Technische Universiteit
**Eindhoven**
University of Technology

# Not-so-supervised learning of algorithms and academics

**Veronika Cheplygina**

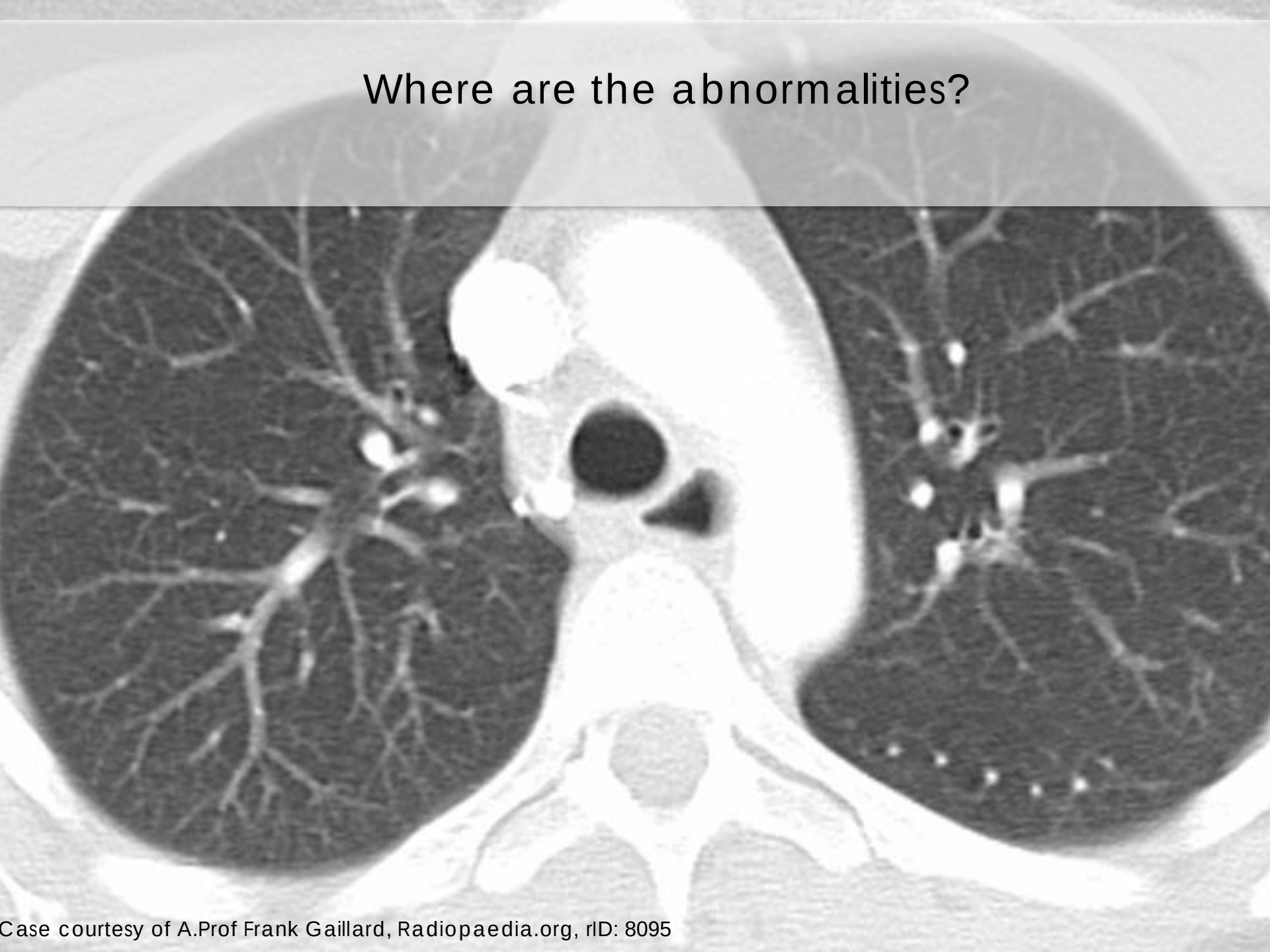
@vcheplygina

http://www.veronikach.com

Big data

Representative
& annotated data

# Where are the abnormalities?



Case courtesy of A.Prof Frank Gaillard, Radiopaedia.org, rID: 8095

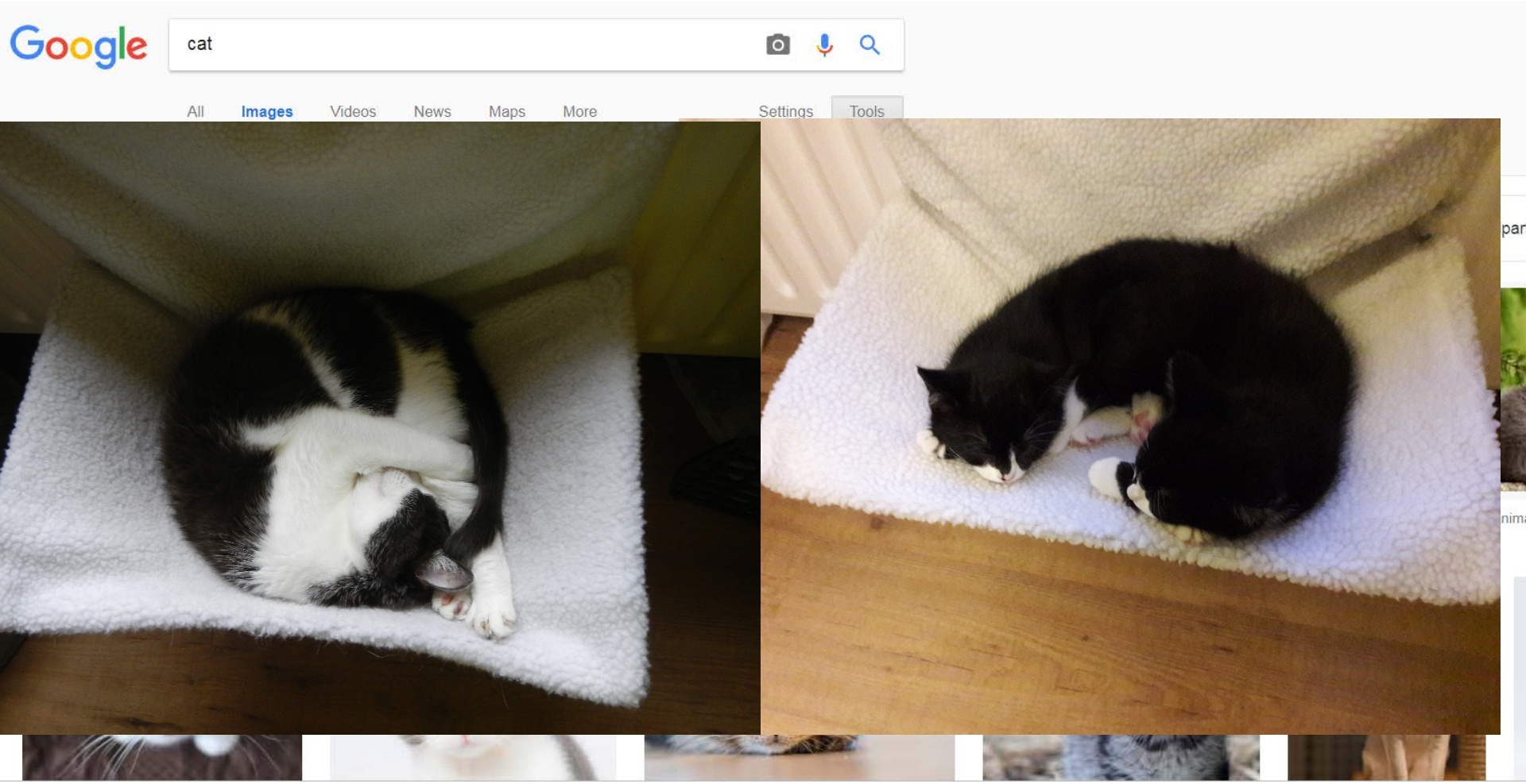
