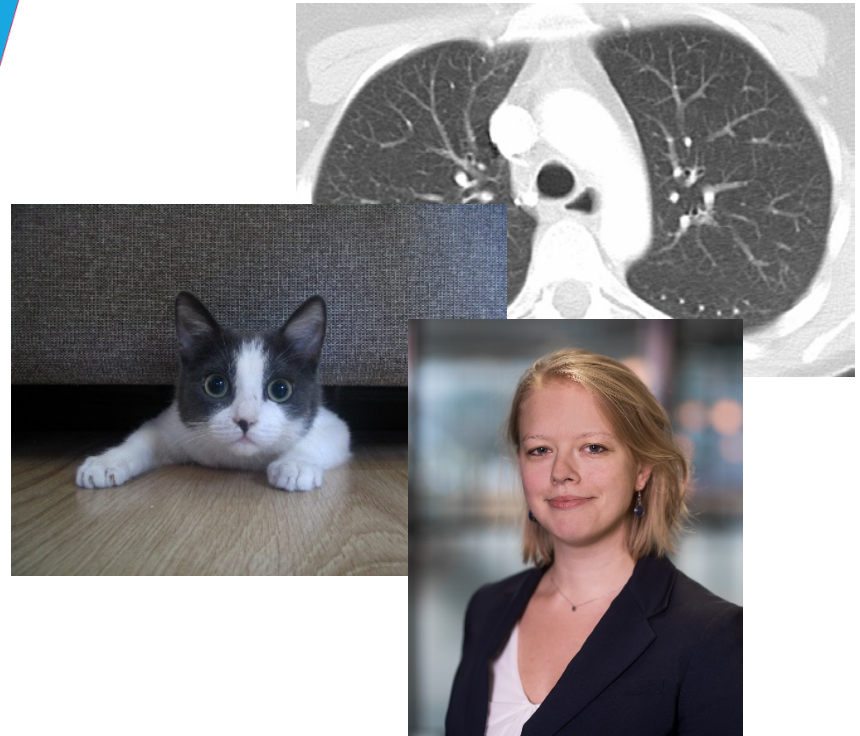


# Not-so-supervised learning of algorithms & academics

Veronika  
Cheplygina



@drveronikach



<http://www.veronikach.com>



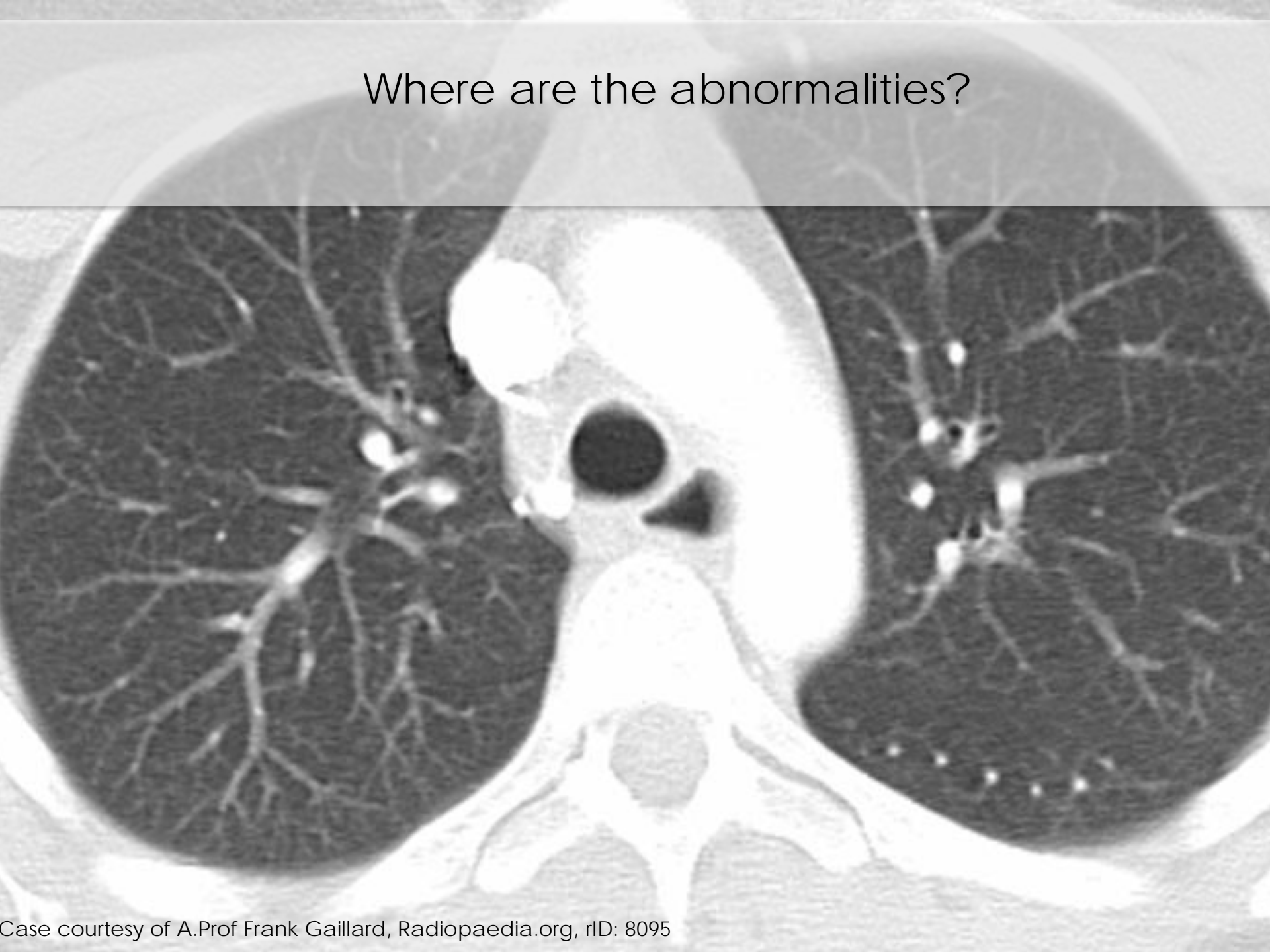




Data?



Where are the abnormalities?

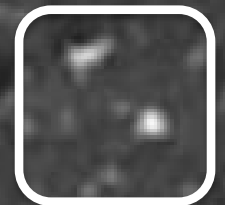
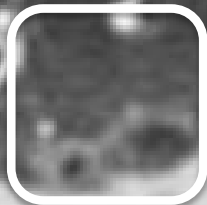
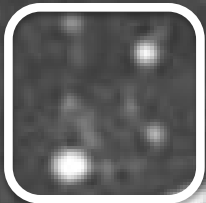


Training data for supervised learning





Training data we get



Training data not representative







Data?

Representative  
& annotated data



# Learning curve



@drveronikach



<http://www.veronikach.com>







**Janelle Shane**

@JanelleCShane

Follow



Does anyone have a picture of sheep in a really unusual place? It's for pranking a neural net.

4:55 PM - 1 Mar 2018

1,329 Retweets 3,698 Likes



466



1.3K



3.7K





**Dr Veronika CH** @DrVeronikaCH · 2 Mar 2018



I've got one of those as well







**Dr Veronika CH** @DrVeronikaCH · 2 Mar 2018



I've got one of those as well



3



45



**Janelle Shane** @JanelleCShane · 2 Mar 2018



I love how it took special care to also get the animal's color wrong: "a little boy holding a brown teddy bear"



1



3



73







This talk: three “solutions”

- Multiple instance learning
- Transfer learning
- Crowdsourcing

Bonus:

- Not-so-supervised career path



# Solution 1: Multiple instance learning

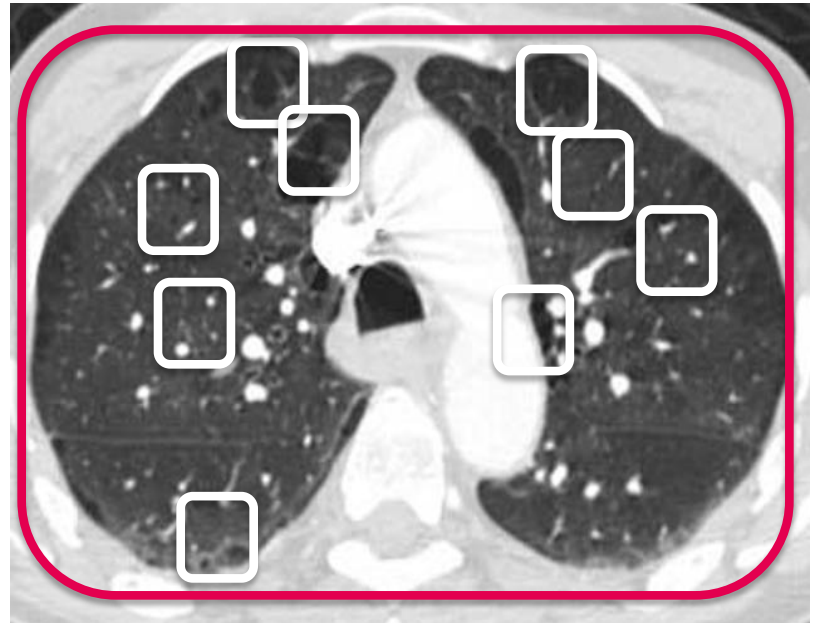
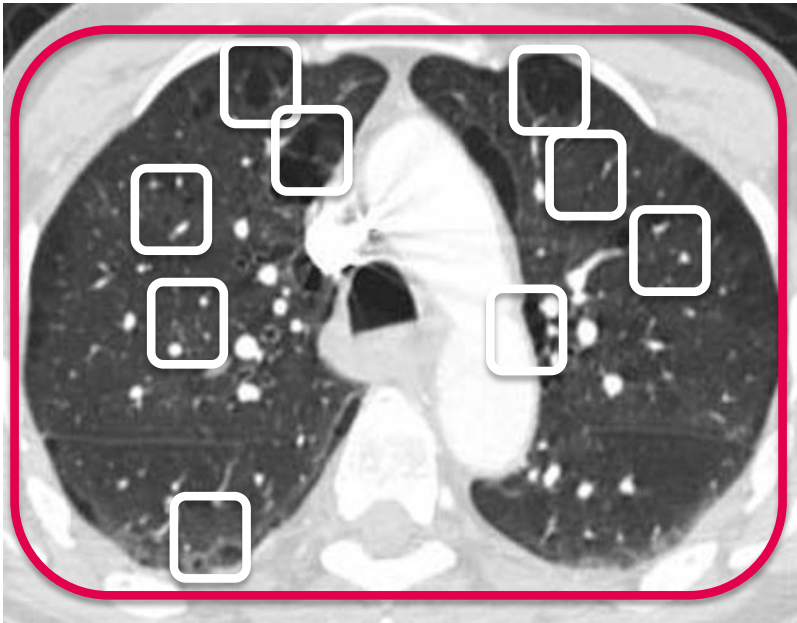
Learn with global information – Carolyn is in both pictures





# Solution 1: Multiple instance learning

Learn with global information



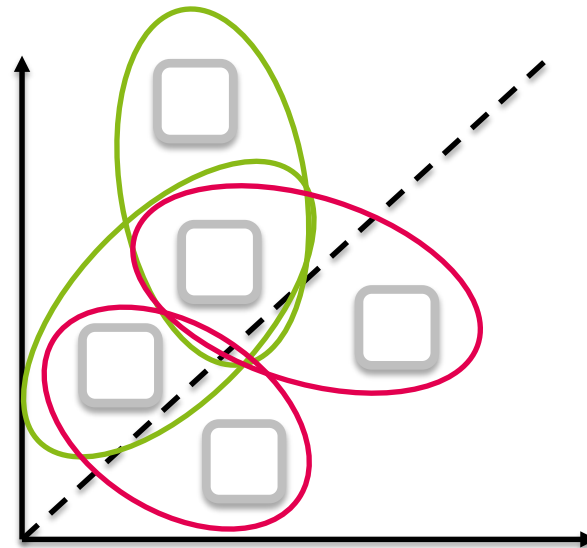
# Solution 1: Multiple instance learning

