Medical (text and image) information retrieval

Henning Müller

Milano, Italy, 18.7.2019

(1) Text and semantics

(in medicine)

(2) Multimodal and visual

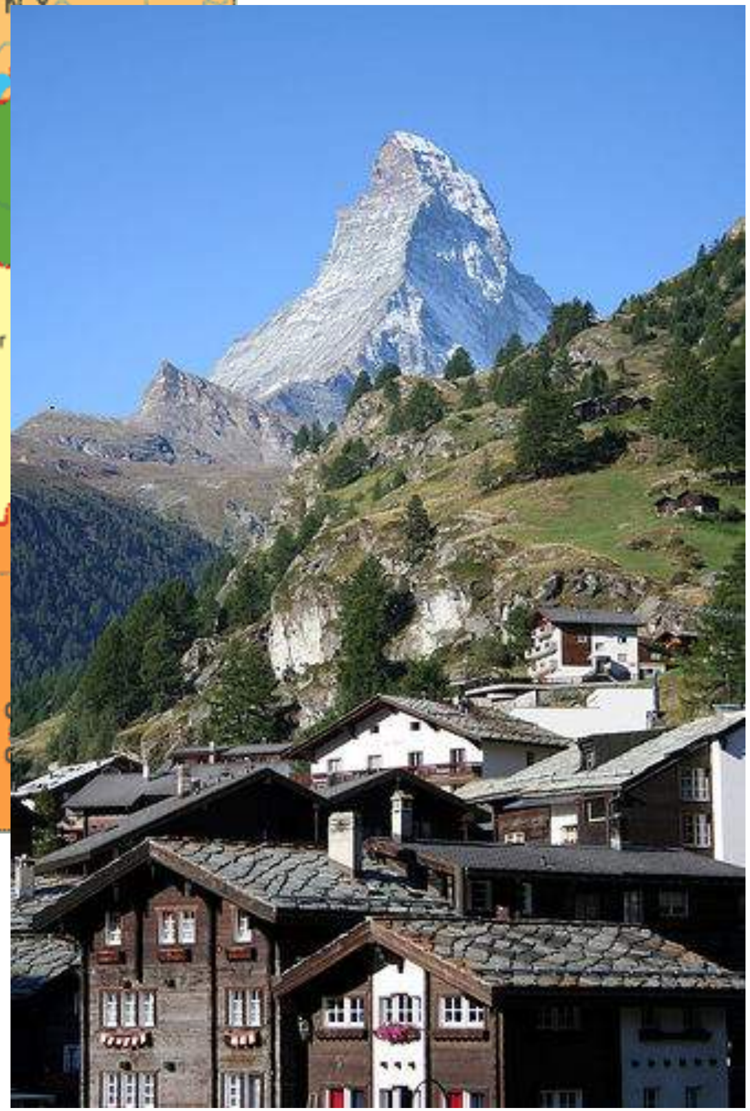
(plus applications)

Who I am

- Medical informatics studies in Heidelberg, Germany (1992-1997)
  - Exchange with Daimler Benz research, USA
- PhD in image processing, image retrieval, Geneva, Switzerland (1998-2002)
  - Exchange with Monash University, Melbourne, AUS
- Professor in radiology and medical informatics at the University of Geneva (2014-)
- Professor in Computer Science at the HES-SO, Sierre, Switzerland (2007-)
  - Visiting faculty at Martinos Center (2015-2016)

Where I am


Switzerland
GERMANY
FRANCE
SWITZERLAND
ITALY
LIECH.
Basel
Baden
Zürich
Winterthur
St.Gallen
Horgen
Lucerne
(Luzern)
Biel
Neuchâtel
Burghdorf
Bern
Steffisburg
Fribourg
Thun
Interlaken
Chur
Lausanne
Montreux
Sierre
Geneva
(Genève)
Martigny
Champex
Locarno
8°
©maps.com



