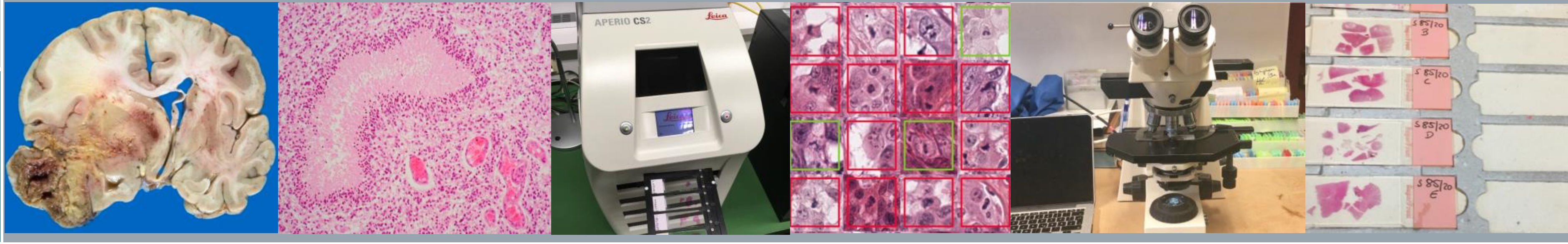
Pattern Recognition Lab
Lehrstuhl für Neuropathologie
Medizinische Fakultät der
FAU Erlangen-Nürnberg
FAU
FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG
FACULTY OF ENGINEERING
+
FACULTY OF MEDICINE

Seminar: Digital Pathology and Deep Learning



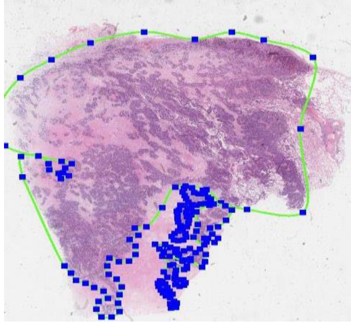
Weakly supervised learning for Digital Pathology

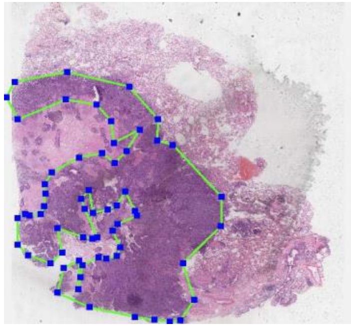
Tuhin Mallick

Agenda

- Motivation
- Introduction
  - The Cost of Annotations
  - Strongly versus Weakly Supervised learning
  - Sources of Information
  - Challenges
- Paper Discussion
  - Literature Review
  - Proposed Methodology
  - Experimental Results
  - Conclusions
- Summary

Motivation





Coarse annotations from pathologist

Cancer regions that are not annotated

Non – Cancer regions that are annotated

Annotations done by pathologist

Challenges for fully supervised learning for Whole Slide Image analysis

- Pixel-wise annotations are prohibitive
- Ambiguous regions cannot be distinguished

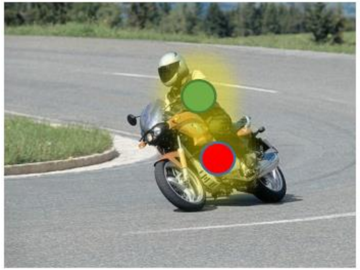
Weakly Supervised Learning is appealing

- Large number of image-level labels
- Small number of coarse annotations are required

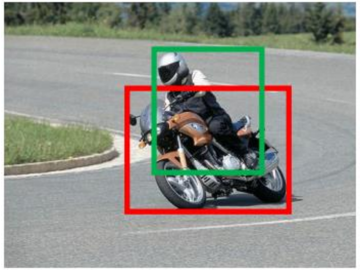
Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

The Cost of Annotations

{Motorbike(image), Person(image-level)}



{Motorbike(point), Person(point)}



{Motorbike(b-box), Person(b-box)}

{Motorbike(pixel labels), Person(pixel labels)}

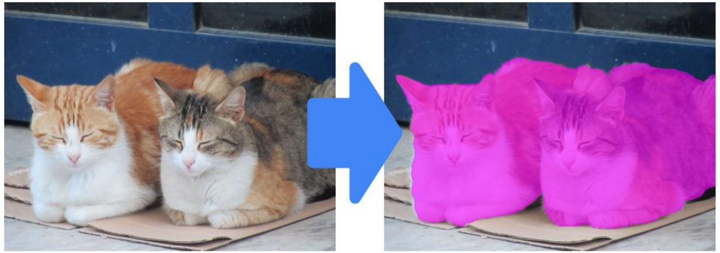
Image classification	**Instance spotting**	**Object detection**	**Semantic segmentation**
1 second per instance	2.4 second per instance	10 second per instance	78 second per instance

Time taken for annotation →

Source: Olga Russakovsky, Amy L. Bearman, Vittorio Ferrari, Fei-Fei Li:
What's the point: Semantic segmentation with point supervision. CoRR abs/1506.02106 (2015)

What is Weakly Supervised Learning ?

Fully supervised → **Directly** supervised



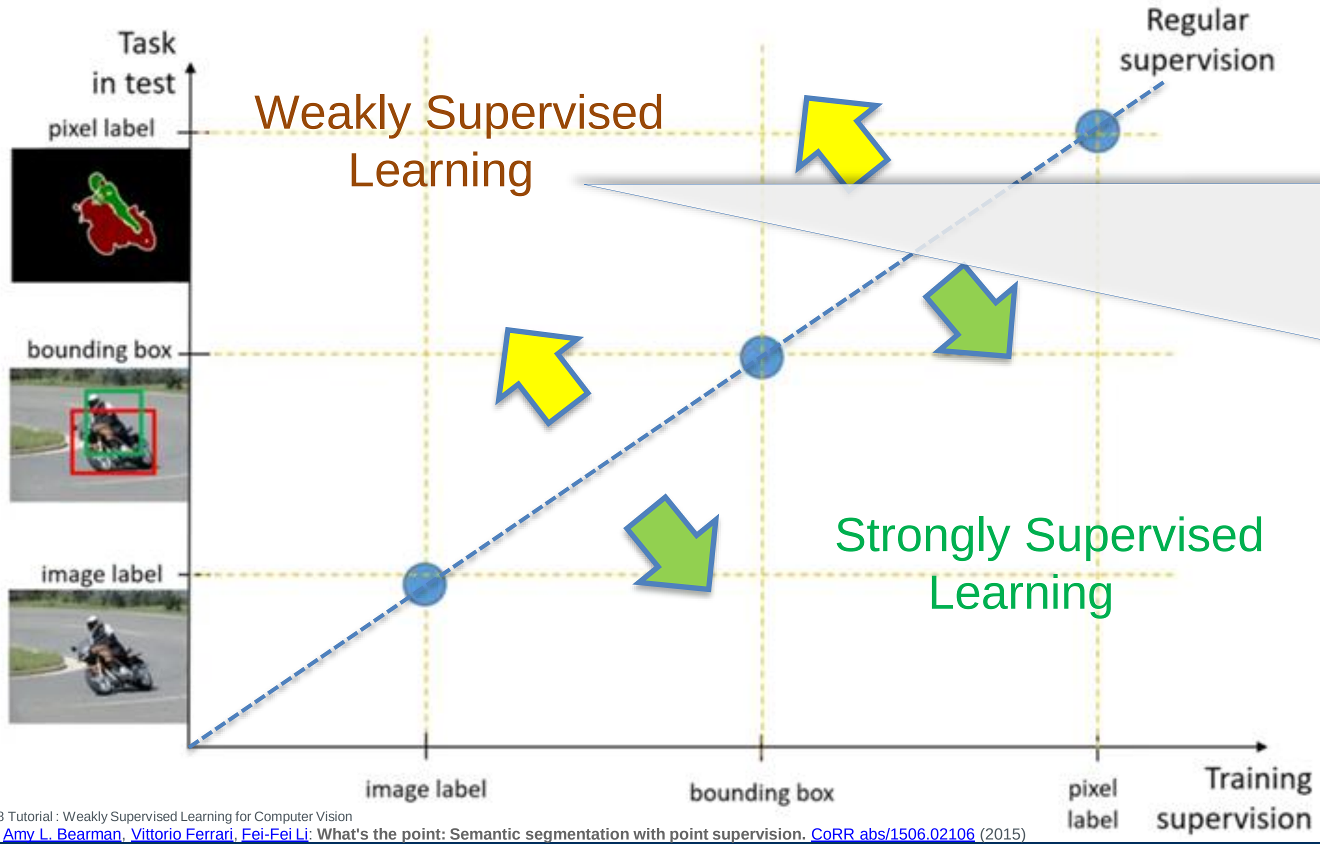
For each input, the desired output is provided

