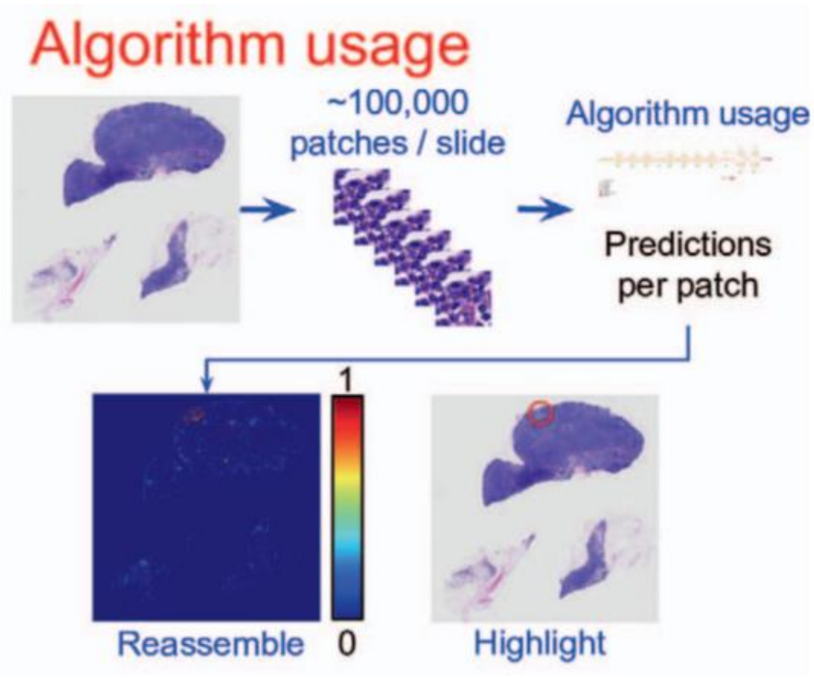
- Training stage:
  – Patch-wise processing with patch-level labels



Source: Y. Liu et al. *Artificial intelligence-based breast ancer nodal metastasis detection*. Arch Pathol Lab Med, 2018.

Cancer Detection

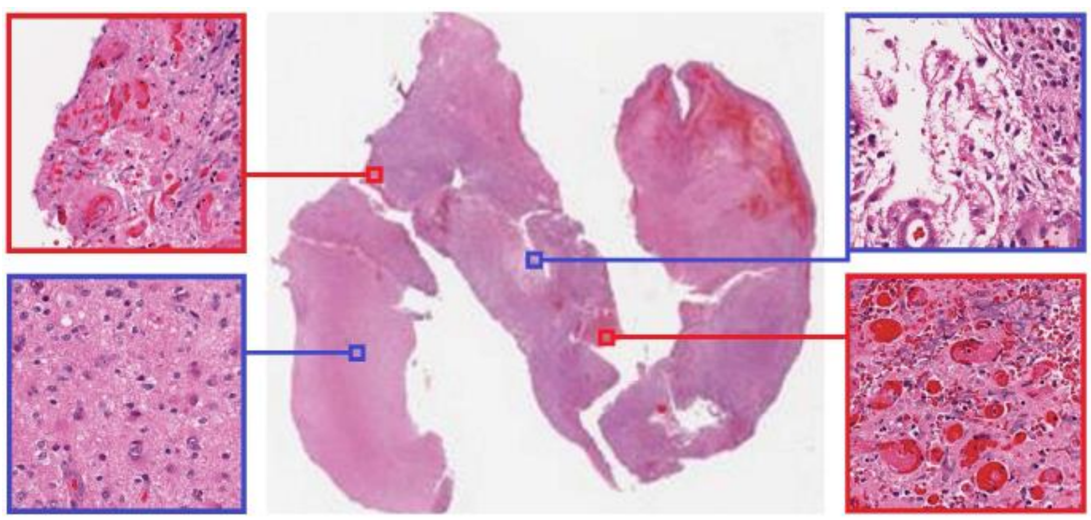
- Testing stage:
  - Patch-wise classification



Source: Y. Liu et al. *Artificial intelligence-based breast ancer nodal metastasis detection*. Arch Pathol Lab Med, 2018.

Real Challenges

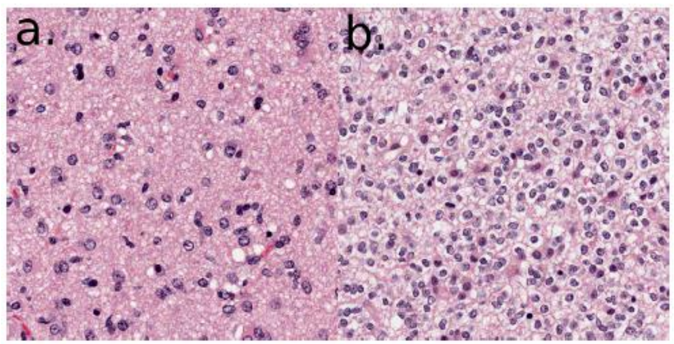
- Challenge I:
  - Large image with image-level label only



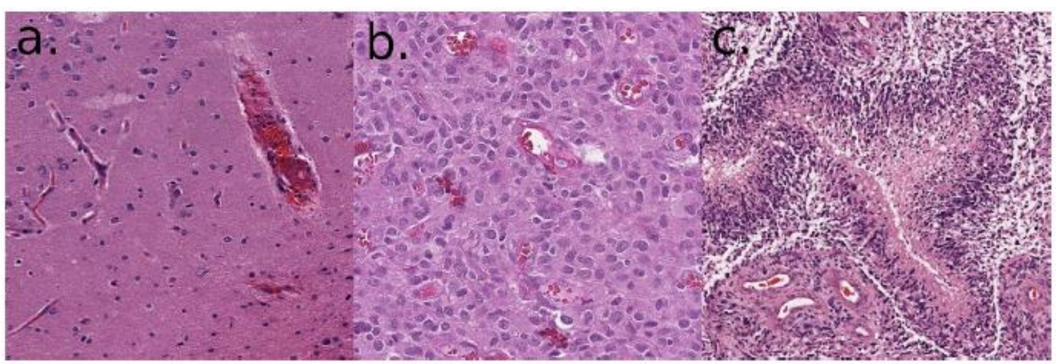
Real Challenges

- Challenge II:
  - Histology heterogeneity (subtypes and regional variations)

LGG

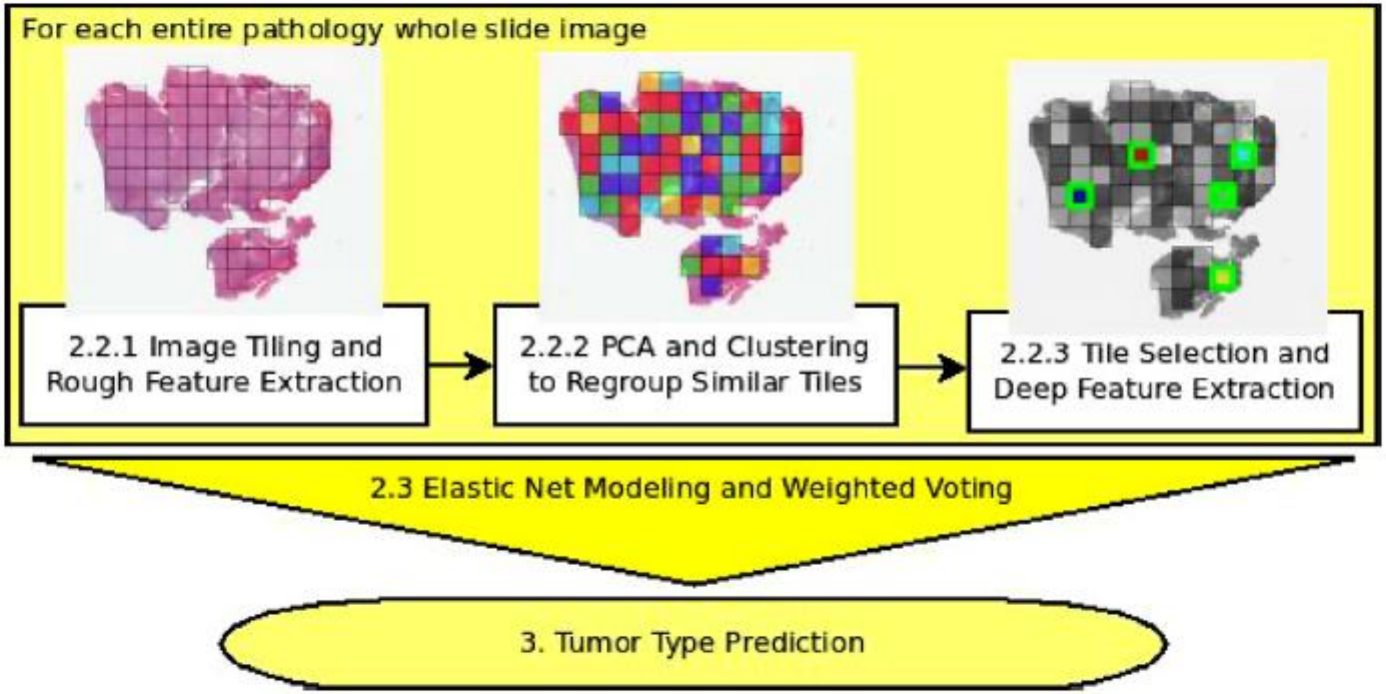


GBM



Clustering-based

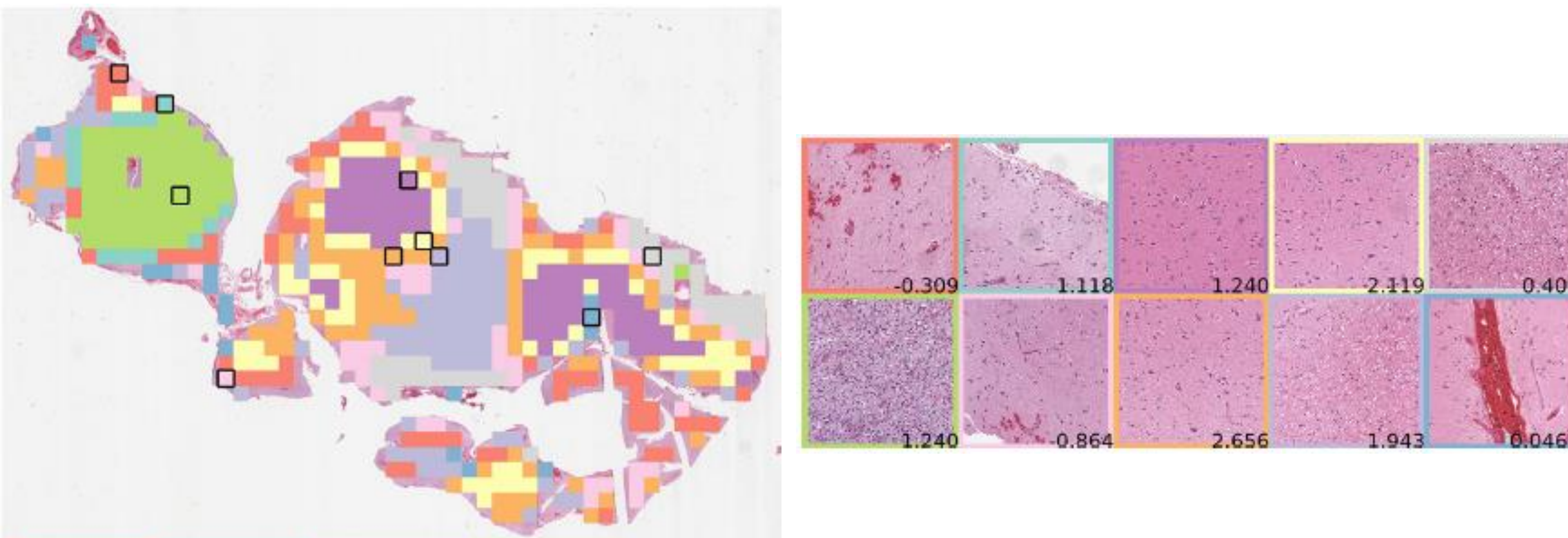
- Coarse and fine feature extraction:



J. Barker *et al.*, "Automated classification of brain tumor type in whole-slide digital pathology images using local representative tiles," *Medical Image Analysis*, 30(1):60-71, 2016.