MIL

Search for positive instances

VS

Classify whole bag

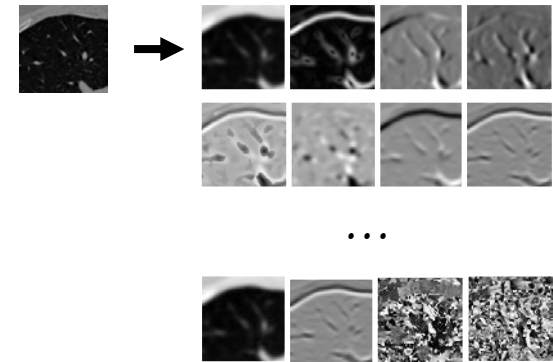


@drveronikach

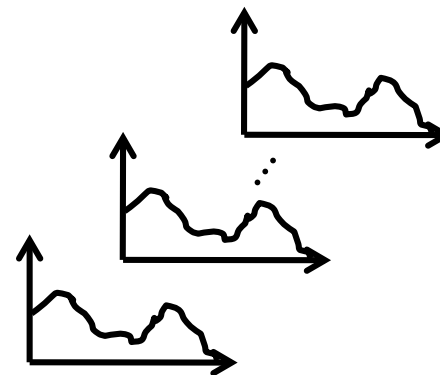


<http://www.veronikach.com>

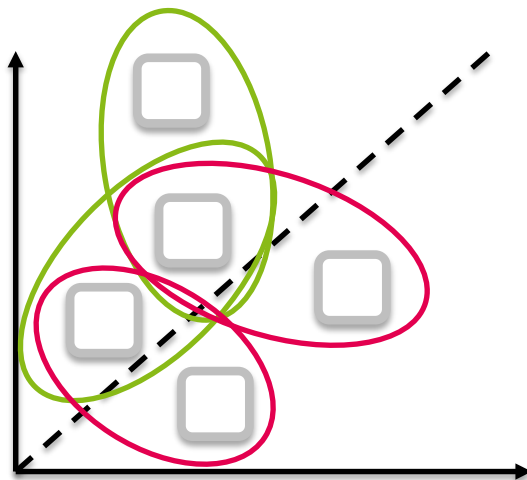




Texture filters



Histograms



MIL

Image = COPD or not (lung function), 50 ROIs



Performance metric: subject-level AUC

Search for abnormalities

VS

Classify whole subject

Classifier	AUC $\mathcal{X}_{val}$	AUC $\mathcal{X}_{te}$
Simple logistic noisy	50.0	50.0
Simple logistic avg	<b>71.9</b>	<b>70.5</b>
Simple $k$ -NN noisy	61.0	65.9
Simple $k$ -NN avg	67.0	67.8
miSVM noisy	<b>69.7</b>	65.4
miSVM avg	<b>74.5</b>	<b>71.7</b>
MILBoost	55.8	61.4
Citation $k$ -NN	65.2	61.5
mean-inst SVM	<b>74.0</b>	<b>74.2</b>
extremes SVM	<b>70.8</b>	68.6
BoW SVM	50.0	50.0
MILES	65.8	68.2
meanmin SVM	70.8	<b>71.3</b>
meanmin $k$ -NN	65.0	<b>69.1</b>
emd SVM	<b>73.7</b>	<b>74.6</b>
emd $k$ -NN	65.1	67.1

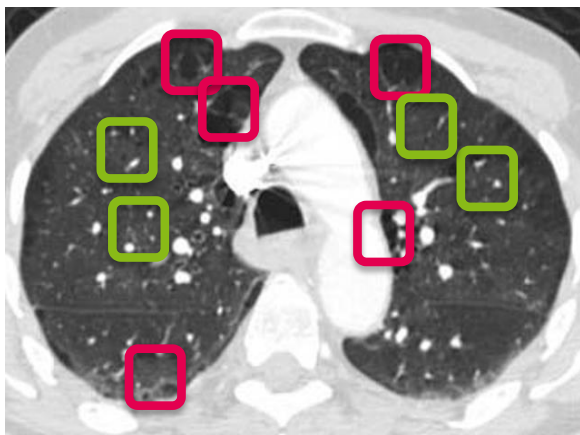
Cheplygina, V., Sorensen, L., Tax, D. M. J., Pedersen, J. H., Loog, M., & de Bruijne, M. (2014). Classification of COPD with multiple instance learning. In International Conference on Pattern Recognition (pp. 1508-1513).

Training



MIL classifier

Test



Ground truth?

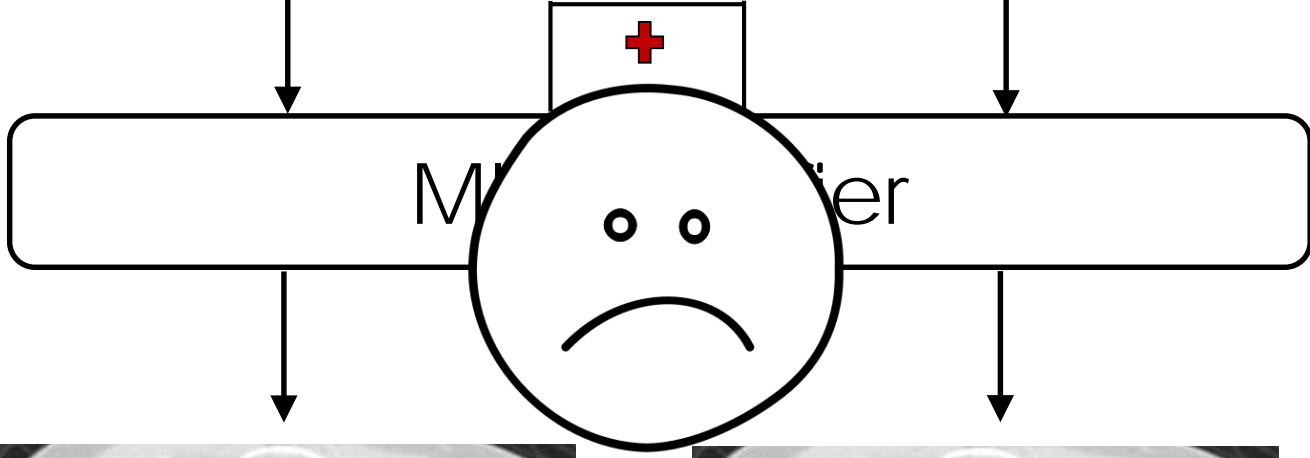
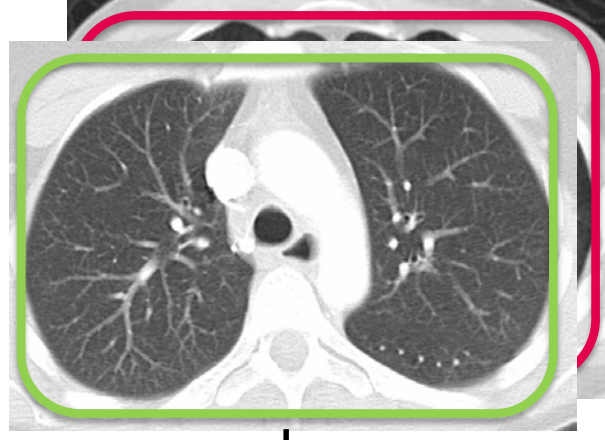
@drveronikach



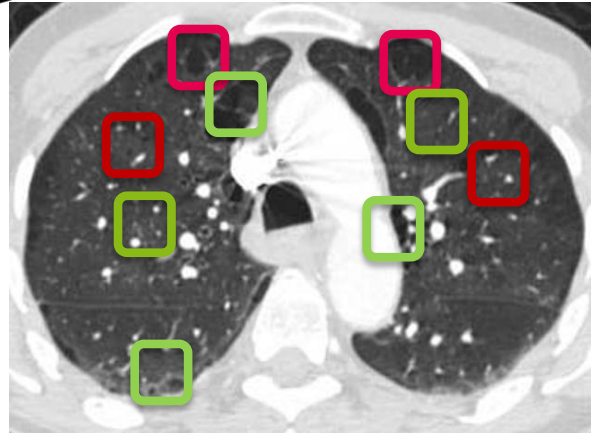
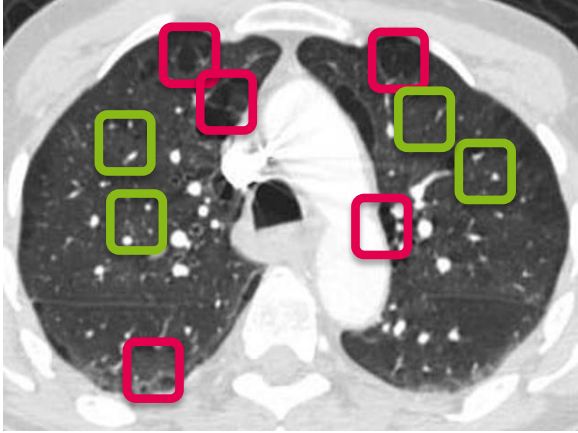
<http://www.veronikach.com>



Training



Test





# Evaluate stability

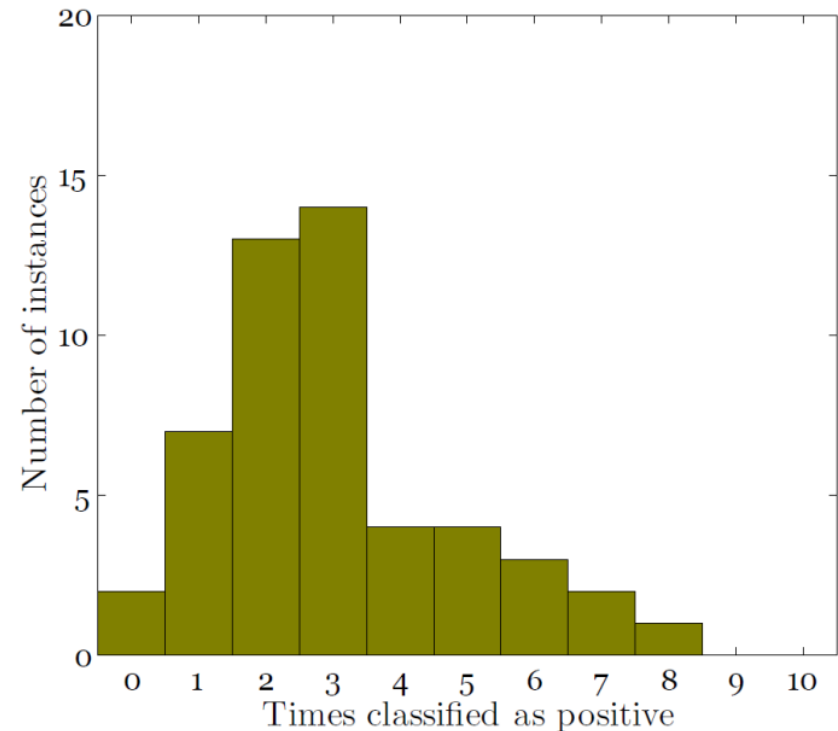
Fraction of agreement on positives

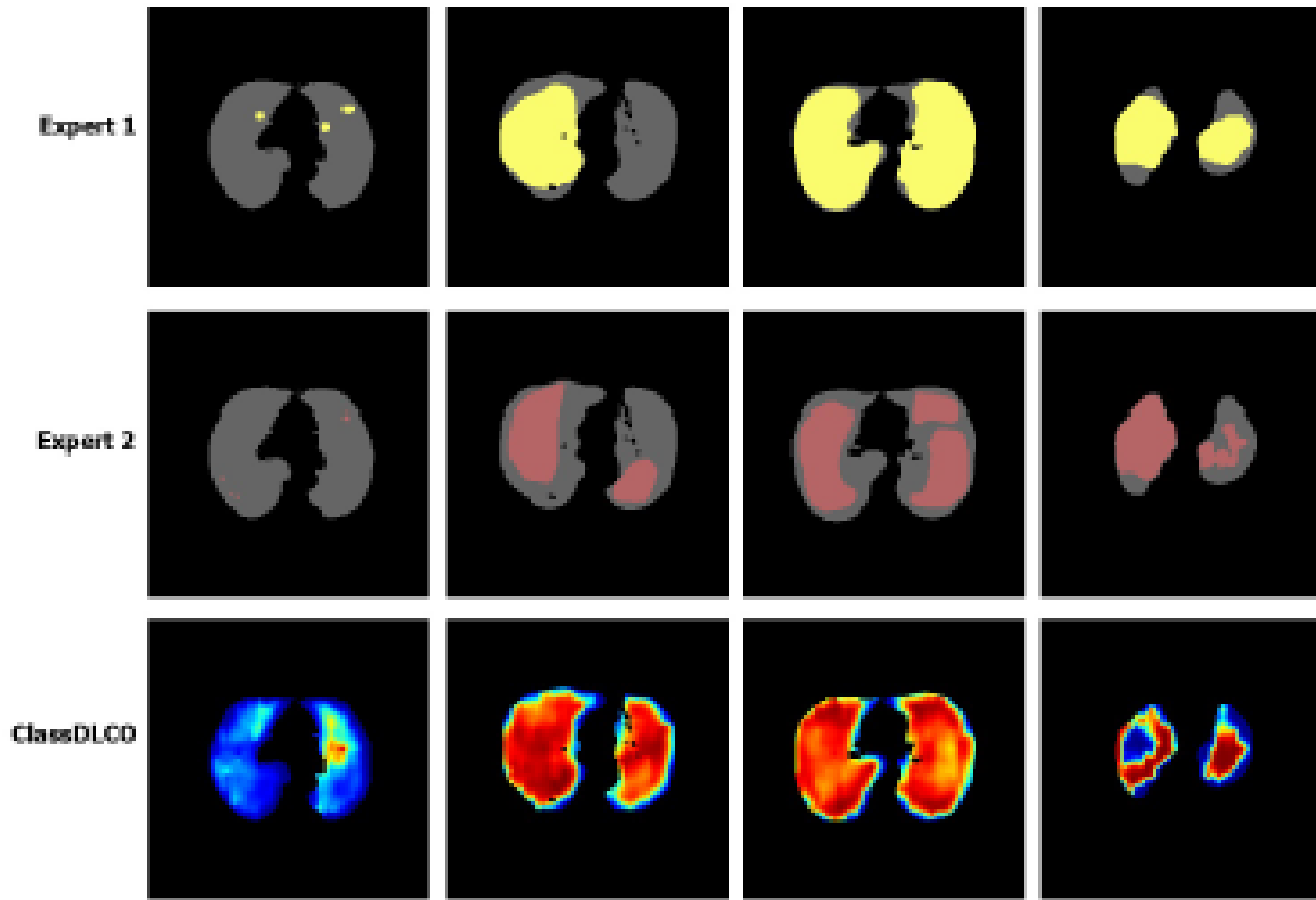
$S_+$

1.0	0.4	0.5	0.3	0.6	0.5	0.6	0.2	0.4	0.5
0.4	1.0	0.4	0.2	0.2	0.6	0.3	0.1	0.3	0.4
0.5	0.4	1.0	0.2	0.4	0.5	0.4	0.1	0.3	0.6
0.3	0.2	0.2	1.0	0.3	0.2	0.4	0.8	0.6	0.2
0.6	0.2	0.4	0.3	1.0	0.3	0.6	0.3	0.3	0.4
0.5	0.6	0.5	0.2	0.3	1.0	0.4	0.1	0.3	0.4
0.6	0.3	0.4	0.4	0.6	0.4	1.0	0.3	0.4	0.4
0.2	0.1	0.1	0.8	0.3	0.1	0.3	1.0	0.5	0.1
0.4	0.3	0.3	0.6	0.3	0.3	0.4	0.5	1.0	0.3
0.5	0.4	0.6	0.2	0.4	0.4	0.4	0.1	0.3	1.0

Any patches always positive?