Weakly supervised → **Indirectly** supervised

Partial annotation of few points are provided

Source : Tsung-Yi Lin, Michael Maire, Serge J. Belongie, Lubomir D. Bourdev, Ross B. Girshick, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, C. Lawrence Zitnick: **Microsoft COCO: Common Objects inontext.** CoRR abs/1405.0312 (2014)

Strongly vs. Weakly Supervised Learning



Lower degree (or cheaper) annotation at train time than the required output at test time

Source: Reproduced from CVPR18 Tutorial : Weakly Supervised Learning for Computer Vision
Source: Olga Russakovsky, Amy L. Bearman, Vittorio Ferrari, Fei-Fei Li: **What's the point: Semantic segmentation with point supervision.** CoRR abs/1506.02106 (2015)

Sources of Information

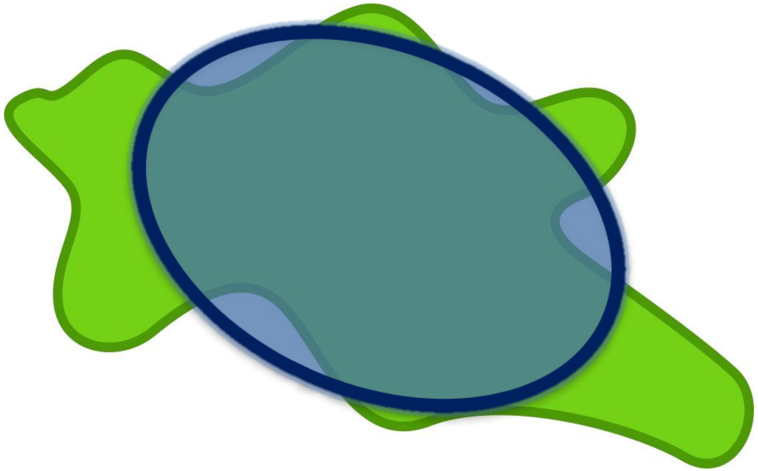
Priors: What is believed to be true independent of any particular image sample

Hints: indirect supervision received for each image

Priors and Hints

Priors : Explicit and Implicit

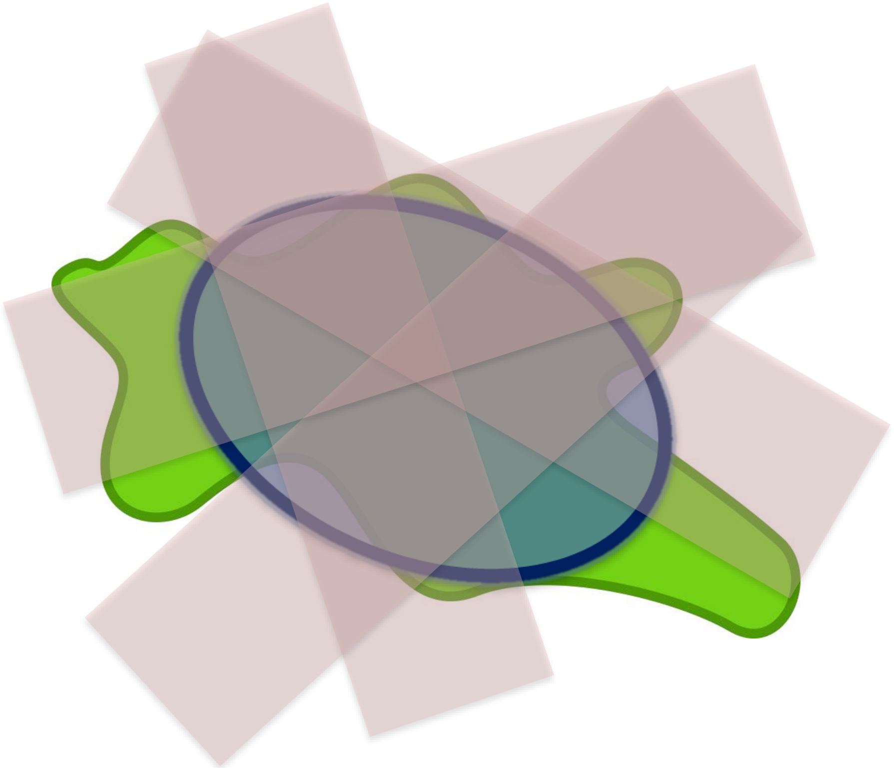
- Size
- Shape
- Location
- Number of instances
- Contrast (boundaries, saliency)
- Class distribution
- Similarity across images
- Similarity with external images



Priors and Hints

Hints

- Image labels
- Image captions
- Video labels
- Transfer across images
- Clicks inside images
- Object bounding boxes
- Scribbles

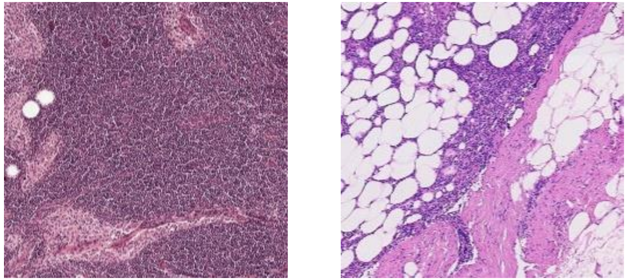


Challenges

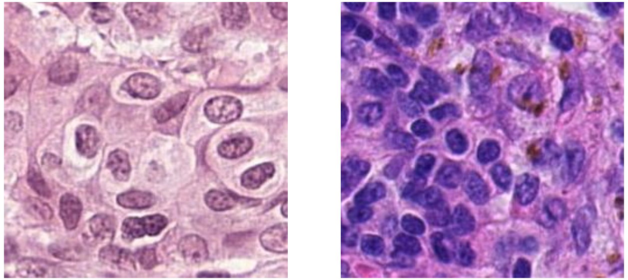
Generic Object Recognition

Intra-class variations

- Appearance
- Viewpoint
- Scale
- Aspect Ratio
- Background clutter
- Occlusions