# Training data we want

Case courtesy of Radswiki, Radiopaedia.org, rID: 11384

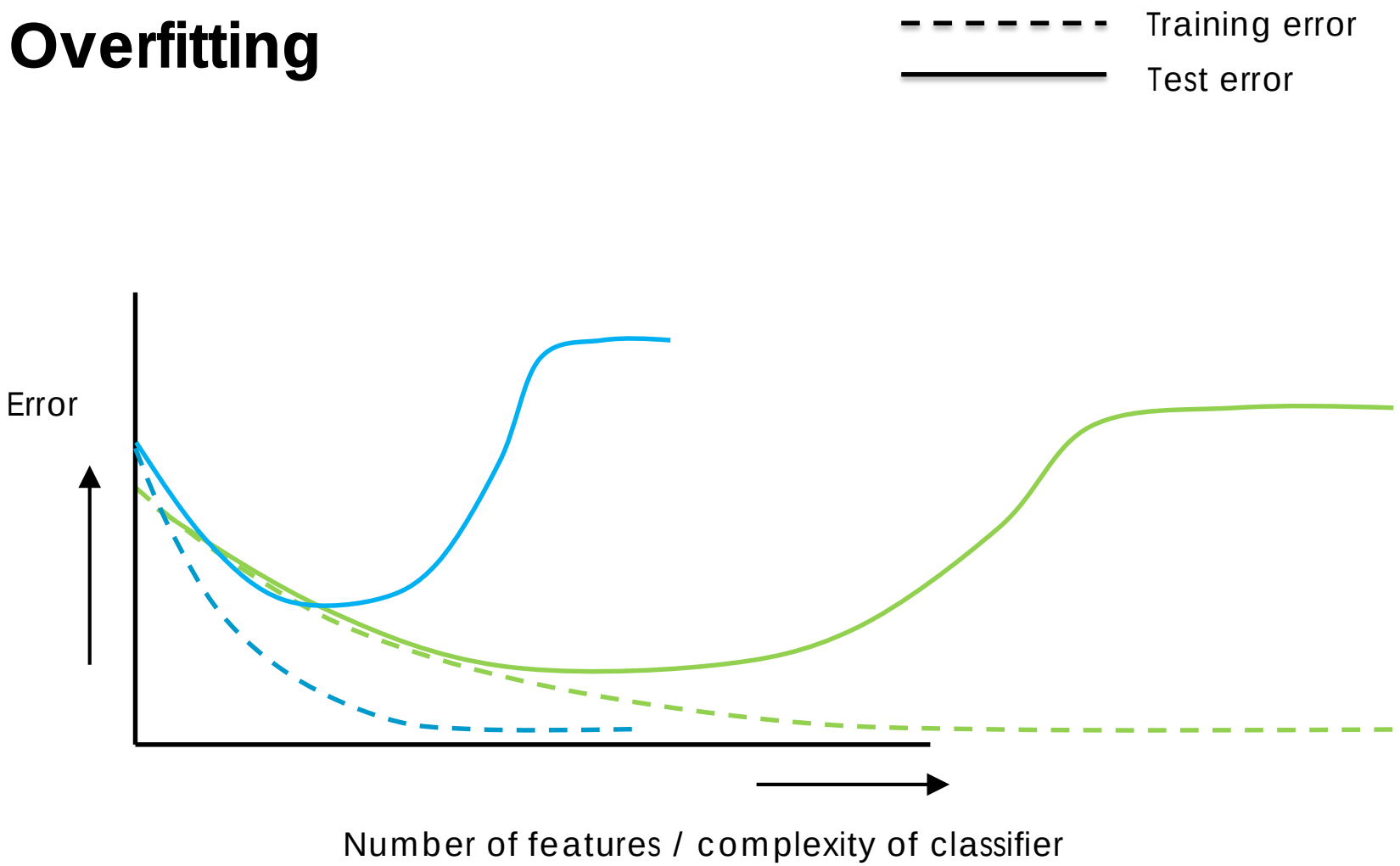
# Training data we get

Case courtesy of Radswiki, Radiopaedia.org, rID: 11384

# Training data not representative



# Overfitting

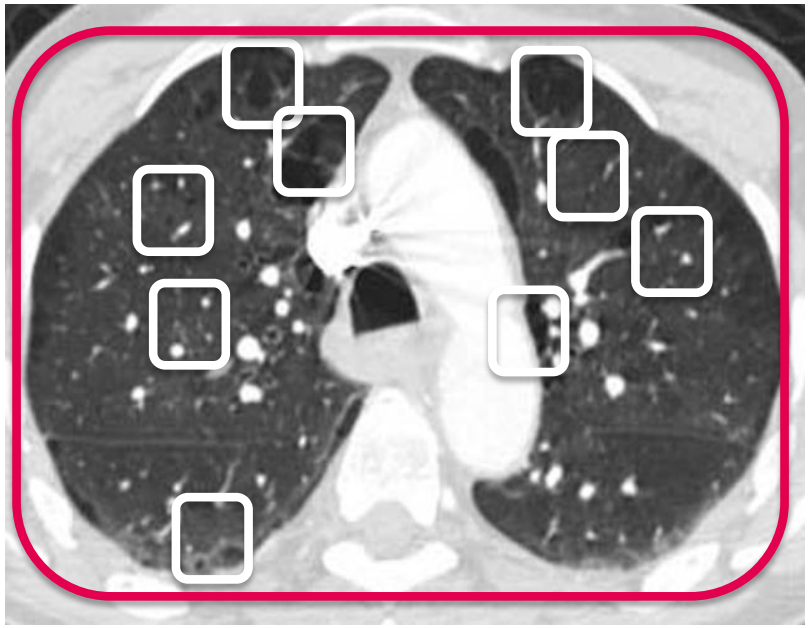


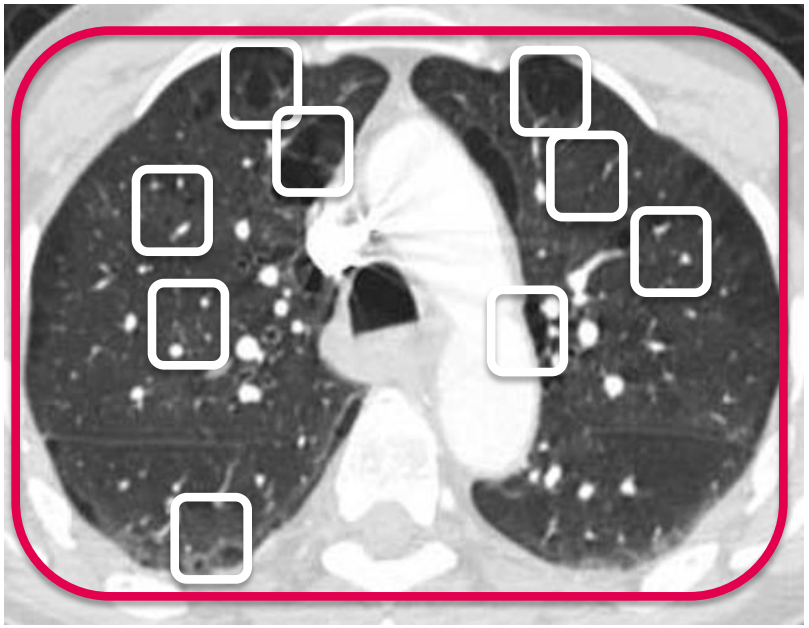
# Solution 1: Multiple instance learning