Where I am

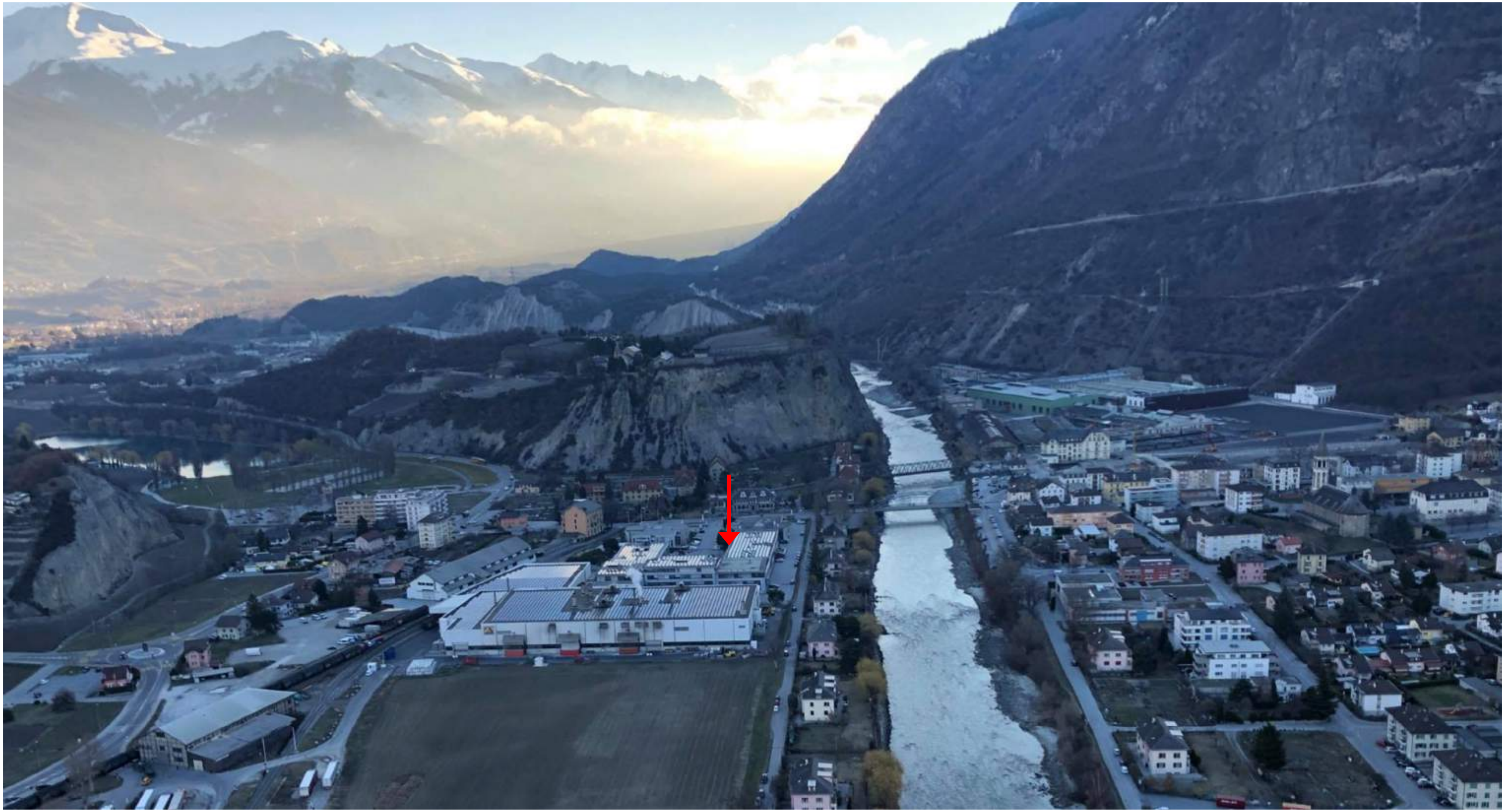


Photo courtesy of Laurent Borella

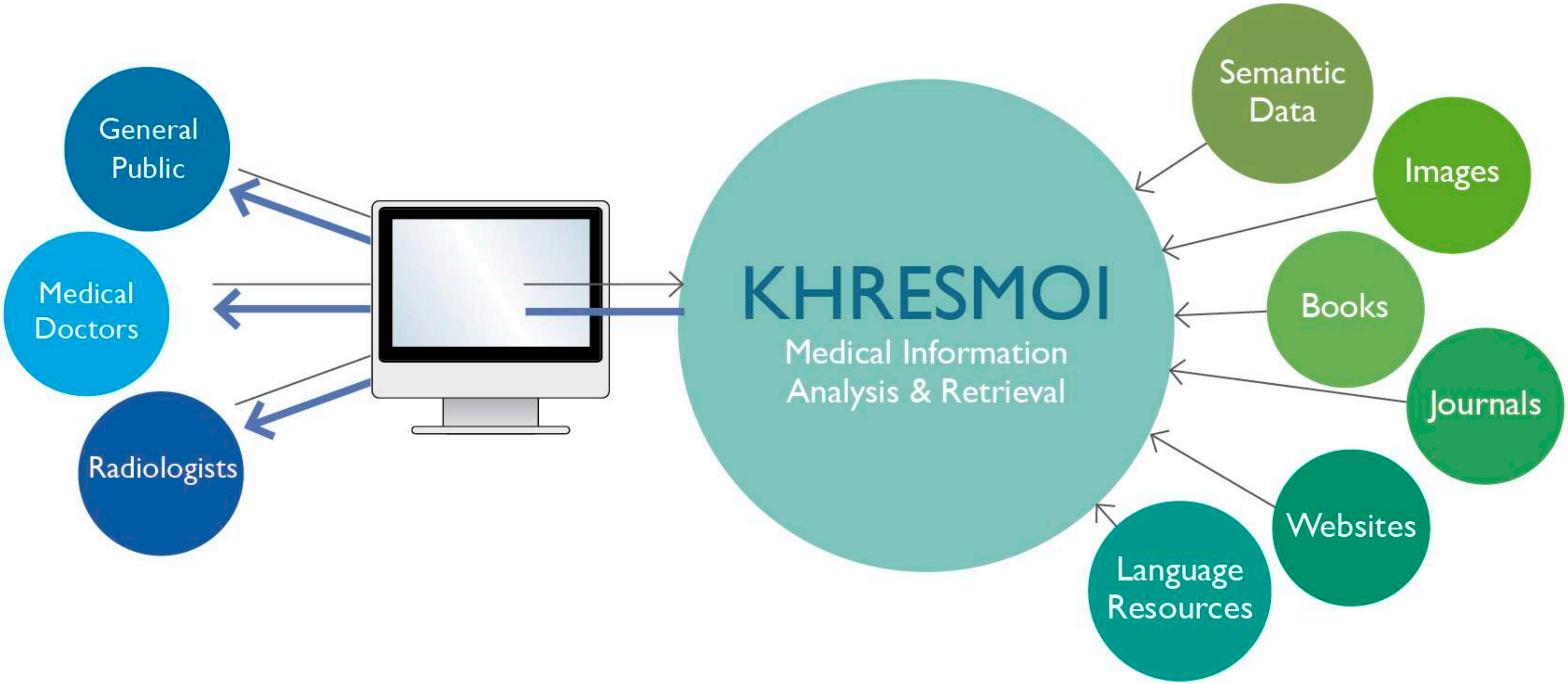
Your background

- Who has a medical background or worked in a medical project?
- Who has worked in image analysis or computer vision?
- Who has a good background in NLP or text retrieval?

Objectives of this course

- Learn about medical information retrieval
  - With a focus on visual information retrieval
  - And the types of information needs that are encountered
- Learn about all the techniques around retrieval applications (pre-treatment, visualization, …)
  - Machine learning and computer vision
- Understand multi-modal retrieval and the links between modalities (images, volumes, signals, …)
- Understand the road towards decision support in medical environments

- Mixing multilingual data from many resources and semantic information for medical retrieval
  - LinkedLifeData.com

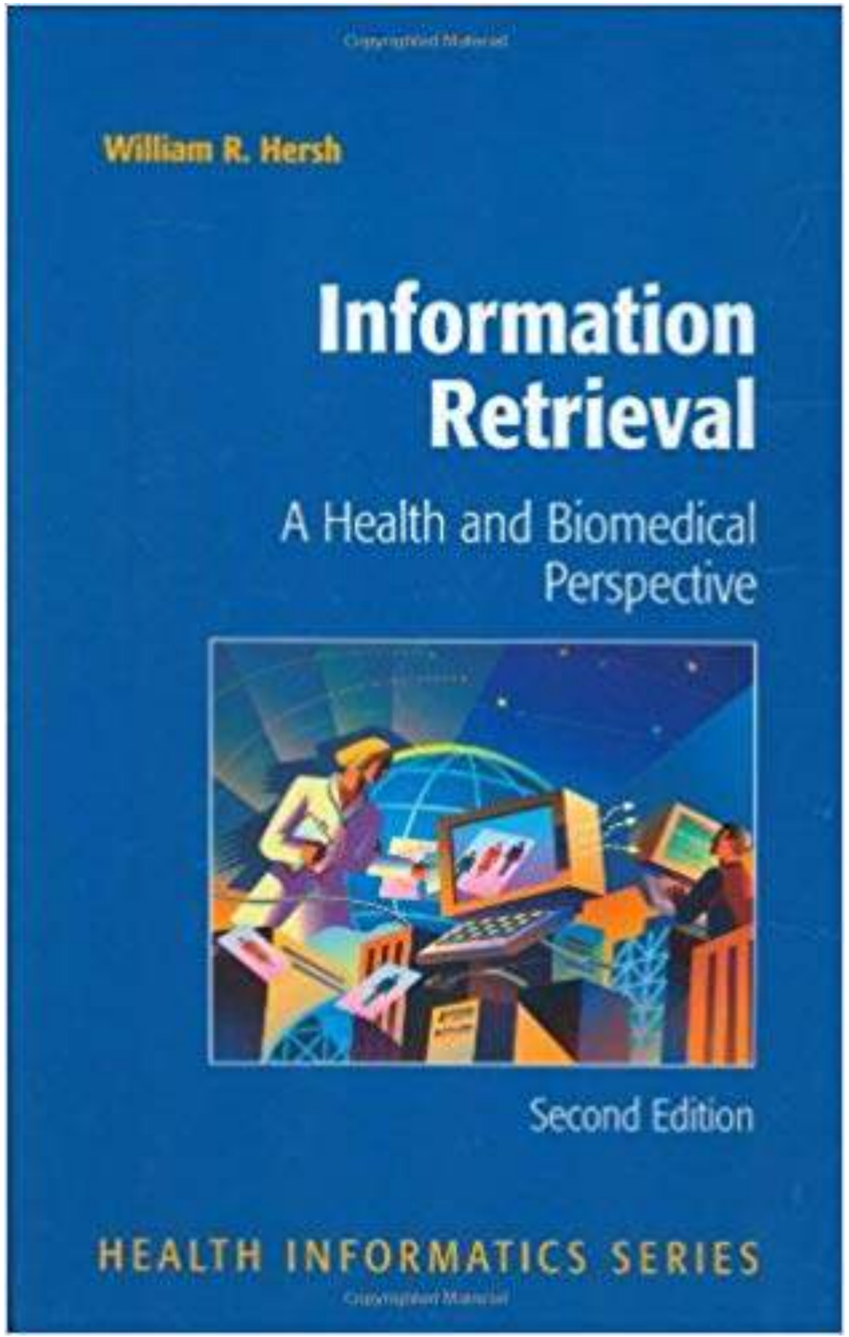


Allan Hanbury, Célia Boyer, Manfred Gschwandtner, Henning Müller, KHRESMOI: Towards a Multi-Lingual Search and Access System for Biomedical Information, Med-e-Tel, pages 412-416, Luxembourg, 2011.