Large variations introduced during the image acquisition process

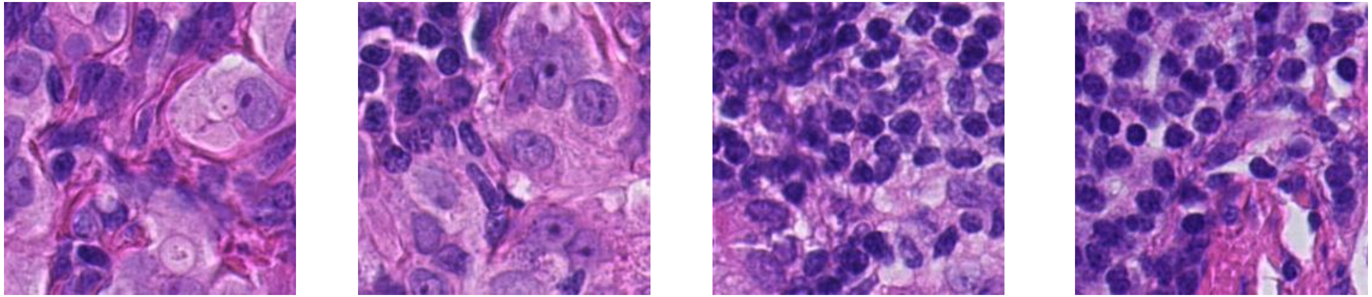


Variations of biological structures and textures

Source : Lin, Huangjing & Chen, Hao & Dou, Qi & Wang, Liansheng & Qin, Jing & Heng, Pheng-Ann. (2018).
ScanNet: A Fast and Dense Scanning Framework for Metastatic Breast Cancer Detection from Whole-Slide Images. 10.1109/WACV.2018.00065.

Challenges

Mimics



Tumor Background Tumor Background

Similarity group 1 Similarity group 2

Source : Lin, Huangjing & Chen, Hao & Dou, Qi & Wang, Liansheng & Qin, Jing & Heng, Pheng-Ann. (2018).

ScanNet: A Fast and Dense Scanning Framework for Metastatic Breast Cancer Detection from Whole-Slide Images. 10.1109/WACV.2018.00065.

Paper Discussion

To maximize the use of WSI labels readily available in clinical practice, *Wang et al*. presented a weakly supervised approach for fast and efficient classification of WSI of lung cancer.

Major Contribution

- Use of image-level labels with small number of coarse annotations
- Fully connected network to retrieve discriminative regions
- Different context-aware block selection and feature aggregation methods

Largest fine-grained lung cancer WSI dataset for comprehensive analysis

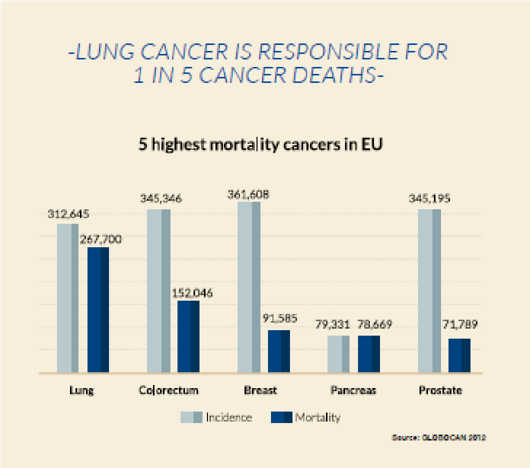
Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

Literature Review

- Lung Cancer is a group of diseases characterized by abnormal growths (cancers) that has started in the lungs

- Lung cancer is literally the biggest cancer killer worldwide causing more deaths than breast and prostate cancer put together all over the world.

- To sum it up every 30 seconds, someone, somewhere in the world dies of lung cancer.



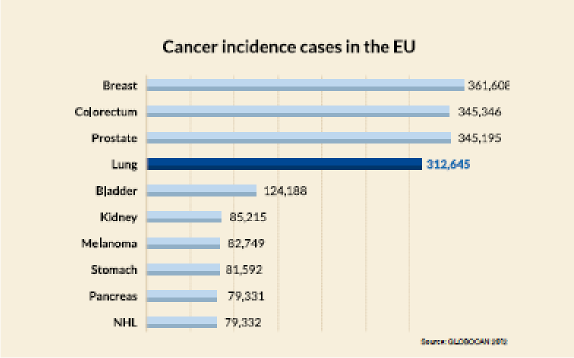
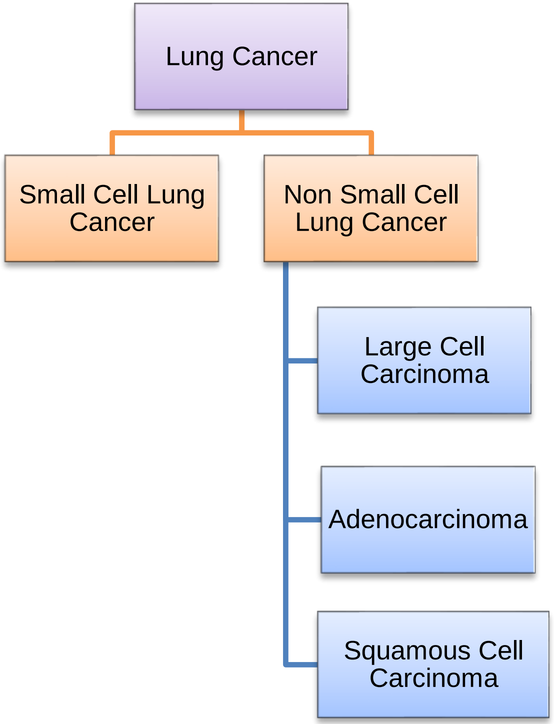
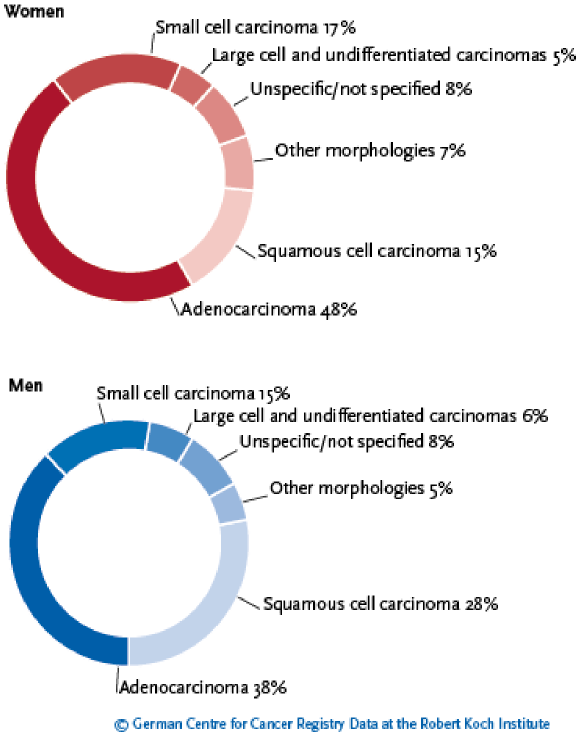


Image Source : https://www.lungcancereurope.eu/lung-cancer/

Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

Types of Lung Cancer



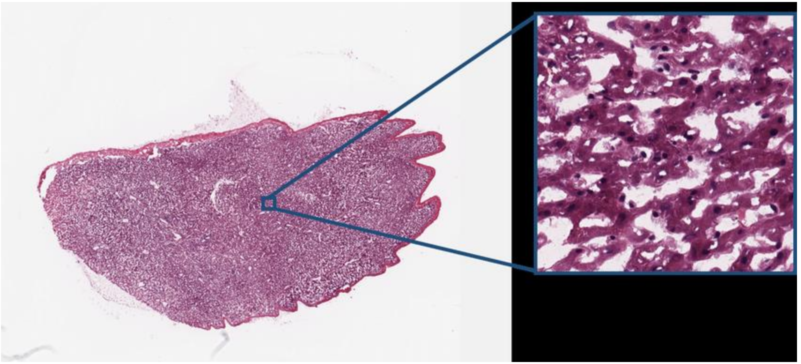


Distribution of malignant neuplasms of the lungs by histological type and sex

Source : Zentrum für Krebsregisterdate

Whole Slide Images

- One WSI may contain more than 10 billion pixels
- Only Image-level labels are available
  - Tumor regions are sparsely distributed (10%~90%) in WSIs.
  - Ground truth labels of individual extracted patches are unknown



A digital pathology whole slide image

Source : Liu, Feng & Hernández-Cabronero, Miguel & Sanchez, Victor & Marcellin, Michael & Bilgin, Ali. (2017).
The Current Role of Image Compression Standards in Medical Imaging. Information. 8. 131. 10.3390/info8040131.