Bag 157

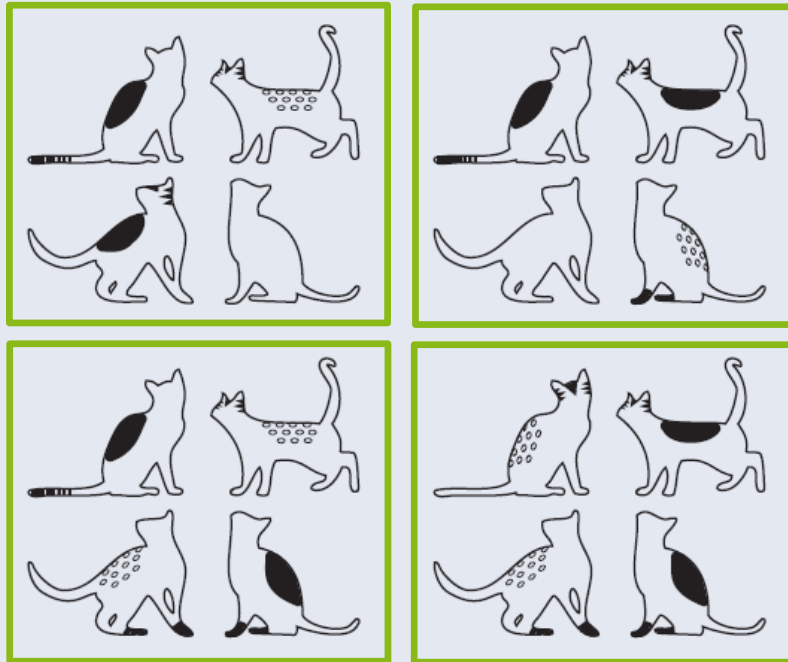




Pena, I. P., Cheplygina, V., Paschaloudi, S., Vuust, M., Carl, J., Weinreich, U. M., ... & de Bruijne, M. (2018). Automatic emphysema detection using weakly labeled HRCT lung images. *PloS one*, 13(10), e0205397.

## Dissimilarity-Based Multiple Instance Learning

What is different about the groups of cats on the front cover, from the groups on the back cover? If you can answer this question, you are probably also able to categorize another, previously unseen to you, group of cats. This thesis is about different applications where similar puzzles may occur, and how some machine learning algorithms approach such problems.



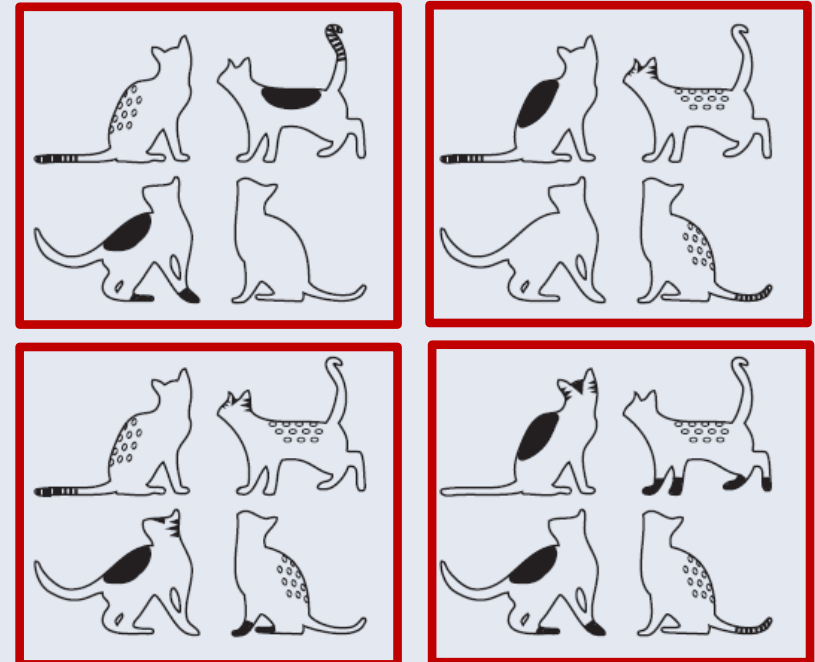
ISBN 978-94-6295-192-1

# Dissimilarity-Based Multiple Instance Learning

Veronika Cheplygina

Dissimilarity-Based Multiple Instance Learning

Veronika Cheplygina





## Solution 2: Transfer learning

Not learning “from scratch”

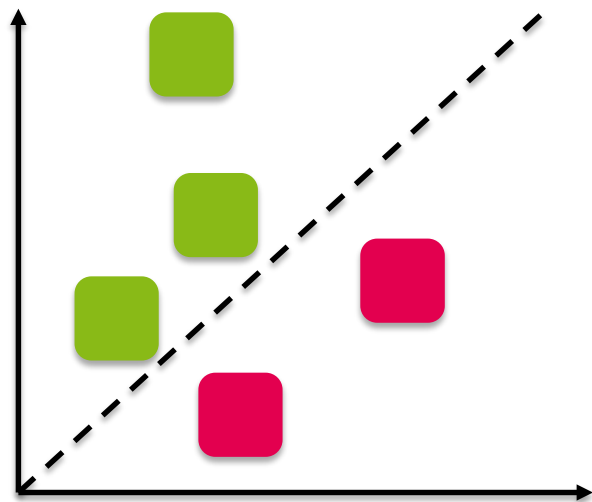


## Use other similar datasets

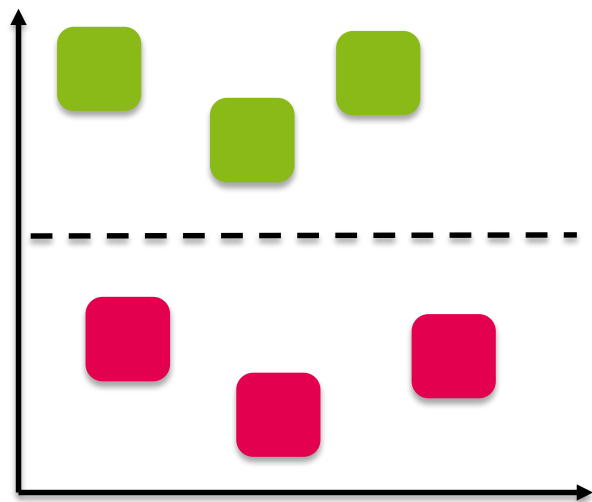
Dataset	Subjects	Age	GOLD (1/2/3/4)	Smoking (c/f/n)	Scanner	Resolution (mm)
DLCST	300 + 300 -	59 [50, 71] 57 [49, 69]	69/28/2/0	77/23/0 74/26/0	Philips 16 rows Mx 8000	0.72×0.72×1 to 0.78×0.78×1
COPDGene1	74 + 46 -	64 [45, 80] 59 [45, 78]	21/18/19/16	17/57/0 23/20/3	Siemens Definition	0.65×0.65×0.75
COPDGene2	42 + 25 -	65 [45, 78] 60 [47, 78]	9/13/7/13	12/30/0 9/11/5	Siemens Definition AS+	0.65×0.65×0.75
Frederikshavn	8 + 8 -	66 [48, 77] 56 [25, 73]	1/3/3/1	1/7/0 1/2/5	Siemens Definition Flash	0.58×0.58×0.6