Clustering-based

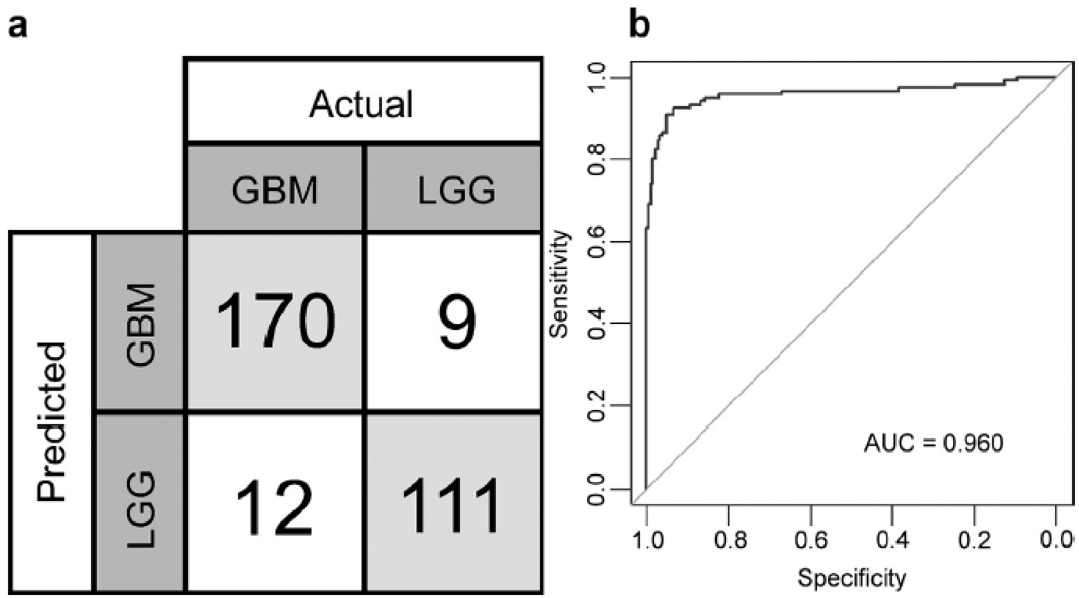
- Coarse and fine feature extraction:
  - Clustering-based representative tile extraction



J. Barker *et al.*, “Automated classification of brain tumor type in whole-slide digital pathology images using local representative tiles,” *Medical Image Analysis*, 30(1):60-71, 2016.

Clustering-based

- Coarse and fine feature extraction:



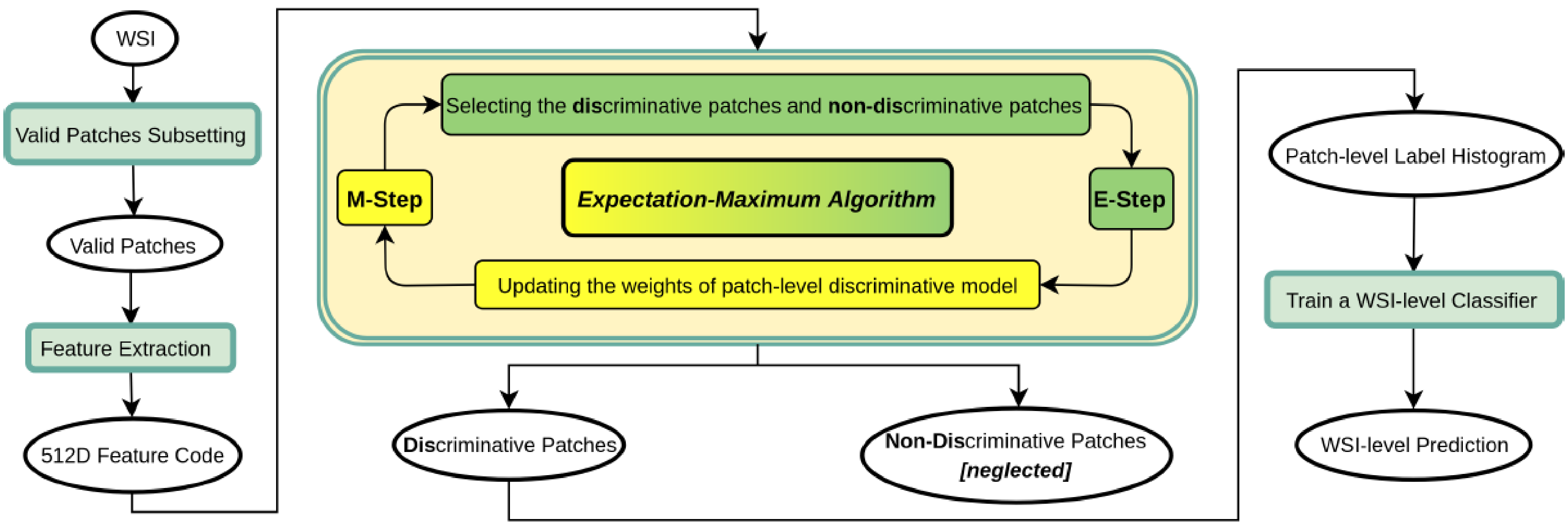
Sensitivity (recall) = TP/P

Specificity = TN/N

J. Barker *et al.*, “Automated classification of brain tumor type in whole-slide digital pathology images using local representative tiles,” *Medical Image Analysis*, 30(1):60-71, 2016.

Pruning-based

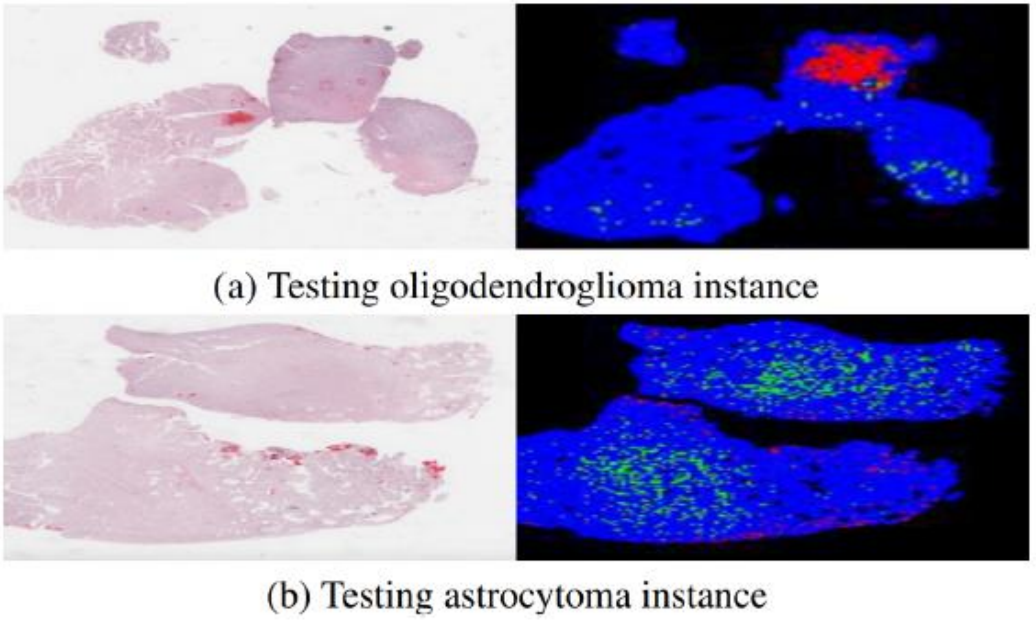
- Discriminative patch-based CNN:
  - EM-based discriminative patch extraction



Source: C. Zhang et al. *Whole slide image classification via iterative patch labelling*. ICIP, 2018.

Pruning-based

- Discriminative patch-based CNN:
  - EM-based discriminative patch extraction



(a) Testing oligodendroglioma instance

(b) Testing astrocytoma instance

Source: C. Zhang et al. *Whole slide image classification via iterative patch labelling*. ICIP, 2018.

Pruning-based

- Results on the CBTC challenge dataset:
  - 32 WSI images, with 16 astrocytoma and 16 oligodendroglioma cases

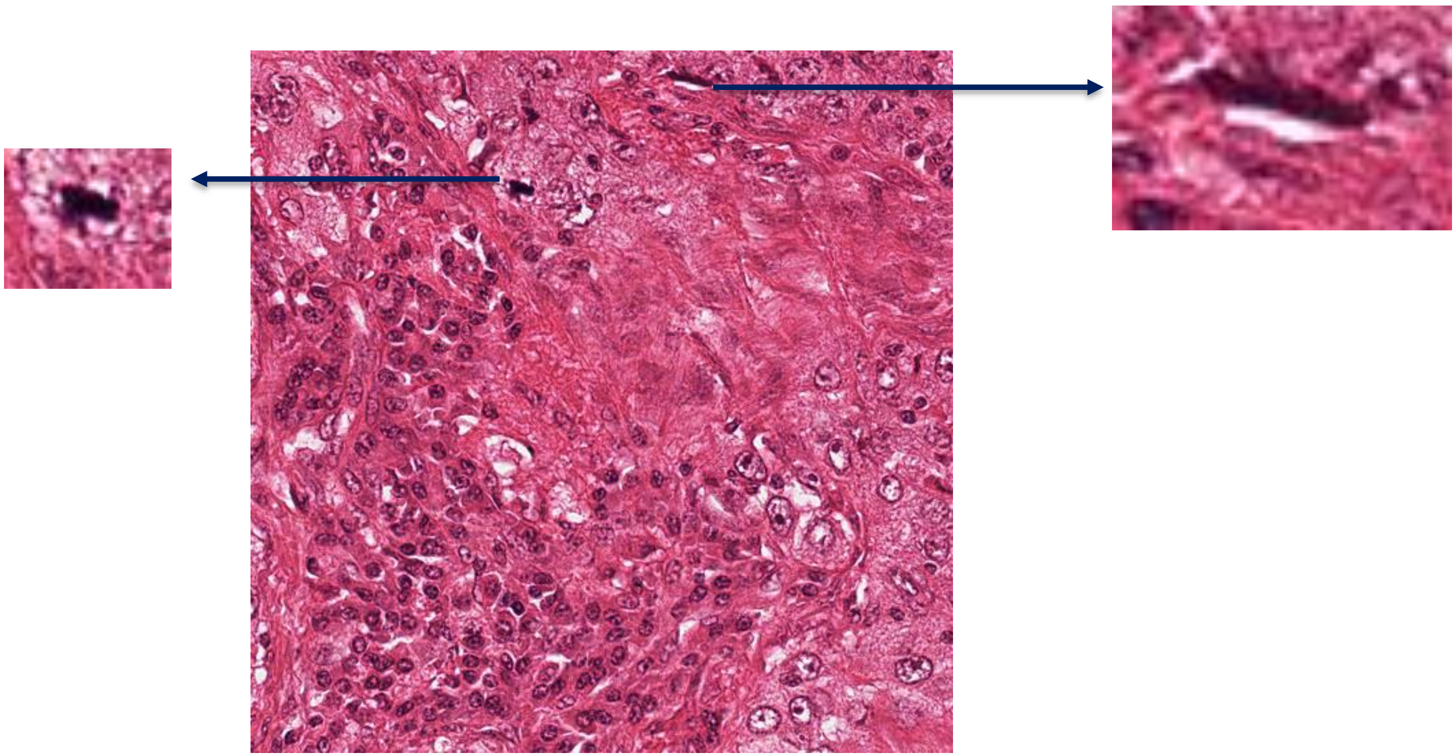
Methods	Acc.
CNN-Feat-SVM	62.50%
Finetune-CNN-Feat-SVM	69.13%
Iter-Finetune-CNN-SVM[Discriminative]	76.62%
Iter-Finetune-CNN-SVM[Both]	**84.38%**

Source: C. Zhang et al. *Whole slide image classification via iterative patch labelling*. ICIP, 2018.