Learn with global information – Carolyn is in both pictures

# Solution 1: Multiple instance learning

Learn with global information



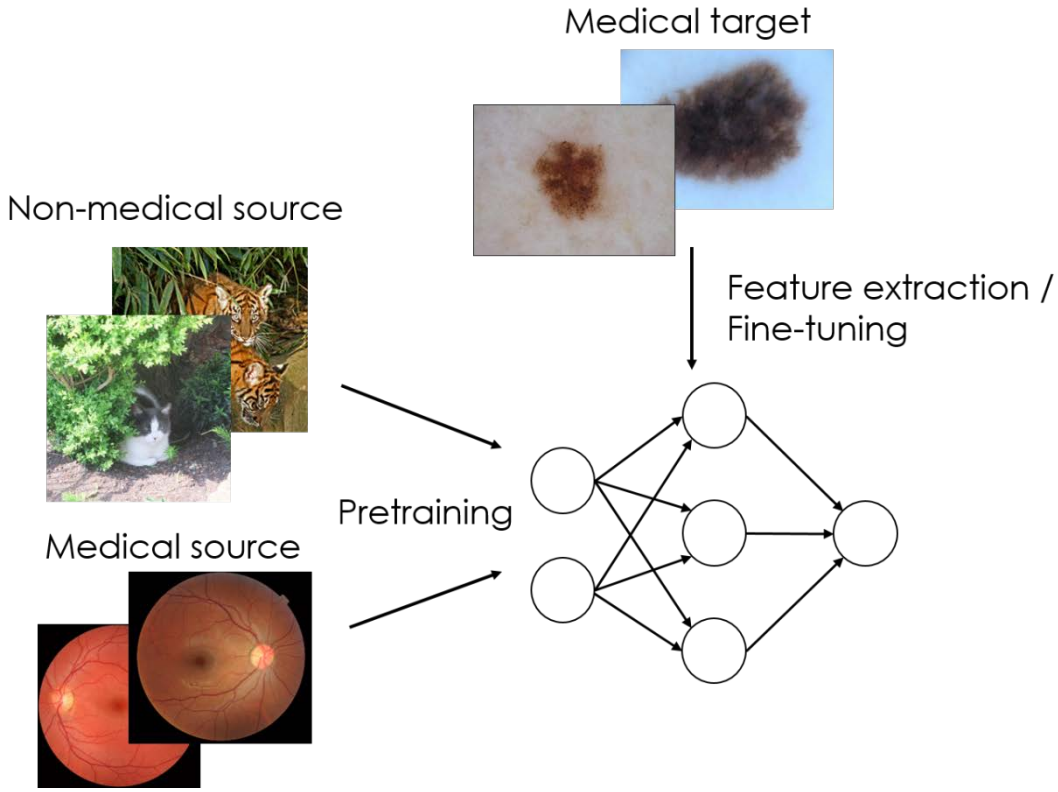


Cheplygina, V., Sorensen, L., Tax, D. M. J., Pedersen, J. H., Loog, M., & de Bruijne, M. (2014). Classification of COPD with multiple instance learning. In International Conference on Pattern Recognition (pp. 1508-1513).

## Solution 2: Transfer learning

Not learning “from scratch”

# … even from entirely different tasks




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

Veronika Cheplygina

https://www.sciencedirect.com/science/article/pii/S2468451118300527

# Solution 2: Transfer learning

Not learning “from scratch”

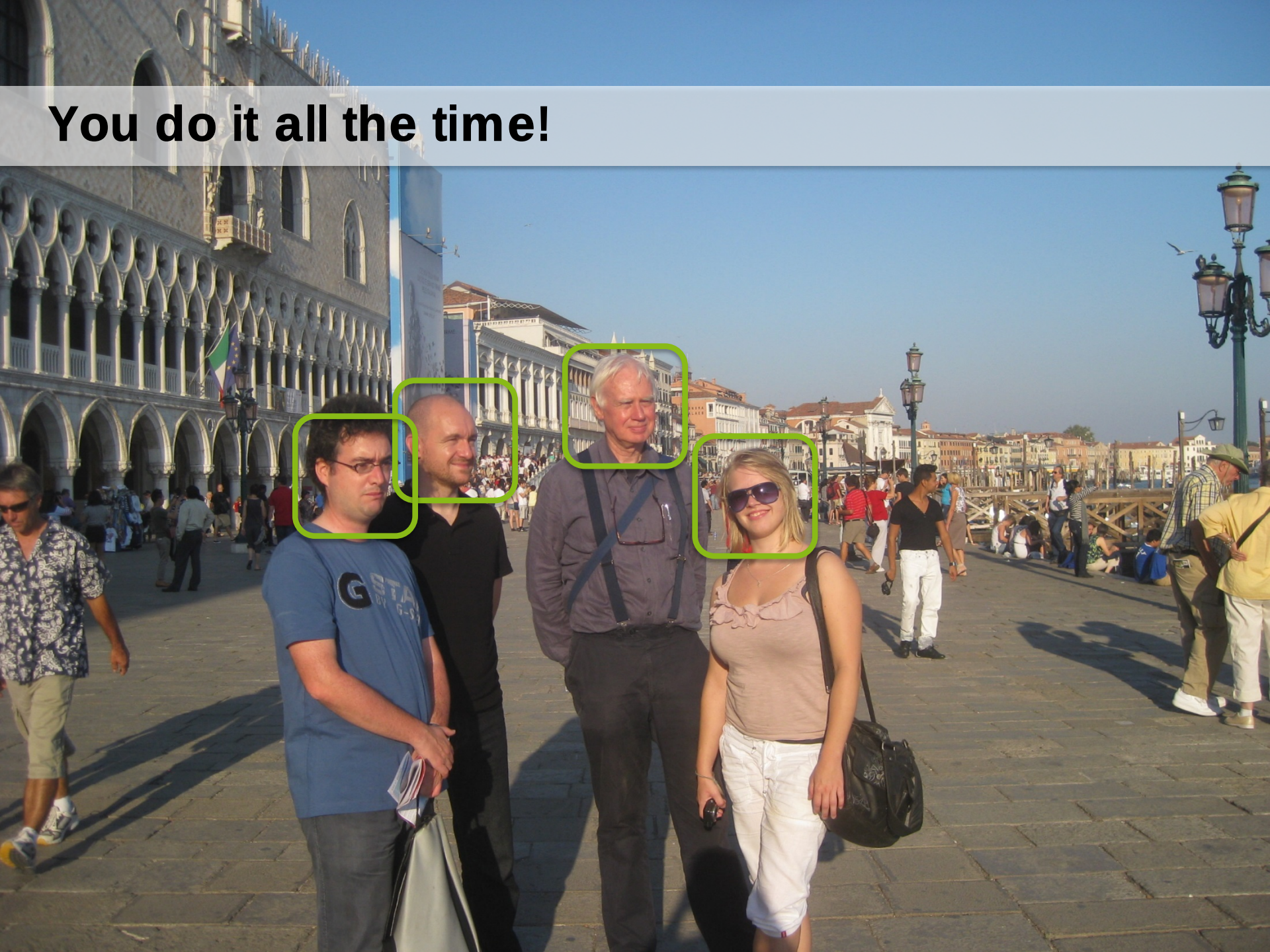
Not-so-supervised:
a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis

Veronika Cheplygina, Marleen de Bruijne, Josien P. W. Pluim

https://arxiv.org/abs/1804.06353

# Solution 3: Crowdsourcing

# You do it all the time!



tvst

DOI: 10.1167/tvst.5.5.6

Article

# The Accuracy and Reliability of Crowdsource Annotations of Digital Retinal Images

Danny Mitry[1], Kris Zutis[2], Baljean Dhillon[3], Tunde Peto[1], Shabina Hayat[4], Kay-Tee Khaw[5], James E. Morgan[6], Wendy Moncur[7], Emanuele Trucco[2], and Paul J. Foster[1] for the UK Biobank Eye and Vision Consortium

IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 35, NO. 5, MAY 2016



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

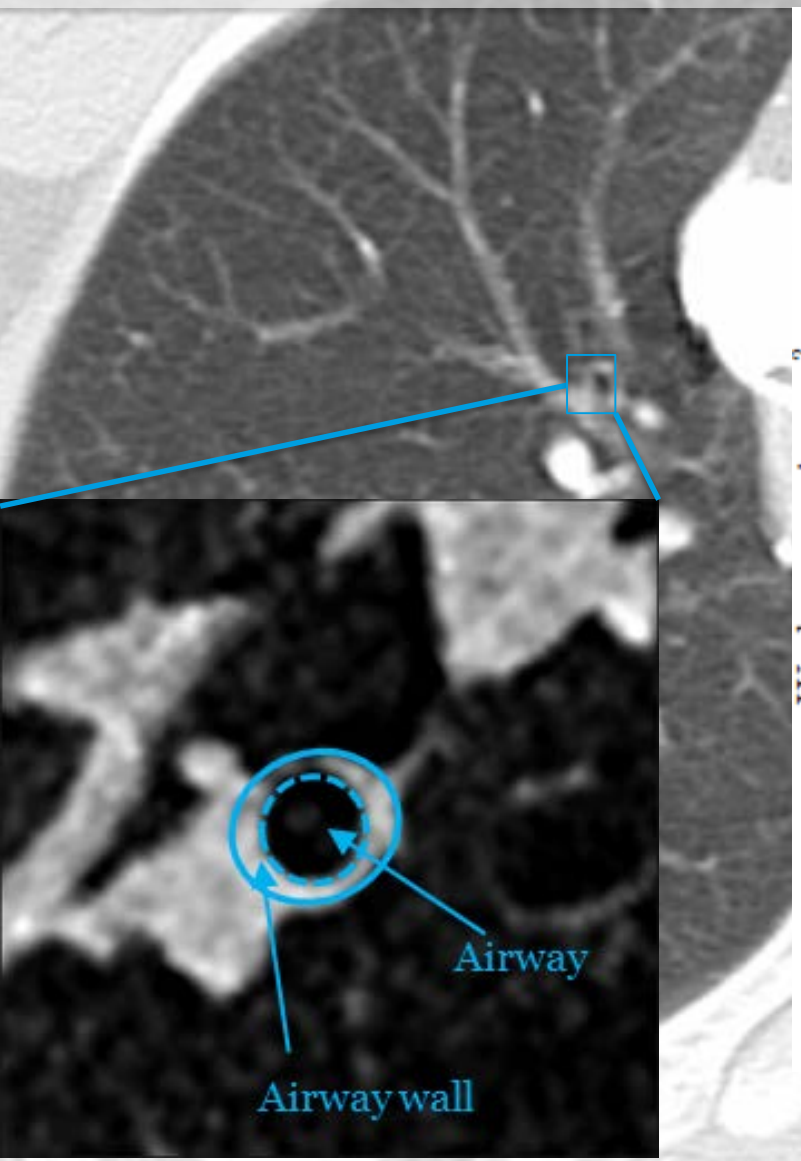
Shadi Albarqouni*, *Student Member, IEEE*, Christoph Baur, Felix Achilles, *Student Member, IEEE*, Vasileios Belagiannis, *Student Member, IEEE*, Stefanie Demirci, and Nassir Navab, *Member, IEEE*

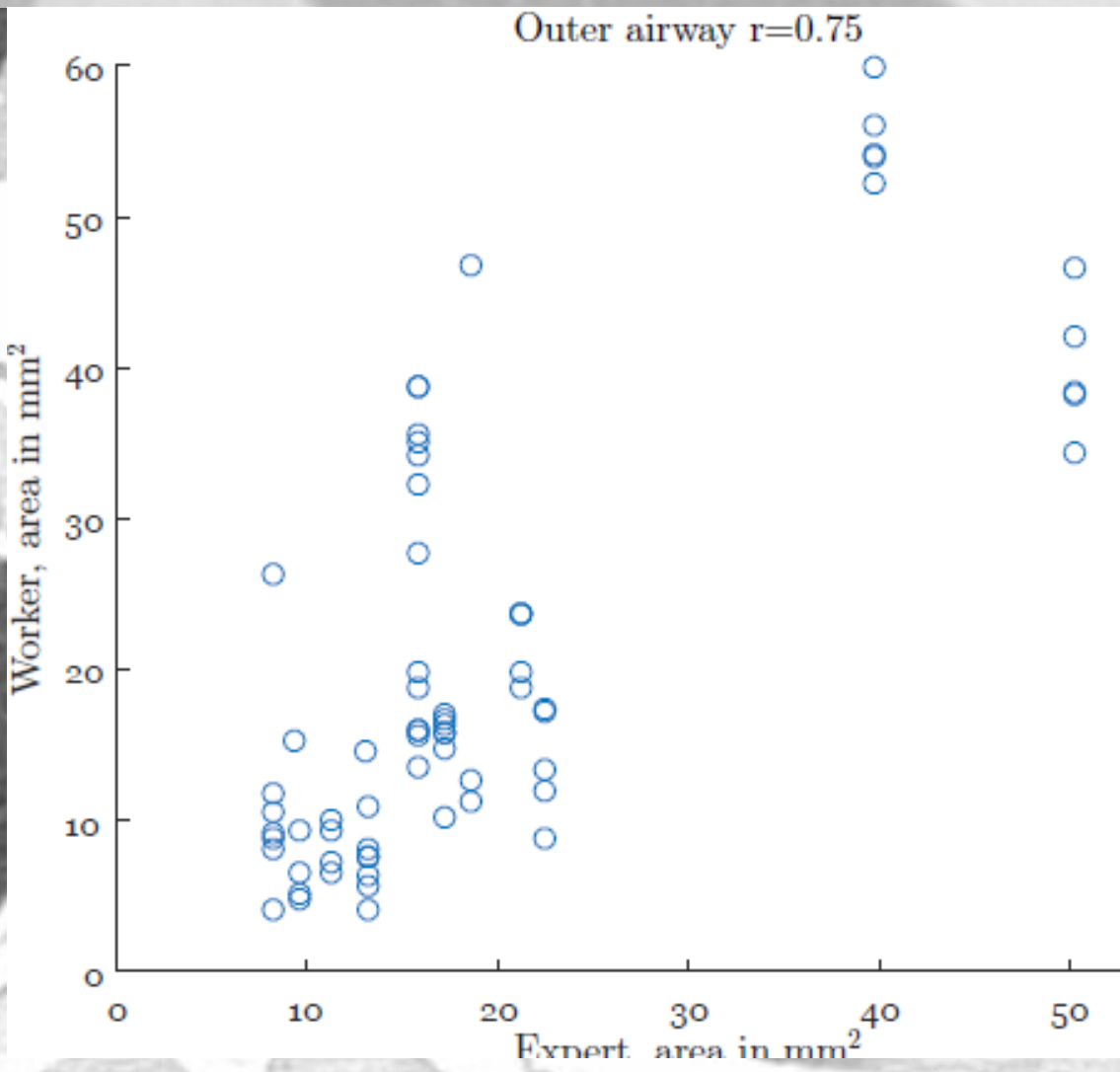
# Can Masses of Non-Experts Train Highly Accurate Image Classifiers?