# Use data from similar datasets

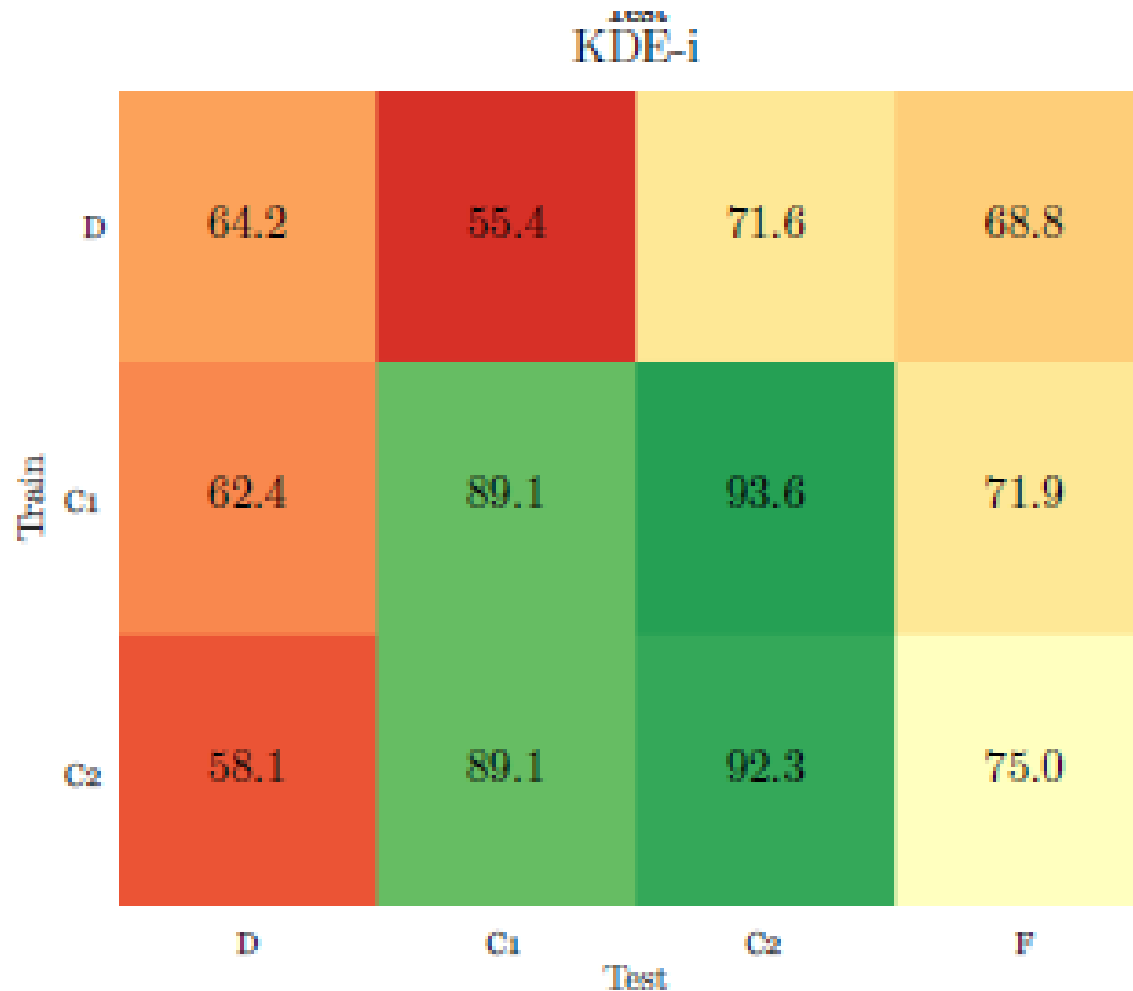
Source



Target



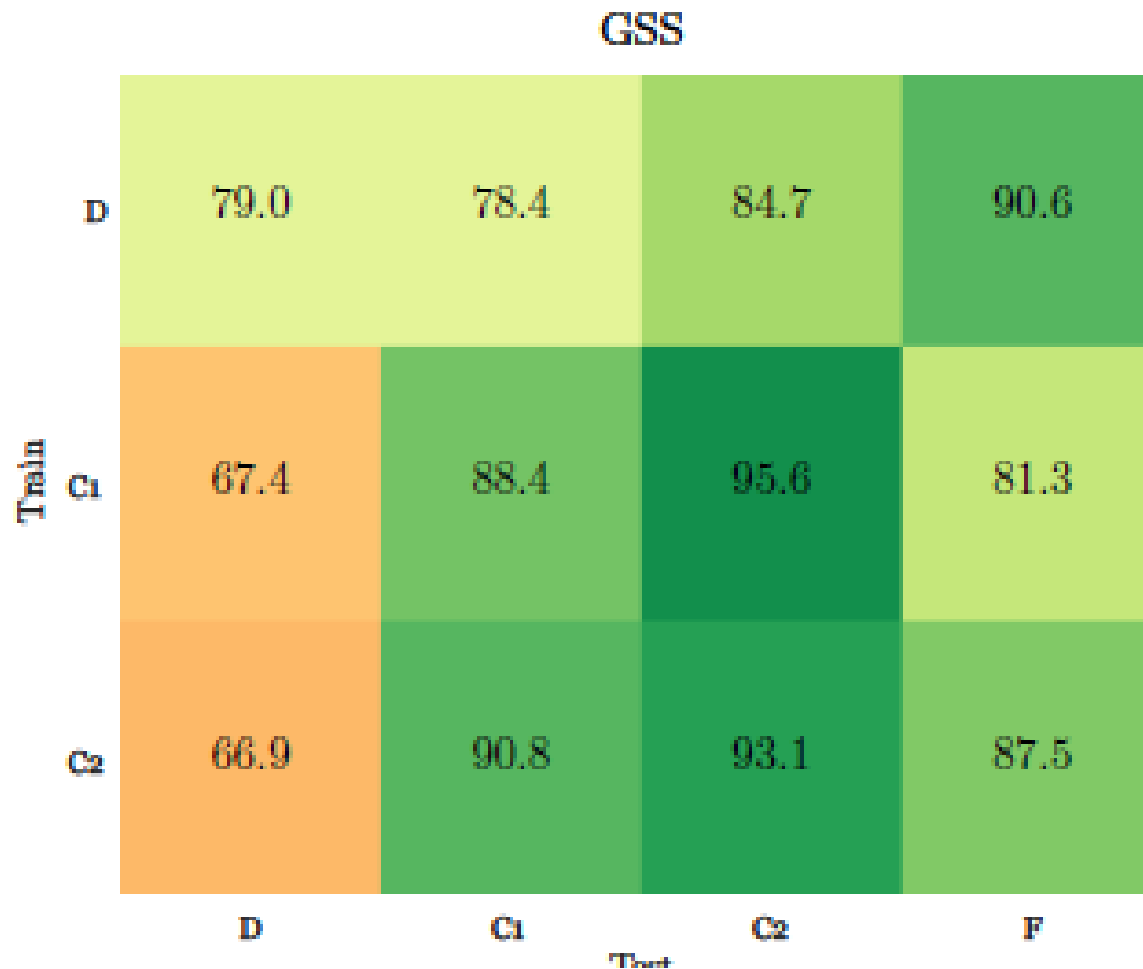
## Use other similar datasets – Intensity features



Cheplygina, V., Pena, I. P., Pedersen, J. H., Lynch, D. A., Sørensen, L., & de Bruijne, M. (2018). Transfer learning for multicenter classification of chronic obstructive pulmonary disease. *IEEE journal of biomedical and health informatics*, 22(5), 1486-1496.

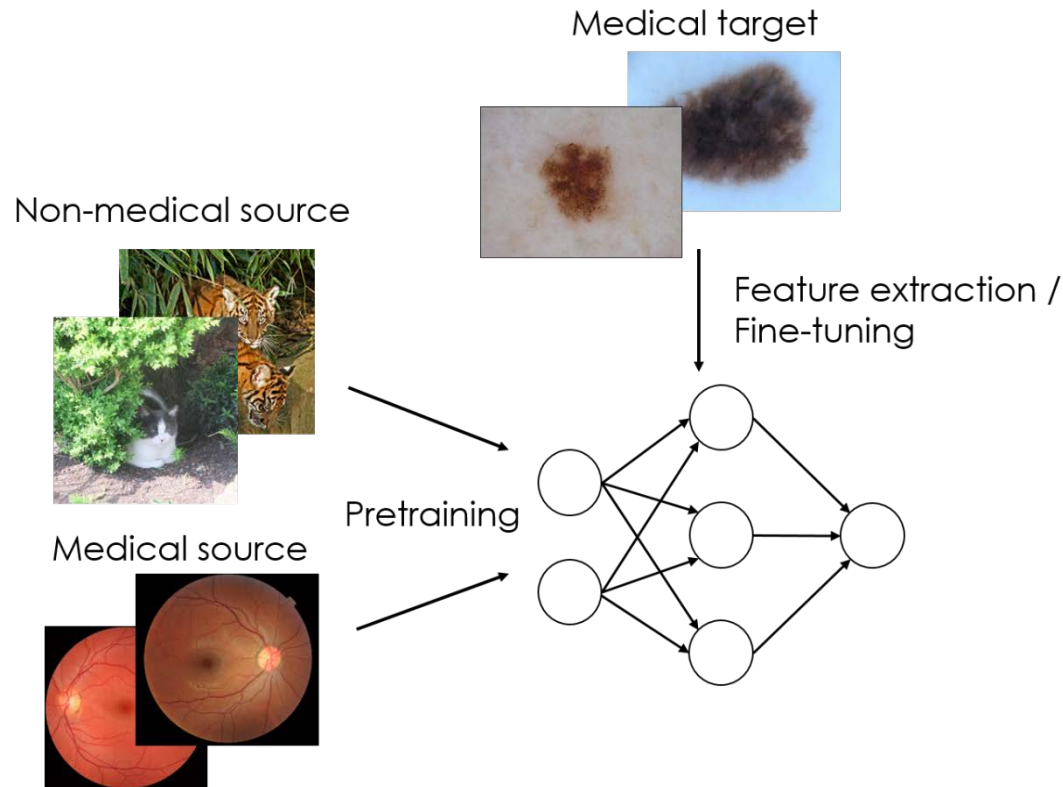


## Use other similar datasets – Texture features



Cheplygina, V., Pena, I. P., Pedersen, J. H., Lynch, D. A., Sørensen, L., & de Bruijne, M. (2018). Transfer learning for multicenter classification of chronic obstructive pulmonary disease. *IEEE journal of biomedical and health informatics*, 22(5), 1486-1496.

... or data that's entirely different



Cats or CAT scans: transfer learning from natural or medical image source datasets?

Cheplygina, V. (2019). Cats or CAT scans: transfer learning from natural or medical image source datasets?. *Current Opinion in Biomedical Engineering*. [URL](#)

# Non-medical or medical data for pretraining?

3 papers: non-medical is better

5 papers: medical is better

2 papers: no differences

3 papers: inconclusive, BUT more data not better

@drveronikach



<http://www.veronikach.com>

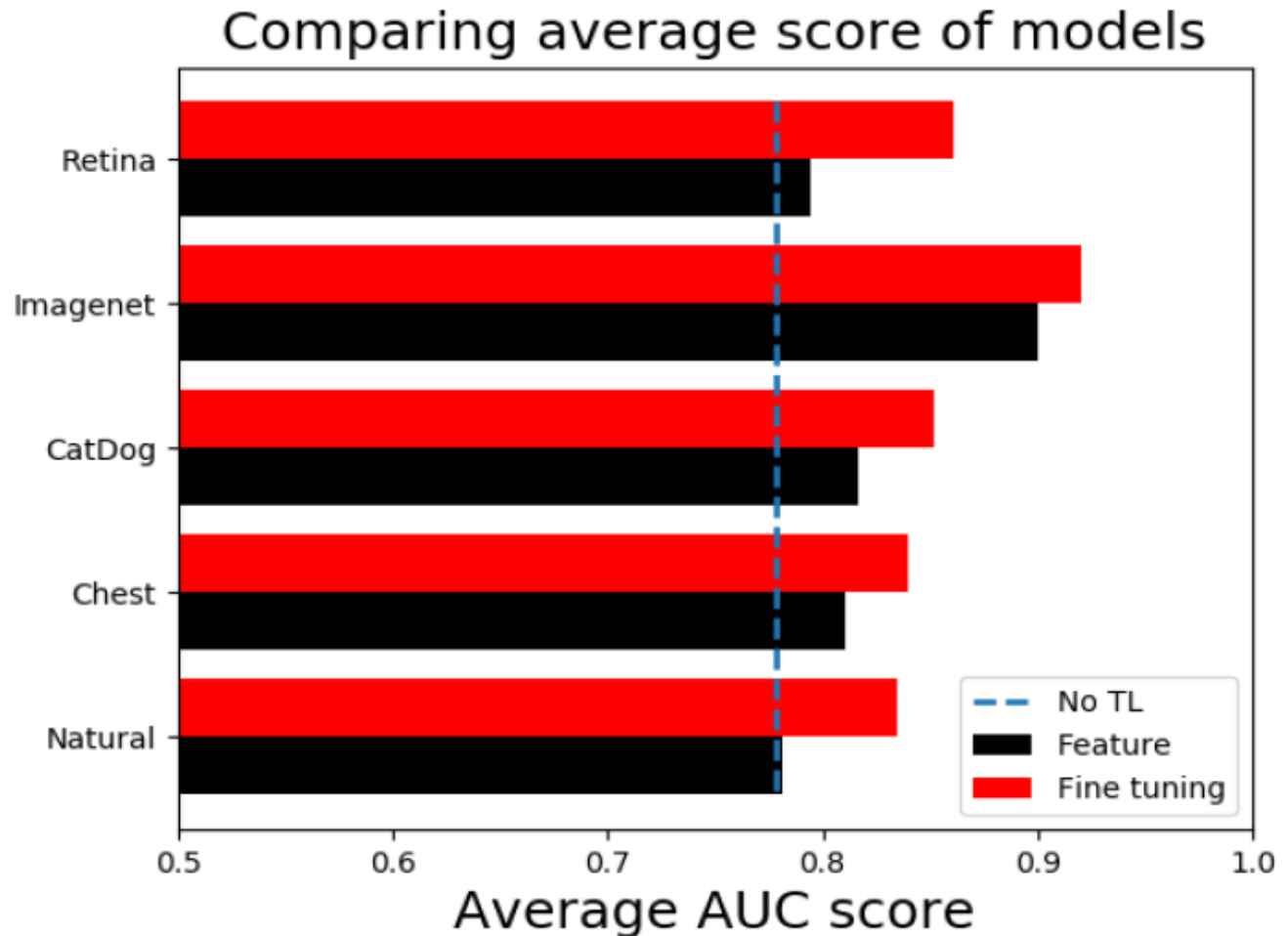


# Systematic comparison?

ImageNet best as source data

BUT

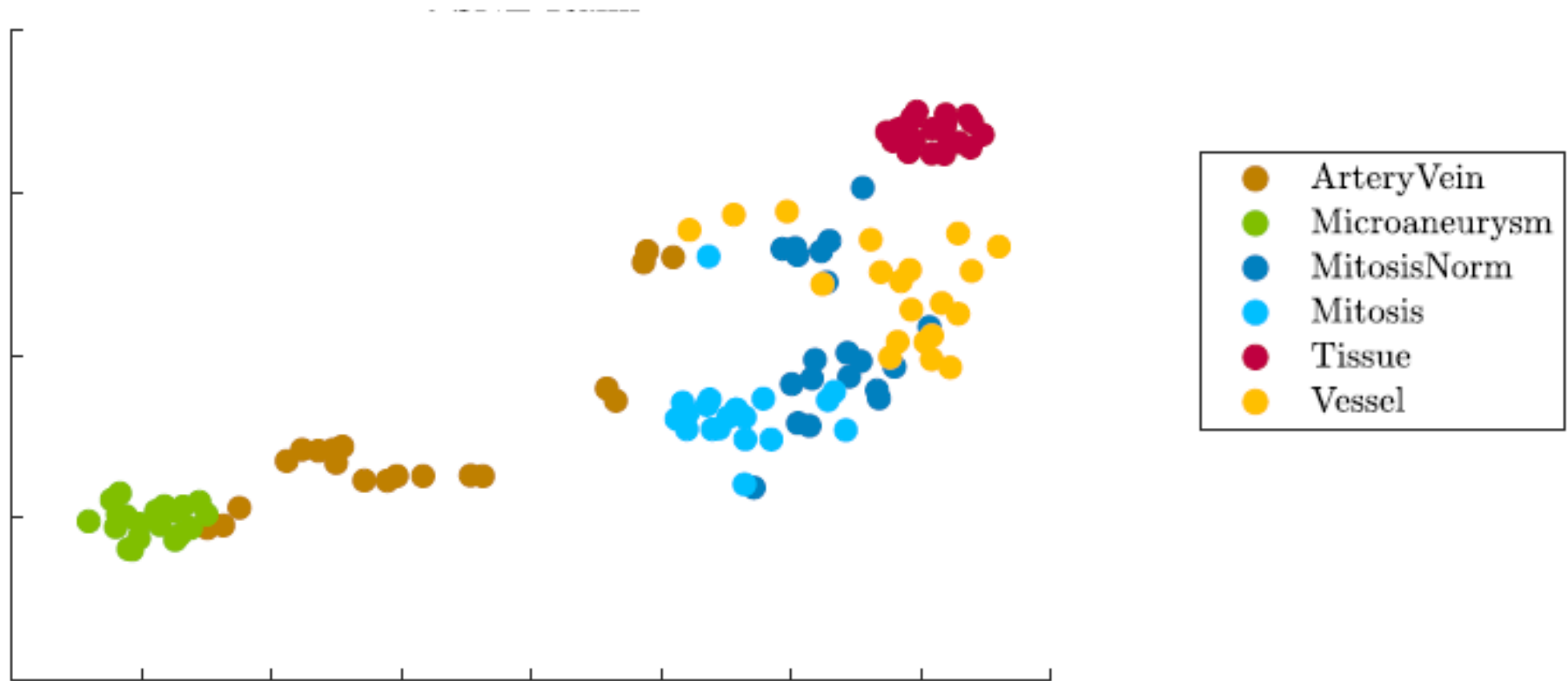
is ImageNet is a much bigger dataset



Work by Floris Fok



# Meta-learning: how to quantify similarity of data?



Cheplygina, V., Moeskops, P., Veta, M., Bozorg, B. D., & Pluim, J. (2017). Exploring the similarity of medical imaging classification problems. In *Large-Scale Annotation of Biomedical Data and Expert Label Synthesis (MICCAI LABELS)* (pp. 59-66)