Mitosis Detection

Aim of Study

- Mitosis detection and classification for cancer diagnosis

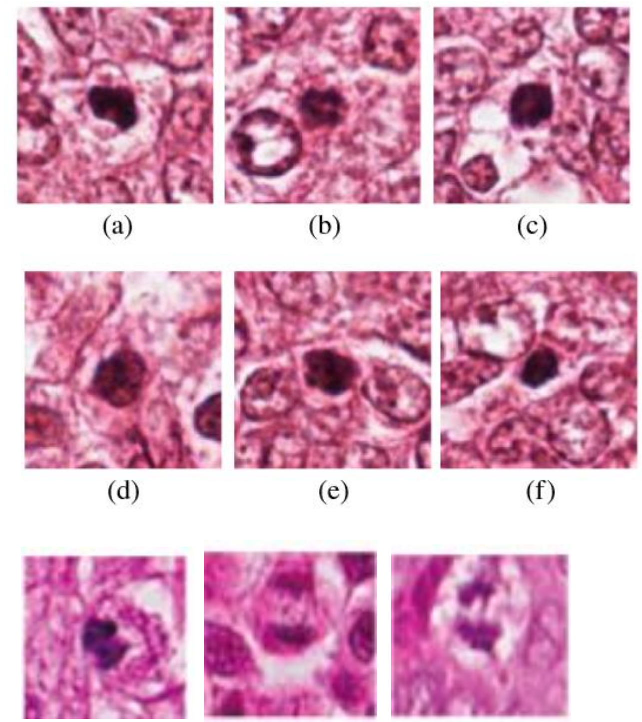


Aim of Study

- Differentiate Mitosis Cells from other Cells
  - Similar appearance
  - Staining
  - Different appearance through different mitosis phases

Mitosis cells

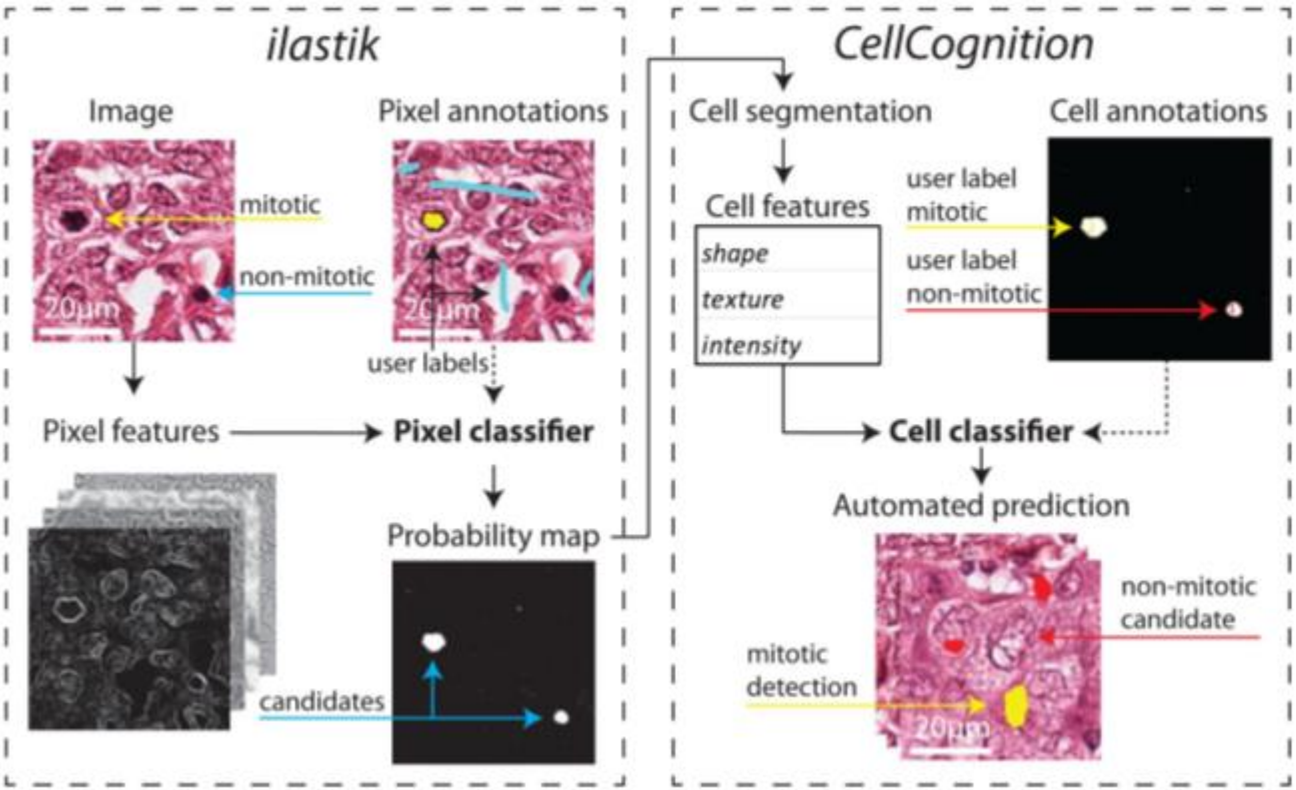
Non Mitosis cells



Appearance Change Through Phases

Handcrafted Features

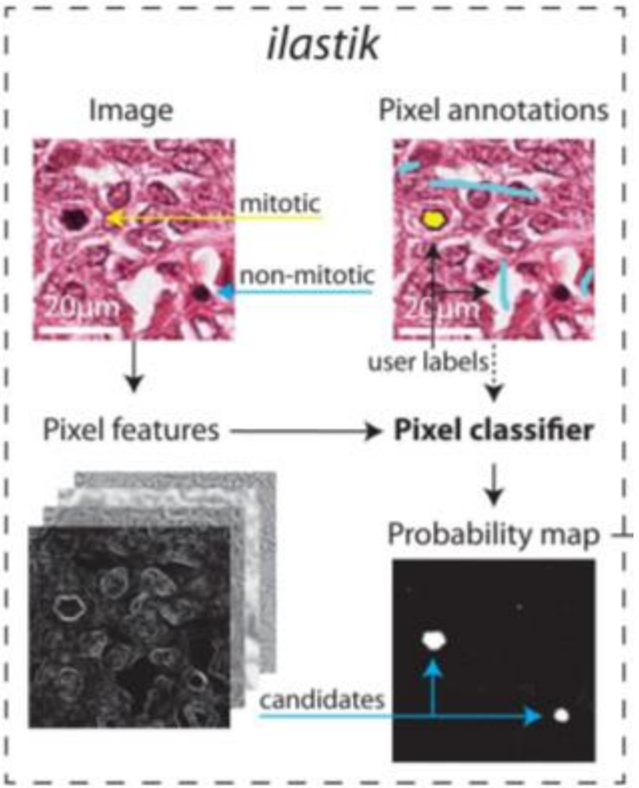
- A two-stage approach



C. Sommer et al., “Learning-based mitotic cell detection in histopathological images”, ICPR, 2012

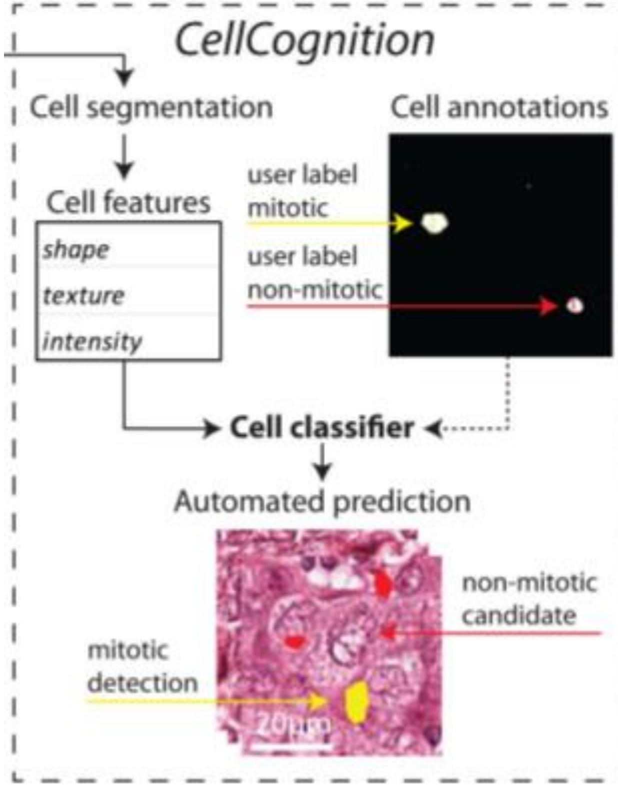
Handcrafted Features

- Stage 1 – segmentation of candidate cells
  - Pixel classification with Gaussian filter-based features, and random forest classifier
  - Then local adaptive thresholding on the classification probability map to produce cell candidates



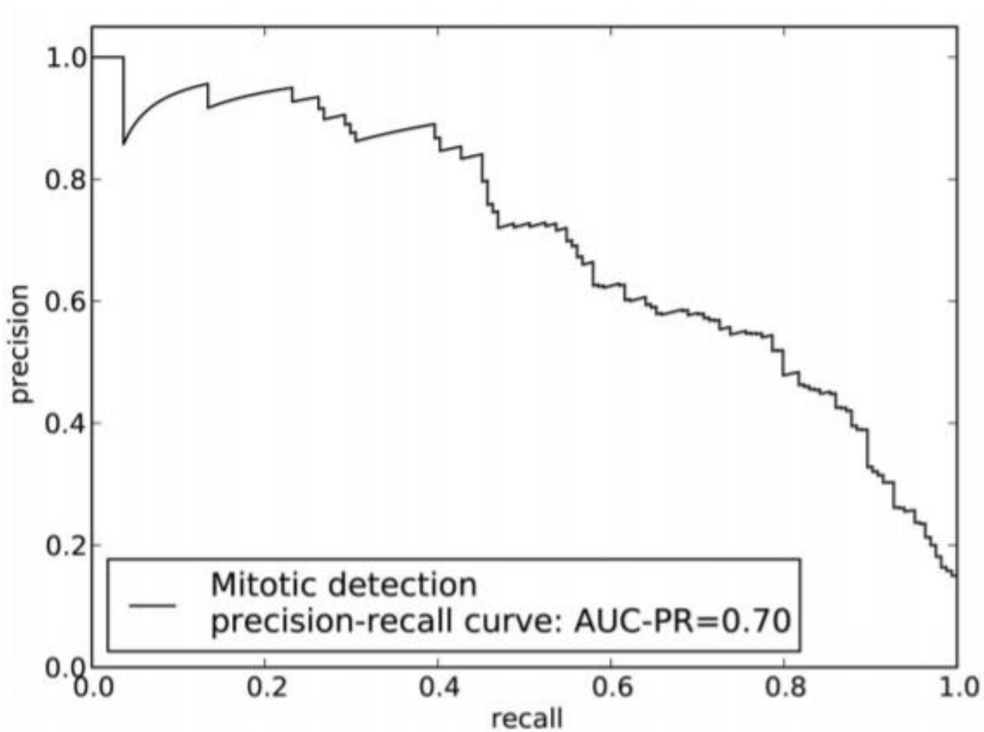
Handcrafted Features

- Stage 2 – mitosis classification
  - Object/cell classification with texture and shape features, and Gaussian-kernel SVM classifier
  - Require object-level ground truth labels