## A Crowdsourcing Approach to Instrument Segmentation in Laparoscopic Images

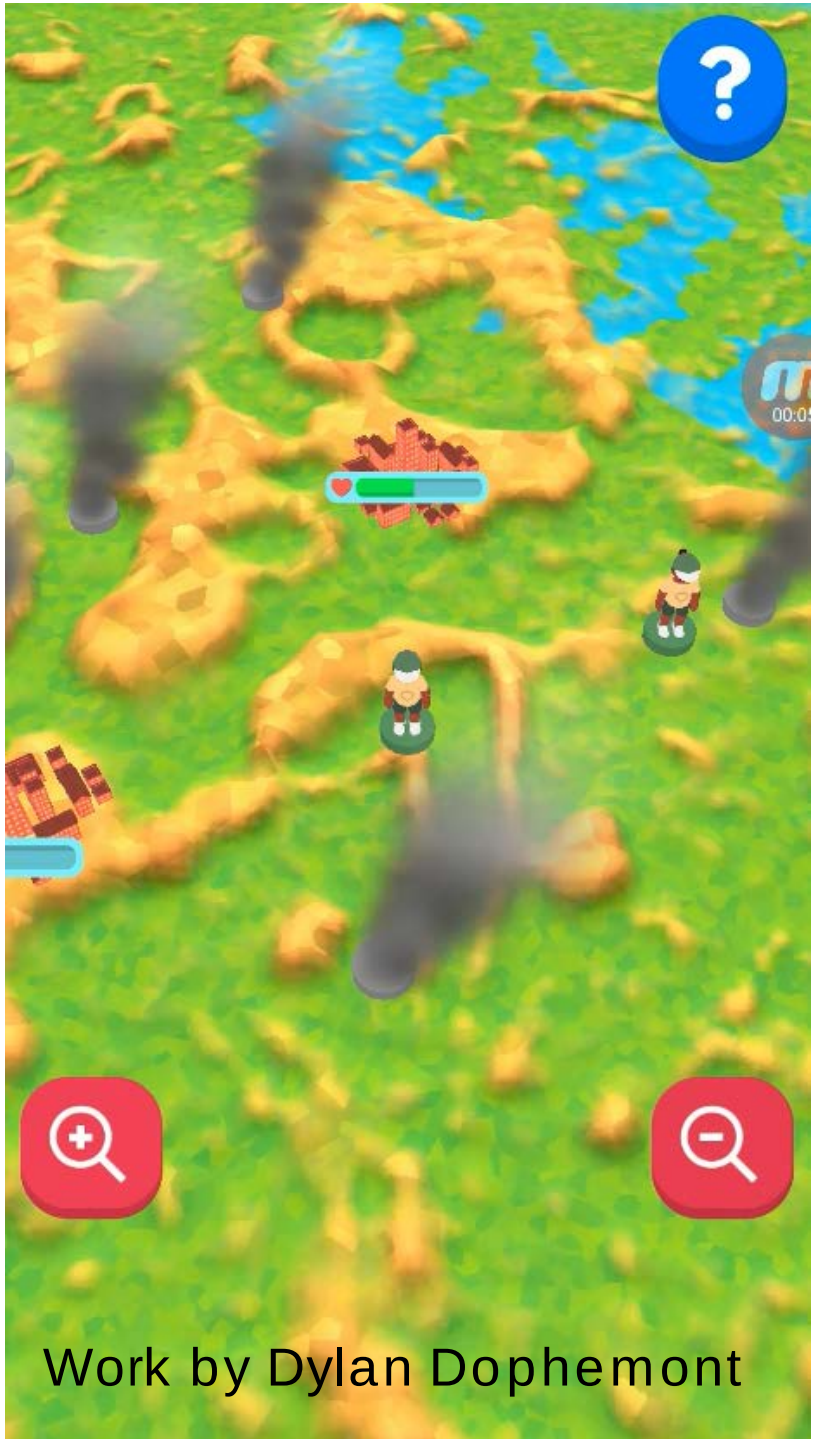
Lena Maier-Hein[1,⋆,⋆⋆], Sven Mersmann[1], Daniel Kondermann[2], Sebastian Bodenstedt[3], Alexandro Sanchez[2], Christian Stock[4], Hannes Gotz Kenngott[5], Mathias Eisenmann[3], and Stefanie Speidel[3]

# Crowdsourcing airway annotations




Cheplygina, V et al. (2016). Early Experiences with Crowdsourcing Airway Annotations MICCAI LABELS 2016