# Solution 2: Transfer learning

Not learning “from scratch”



Cheplygina, V., de Bruijne, M., & Pluim, J. P. W. (2019). Not-so-supervised: a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis. *Medical Image Analysis*. [URL](#)

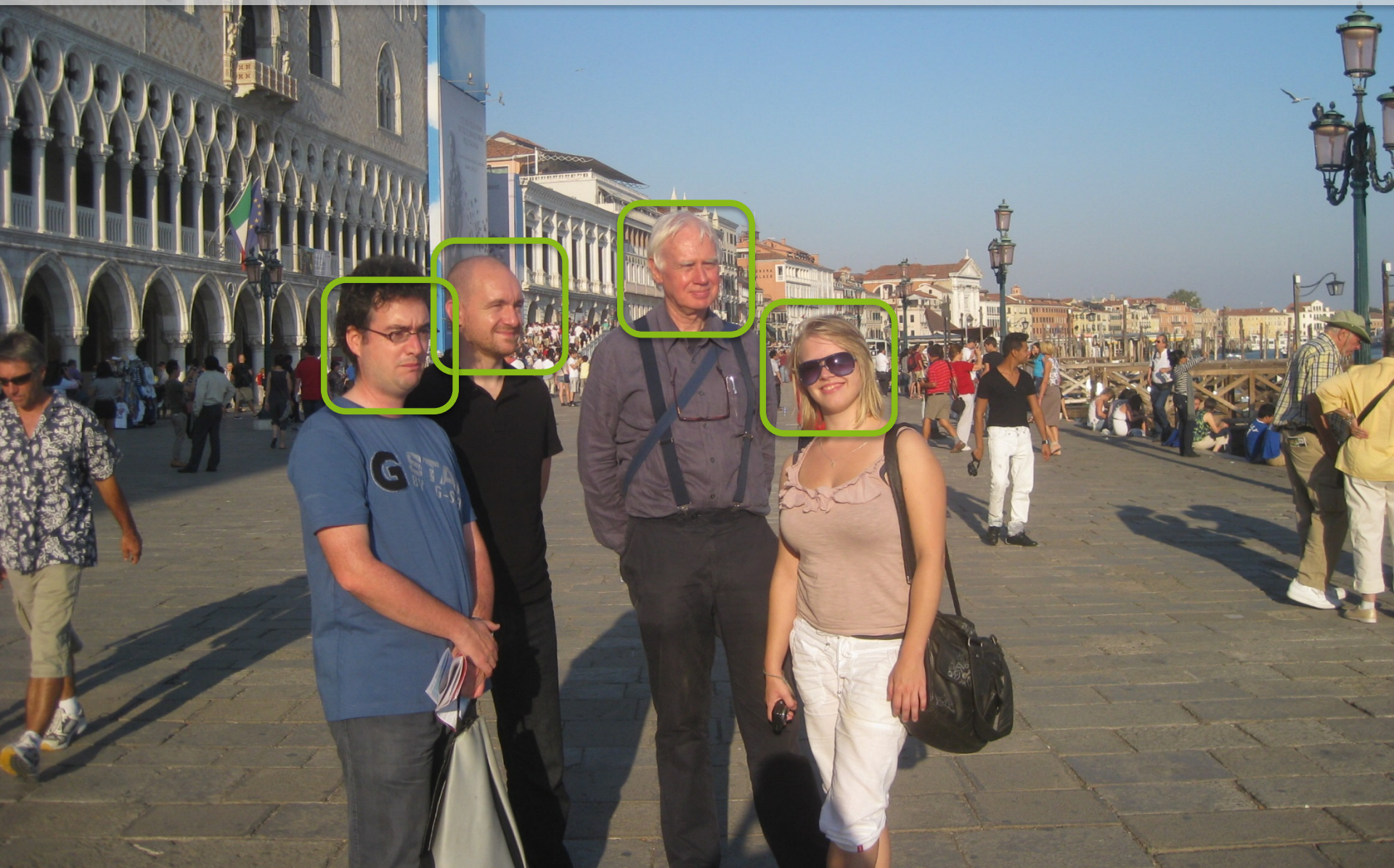


## Solution 3: Crowdsourcing





You do it all the time!





## 2009: ImageNet



[Main](#)
[Instructions](#)
[Unsure? Look up in Wikipedia](#)
[Google](#)
[\[ Additional input \]](#)
[No good photos? Have expertise? comments? Click here!](#)

**First time workers please click here for instructions.**

Click on the photos that contain the object or depict the concept of : **delta**: **a low triangular area of alluvial deposits where a river divides before entering a larger body of water; "the Mississippi River delta"; "the Nile delta"** .(PLEASE READ DEFINITION CAREFULLY)  
Pick as many as possible. **PHOTOS ONLY, NO PAINTINGS, DRAWINGS, etc.** It's OK to have other objects, multiple instances, occlusion or text in the image.  
Do not use **back or forward button of your browser**. **OCCASIONALLY THERE MIGHT BE ADULT OR DISTURBING CONTENT.**

Below are the photos you have selected FROM THIS PAGE ONLY ( they will be saved when you navigate to other pages ). Click to deselect.

what's this?

select all

deselect all

<

page 1

of 6

>

Submit

PREVIEW MODE. TO WORK ON THIS HIT, ACCEPT IT FIRST.

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on* (pp. 248-255). IEEE.

# Can Masses of Non-Experts Train Highly Accurate Image Classifiers?

## A Crowdsourcing Approach to Instrument Segmentation in Laparoscopic Images

Lena Maier-Hein<sup>1,\*,\*\*</sup>, Sven Mersmann<sup>1</sup>, Daniel Kondermann<sup>2</sup>,  
Sebastian Bodenstedt<sup>3</sup>, Alexandro Sanchez<sup>2</sup>, Christian Stock<sup>4</sup>,  
Hannes Gotz Kenngott<sup>5</sup>, Mathias Eisenmann<sup>3</sup>, and Stefanie Speidel<sup>3</sup>

The screenshot displays the Amazon Mechanical Turk interface for a HIT titled "Save lives by adjusting the outline of a tool!". The interface includes a header with the Amazon Mechanical Turk logo and navigation links. The main content area shows the HIT details, including the requester's name, the reward (\$0.12 per HIT), and the duration (15 minutes). The HIT description states: "The polygon in the bottom left corner contains a medical tool. Improve this polygon by adding and moving points until its shape perfectly matches the tool. Controls: • Zoom using your mouse wheel or the zoom slider. • Double click to add or remove points on the polygon. • Click and drag to stretch or move the polygon, individual points or to pan the image. Once you are finished, use the form above to send us the results." The interface also shows a "Submit Results" button and a "Change contrast" slider. The bottom part of the screenshot shows two side-by-side laparoscopic images with a red polygon overlaid on a medical tool. The left image is labeled "(Bad)" and the right image is labeled "(Good)". The zoom level is set to 1.

amazonmechanicalturk  
Artificial Intelligence

Your Account | HITs | Qualifications | 462,821 HITs available now

All HITs | HITs Available To You | HITs Assigned To You

Find HITs:  containing  Start pay at least \$  0.00

Timer: 00:00:29 of 15 minutes

Finished with this HIT? Let someone else do it?

☐ Automatically accept the next HIT

Save lives by adjusting the outline of medical tool! [VIDEO]

Requester: [sven.mersmann](#)

Reward: \$0.12 per HIT

HITs Available: 235

Duration: 15 minutes

Qualifications Required: HIT approval rate (%) is not less than 95

Save lives by adjusting the outline of a tool!

You can leave your feedback here (Optional)

The polygon in the bottom left corner contains a medical tool. Improve this polygon by adding and moving points until its shape perfectly matches the tool. Controls:

- Zoom using your mouse wheel or the zoom slider.
- Double click to add or remove points on the polygon.
- Click and drag to stretch or move the polygon, individual points or to pan the image.

Once you are finished, use the form above to send us the results.

Change contrast:  Status:

Transform this:

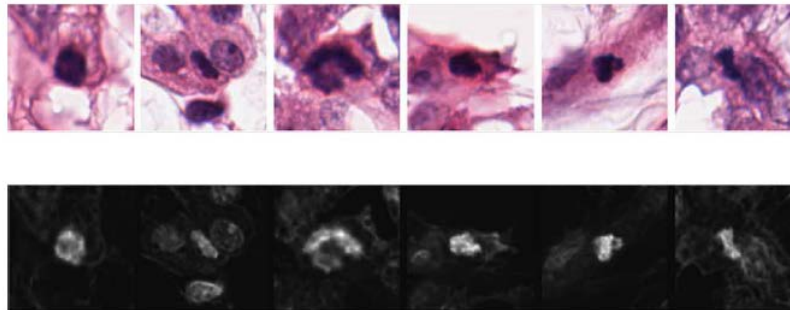
(Bad) (Good)

Zoom: 1

# AggNet: Deep Learning From Crowds for Mitosis Detection in Breast Cancer Histology Images

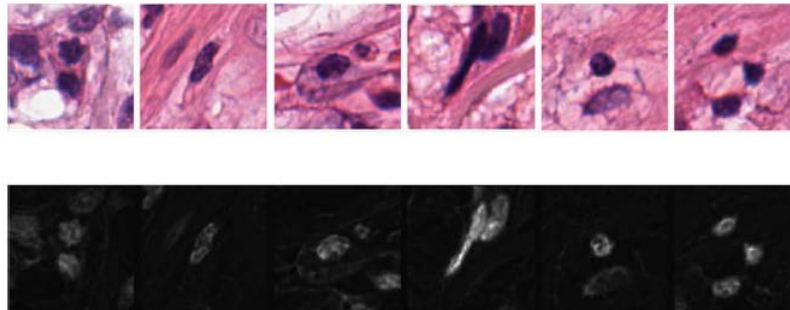
Shadi Albarqouni\*, *Student Member; IEEE*, Christoph Baur, Felix Achilles, *Student Member; IEEE*, Vasileios Belagiannis, *Student Member; IEEE*, Stefanie Demirci, and Nassir Navab, *Member; IEEE*

Mitosis:



The second row shows the corresponding so called "blueRatio" representation of the mitotic figures. Note how they have very bright spots!

Non-Mitosis

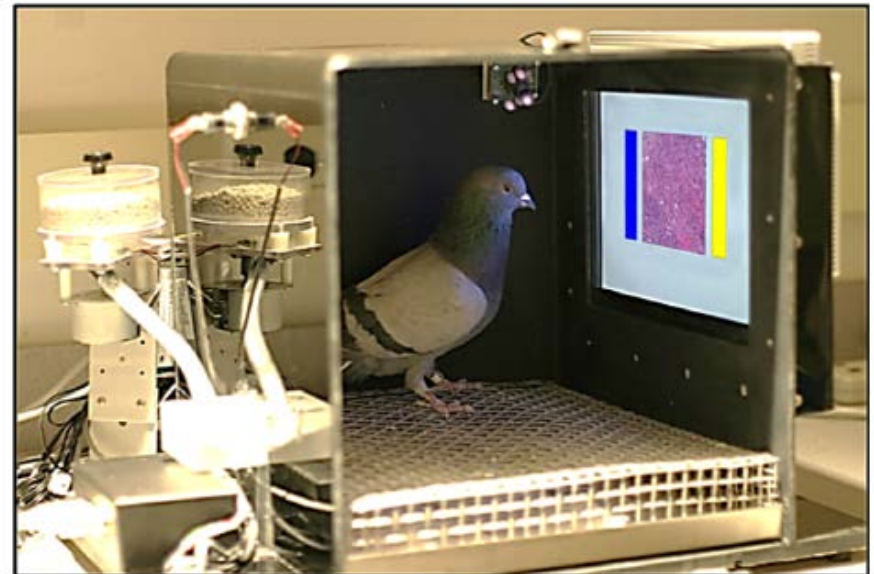
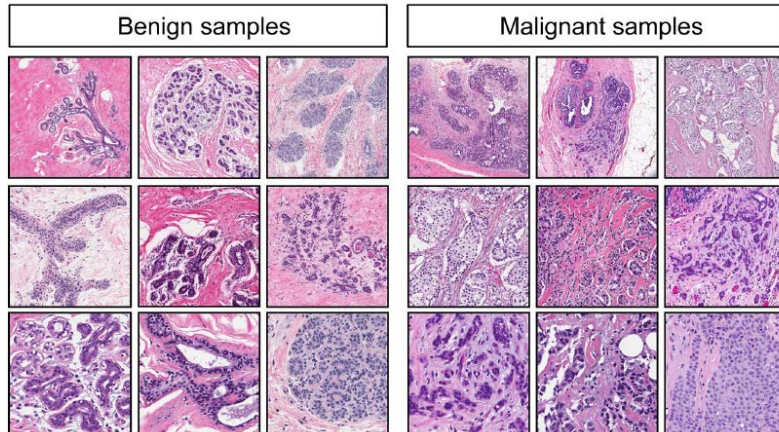


The second row shows the corresponding so called "blueRatio" representation of the non-mitotic figures. Note how they do not have such bright spots as the mitotic blue ratio representations!