The informed patient

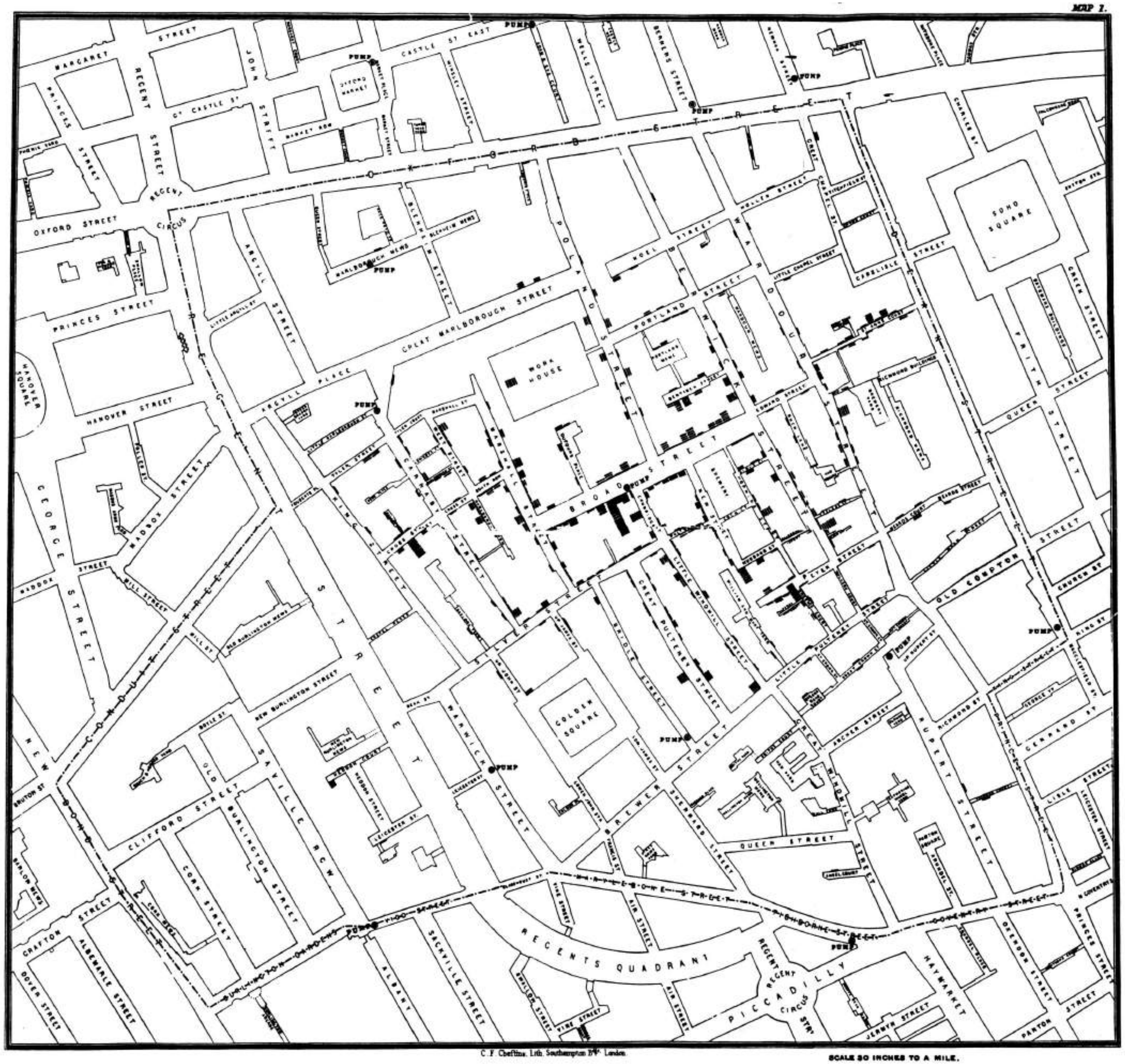
"I'M SORRY DOCTOR, BUT AGAIN I HAVE TO DISAGREE."

A good book!

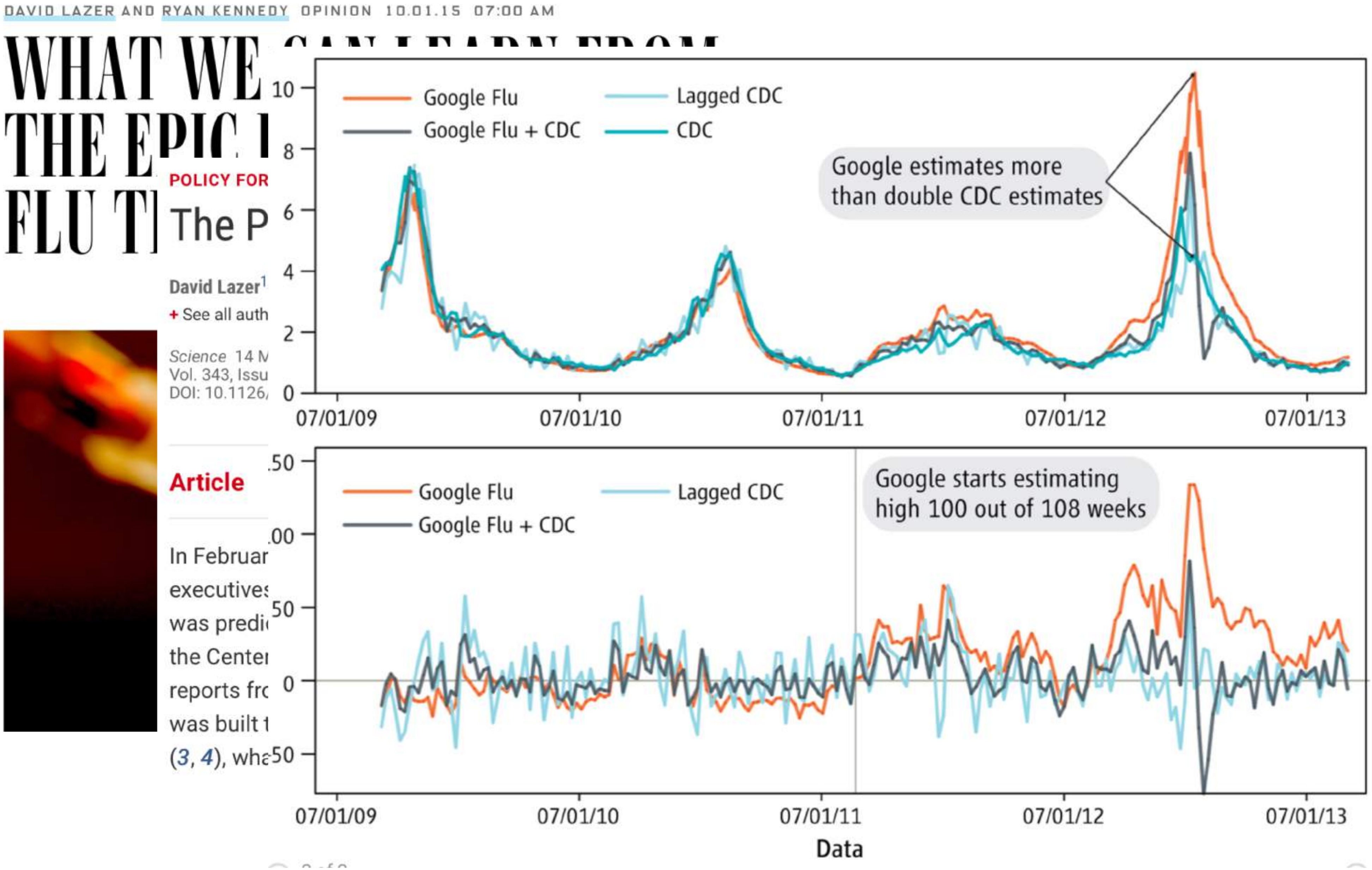

William R. Hersh
Information Retrieval
A Health and Biomedical Perspective
Second Edition
HEALTH INFORMATICS SERIES

Systematic medical data analysis

- Broad Stre
- Prevailing o
  transmitted
- Physician Jo
  water or ot
  - Water came
    filtered at th
  - He noted all



Google Flu Trends



A short history of medical IR

- 1836: Library of the Surgeon General's Office
- 1879: Publication of the Index Medicus
- 1950: MeSH was introduced
- Mid 1960s: MEDLINE/MEDLARS, a digital collection with the medical literature started
  - Structured data with free text abstract
- 1997: PubMed was offered as a free web search engine
- 2000: PubMed Central made the biomedical open access literature accessible

Many different types of users

- Clinicians looking for the latest research
  - Half-life of medical information is estimated to be between 5 and 10 years
  - Impossibility to be up to date even in an extremely narrow domain
- Patients search for information on their disease or for friends/family
  - Hard to get information on rare diseases
- Pharma industry may search for clinical trials
  - And IPR questions
- Radiologists search in images