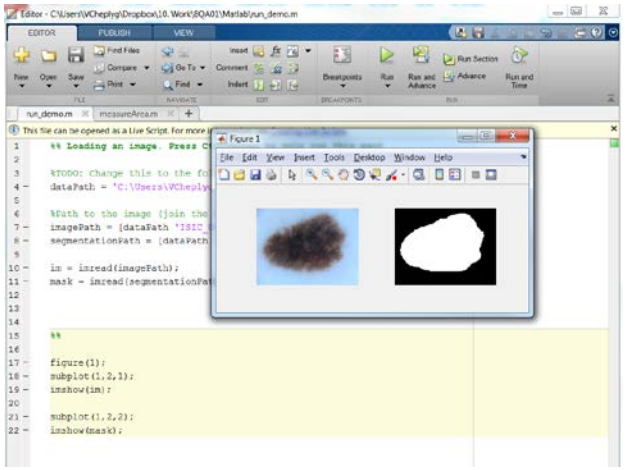

?
00:05
Work by Dylan Dophemont


?
03:08
1
2
OK

# Image analysis project for 1st year students

Technische Universiteit Eindhoven University of Technology

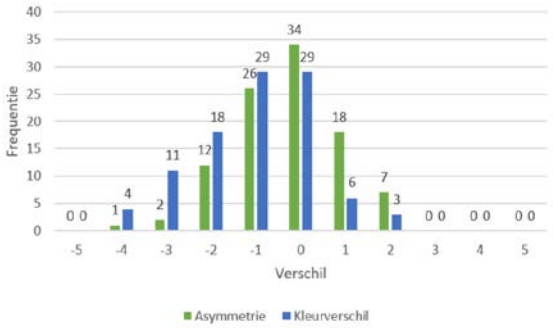
## 1. Measure features with algorithms



## 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

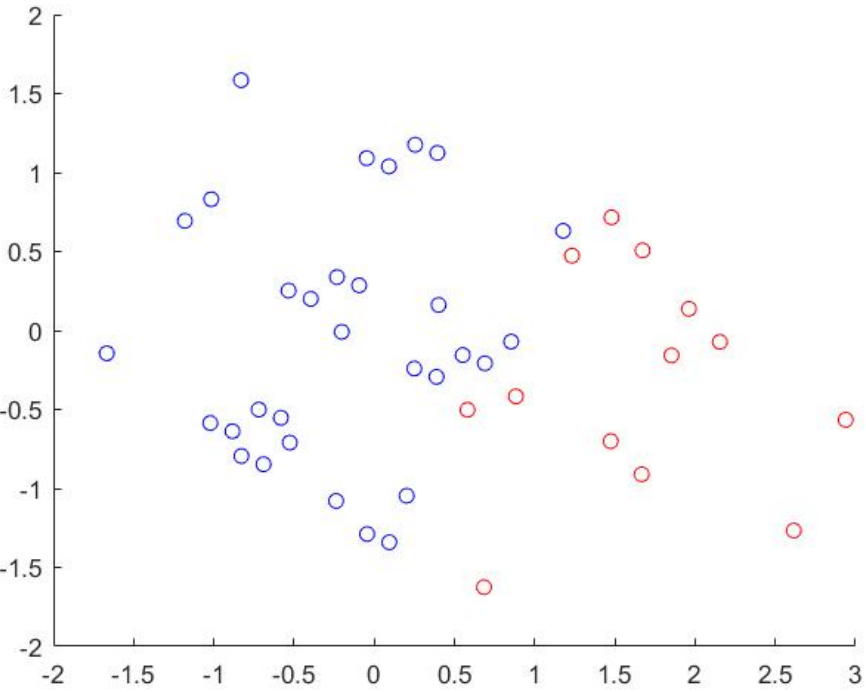


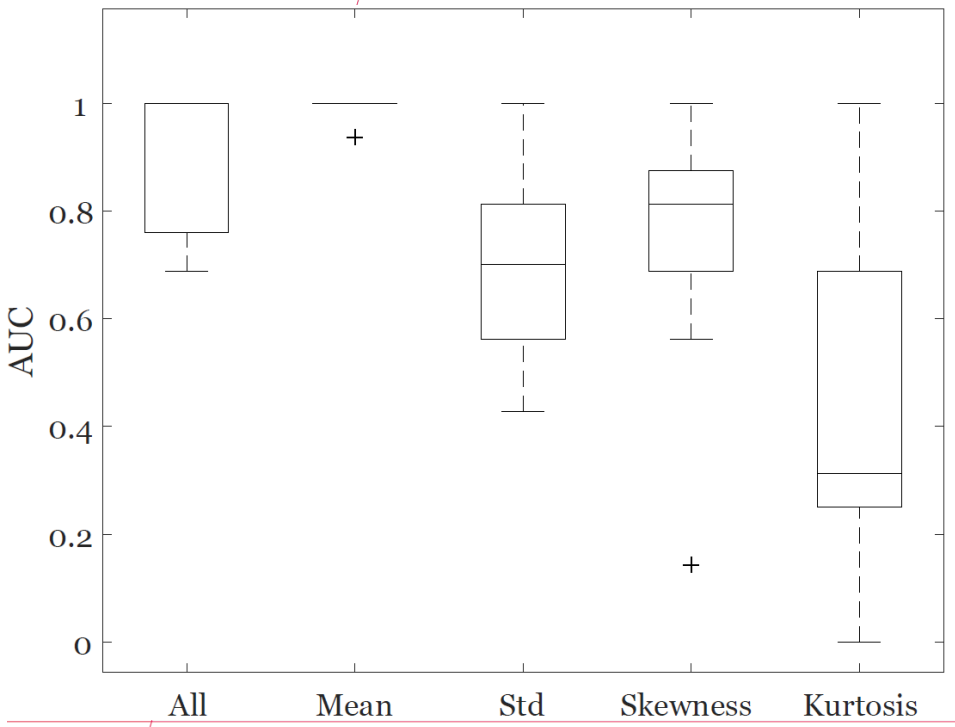

*Grafiek 1: De frequenties van de verschilwaardes tussen de metingen in Matlab en de metingen op het oog*

**Crowdsourcing!**

**Analyzing 100 annotated images**

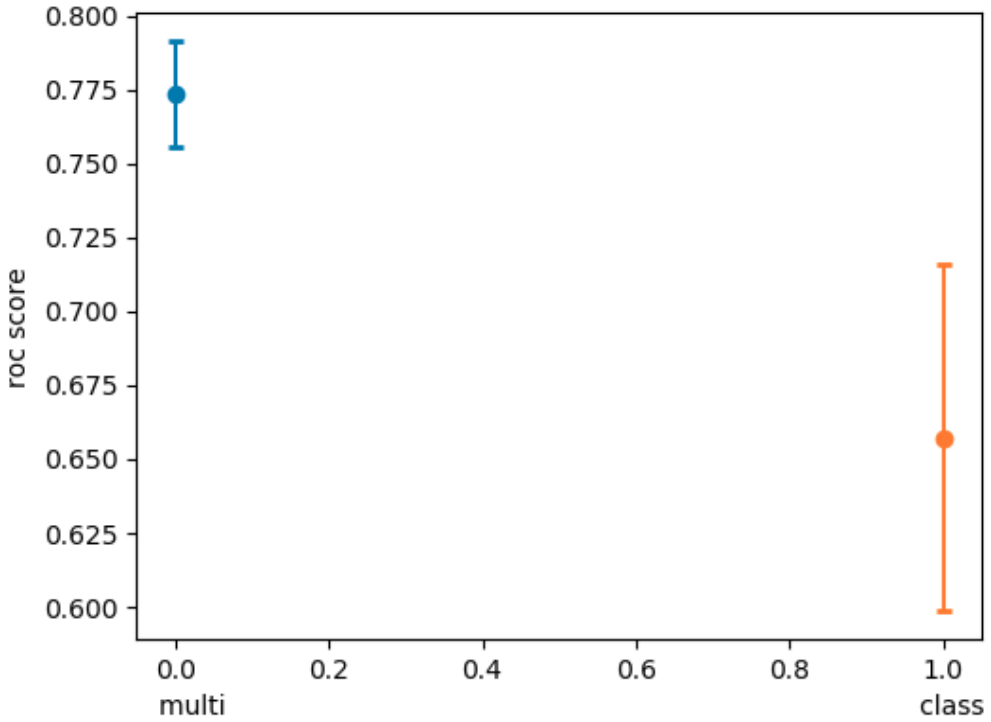
- 5 features: asymmetry etc
- 6 annotators
- Predict healthy vs melanoma **without** any image data