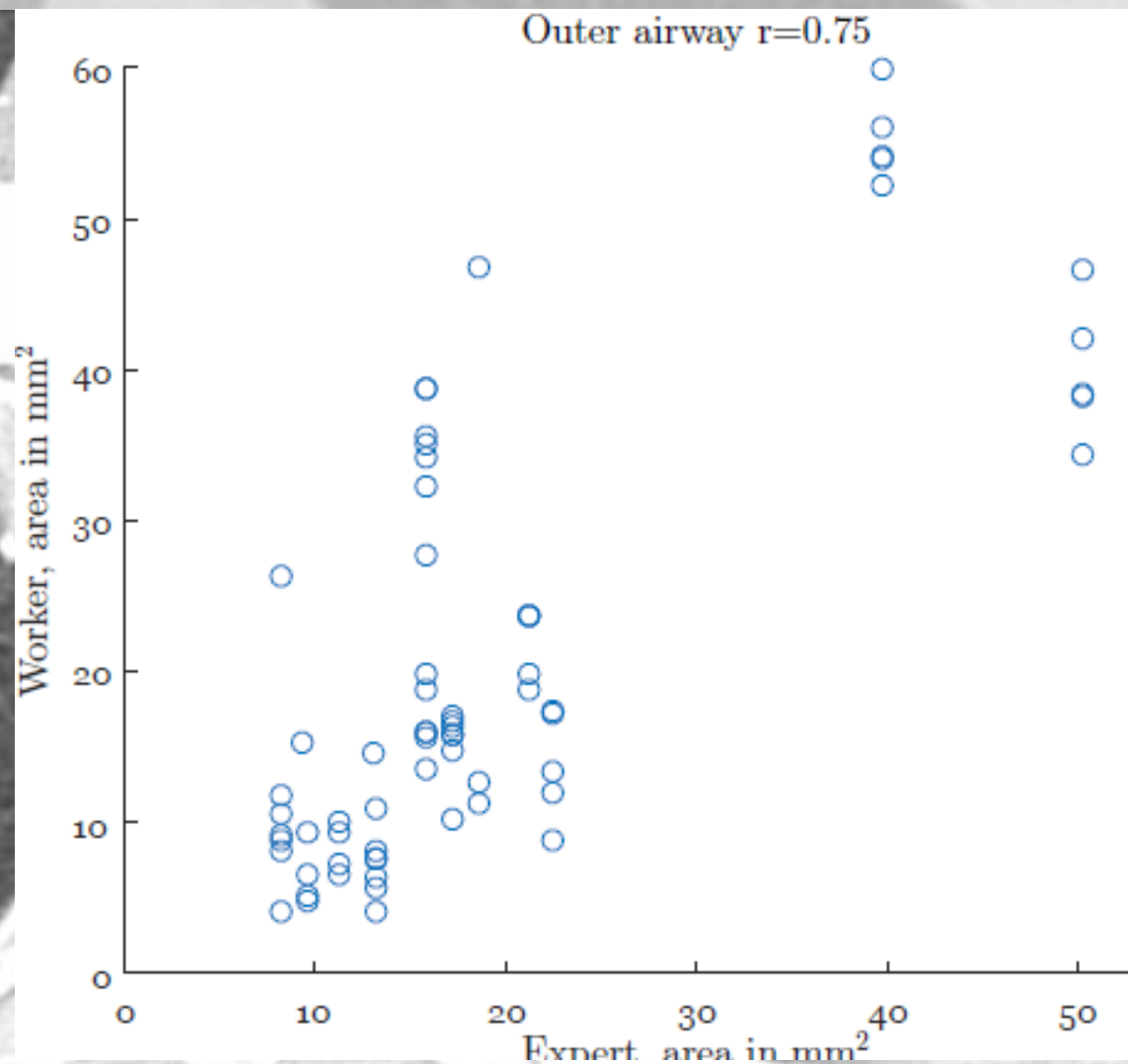
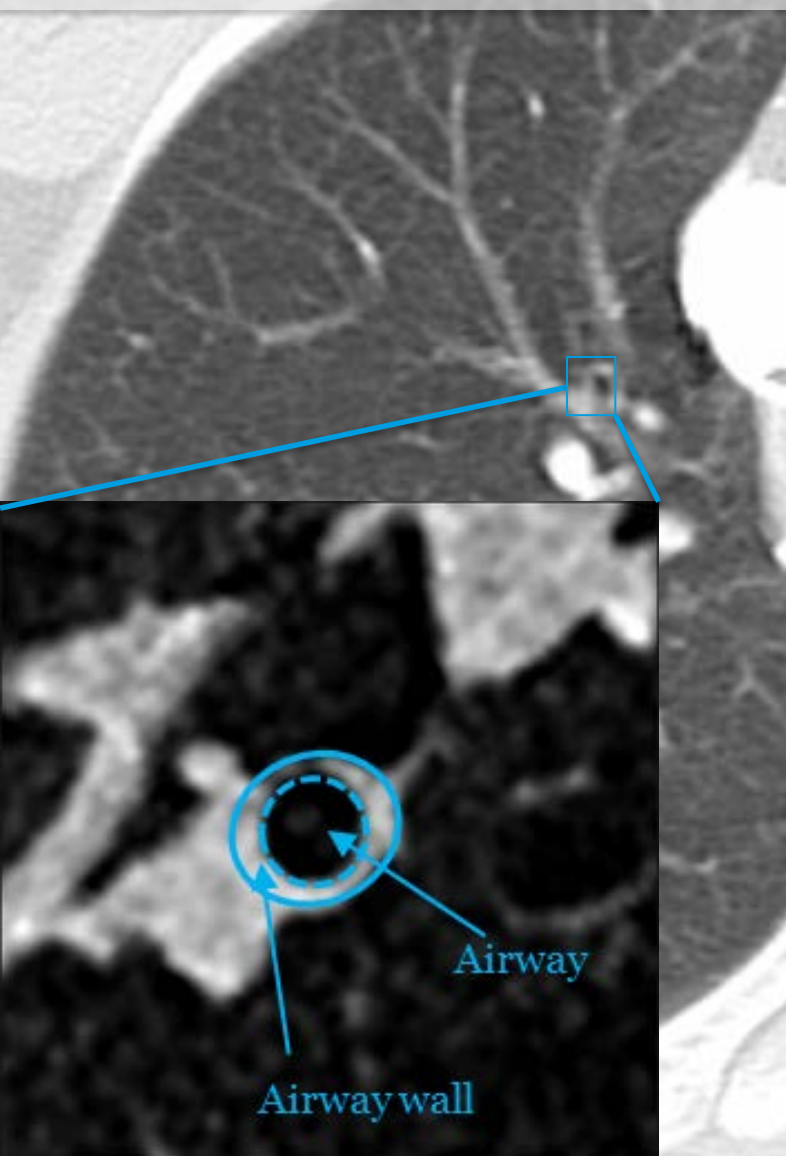
# Pigeons (*Columba livia*) as Trainable Observers of Pathology and Radiology Breast Cancer Images

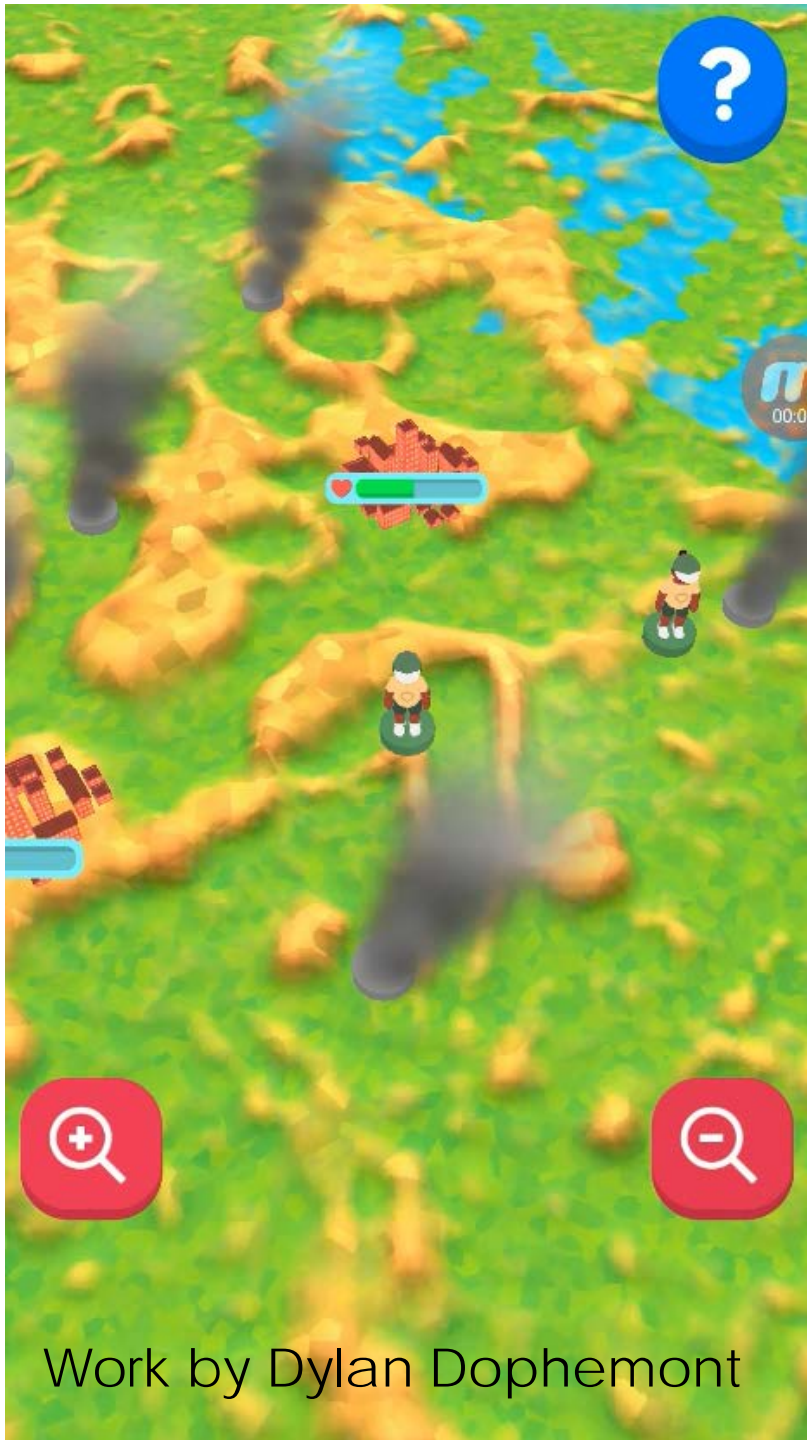
Richard M. Levenson<sup>1\*</sup>, Elizabeth A. Krupinski<sup>3</sup>, Victor M. Navarro<sup>2</sup>, Edward A. Wasserman<sup>2\*</sup>





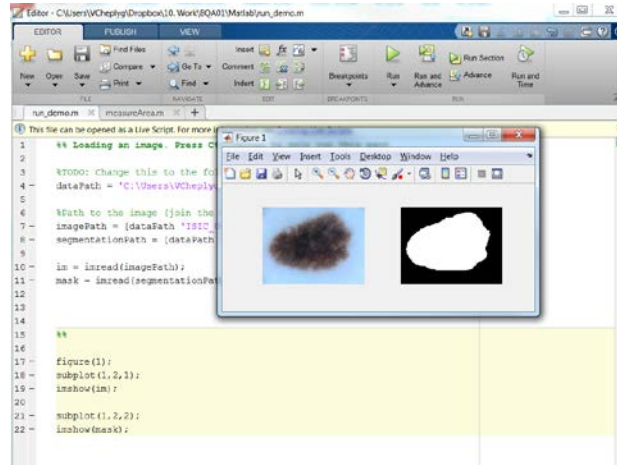
# Crowdsourcing airway annotations





Work by Dylan Dophemont

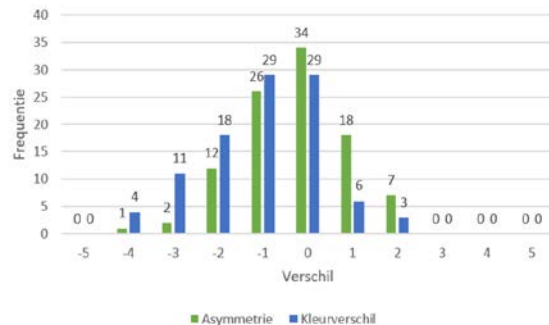
## 1. Measure features with algorithms



## 2. Measure features yourself

	A	B	C	D
1	ID	Asymmetry_7_1	Color_7_1	Border_7_1
2	ISIC_0000549	2	4	1
3	ISIC_0000550	1	3	1
4	ISIC_0000551	2	2	1
5	ISIC_0000552	1	4	1
6	ISIC_0000554	2	3	1
7	ISIC_0000555	2	3	1
8	ISIC_0001100	2	5	1
9	ISIC_0001102	2	5	1
10	ISIC_0001103	1	5	1
11	ISIC_0001105	0	2	1
12	ISIC_0001118	2	5	1
13	ISIC_0001119	2	3	1
14	ISIC_0001126	2	2	1
15	ISIC_0001128	1	3	1
16	ISIC_0001131	1	5	1
17	ISIC_0001133	1	5	1
18	ISIC_0001134	2	3	1
19	ISIC_0001140	2	2	1
20	ISIC_0009923	1	2	1
21	ISIC_0009925	2	2	1
22	ISIC_0009929	1	2	1
23	ISIC_0009930	1	2	1
24	ISIC_0009931	1	3	1
25	ISIC_0009932	2	3	1
26	ISIC_0009933	1	2	1
27	ISIC_0009935	1	3	1
28	ISIC_0009936	1	2	0