Related Works

Bejnord et al. 2015

Features

- Features : local binary patterns, HSD color histogram
- Classifier : Random forest classifier

Drawbacks

- Handcrafted features like color histogram from manually selected ROIs
- Inability to learn discriminative patterns automatically from WSIs

Source : B. E. Bejnordi *et al.*, "Context-aware stacked convolutional neural networks for classification of breast carcinomas in whole slide histopathology images," *J. Med. Image.*, vol. 4, no. 4, 2017, Art. no. 044504.

Related Works

Hou et al. 2016

Features

- Features : EM based with patch-level CNN,
  statistic features from discriminative regions
- Classifier : Count based feature fusion model

Drawbacks

- Computationally expensive

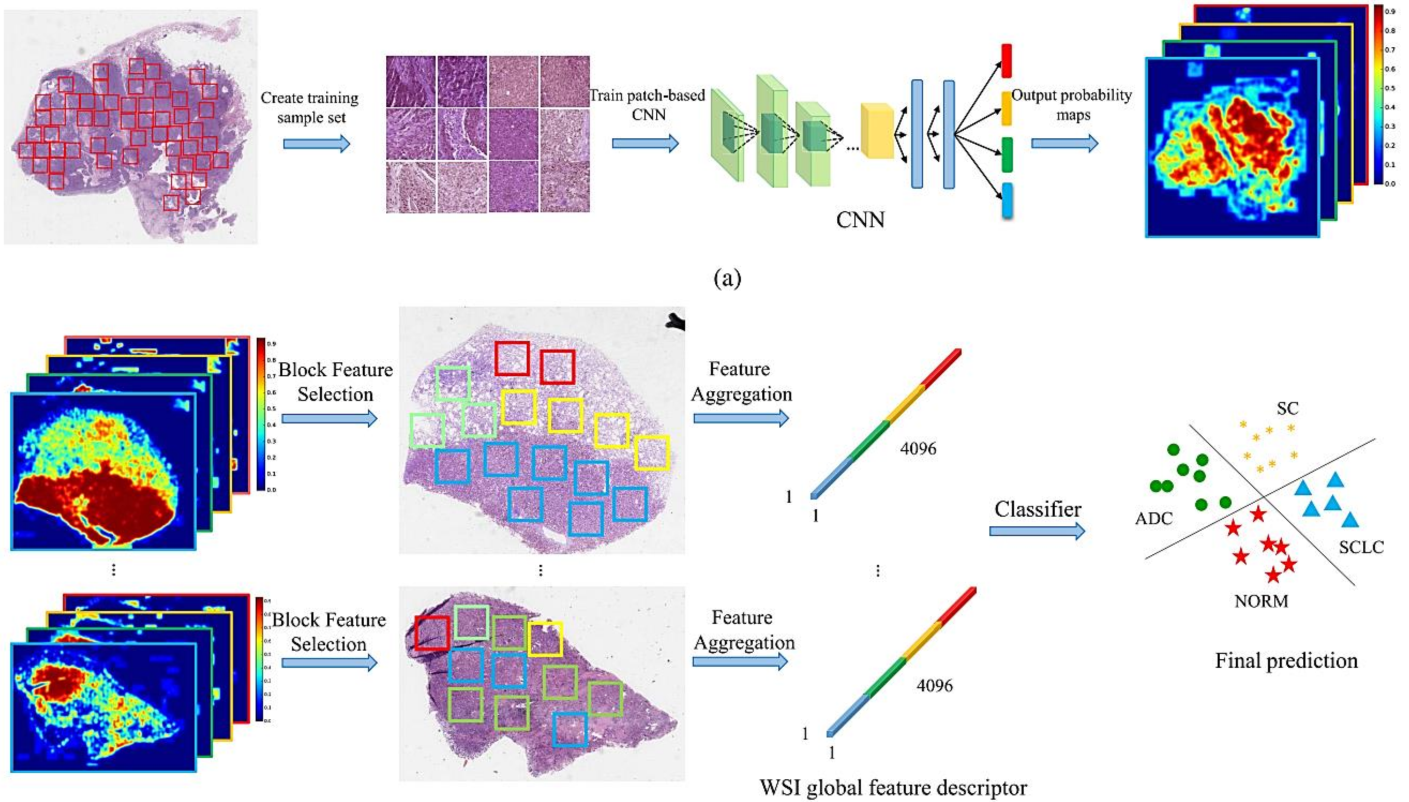
Source : L. Hou, D. Samaras, T. M. Kurc, Y. Gao, J. E. Davis, and J. H. Saltz, “Patch-based convolutional neural network for whole slide tissue image classification,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016,.