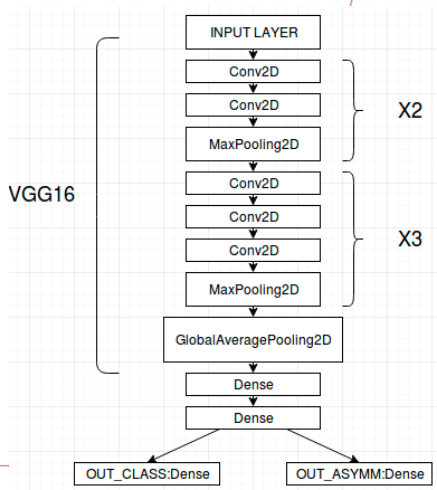


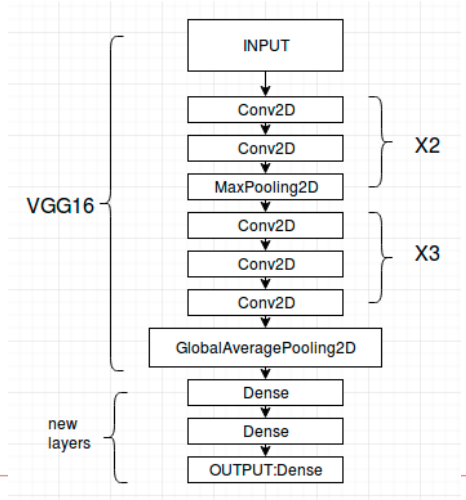



# Work by Elif Kubra Contar

Multi-task approach outperforms classification-only approach



# 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!

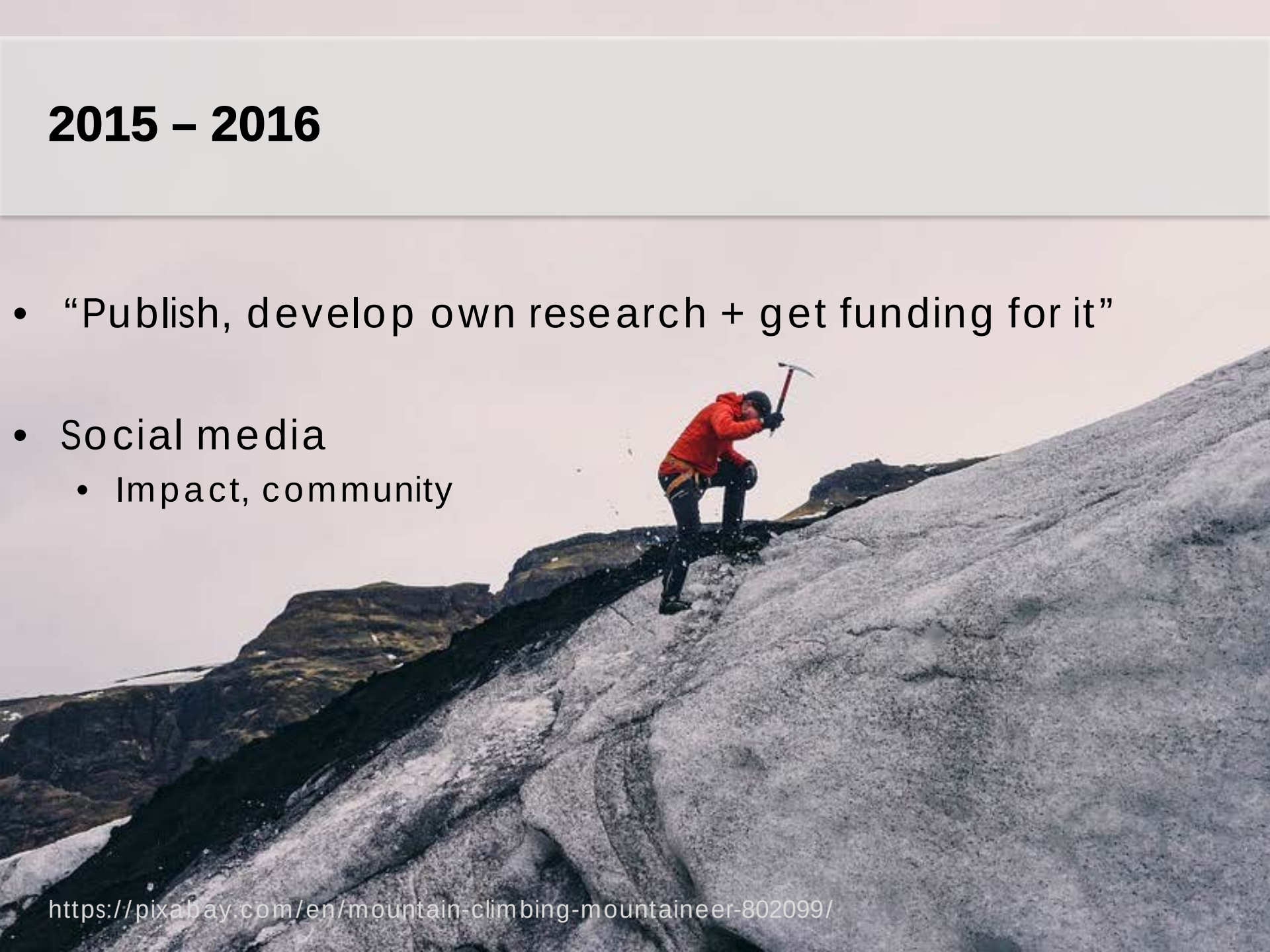
2015

TU Delft Delft University of Technology

Erasmus MC Universitair Medisch Centrum Rotterdam

# 2015 – 2016

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



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

# 2017?


## 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 t started looking for my next position halfway through my postdoc, which was a job in its not reflect well on my postdoc project. A few things were not really going well for me ir 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 reassure myself by thinking that if I'm an impostor and they are the the real deal, they s 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" whil "good, but not excellent CV" (citing reviews on some of my rejected grant applications have to deal with hundreds of rejections – I applied to four jobs, interviewed for three, a 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 bi