## 3. Evaluate

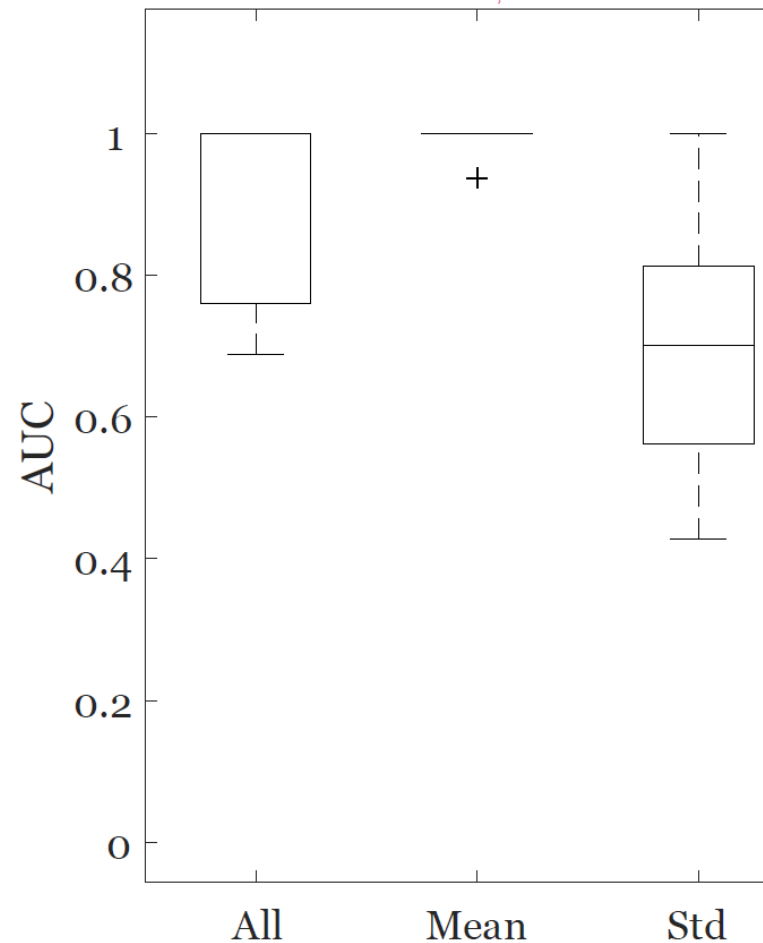


Grafiek 1: De frequenties van de verschilwaardes tussen de metingen in Matlab en de metingen op het oog

**Crowdsourcing!**

## 100 annotated images

- 5 features annotated by 6 people = 30 features
- Predict healthy vs melanoma **without** images
- Mean of annotators best, but “disagreement” (standard deviation) also informative



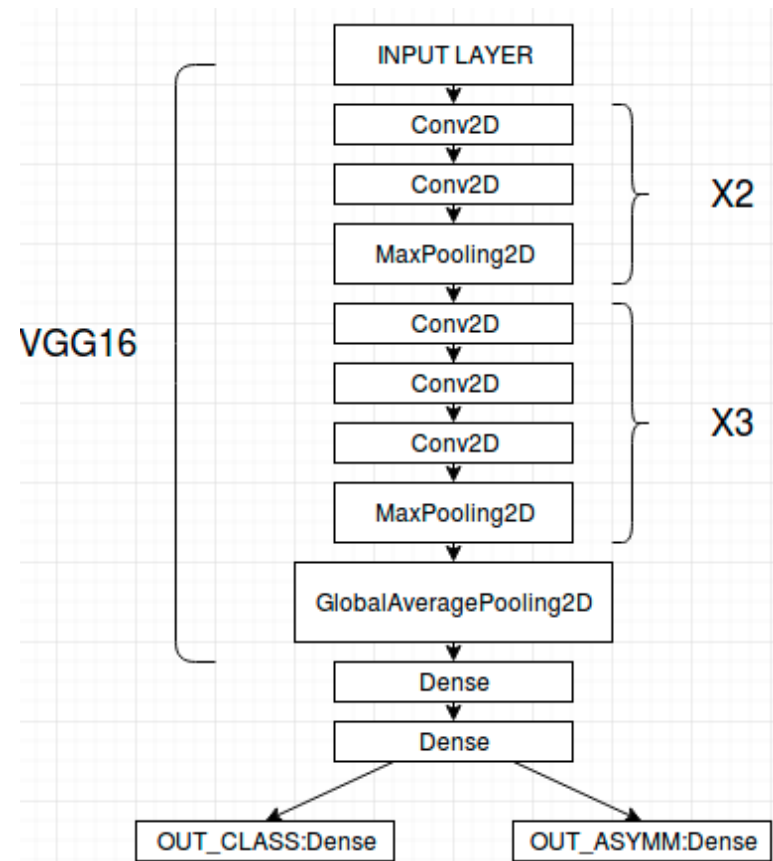
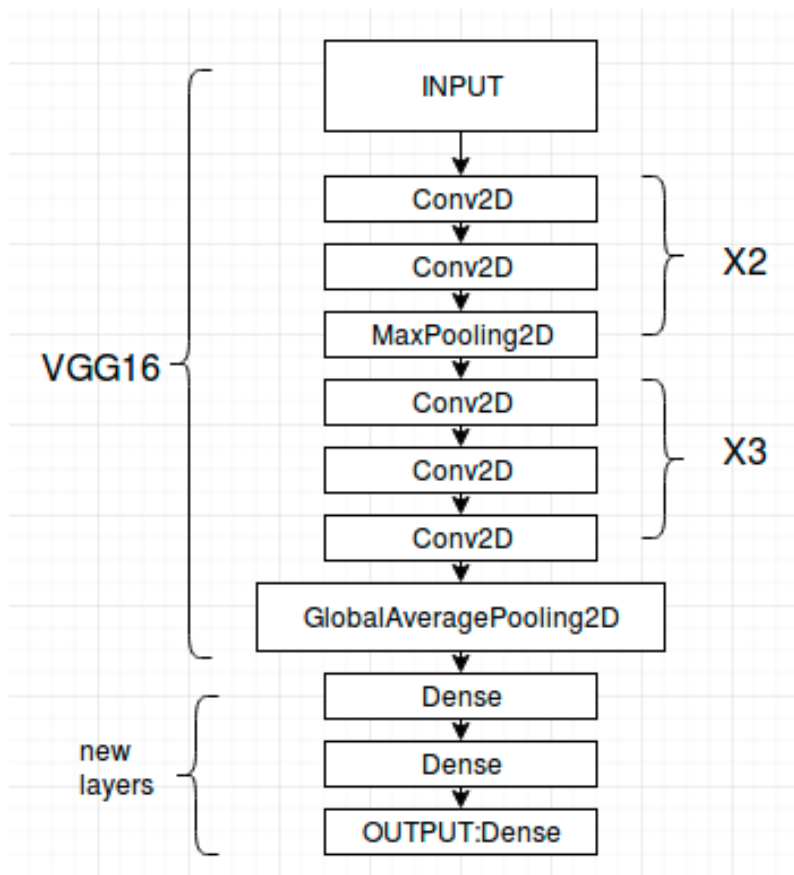
Cheplygina, V., & Pluim, J. P. W. (2018). Crowd disagreement about medical images is informative. In *Intravascular Imaging and Computer Assisted Stenting and Large-Scale Annotation of Biomedical Data and Expert Label Synthesis* (pp. 105-111).



# Work by Elif Kubra Contar

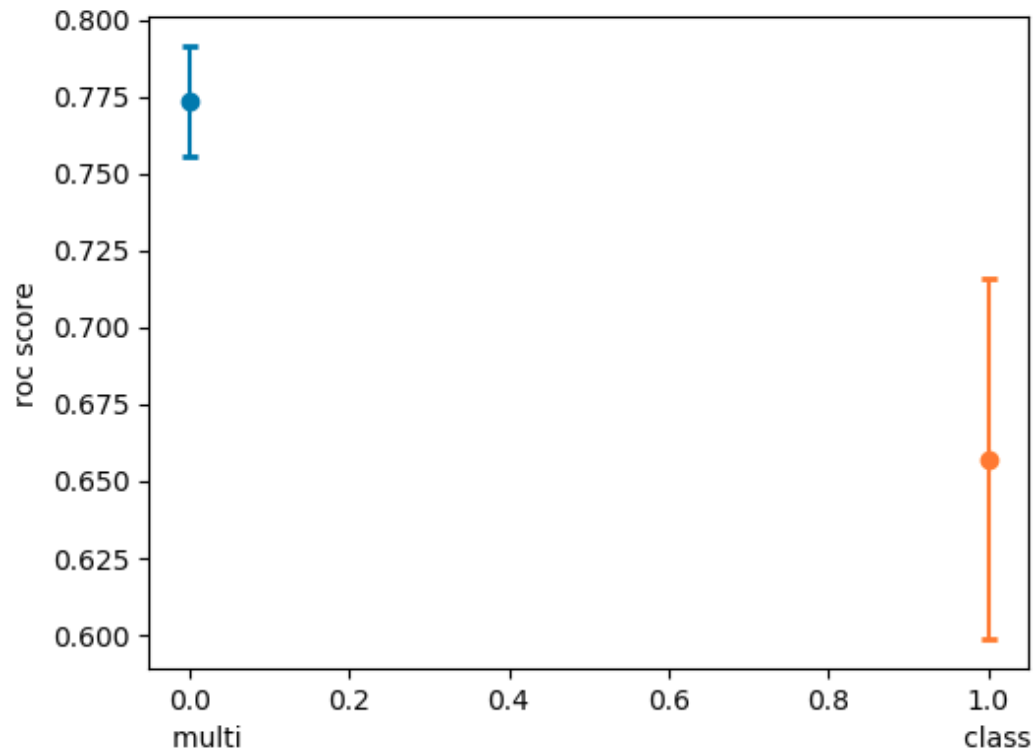
Same network

- Single-task with class label
- Multi-task with class label and asymmetry



# Work by Elif Kubra Contar

Multi-task network outperforms single-task network



@drveronikach



<http://www.veronikach.com>





# A Survey of Crowdsourcing in Medical Image Analysis

Silas Ørting<sup>1</sup>✉, Andrew Doyle<sup>2,\*</sup>, Matthias Hirth<sup>3,\*</sup>, Arno van Hilten<sup>4,\*</sup>, Oana Inel<sup>5,7,\*</sup>, Christopher R. Madan<sup>6,\*</sup>, Panagiotis Mavridis<sup>7,\*</sup>, Helen Spiers<sup>8,9,\*</sup>, and Veronika Cheplygina<sup>10</sup>✉

<sup>1</sup> University of Copenhagen, Copenhagen, Denmark

<sup>2</sup> McGill Centre for Integrative Neuroscience, Montreal, Canada

<sup>3</sup> Technische Universität Ilmenau, Ilmenau, Germany

<sup>4</sup> Erasmus Medical Center, Rotterdam, The Netherlands

<sup>5</sup> Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

<sup>6</sup> University of Nottingham, Nottingham, United Kingdom

<sup>7</sup> Delft University of Technology, Delft, The Netherlands

<sup>8</sup> University of Oxford, Oxford, United Kingdom

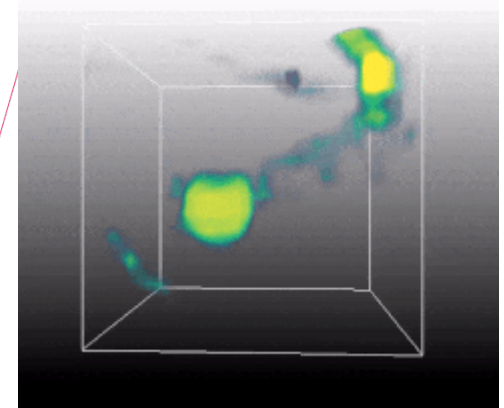
<sup>9</sup> Zooniverse, University of Oxford, Oxford

<sup>10</sup> Eindhoven University of Technology, Eindhoven, The Netherlands

Ørting, S., Doyle, A., van Hilten, A., Hirth, M., Inel, O., Madan, C. R., Mavridis, P., ... & Cheplygina, V. (2019). A survey of crowdsourcing in medical image analysis. *arXiv preprint arXiv:1902.09159*

## Next

- Other features / annotators (1400+ images)
- Crowdsourcing for lung nodules (3D)
- Evaluate the evaluation





# Not-so-supervised academics



# 2011-2014 PhD

“Publish papers”

- Science vs prestige

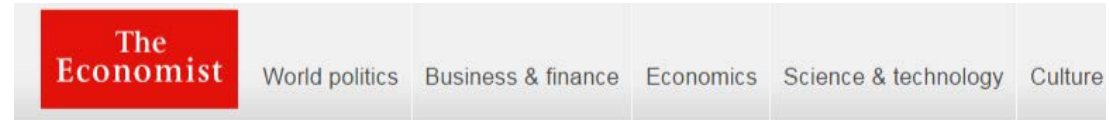
“Good for your CV”

- Experience vs time



# Maybe academic career?

- “You have to go abroad”
- Not good enough?
- Mentors!