Medical information

- Textbooks
- Scientific articles
  - Semi-structured
- Medical content of all sorts on the web
  - For patients, for professionals and much incorrect or controversial information
- Clinical texts
  - Structured data and many reports (semi-structured)
  - In many languages
  - Search in a single patient and across patients
  - Clinical trials

Many resources are available

- NLM maintains many language resources
  - MeSH – Medical Subject Headings
  - UMLS – Unified Medical Language System
  - PMC and PubMed data sets
  - RadLex is maintained by the RSNA
- NLM maintains software and web interfaces
  - OpenI, PubMed.gov, PubMed Central
  - MedLine Plus – a resource for patients
  - NegBio, MetaMap
  - …

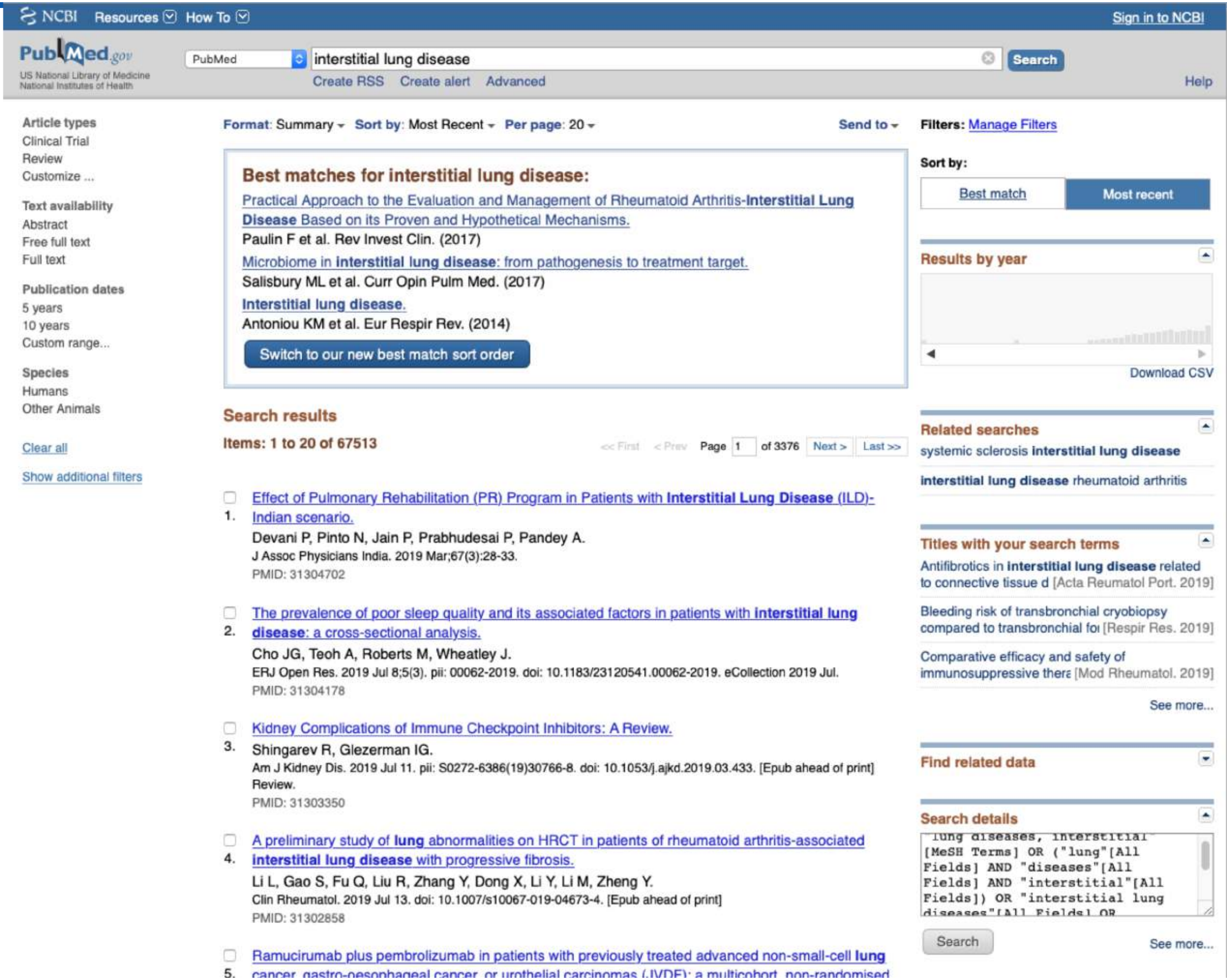
… but there are challenges

Hes·so// VALAIS WALLIS
School of Management & Tourism

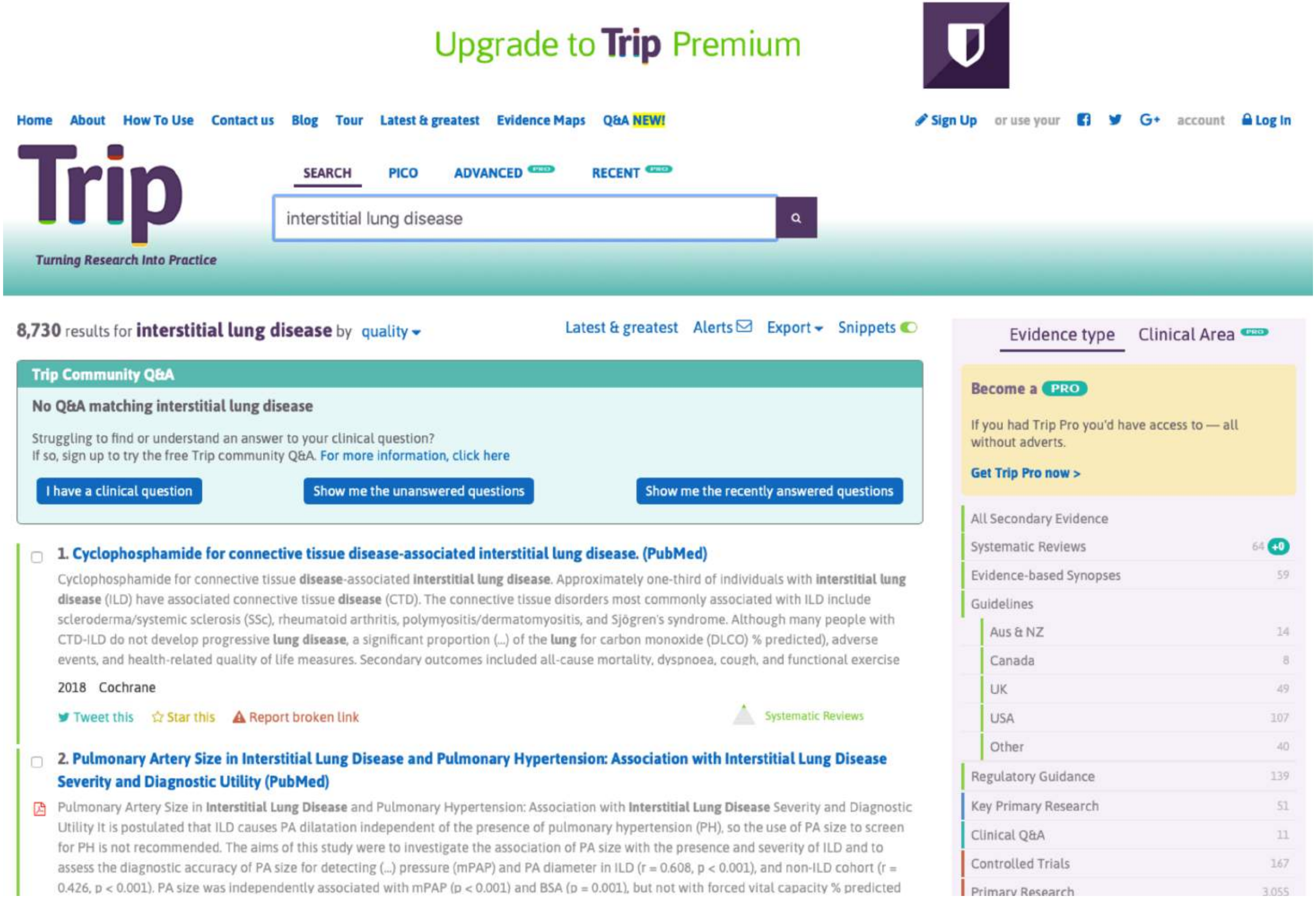
- Non-standard abbreviations
- Spelling mistakes
  - Quickly written
- Technical language
  - Latinized terms, synonyms
- Nested, complex phrases
- Negation …
  - Several levels ("little evidence of")
  - Not clear what terms they refer to, double negations