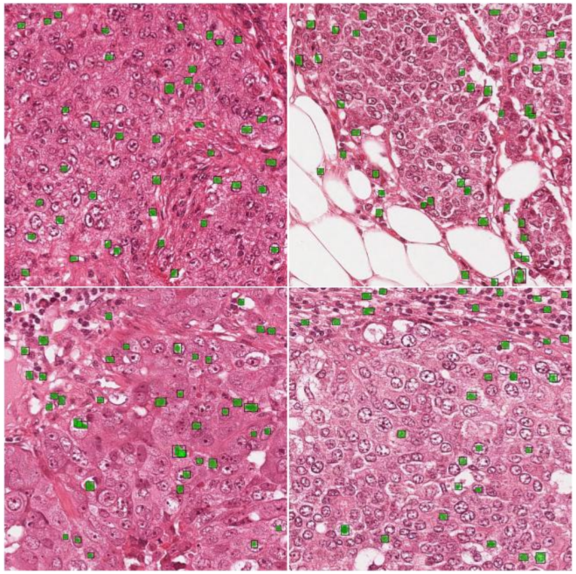
Handcrafted Features

- Evaluation – PR curve
  - Five-fold cross-validation
  - 35 images



Deep Learning

- Problem formulation as **Object Detection** and **Instance Segmentation**



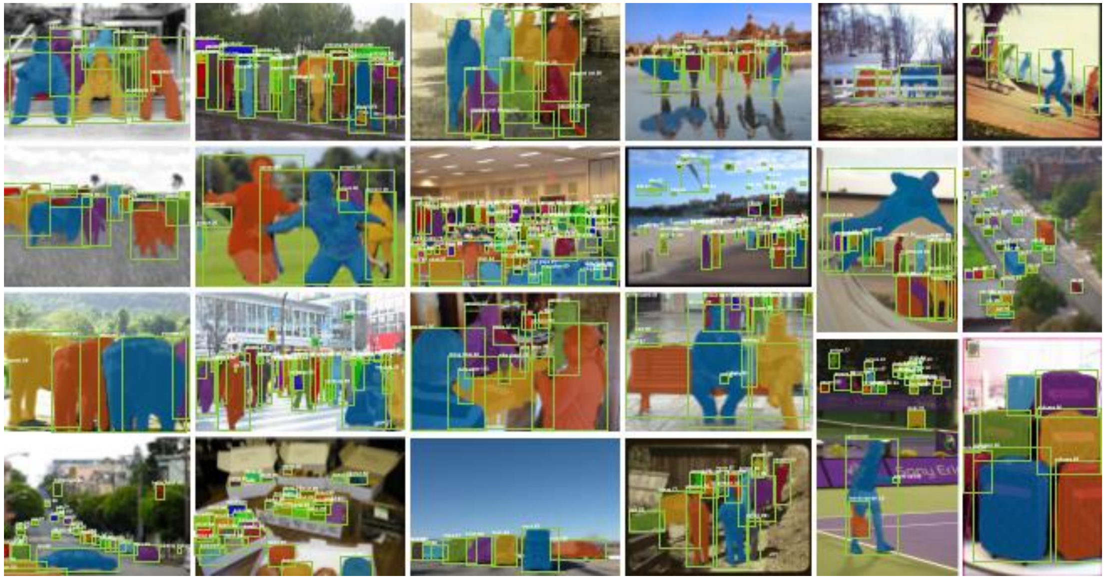
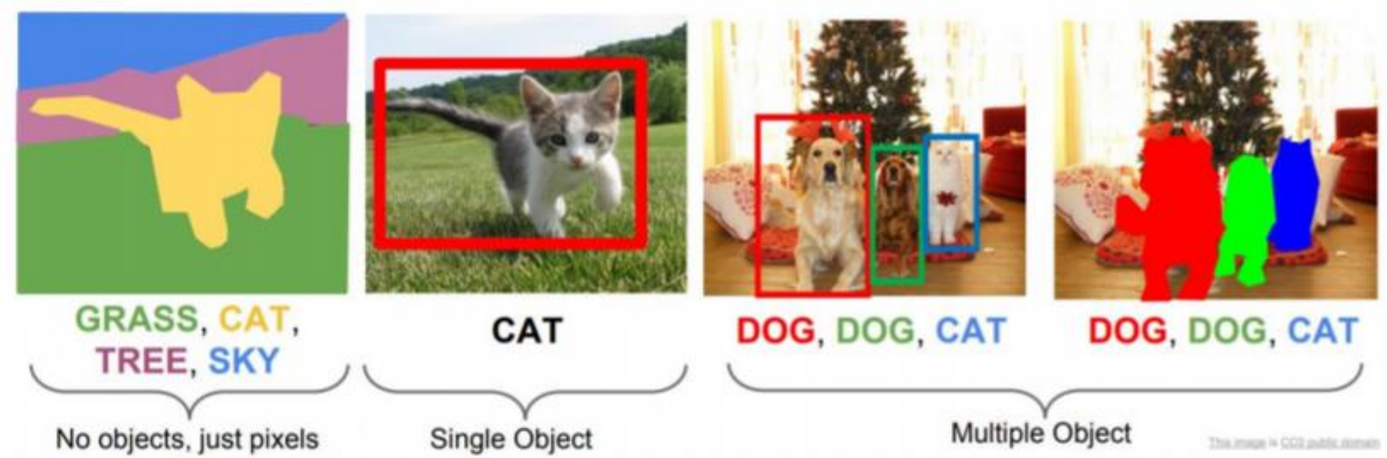


Image from He et al., "Mask RCNN", ICCV, 2017

Mask R-CNN

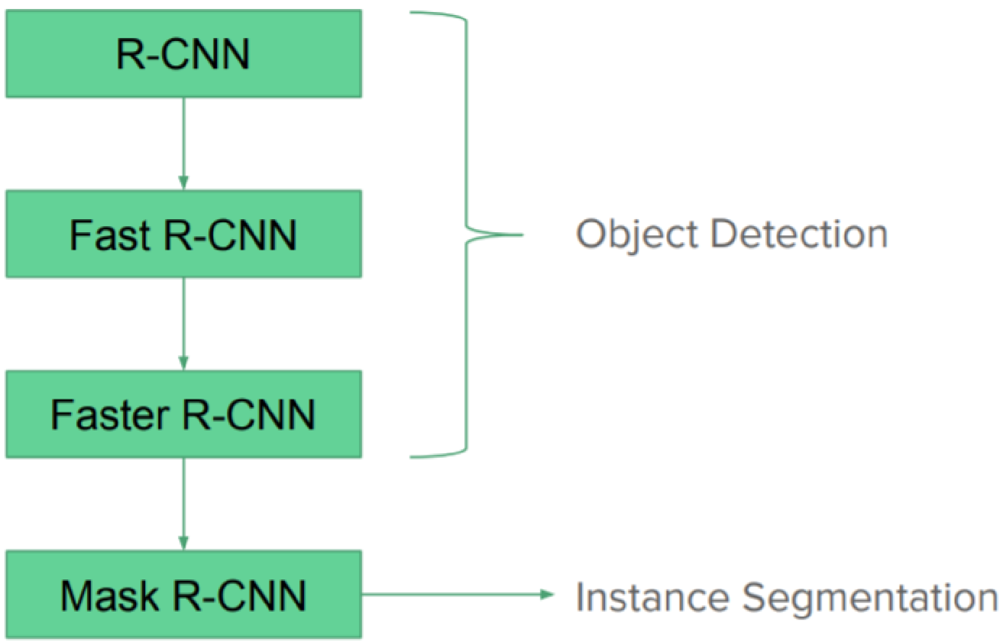
- Background:
  - Semantic segmentation
  - Single object detection
  - Multiple objects detection
  - Instance segmentation



https://cseweb.ucsd.edu/classes/sp18/cse252C-a/CSE252C_20180509.pdf

Mask R-CNN

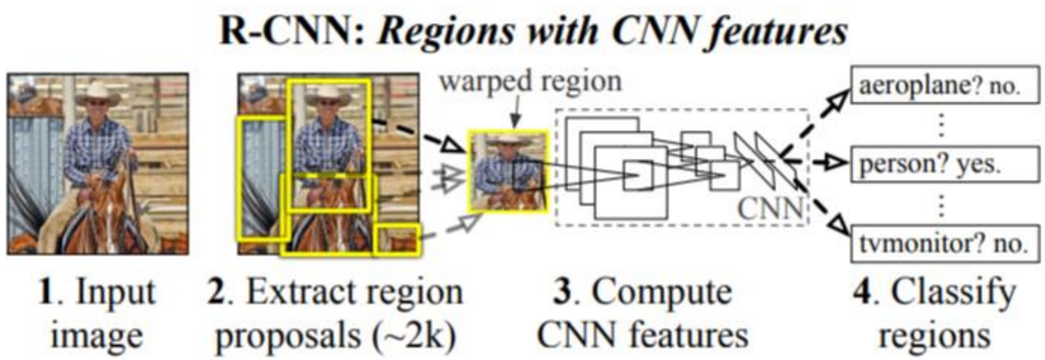
- The R-CNN family



https://cseweb.ucsd.edu/classes/sp18/cse252C-a/CSE252C_20180509.pdf

Mask R-CNN

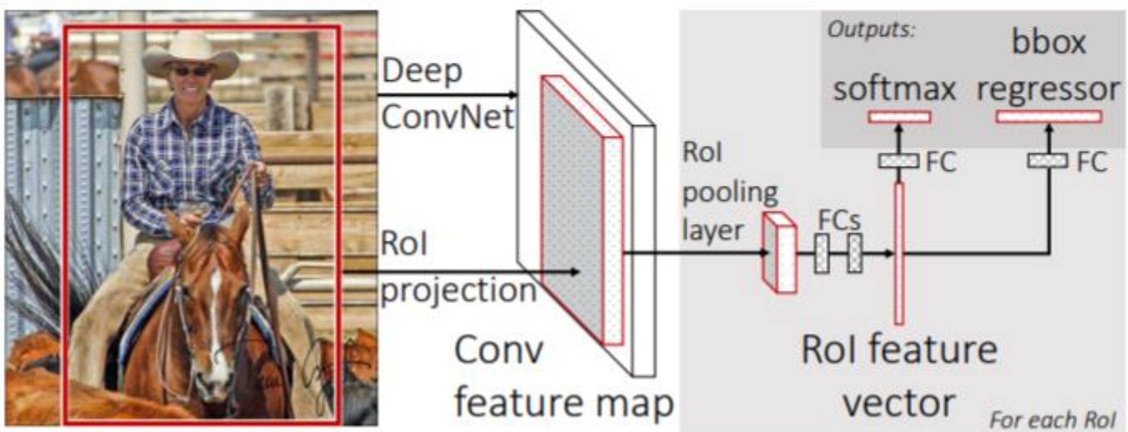
- R-CNN
  - Training is expensive and slow because of selective search and lack of shared computation



Girshick et al., Rich feature hierarchies for accurate object detection and semantic segmentation, CVPR, 2014.