Doctoral degrees

## The disposable academic

Why doing a PhD is often a waste of time



PROFESSIONAL JOBS SUMMITS RANKINGS

## Too many PhDs, not enough tenured positions

European study reveals stress suffered by doctoral holders over insecurity of academic careers



2015





# 2015 – 2016

- “Publish, develop own research + get funding for it”
- Social media
  - Impact, community



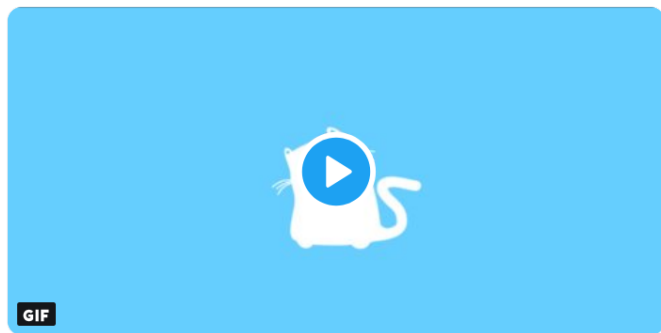
# 2017



**Dr Veronika Cheplygina**  
@vcheplygina



Excited to announce that I'll join the Medical Image Analysis group @TUeindhoven as assistant professor in Feb 2017!  
[veronikach.com/news/](http://veronikach.com/news/)



5:23 PM - 26 Nov 2016

2 Retweets 32 Likes



13 2 32



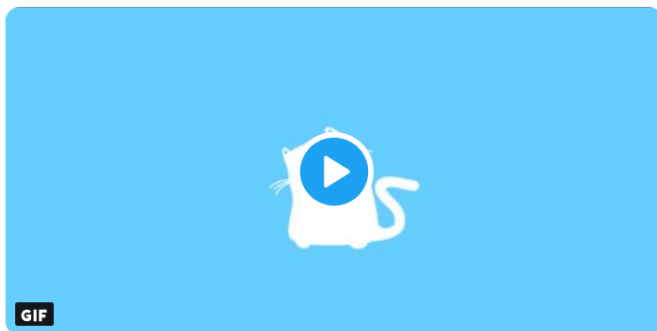
# 2017



**Dr Veronika Cheplygina**  
@vcheplygina



Excited to announce that I'll join the Medical Image Analysis group @TUeindhoven as assistant professor in Feb 2017!  
[veronikach.com/news/](http://veronikach.com/news/)



5:23 PM - 26 Nov 2016

2 Retweets 32 Likes



13 2 32

## 1. Excitement

I get to do research and teach and learn from others for the next 5 years! How amazing have so many ideas, I can't wait!

## 2. Relief

I get to have a job for 5 years and don't have to apply for positions for like, a very long time. I started looking for my next position halfway through my postdoc, which was a job in itself that did not reflect well on my postdoc project. A few things were not really going well for me in the end, so the news about the position couldn't have come at a better time.

## 3. Fear

I worry they will discover I'm an impostor and they should have hired somebody else. I try to reassure myself by thinking that if I'm an impostor and they are the the real deal, they should have figured out that I was one already. But I also worry about just being able to handle it all.

## 4. Guilt

As many other researchers are forced out of academia, I feel guilty for "surviving" while others are not. I have a "good, but not excellent CV" (citing reviews on some of my rejected grant applications) and I have to deal with hundreds of rejections – I applied to four jobs, interviewed for three, and was offered one. Sure, I worked hard, but I think luck and privilege played a big role.

## 5. Hope

I get to be one step closer to maybe one day being able to change things, just a little bit.

# 2017+ Not without challenges



<https://pixabay.com/en/mountain-climbing-mountaineer-802099/>



# 2017+ Find support



<https://unsplash.com/photos/iuqmGmst5Po>

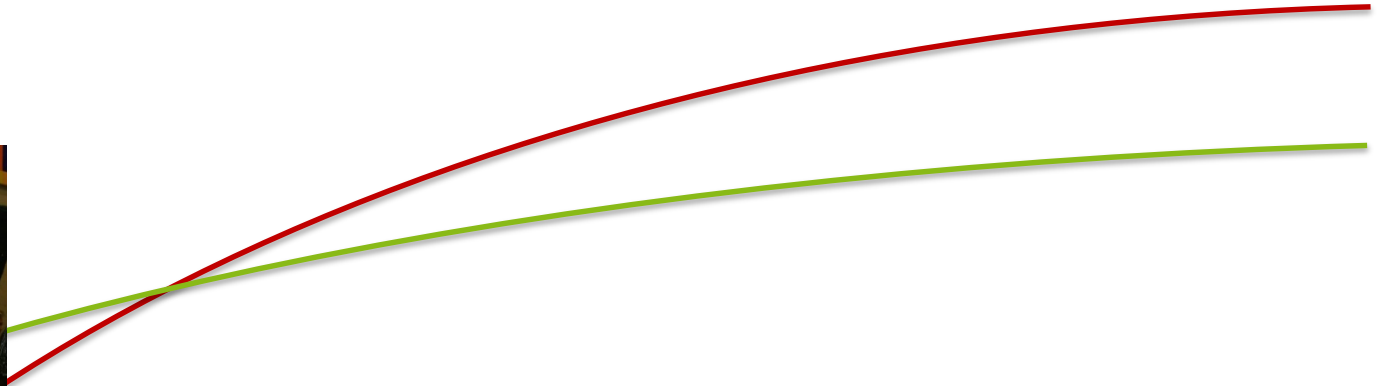
# Academia as supervised learning?

- Input = CV at time  $t$
- Output = Success / failure at  $t+1$
- Successes at  $t+1$  define “decision tree”

## But CVs $\neq$ true data distribution

- Input space is much larger (Shadow CV)
- Output space is much larger (Impact, being happy)
- Noisy labels, many unlabelled inputs
- Overfitting!

# Learning curve





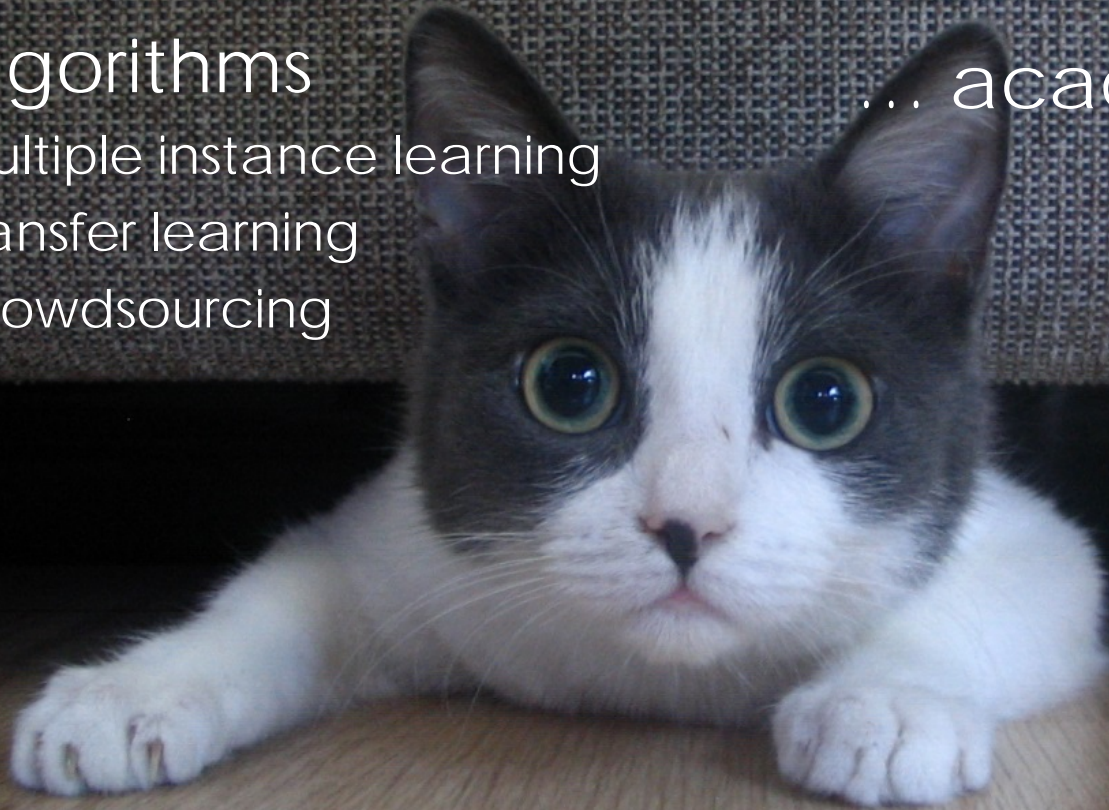
Not-so-supervised  
learning of

... algorithms

- Multiple instance learning
- Transfer learning
- Crowdsourcing

Not-so-supervised  
learning of

... academics



@drveronikach



<http://www.veronikach.com>







@drveronikach

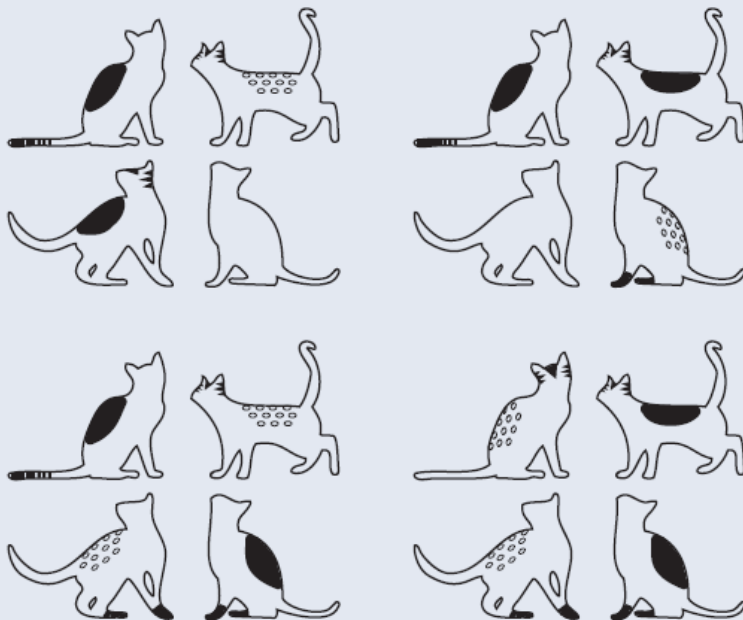


<http://www.veronikach.com>



## Dissimilarity-Based Multiple Instance Learning

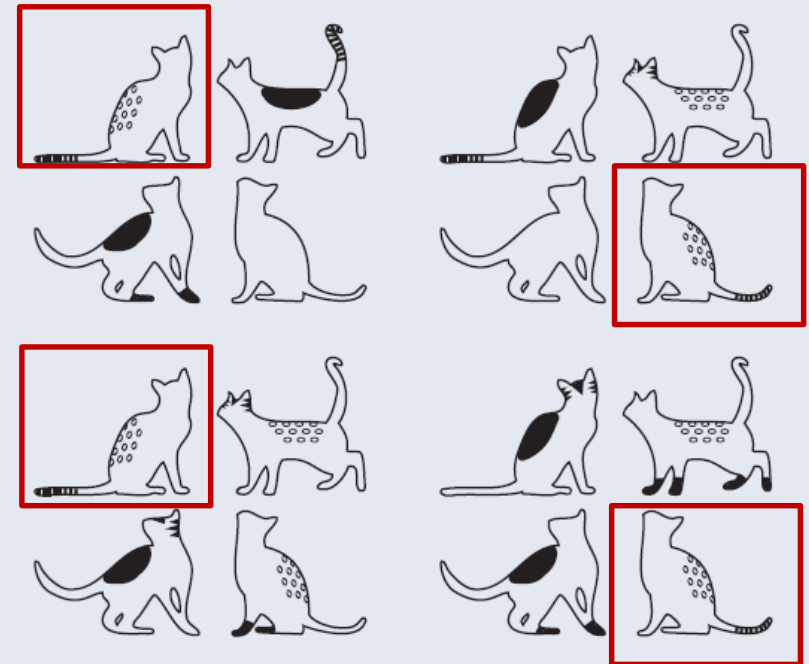
What is different about the groups of cats on the front cover, from the groups on the back cover? If you can answer this question, you are probably also able to categorize another, previously unseen to you, group of cats. This thesis is about different applications where similar puzzles may occur, and how some machine learning algorithms approach such problems.



ISBN 978-94-6295-192-1

## Dissimilarity-Based Multiple Instance Learning

Veronika Cheplygina



Dissimilarity-Based Multiple Instance Learning

Veronika Cheplygina