Proposed Method

(a) Discriminative patch prediction

(b) Context-aware feature selection and aggregation

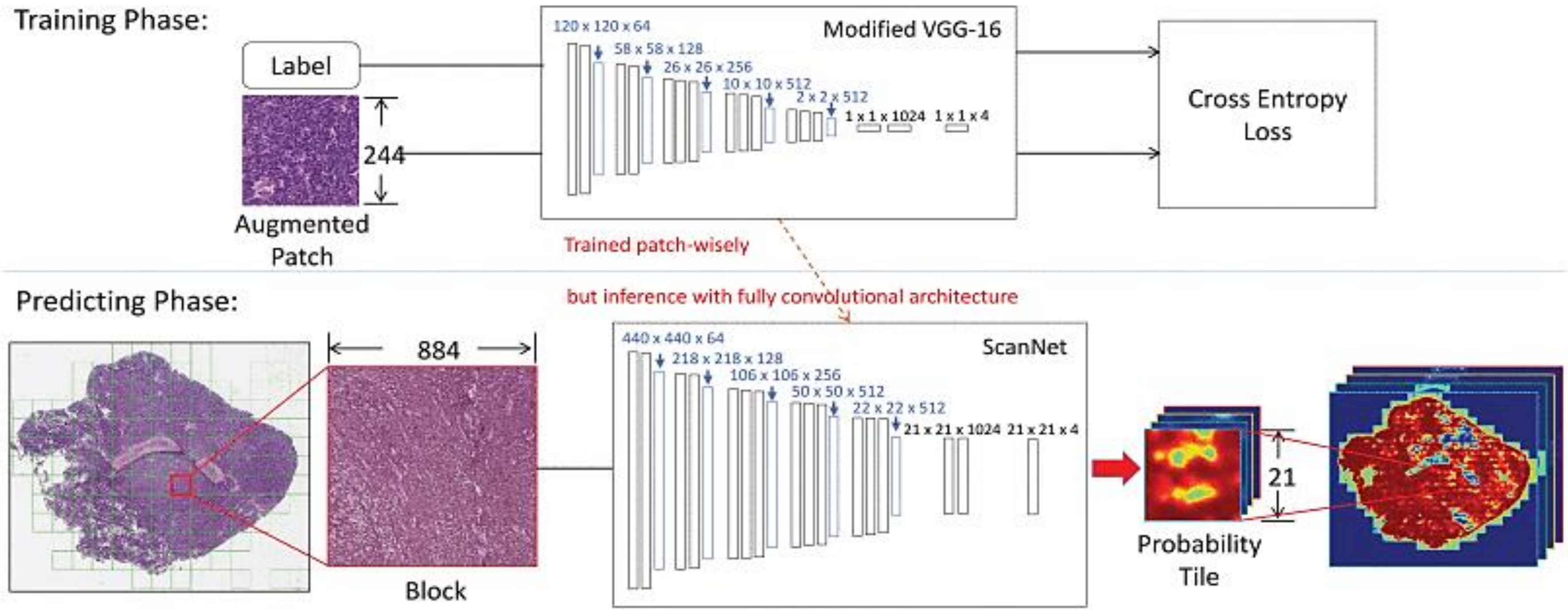
(c) WSI- level prediction



Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

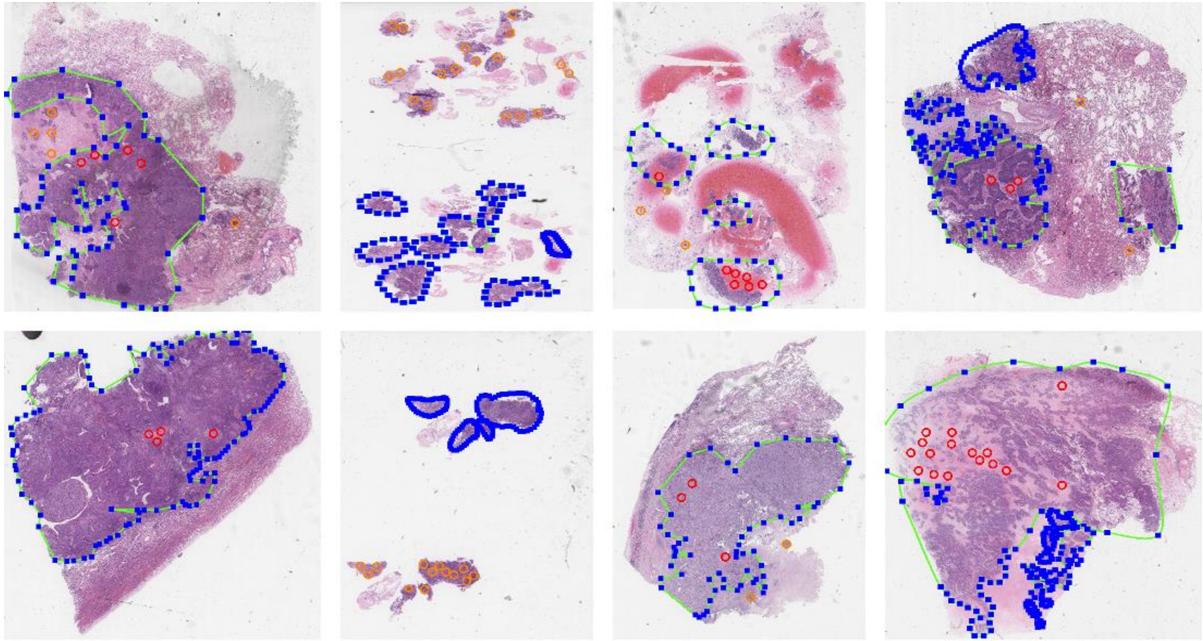
Discriminative patch prediction

Fast prediction via FCN



Source : Lin, Huangjing & Chen, Hao & Dou, Qi & Wang, Liansheng & Qin, Jing & Heng, Pheng-Ann. (2018).
ScanNet: A Fast and Dense Scanning Framework for Metastatic Breast Cancer Detection from Whole-Slide Images. 10.1109/WACV.2018.00065.

Coarse Annotations



Illustrations of non − exhaustive annotations by experienced pathologist

Coarse annotations from pathologists

Cancer regions that are not annotated

Non-cancer regions that are annotated

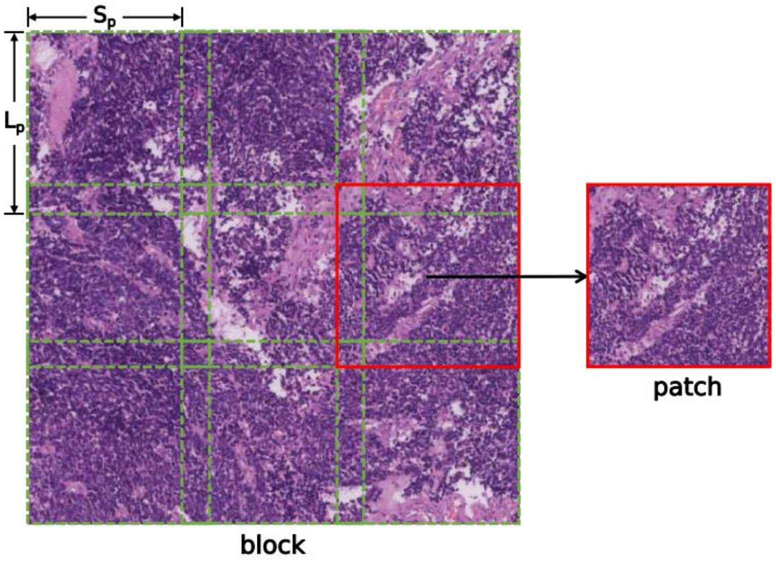
Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

Context- aware Block Selection

All previous works uses Discriminative patches

Drawbacks:

- Feature redundancy
- Outliers/mimics



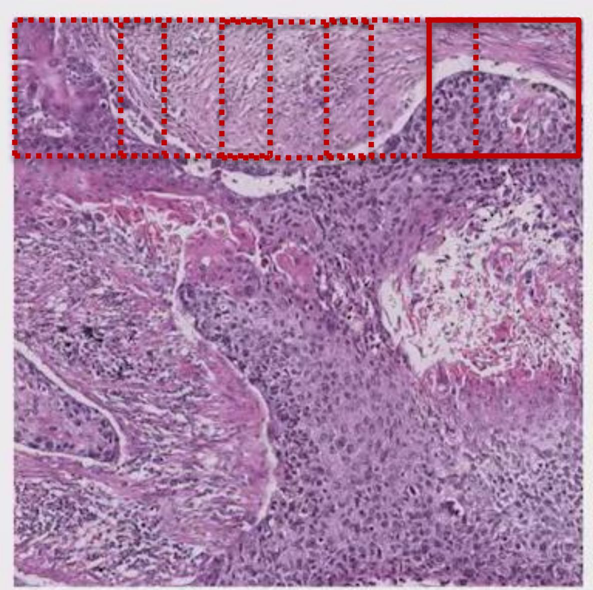
Simplified example of a block with 3 × 3 overlapped patches.

Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

Context- aware Block Selection

Rich Contextual information is quite important :

- Tumor area is generally larger than patch
- A block refers to a number of overlapped patches
- Each patch owns a probability vector from the last layer
- For each class, a block is considered discriminative if it exceeds a certain threshold τ


block

Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

Feature Aggregation

Goal : To find a good holistic feature descriptor for each WSI

- **Positive** evidence :

  It support the **existence** of cancer/non-cancer class that is consistent with the ground truth

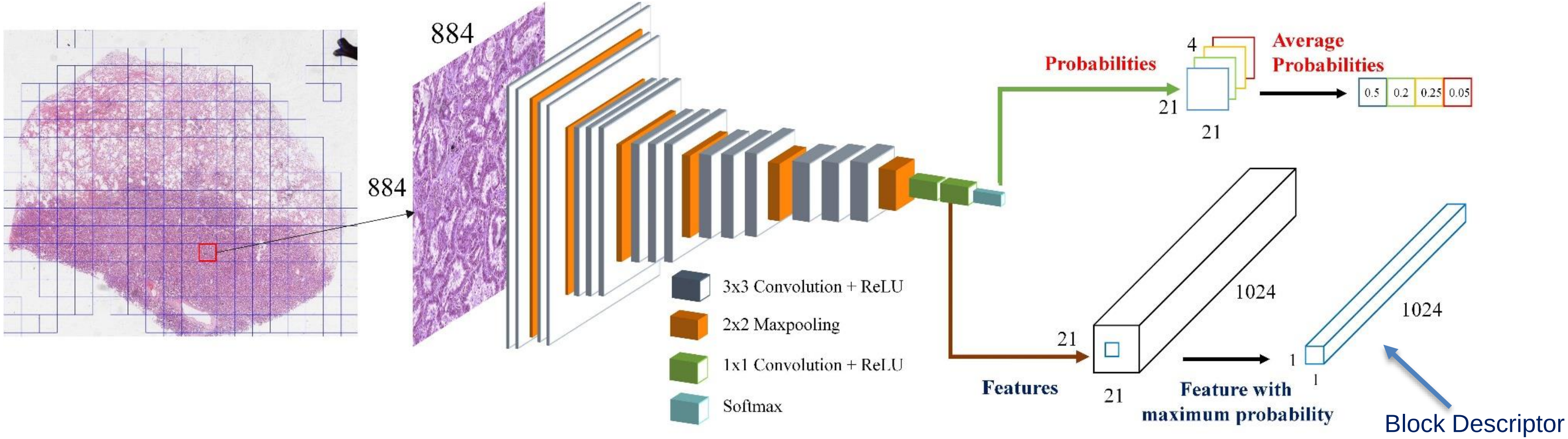
- **Negative** evidence :

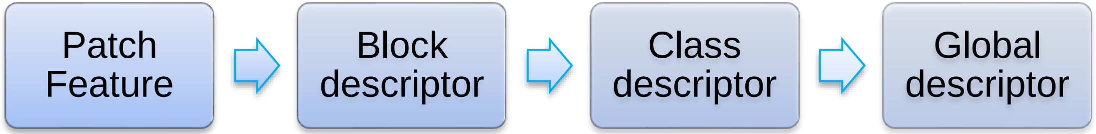
  It manifests the **absence** of any other classes

Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

WANG et al.: WEAKLY SUPERVISED DEEP LEARNING FOR WHOLE SLIDE LUNG CANCER IMAGE ANALYSIS

Feature Aggregation

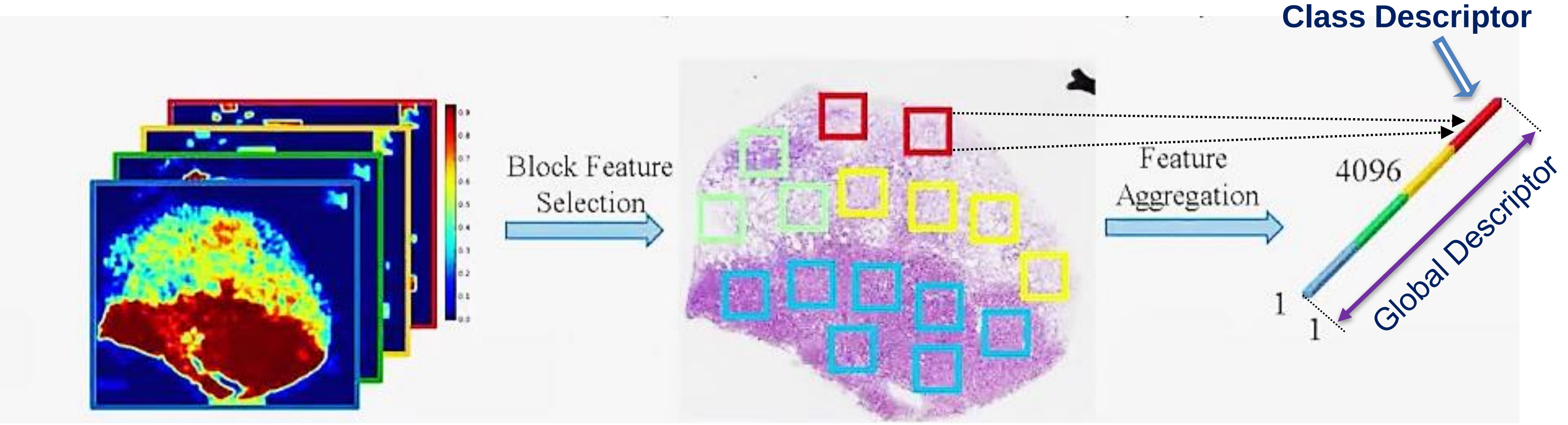




Three-stage feature aggregation

Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

Feature Aggregation



Patch Feature → Block descriptor → Class descriptor → Global descriptor

Three-stage feature aggregation

Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

Strategies for Feature Aggregation

$MaxFeat$

$$D_b = f_{max(i,j)}$$

$AvgFeat$

$$D_b = \frac{\sum_b f_{i,j}}{N^2}$$

$WeightedFeat$

$$D_b = \sum_{i,j} p_{i,j}\, f_{i,}$$

Different pooling strategies for Block descriptor

$Mean - pool$

$$D_c = \frac{\sum_b D_b}{N_{block}}$$

$Norm3$

$$D_c = \left(\sum_{b}^{N} D_b^3\right)^{1/3}$$

Different pooling strategies for Class descriptor

$Global\ Descriptor$

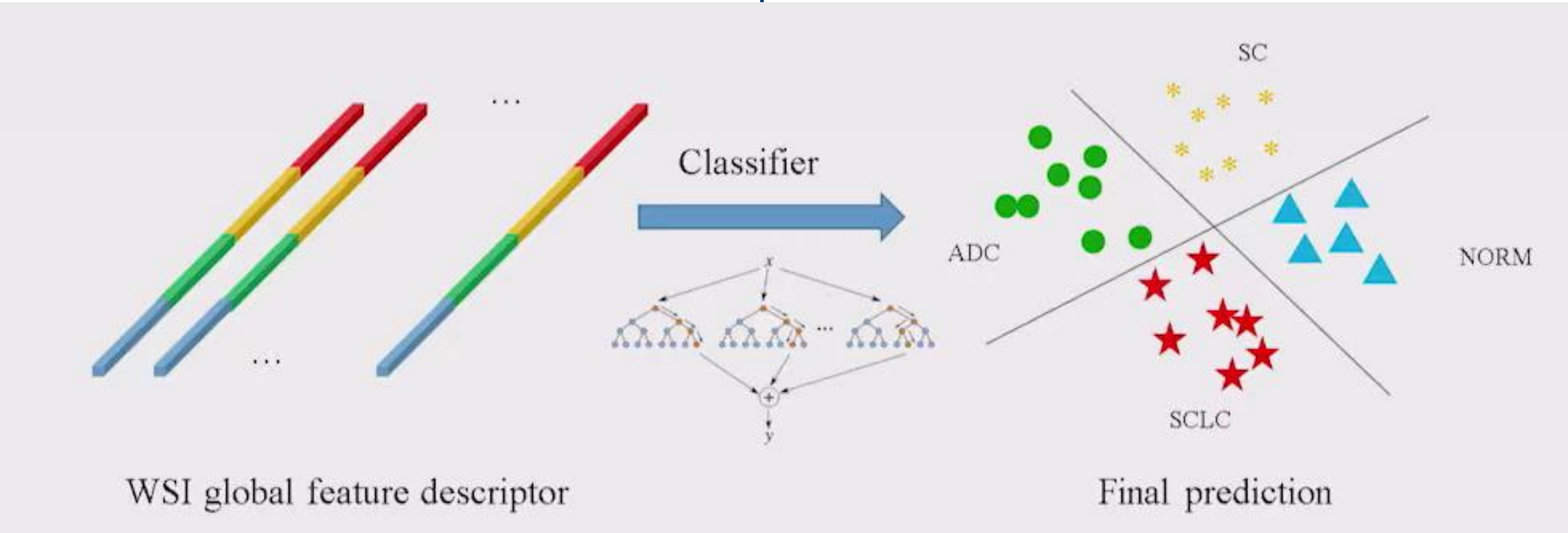
$$D_g = concat\{D_{c_1}, D_{c_2}, \ldots D_{c_c}\}$$