Mask R-CNN

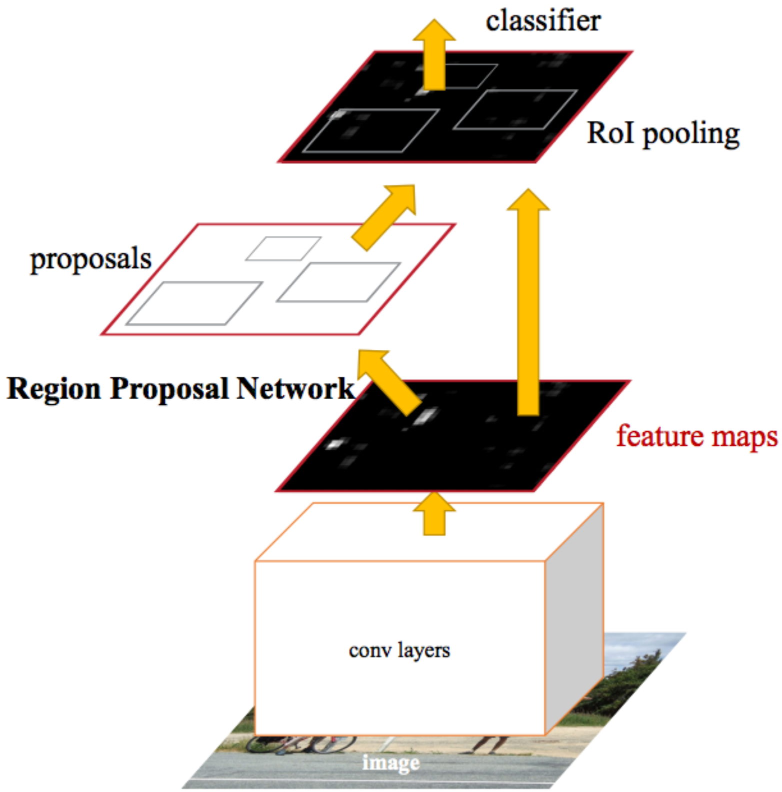
- Fast R-CNN
  - Shared computation of convolutional layers between proposals as a result of ROI pooling
  - Improvement in speed is not large because the region proposals are generated separately by another model



Girshick, Fast R-CNN, ICCV, 2015.

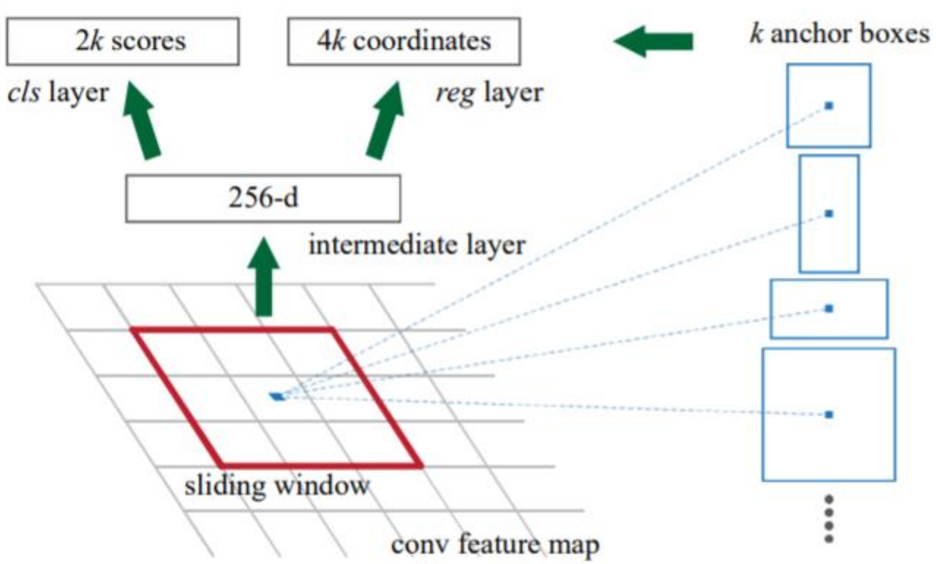
Mask R-CNN

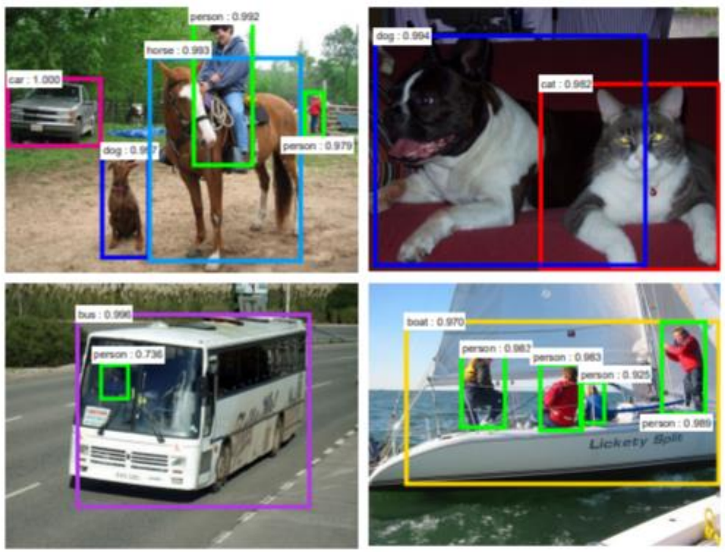
- Faster R-CNN
  - Fast R-CNN + Region Proposal Network (RPN)



Mask R-CNN

- Faster R-CNN
  - Fast R-CNN + Region Proposal Network (RPN)

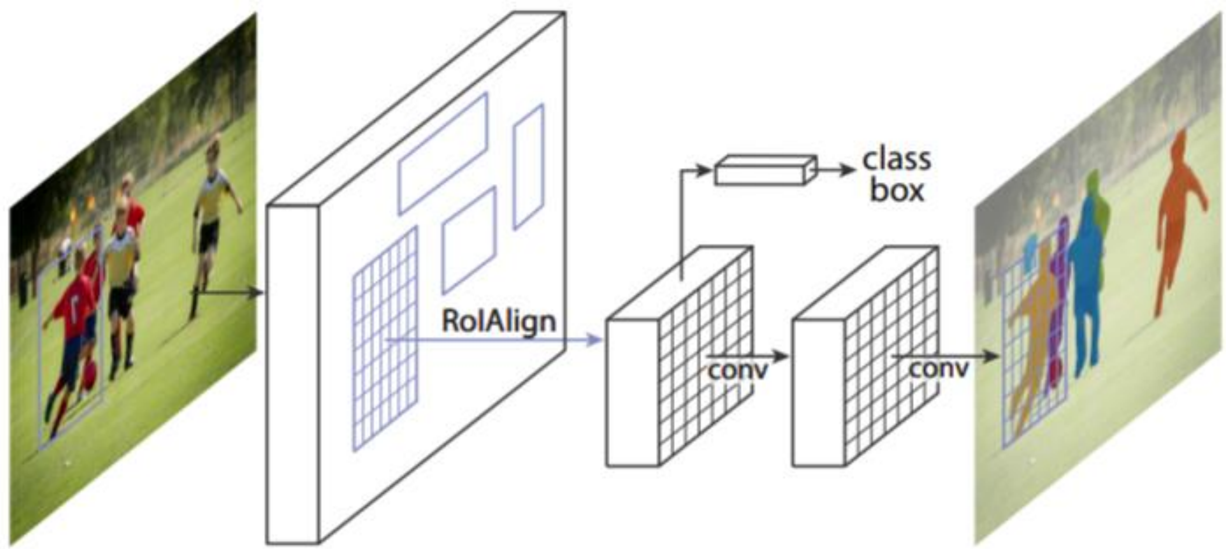




Ren et al., Faster R-CNN: towards real-time object detection with region proposal networks, NeurIPS, 2015.

Mask R-CNN

- Mask R-CNN
  - Convolutional backbone + RPN
  - Parallel heads for box regression
  - RoIAlign



He et al., Mask R-CNN, ICCV, 2017.

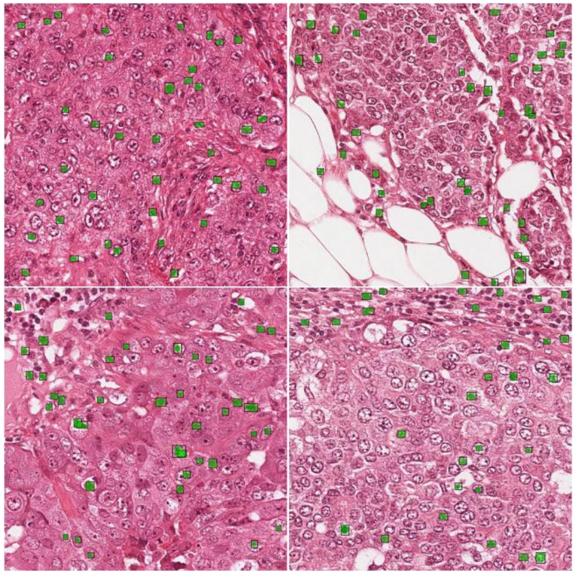
Mask R-CNN

- Mask R-CNN
  - Instance segmentation

He et al., Mask R-CNN, ICCV, 2017.

Mitosis Detection

- Problem formulation as **Object Detection** and **Instance Segmentation**



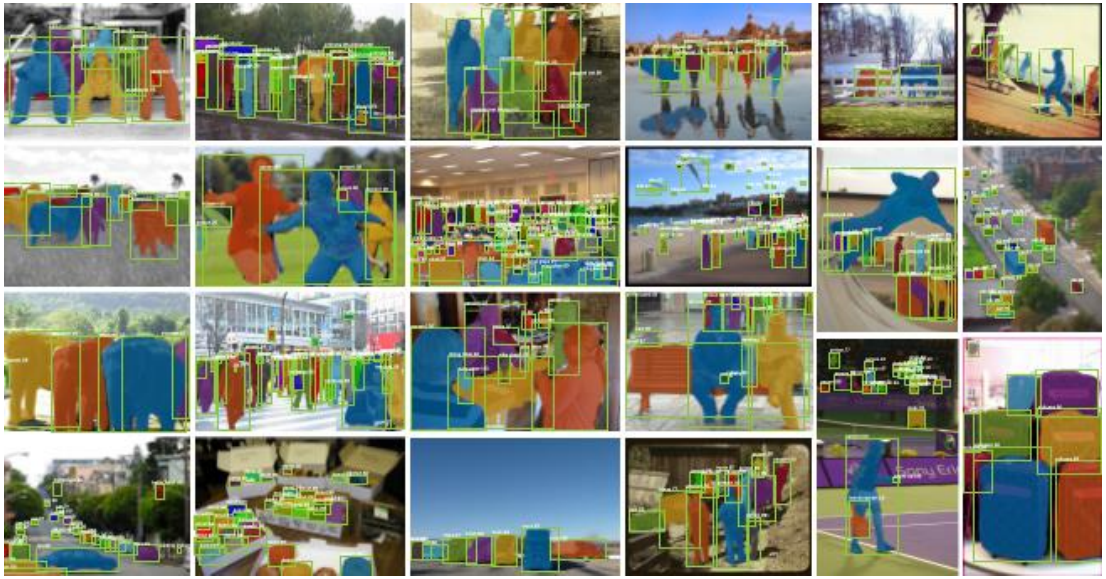
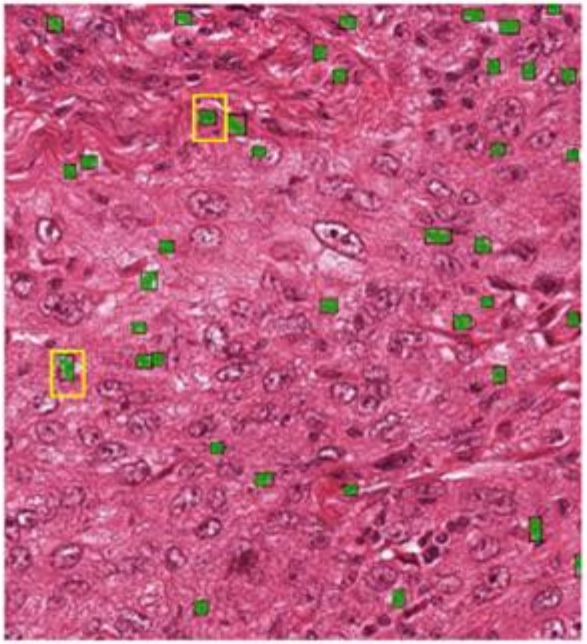