A/B	acid-base ratio
ab	abdomen, abdominal abortion
Ab	antibody
AB	abortion, AB Blood Type
ABC	airway, breathing, circulation aspiration biopsy cytology
ABCD	airway, breathing, circulation, disability asymmetry, borders, color, diameter (feature ABCD rating (a staging system for prostate c
ABCs ABCDs ABCDEs	airway, breathing, circulation, etc. Refers to p recurrent.
ACA	acinic cell carcinoma Affordable Care Act
Abd	abdomen abdominal[abduction]
ABD	army battle dressing

Interface for PubMed



Interface for the TripDatabase

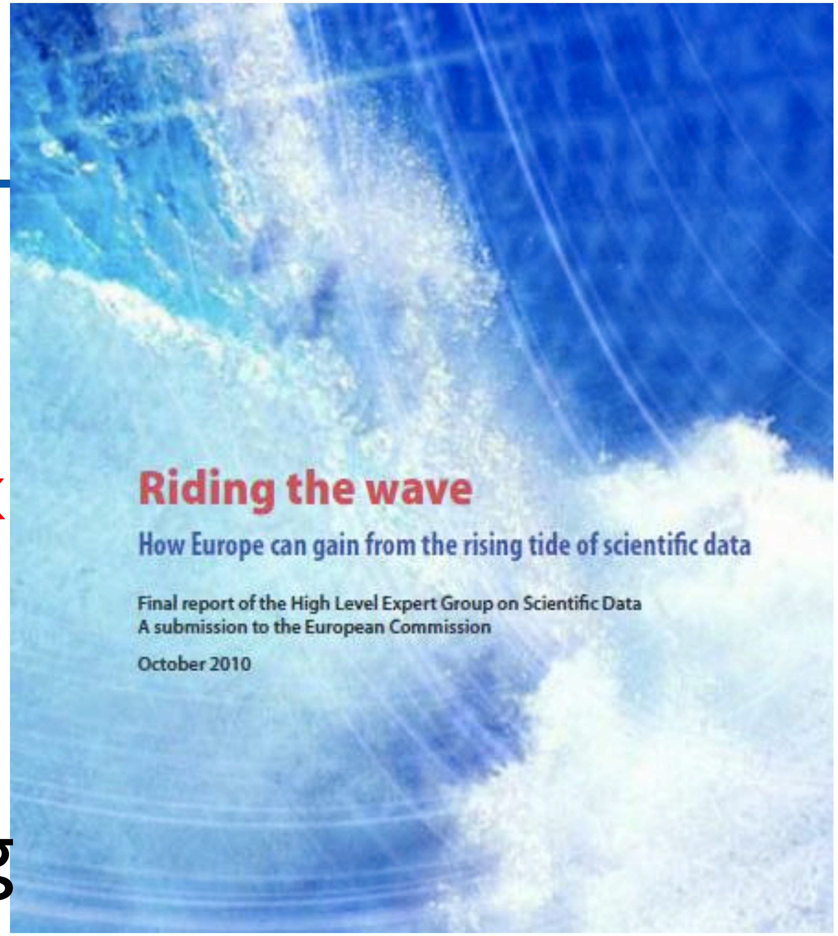


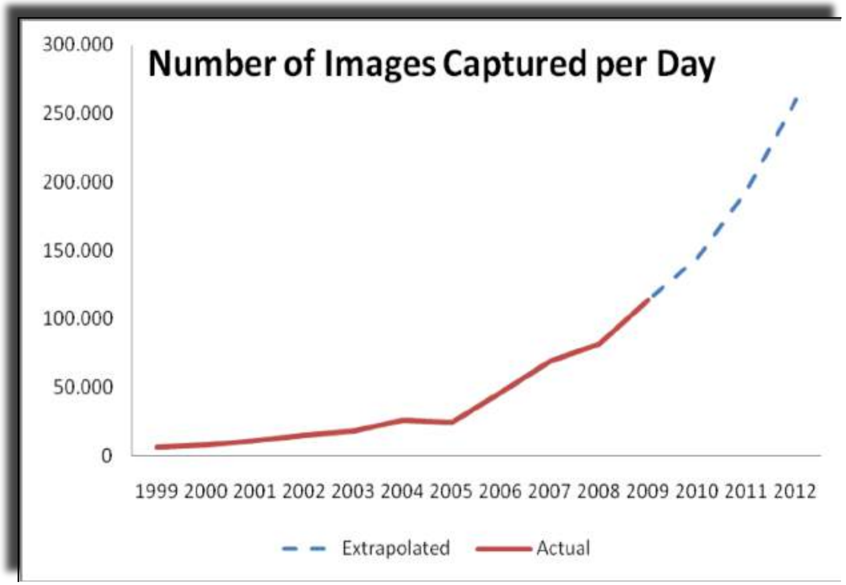
Retrieval vs. classification

- Retrieval is to fulfil an information need with a relevance definition (no fixed classes)
  - Subjective, user- and time-dependent
- Classification means that a fixed number of classes exist (exclusive or non-exclusive)
  - Can be linked to object recognition
- Localization means to find the region in an image (with an object, a lesion, …)
- Detection means to find whether a pattern is present or not
- Recognition, …

Motivation for medical imaging

- Many images are produced
- Imaging data are very complex
  - And getting more complex
- Imaging is essential for diagnosis & treatment planning
- Images out of their context loose most of their sense
  - Clinical data are necessary
- Towards precision medicine
  - Using all data sources





Visual information vs. text retrieval

- Visual retrieval aims at image content and not at text documents
  - Many images do not have any attached metadata
- Formulation of most information needs is easier in text, as we are much more used to it
  - … but annotations vary across persons and time
  - Formulations and annotations are often ambiguous
  - But something like "images that look like but are not tuberculosis" is almost impossible with text alone
- Context needs to be taken into account
  - Anamnesis in case of a medical visual info retrieval