D_b= block descriptor
D_c= class descriptor
D_g= global descriptor
N = no. of discriminative blocks
p = probability
f = feature

Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

WSI Level Prediction

Standard random forest for WSI-level prediction



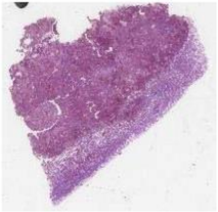
Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

Dataset

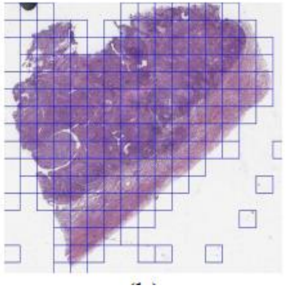
939 lung cancer WSIs containing 4 classes collected from Sun Yat-Sen University Cancer Centre (SUCC) :

- Normal (NORM: 68)
- Squamous Carcinoma (SC: 361)
- Adenocarcinoma (ADC : 390)
- Small Cell Lung Carcinoma (SCLC : 120)

Only 59 WSIs have coarse annotations



Raw WS



Segmented tissue regions denoted by blue rectangles

Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

Configuration to training datasets

- M1 : D1 and D3 (59 C and 53 NC) + typical cross entropy loss function
- M2 : D1 and D3 (59 C and 53 NC) + weighted loss function
- M3 : D1, D2 and D3 (701 C and 53 NC) + typical cross entropy loss function
- M4 : D1, D2 and D3 (701 C and 53 NC) + weighted loss function

C: Cancer images

NC: Non-cancer images

		Squamous Carcinoma	Small Cell	Adenocarcinoma	Small Cell Lung Carcinoma	Normal
Training	*D1*	59	21	20	18	–
	D2	642	267	293	82	–
	D3	–	–	–	–	53
Testing	*D4*	170	73	77	20	15

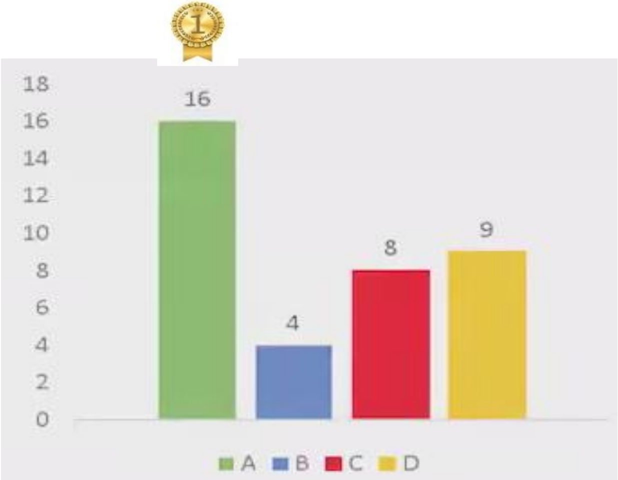
Data Distribution in database

Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

Configuration of Feature Aggregation Methods

1. Majority Pooling

- Obtain a score map by employing CNN on testing WSI
- Prediction of each location votes to four classes
- Category with majority vote is taken as prediction of image



Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

Configuration of Feature Aggregation Methods

2. **Average Pooling**

- Calculate the average probability of the locations on the test WSI
- Score map for each class channel
- Category with highest average probability is taken as image-level prediction

0.54	0.13	0.08	0.25

Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

Configuration of Feature Aggregation Methods

3. Max Pooling

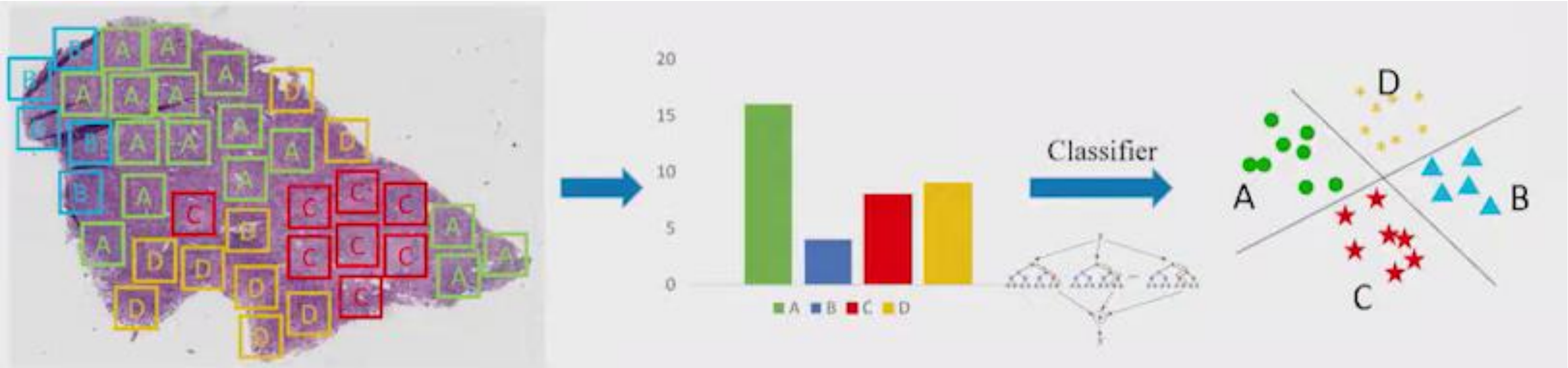
- Select the maximum probability for each class channel
- Score map for each class channel
- Category with highest max-pooling probability is taken as image-level prediction