Image from He et al., “Mask RCNN”, ICCV, 2017

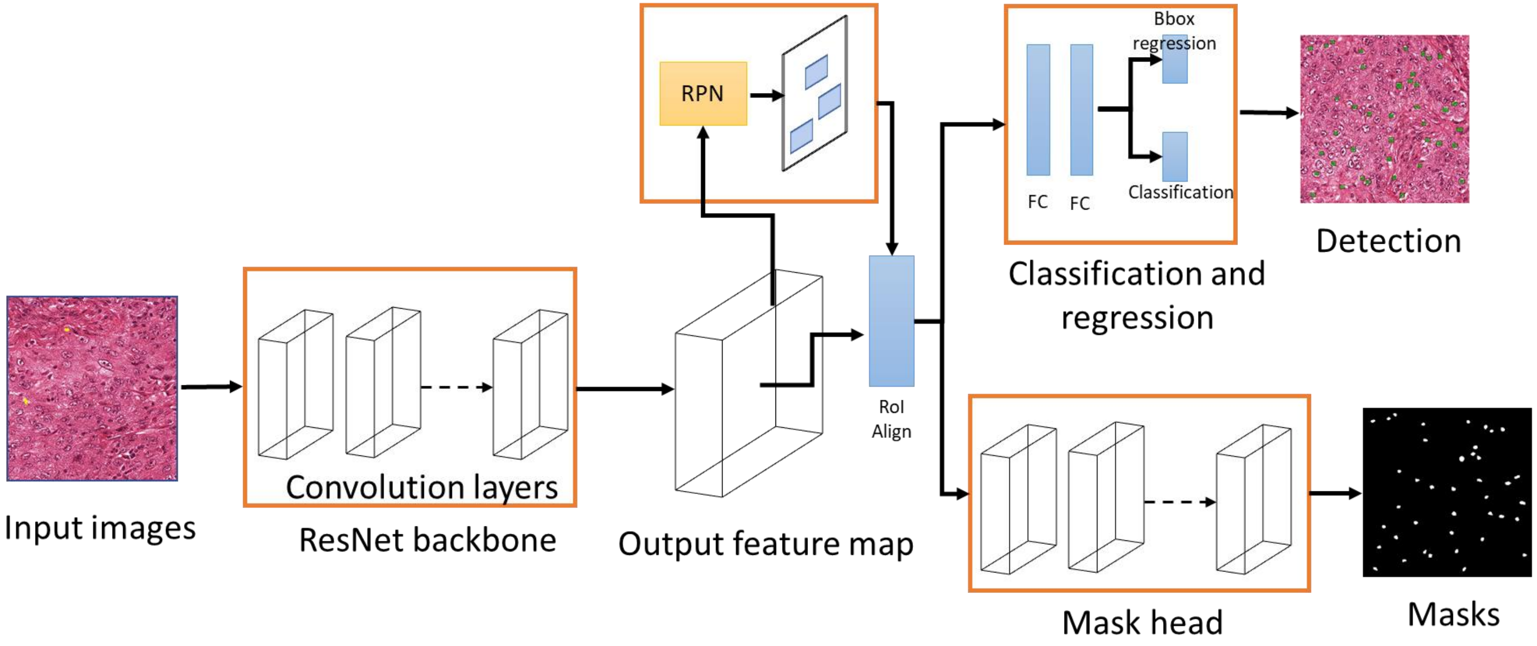
Mitosis Detection

- Two-stage approach:
  - Candidate detection stage
  - Mitosis classification stage



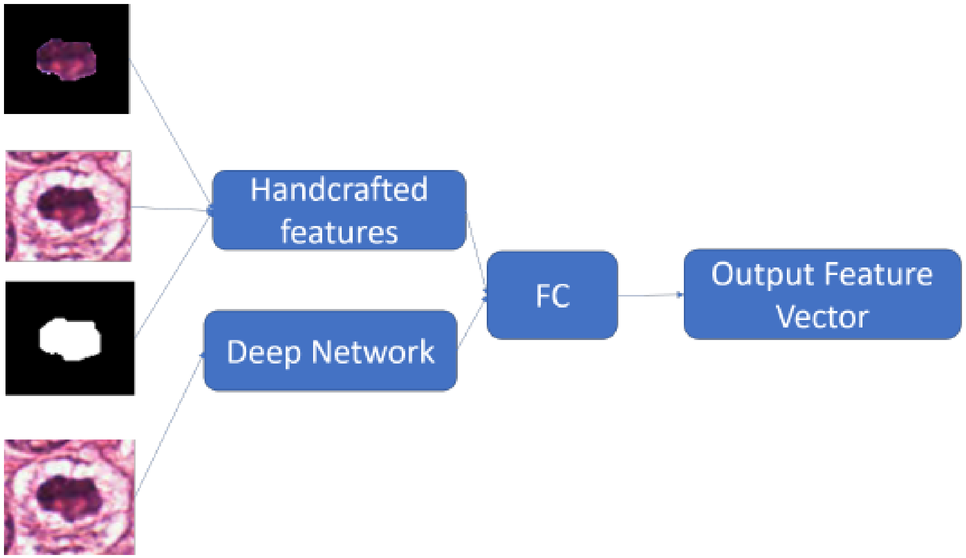
Mitosis Detection

- Stage 1: candidate detection
  - Adapted Mask R-CNN for instance segmentation
  - High recall and low precision output



Mitosis Detection

- Stage 2: mitosis classification
  - Classification using combined deep network features and handcrafted features



Dodballapur et al., Mask-driven mitosis detection in histopathology images, ISBI, 2019.

Mitosis Detection

- Results on the ICPR 2012 dataset
  - 226 mitotic cells in training set and 101 in test set

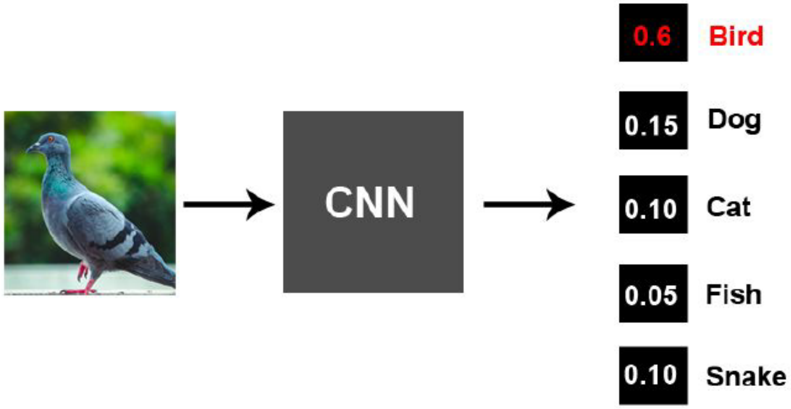
Method	Precision	Recall	F1-Score
DeepMitosis	0.85	**0.81**	0.83
HC+CNN	0.84	0.65	0.73
RCNN	0.78	0.79	0.78
Ciresan et al	0.88	0.70	0.78
Proposed (VGG-16)	0.87	**0.81**	0.84
Proposed (Xception)	**0.94**	0.80	**0.87**

Dodballapur et al., Mask-driven mitosis detection in histopathology images, ISBI, 2019.

Interpreting Deep Learning

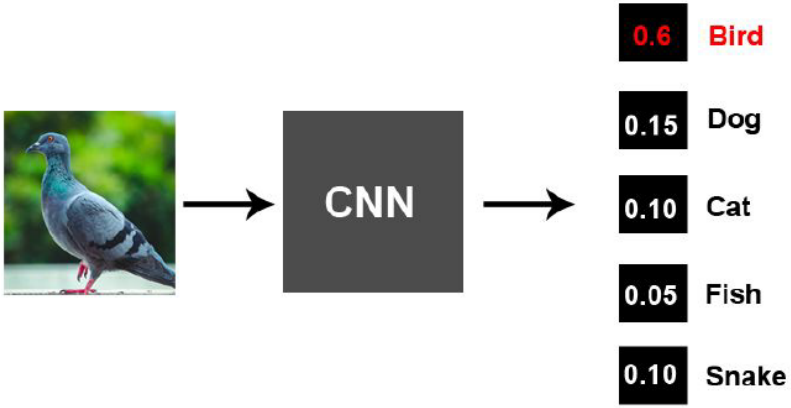
Background

- How do we determine what the subject of an image is?
  - Train a CNN
  - Feed the image into the CNN
  - Choose the output with the highest score



Background

- How do we understand the network's reasoning?
  - Look at the network's weights => there are millions!
  - Give up and treat it as a black box?



Interpretability

- Interpretability explains the reasoning behind an output
- Interpretability is important:
  - Increase trust
  - Lead to improved network design
- Desirable for:
  - High-risk industry, e.g. medicine, law, self-driving cars

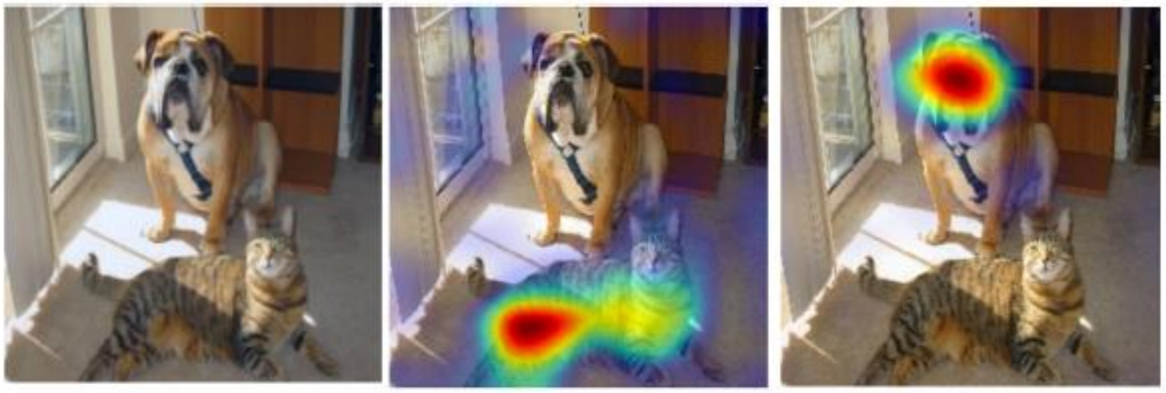
Interpretability

- Visualisation as an interpretability method
  - Where is the network looking?
- Quantitatively comparing visualisations:
  - Train the network as a classifier
  - See where the network looks for each of the target classes
  - Compare against a dataset where object locations are known => object localisation

Original Image | Looking here for 'Cat' | Looking here for 'Dog' | Localisation

Interpretability Methods

- Grad-CAM:
  - A gradient-based approach
  - Produces class-discriminative, high-resolution heat-maps from convolutional layers
  - Works with all CNN structures

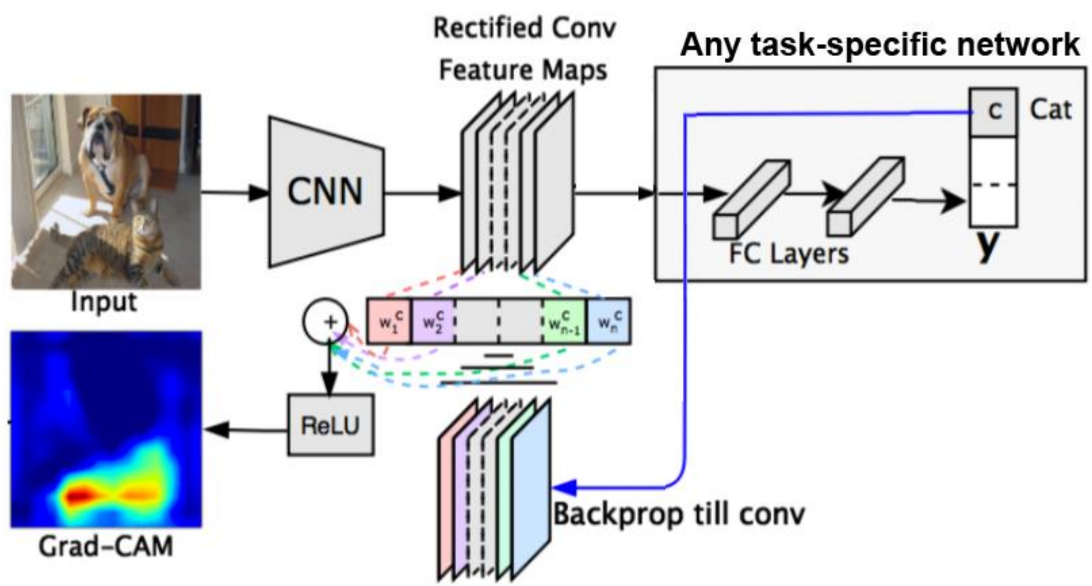


Original Image | Grad-CAM 'Cat' | Grad-CAM 'Dog'

Selvaraju et al., Grad-CAM: Visual explanations from deep networks via gradient-based localisation, ICCV, 2017.