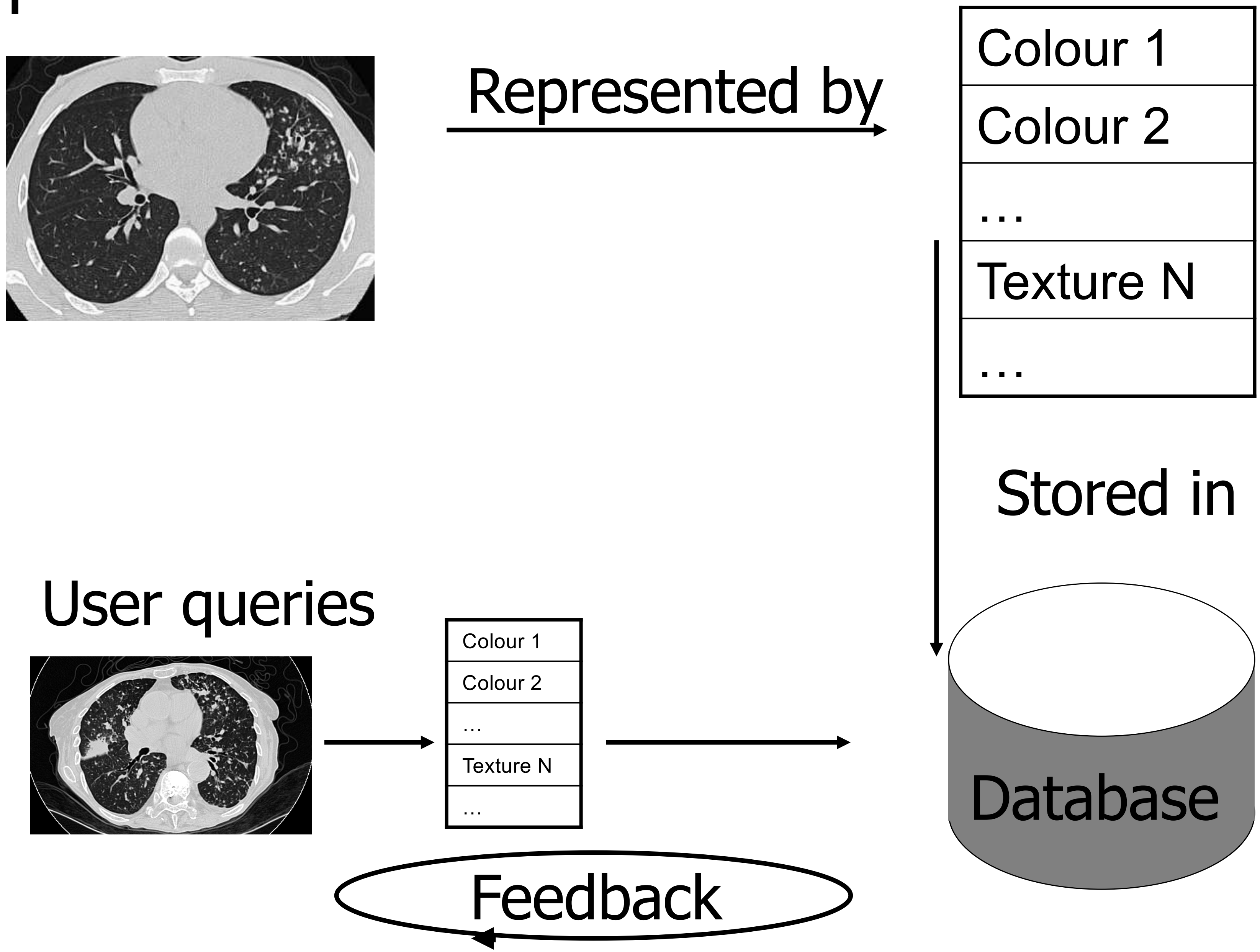
Differences between text and visual

- Text has words with meanings
  - Visual retrieval usually uses low level visual features to represent the data, extracted from the pixel data
  - Learning is required to group features or learn complete representations (as with deep learning)
  - Data-driven vs. model-driven approaches
- Both can be mapped onto semantics
  - But quality of visual mapping is usually lower
- NLP has much more experience than visual feature engineering for retrieval
  - Best practices, and many open source tools …

Gaps in image retrieval

- Sensory gap
  - Images taken mean a data loss compared to reality (limited resolution, no 3D, …)
- Semantic gap
  - Features automatically extracted may not correspond to human semantic search categories
    - Colors, textures, shapes, often global features
- Page zero problem of query formulation
  - How can an image for QBE be found?
- Conclusion: use text wherever available!
  - Or multimodal retrieval to combine text and visual info

Simplified overview



Perception of visual retrieval

Medical vs. non-medical

- An image is usually always present for medical cases, so no page zero problem
  - An image without meta data is pretty much unusable
  - Depersonalization can make data less usable
- Usage of large-scale medical data is legally challenging (privacy constraints)
  - … and annotations are usually very expensive
- Several types of information are always available (clinical data, images)
  - Unless images of the biomedical literature or of teaching files are used, when some of the context is lost

Content vs. context: healthy patients

25 year old man:

- homogeneous tissue

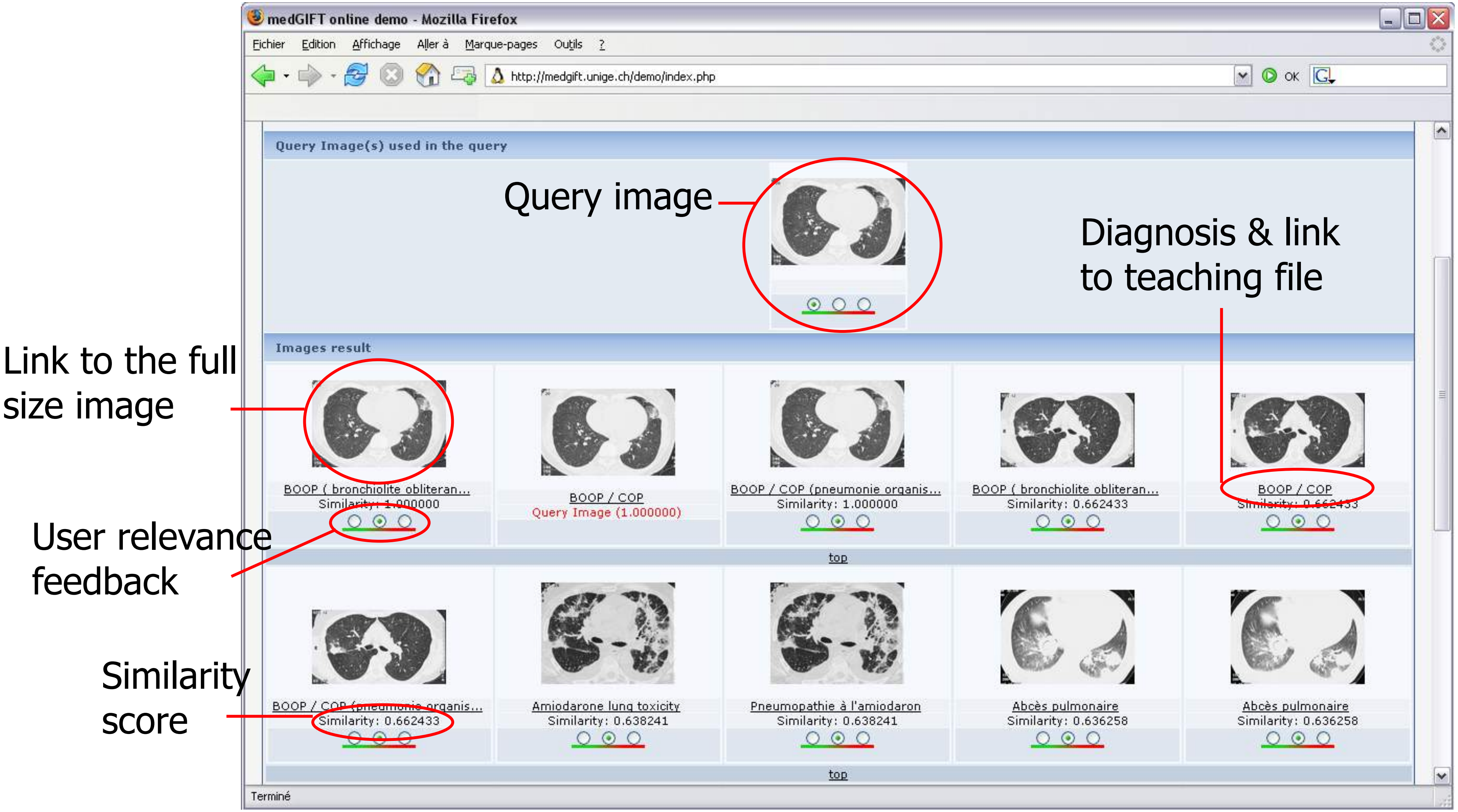
88 year old man:

- lower mean density
- pre-fibrotic lesions

What is in the image and what is it about?