0.92 | 0.82 | 0.57 | 0.95

Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

Configuration of Feature Aggregation Methods

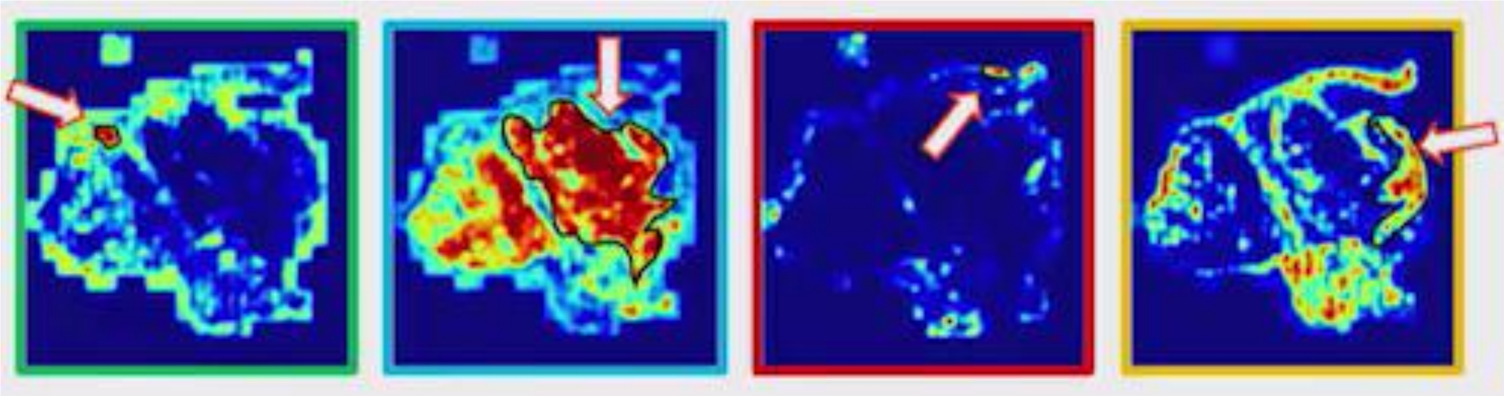
4. **Count-based RF**
   - Count the number of all cancer and non-cancer type prediction in test WSI
   - Score map to form a prediction histogram of classes
   - Four bit histogram is fed into an RF classifier for the image-level prediction



Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

Configuration of Feature Aggregation Methods

5. **Component-Based RF**
   - Connected component with the largest area for each class is chosen as the ROI
   - Obtain different features of this ROI for each test WSI score map
   - RF classifier takes the feature vector as input to get the final prediction



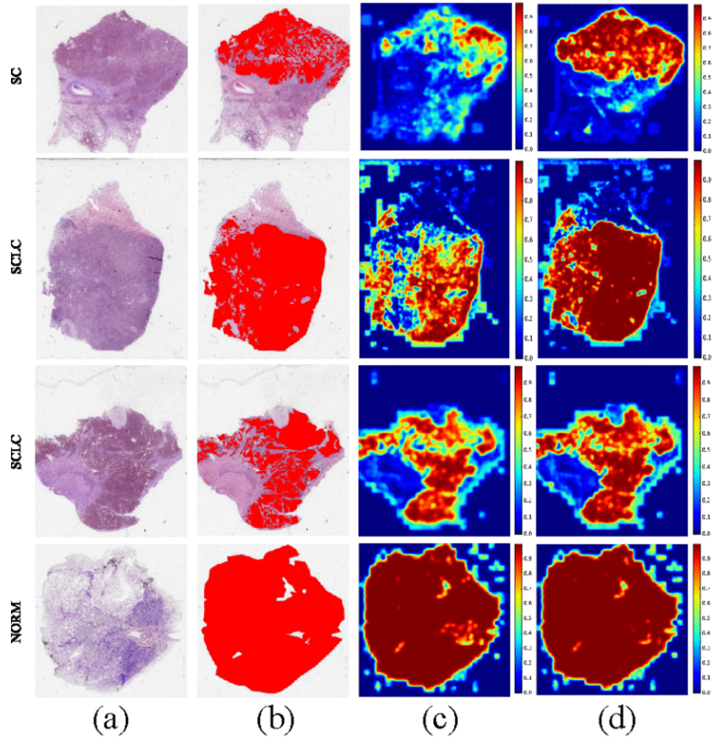
Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

Qualitative evaluation

$Method$	$M1$	$M2$	$M3$	$M4$
$MajorityVoting$	0.708	0.719	0.665	0.697
$AveragePooling$	0.730	0.735	0.676	0.703
$MaxPooling$	0.530	0.681	0.616	0.627
$Count - based\ RF$	0.770	0.783	0.875	0.930
$Component - based\ RF$	0.748	0.759	0.909	0.935
$CNN - AvgFeat - MeanPool - based\ RF$	0.786	0.812	0.928	0.955
$CNN - MaxFeat - MeanPool - based\ RF$	0.732	0.824	0.953	0.971
$CNN - WeightedFeat - MeanPool - based\ RF$	0.767	0.858	0.932	0.960
$CNN - AvgFeat - Norm3 - based\ RF$	0.816	0.843	0.943	0.962
$CNN - MaxFeat - Norm3 - based\ RF$	0.778	0.827	0.931	0.965
$CNN - WeightedFeat - Norm3 - based\ RF$	0.789	0.811	0.941	0.973

$Ablation\ Study:\ Results\ from\ different\ training\ datasets\ and\ feature\ selection\ method$

Qualitative evaluation



Visualization of discriminative region detection. (a) WSI. (b) Ground Truth. (c) M3: Heatmap. (d) M4: Heatmap

Source : Wang, Xi & Chen, Hao & Gan, Caixia & Lin, Huangjing & Dou, Qi & Tsougenis, Efstratios & Huang, Qitao & Cai, Muyan & Heng, Pheng-Ann. (2019). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics. PP. 1-13. 10.1109/TCYB.2019.2935141.

Conclusions

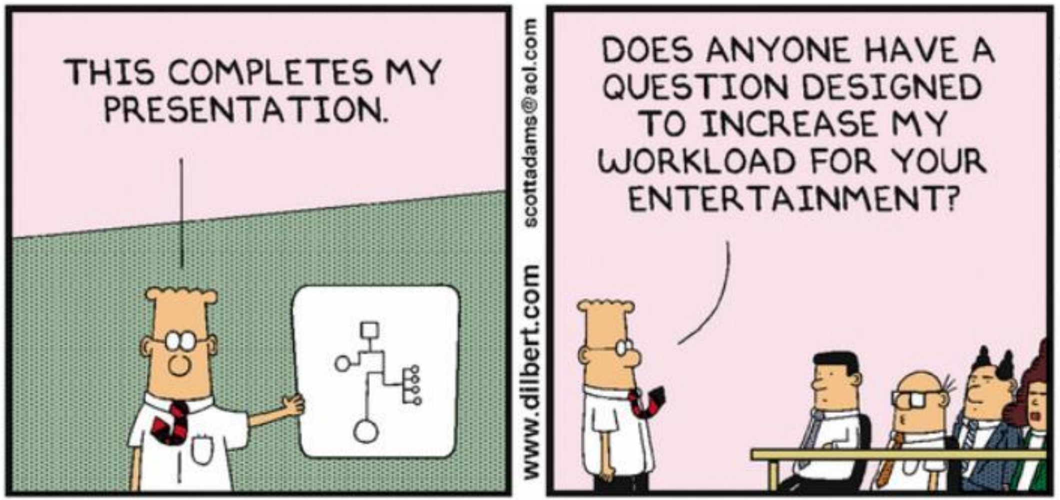
- Weakly supervised learning method to address the whole slide lung cancer image classification problem with minimum annotation effort.
- Proposed weighted loss function for the CNN
- Explored different context aware block selection and feature aggregation methods
- Constructed and validated the performance based on SUCC dataset.

Future scope

- Automated **feature selection** and aggregation by **adaptive** learning
- Replace the **RF classifier** with **MLP classifier** to make it end-to-end
- Use **more** lung cancer datasets to **validate** the generalization capability

Summary

- Computer vision requires lots of data; data is expensive
- Weakly supervised learning method can be used to address the whole slide lung cancer image classification problem with
  - **Minimum** annotation effort
  - Features extracted by CNN are the **ideal substitute** of handcrafted features
  - Alleviate the bottleneck of **expert annotation cost**
  - **Advance** the progress of computer-aided histology image analysis



Thank you for your attention!