https://veronikach.com/progress-reports/getting-tenure-track-

# 2017+ Not without challenges

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

# 2017+ Find others to support you

https://unsplash.com/photos/iuqmGmst5Po

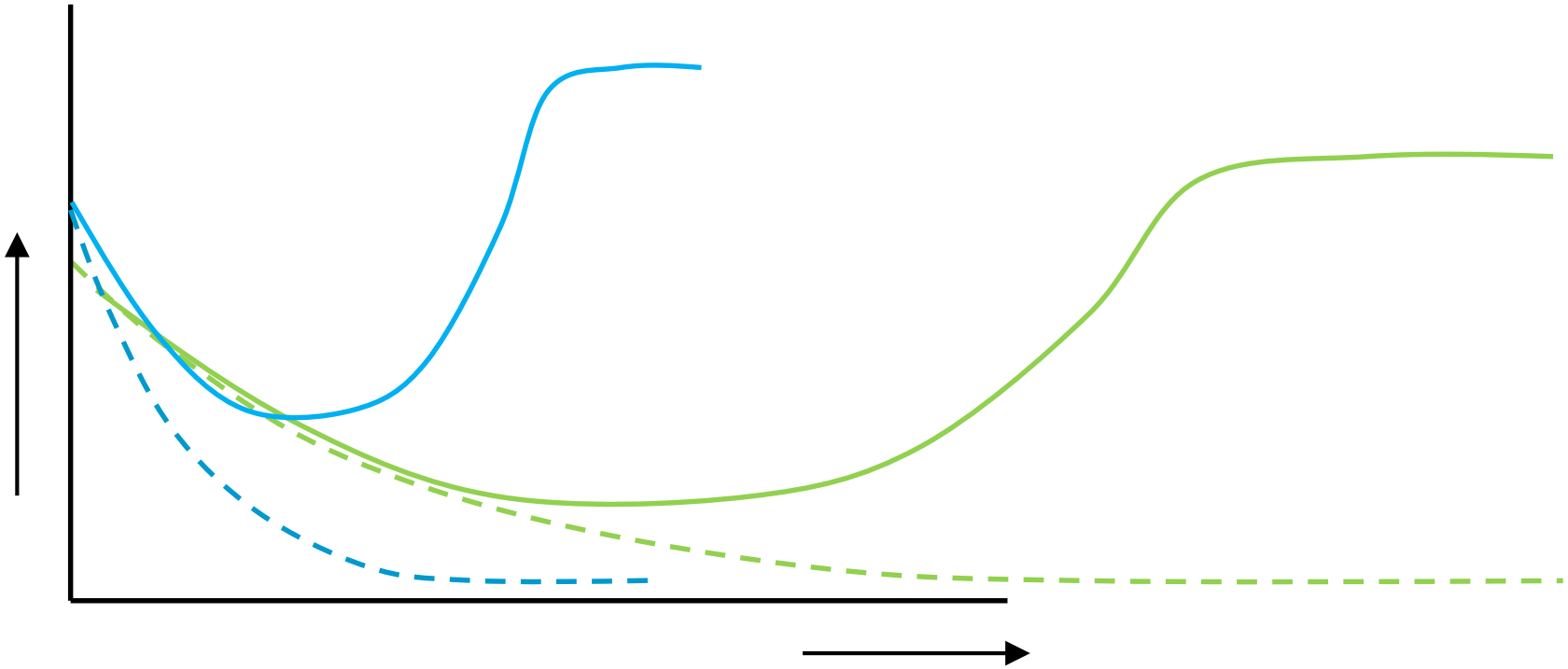
# Academia as supervised learning?

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

## **But** CVs ≠ true data distribution

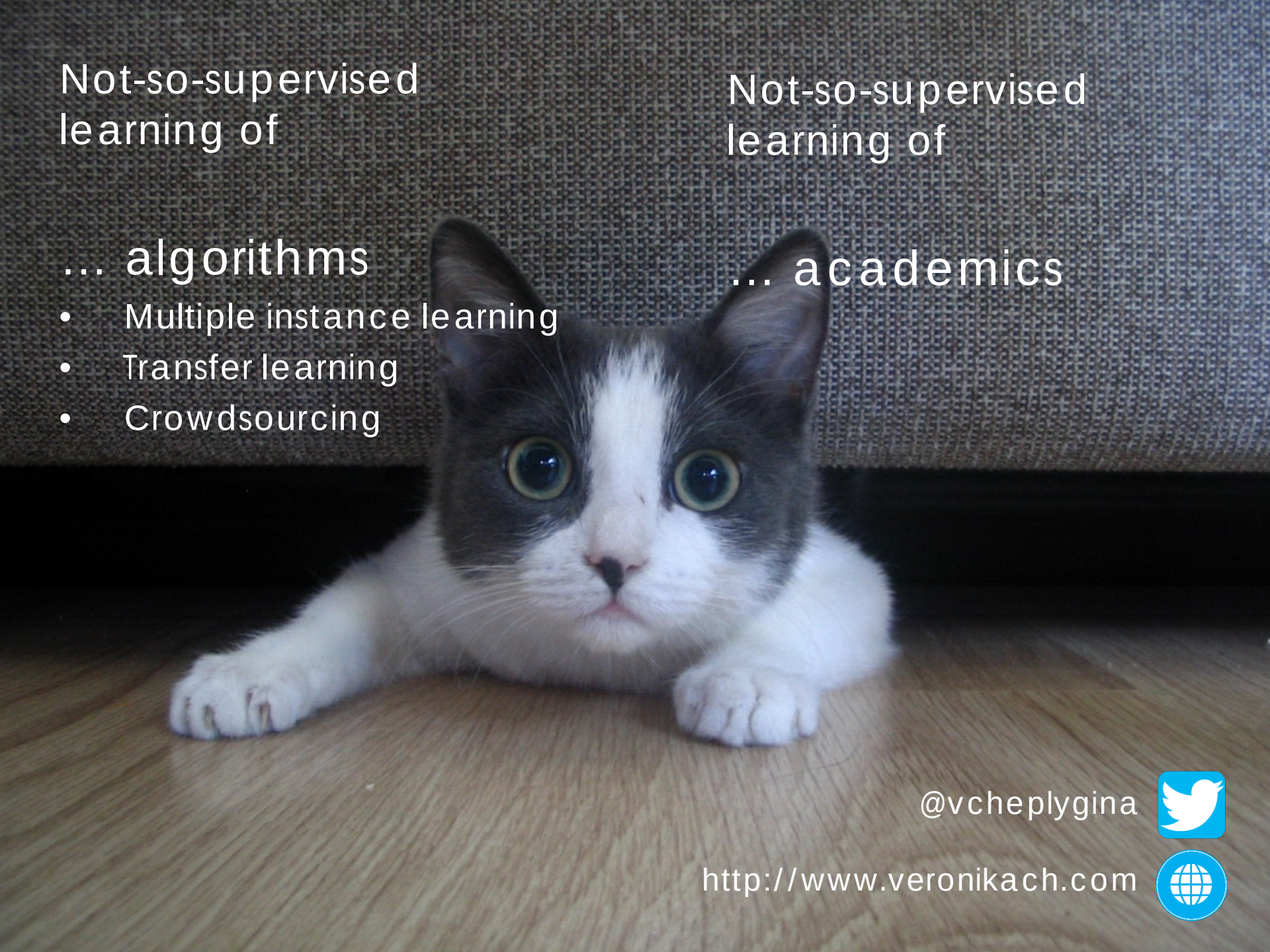
- Input space is much larger (Shadow CV)
- Output space is much larger (Impact, being happy)
- Noisy labels, many unlabelled inputs

# Overfitting



## Not-so-supervised advice

- People vs projects
- Examples vs rules
- Explore vs repeat


Not-so-supervised learning of
... algorithms
• Multiple instance learning
• Transfer learning
• Crowdsourcing
Not-so-supervised learning of
... academics
@vcheplygina
http://www.veronikach.com


David Tax
Marco Loog
PRB
Aasa Feragen
Chloe Azencott
Marleen de Bruijne
BIGR
Josien Pluim
Mitko Veta
IMAG/e
Bettina Speckmann
MICCAI LABELS
My students

Buffy
Wocky!
Ariston'80 Ladies 2
Mattias Hansson
Spinazie
Waardeloze Wetenschappers
AssistantProfSupport
MICrowd
How I Fail
Spoonies
Heartbeats
Pixel and Dot

@vcheplygina

http://www.veronikach.com