Hes·so // VALAIS WALLIS
School of Management & Tourism

An early user interface



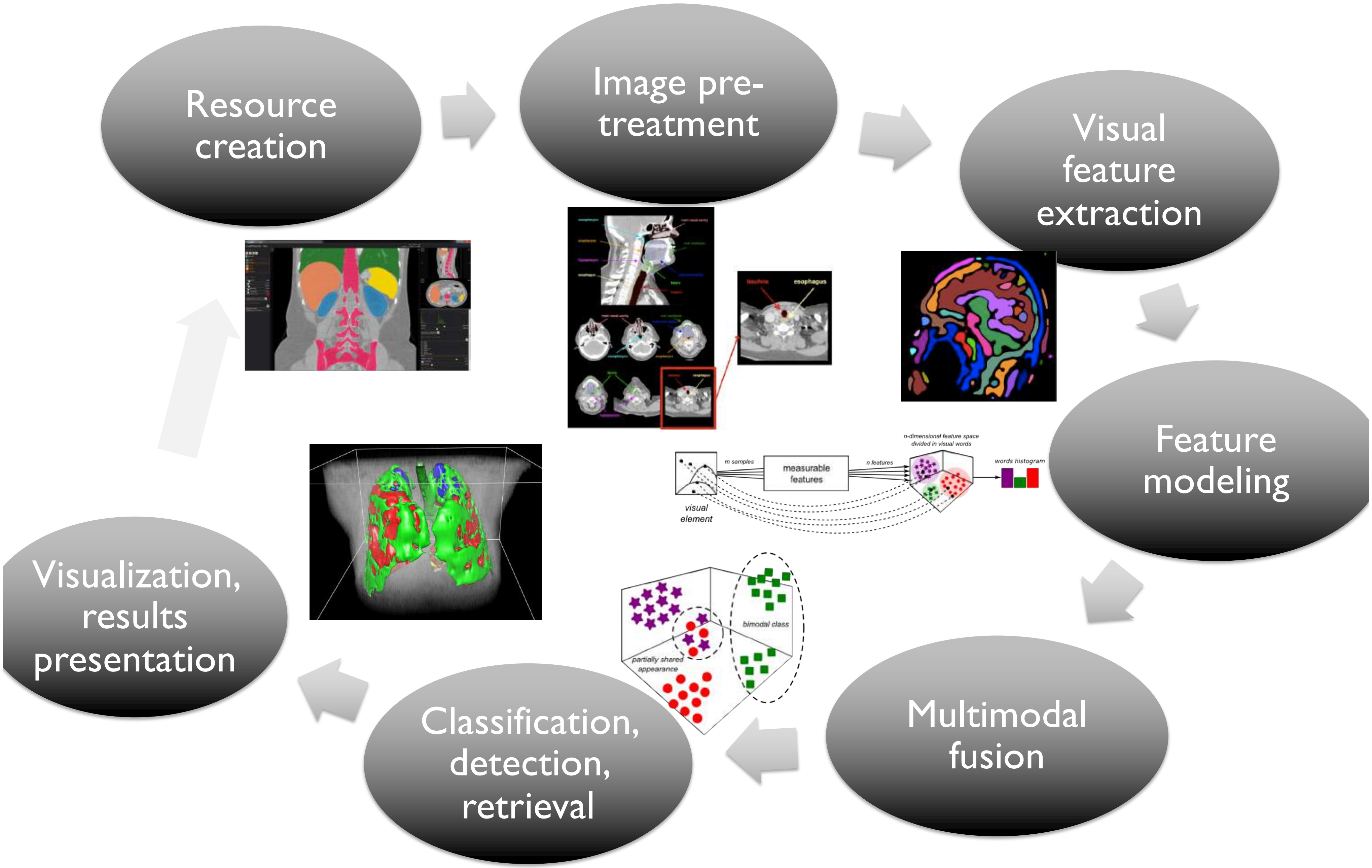
Case-based retrieval

- Find similar cases in a database of cases (teaching file or clinical archive)
  - Using all available data
  - Possibly even longitudinal data
- Challenge to integrate sources and define importance of each source
  - Learning such complex models requires to judge the performance of the system for a large number of cases
- More realistic in many scenarios than image-based retrieval
  - ROI-retrieval might be even more important

Visual Question Answering (VQA)

- Specific questions regarding an existing image
  - In this case on medical images
- Relies on the availability of images and training data with questions and answers to learn models that generalize to new questions
  - Much more complex than textual QA
- 2018 started a task on medical VQA (ImageCLEF)
  - 8 groups submitted results
  - Interest was strong for a challenging task
  - For 2019 a better corpus is available