Interpretability Methods

- Grad-CAM:
  - Backpropagate to a convolutional layer, and sum activation maps weighted by global average pooled gradient values



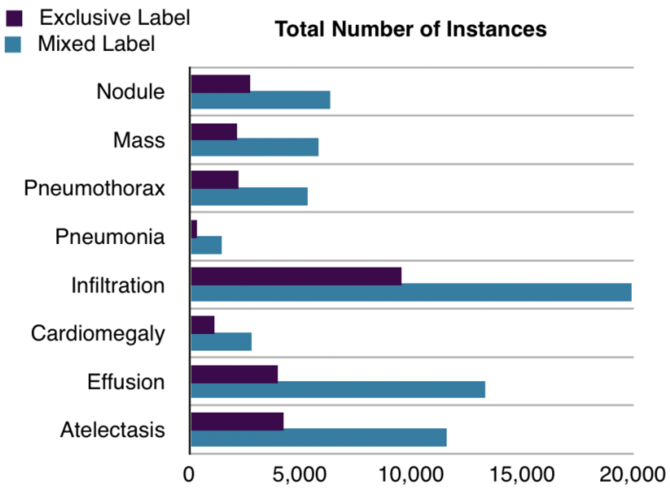
Selvaraju et al., Grad-CAM: Visual explanations from deep networks via gradient-based localisation, ICCV, 2017.

Objective

- Train a CNN model on a Chest X-ray dataset, to effectively classify a lung disease from healthy control
- Experiment with Grad-CAM on the learned CNN model
- Evaluate if the CNN model has learned interpretable features, i.e. able to localise discriminative regions that are representative of a particular disease

Chest X-Ray Dataset

- 112,000 labelled images:
  - 1024 x 1024 image + patient info
  - 14 diseases, 1 'No Finding' category (54% of the data!)
  - Labels are mixed
- 985 bounding boxes: 8 categories (70-180 per class)



Classification CNN

- Experimentation Process:
  - Images rescaled to 128 x 128 for efficiency
    - 256 x 256 was attempted (any higher used too much RAM)
  - Contrast stretching attempted → no noticeable impact
- Approach 1 - Classify into one of the 15 classes
  - 54% accuracy! … but everything classified as 'No Finding'
  - Class balancing → 25% test set accuracy
    - 10% - 60% per-class accuracy

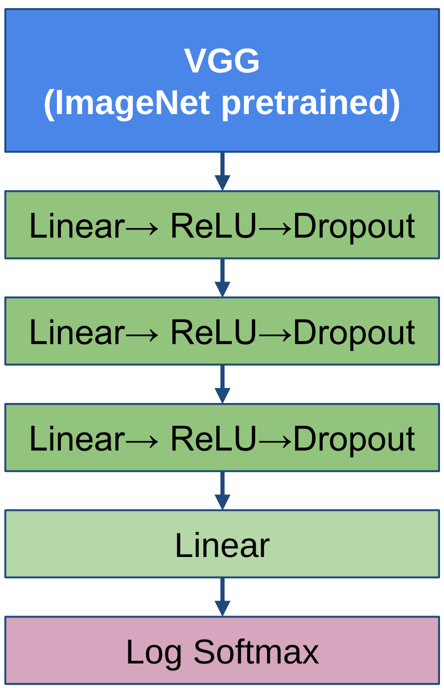
```
At:      0.2
Cardio:  0.6
Eff:     0.2
Inf:     0.1
Mass:    0.5
No_Fi:   0.0
Nod:     0.1
Pneuom:  0.3
0.2529569892473118
```

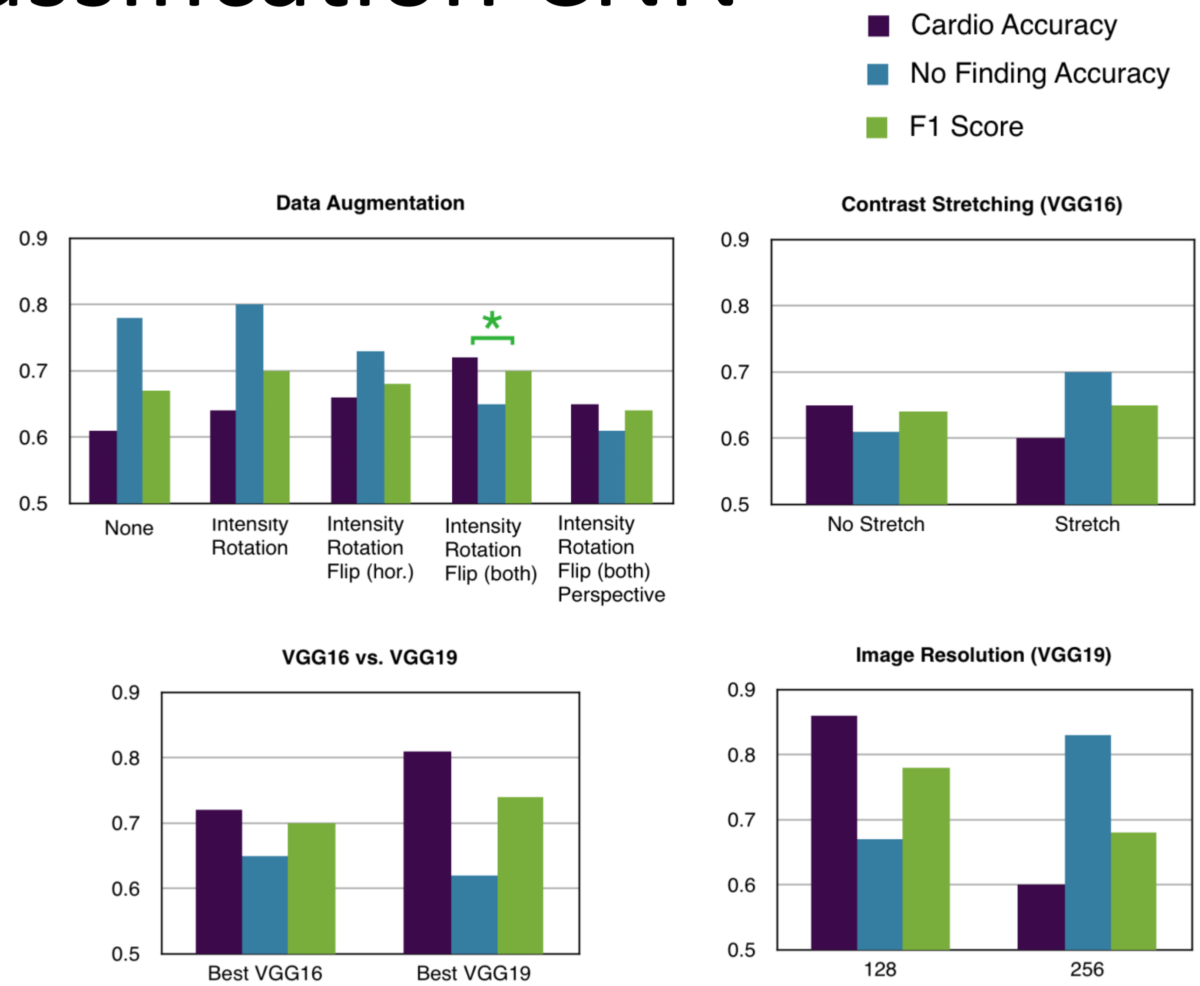
Classification CNN

- Approach 2 - Binary classification:
  - Attempted Mass, Nodule, Cardiomegaly
    - Cardiomegaly had best accuracy + easiest to localise on small images
  - **Cardiomegaly**: *"abnormal enlargement of the heart"*
    - 146 bounding boxes
    - 2776 mixed labels
    - **1093 exclusive labels** → no correlations = better localisation?
      - Of the 2776 Cardiomegaly, 38% also have effusion

Classification CNN

Transfer Learning VGG:



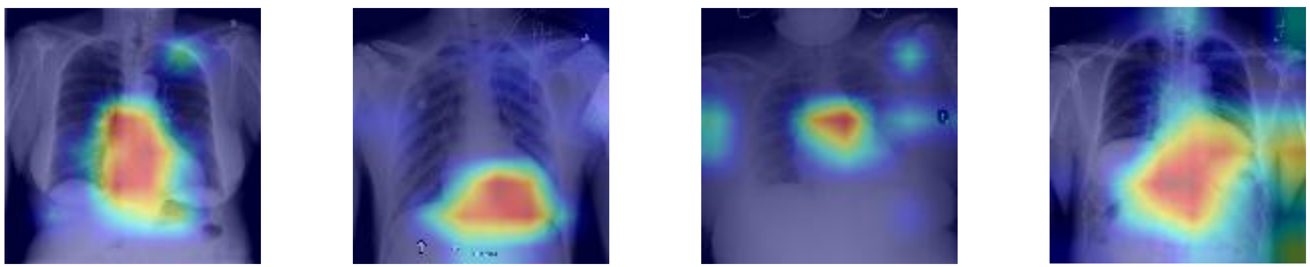


Classification CNN

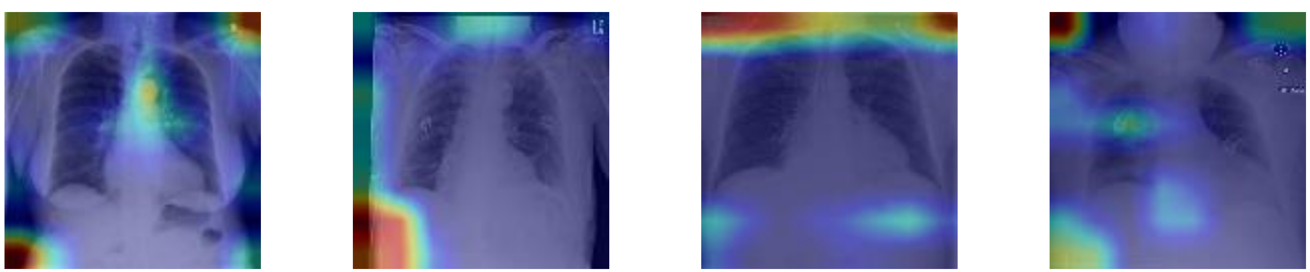
- Best Network: 81% cardio, 71% no finding, F1-score 0.77
- Other Network Types:
  - Zhang et al. [term 3] avoids **ResNet** when implementing their method
  - Vanilla **DenseNet121** achieved similar performance to VGG but is more complex
  - There are interesting approaches in literature for modifying DenseNet to achieve better accuracy **[1]**, **[2]**

Grad-CAM

- Ideal heatmap: Focused on heart (or at least inside the body)
- Some good examples:

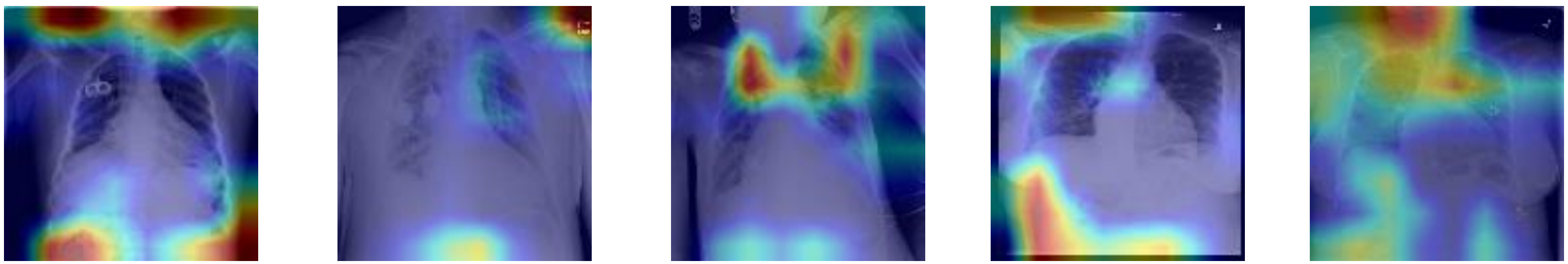


- Some bad examples:

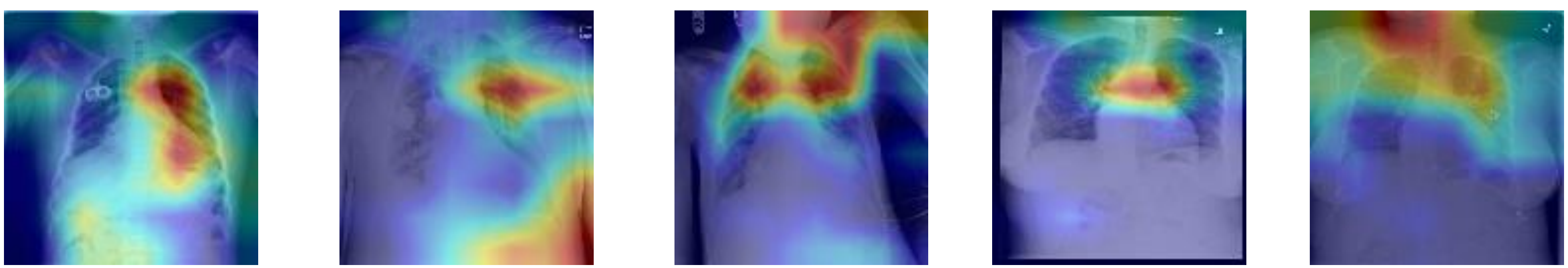


Grad-CAM

- VGG16, Augmented (rotation, intensity):

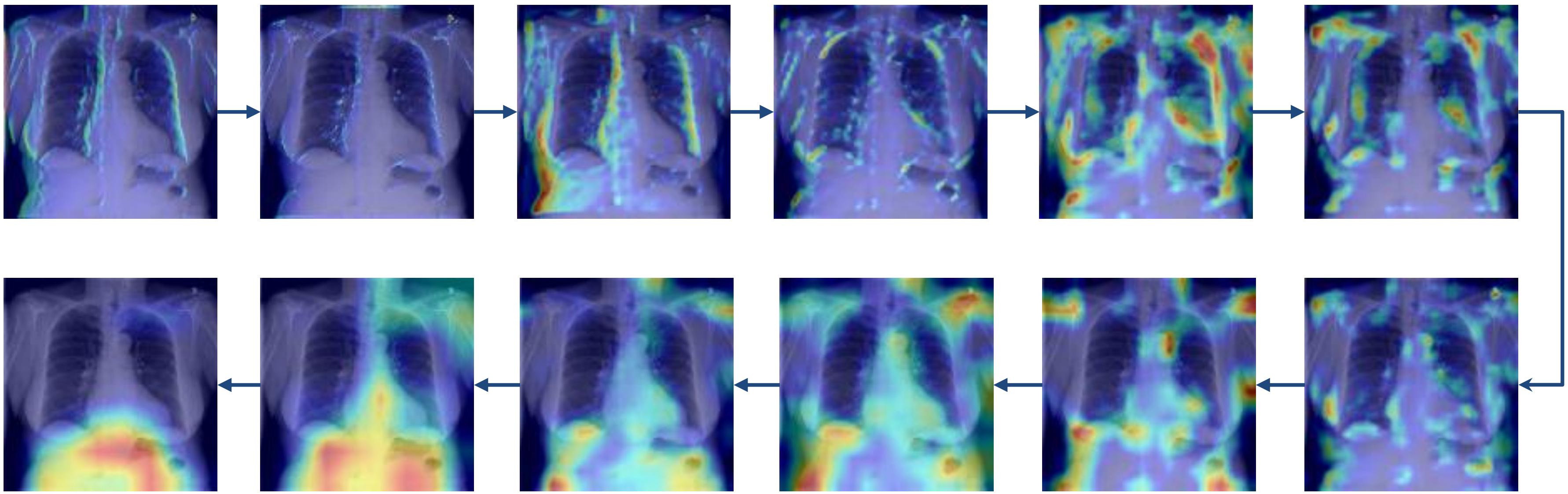


- VGG16, Augmented (rotation, intensity, flipping):



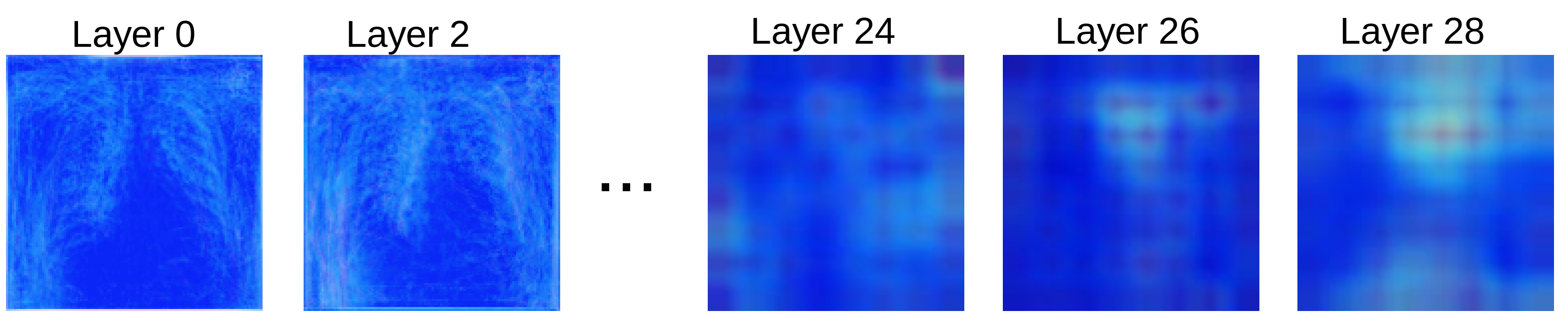
Grad-CAM

- Layer-wise examples:



Grad-CAM

- Mean layer heatmaps (VGG16, augmented intensity, rotation, flipping)

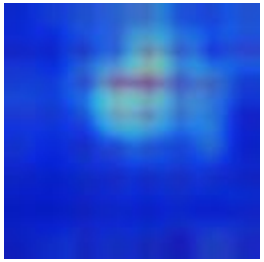


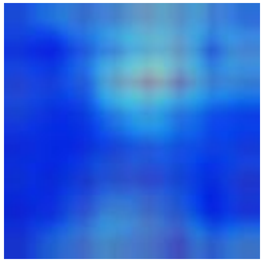
Grad-CAM

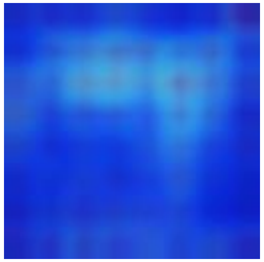
- Final layer of some networks


VGG16 (no augments)


VGG16 (all train data)


VGG16 (all augments)


VGG16 (fewer augments)


VGG19

Grad-CAM

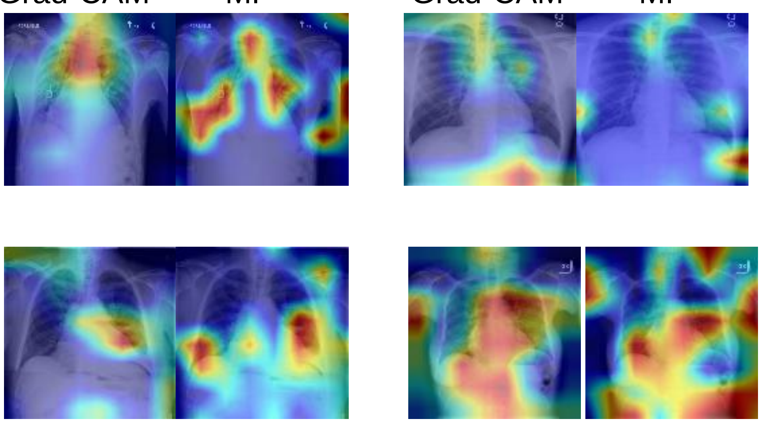
- Reflection:
  - Can target each layer individually
  - No parameters
  - New model topology requires changing implementation
  - Can be used to evaluate performance of different network designs
  - Can be used to compare similarity between models

Other Methods

- Grad-CAM vs. Meaningful Perturbation (MP):

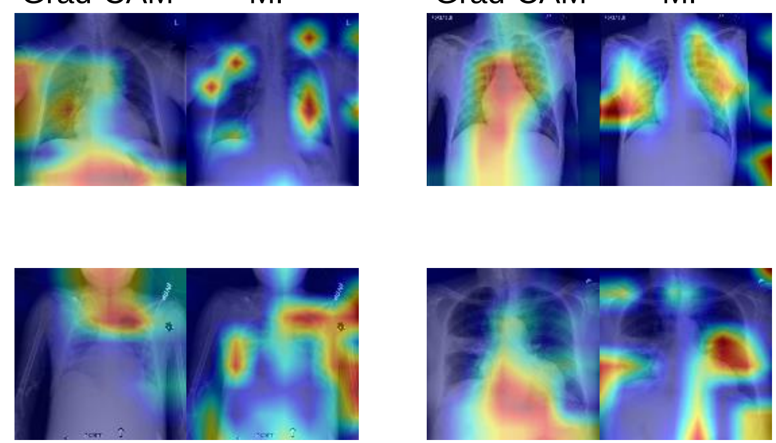
Similar Result

Grad-CAM MP Grad-CAM MP



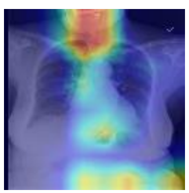
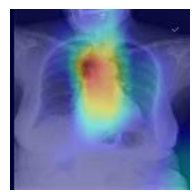
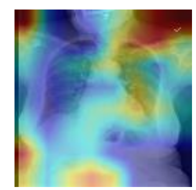
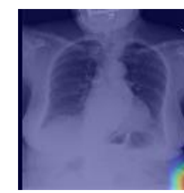
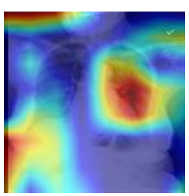
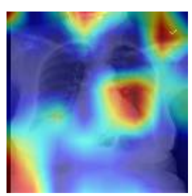
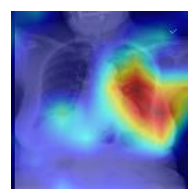
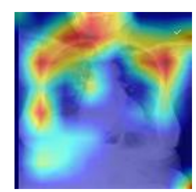
Different Result

Grad-CAM MP Grad-CAM MP



Other Methods

- Grad-CAM vs. Meaningful Perturbation (MP):

	VGG16 (intensity, rotation)	**VGG16 (+ flipping)**	**VGG19 (+ perspective)**	**VGG19 (+ flipping)**	**VGG16 (all weights trained)**
Grad-CAM					
MP					

Summary

- Case studies of:
  - WSI analysis
  - Mitosis detection
  - Interpreting deep learning

- We are recruiting research students in computer vision / deep learning

Acknowledgement

- Adopted some slides from
  - Chaoyi’s study on WSI analysis (ICIP 2018)
  - Veena’s study on mitosis detection (ISBI 2019)
  - Ari’s Honours thesis study