Steps in visual decision support

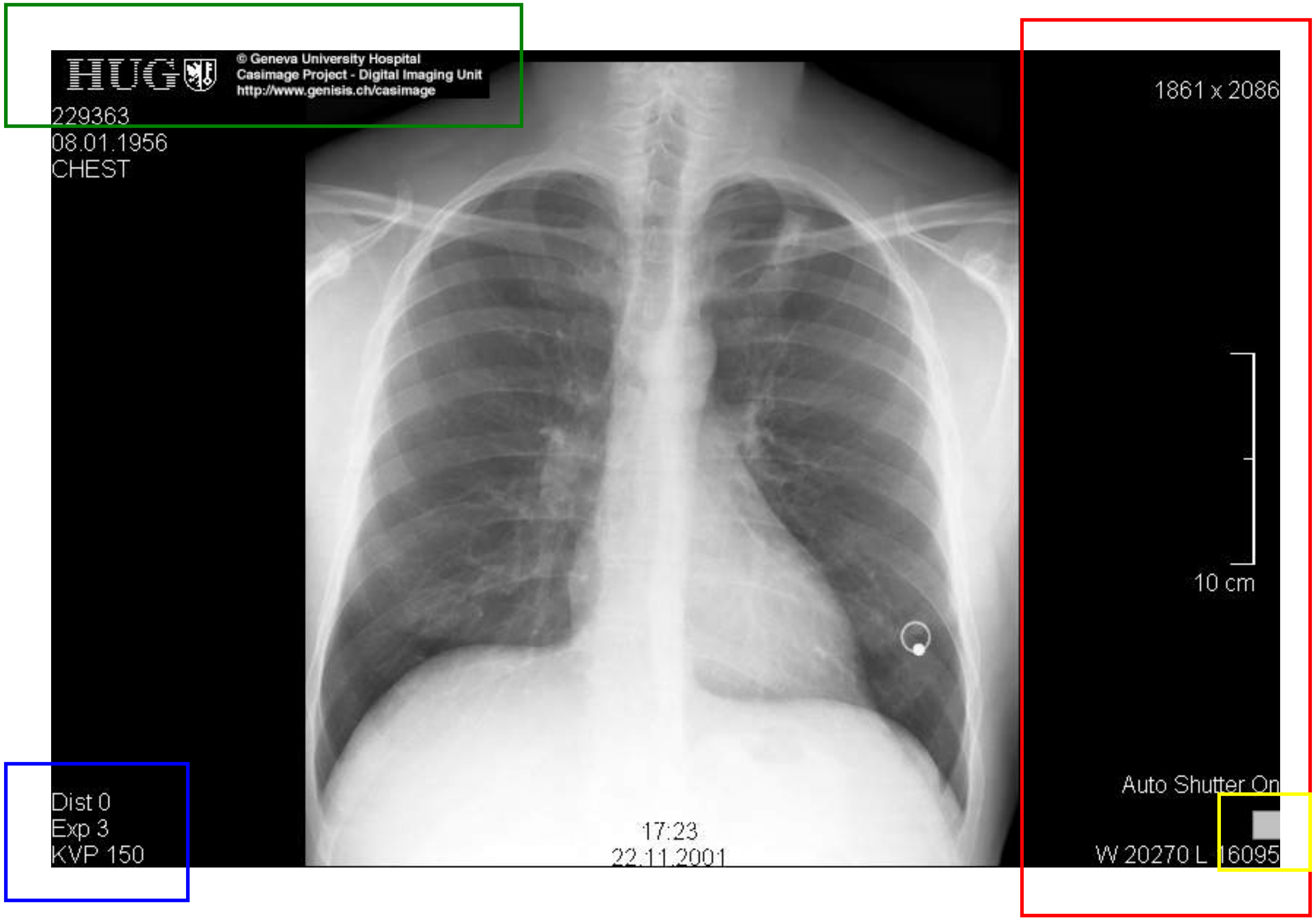


Visual challenges

Hes·so VALAIS WALLIS
School of Management & Tourism

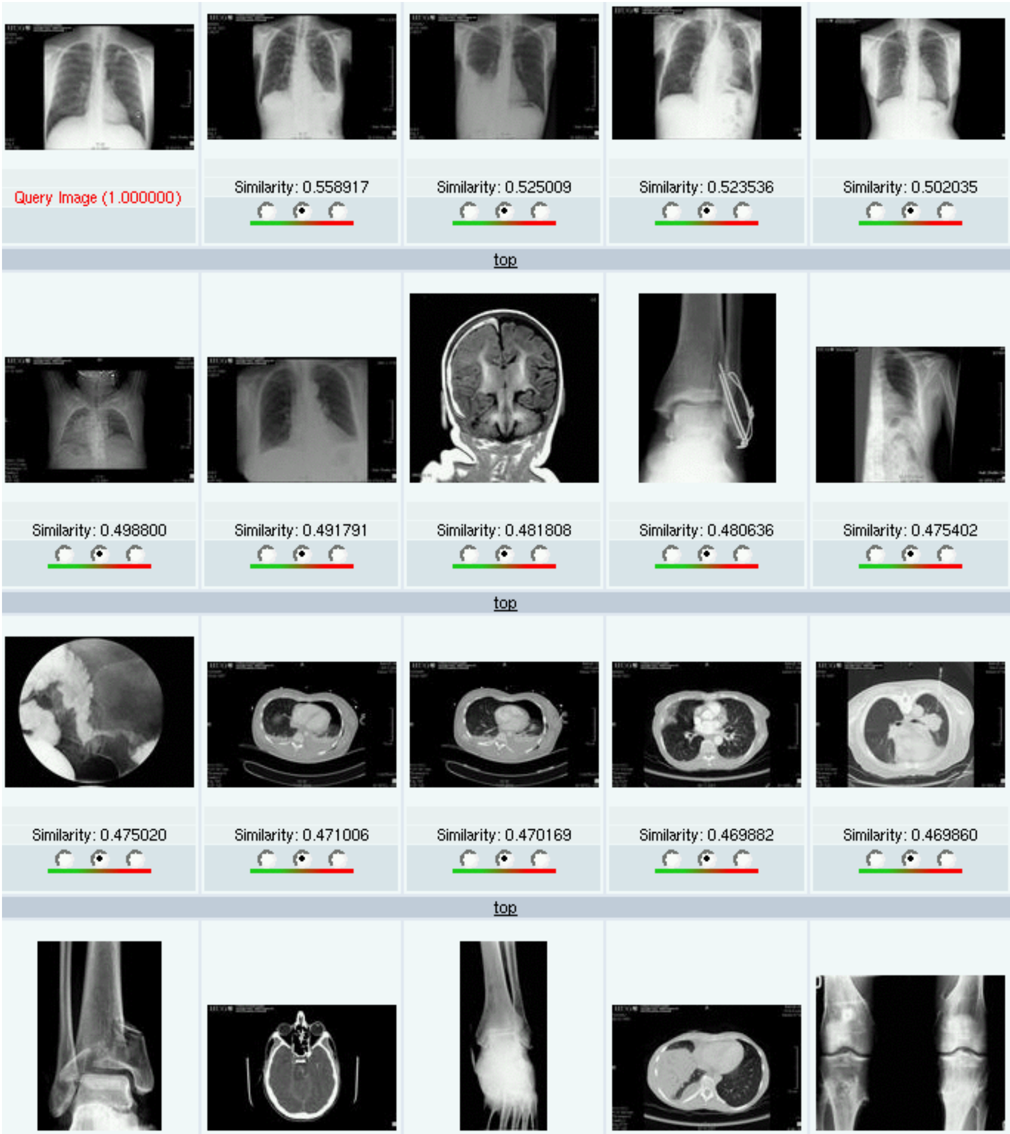
logo

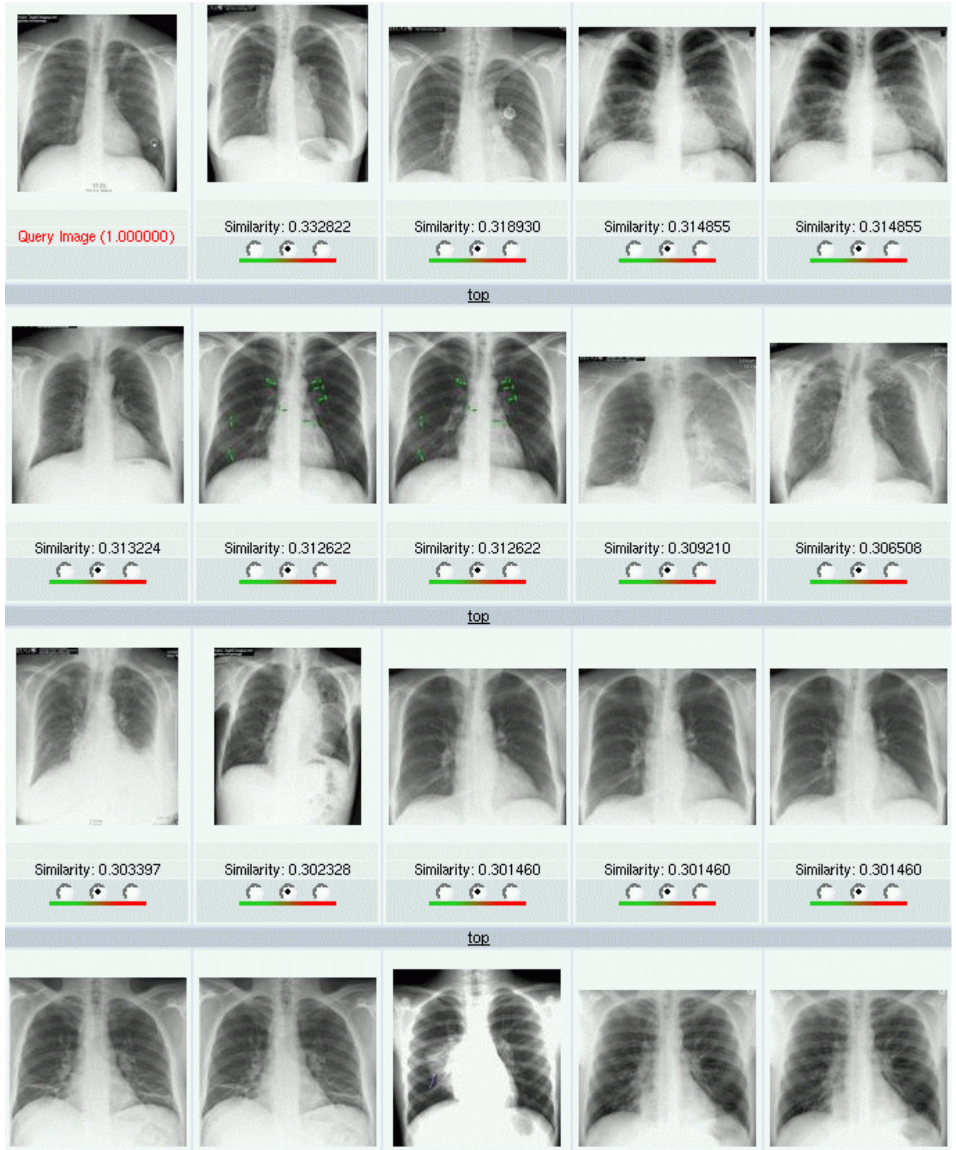
text



specific problems

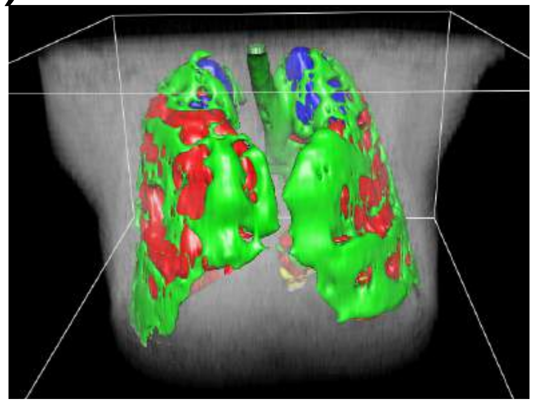
Large parts without information


Query Image (1.000000)
Similarity: 0.558917
Similarity: 0.525009
Similarity: 0.523536
Similarity: 0.502035
top
Similarity: 0.498800
Similarity: 0.491791
Similarity: 0.481808
Similarity: 0.480636
Similarity: 0.475402
top
Similarity: 0.475020
Similarity: 0.471006
Similarity: 0.470169
Similarity: 0.469882
Similarity: 0.469860
top

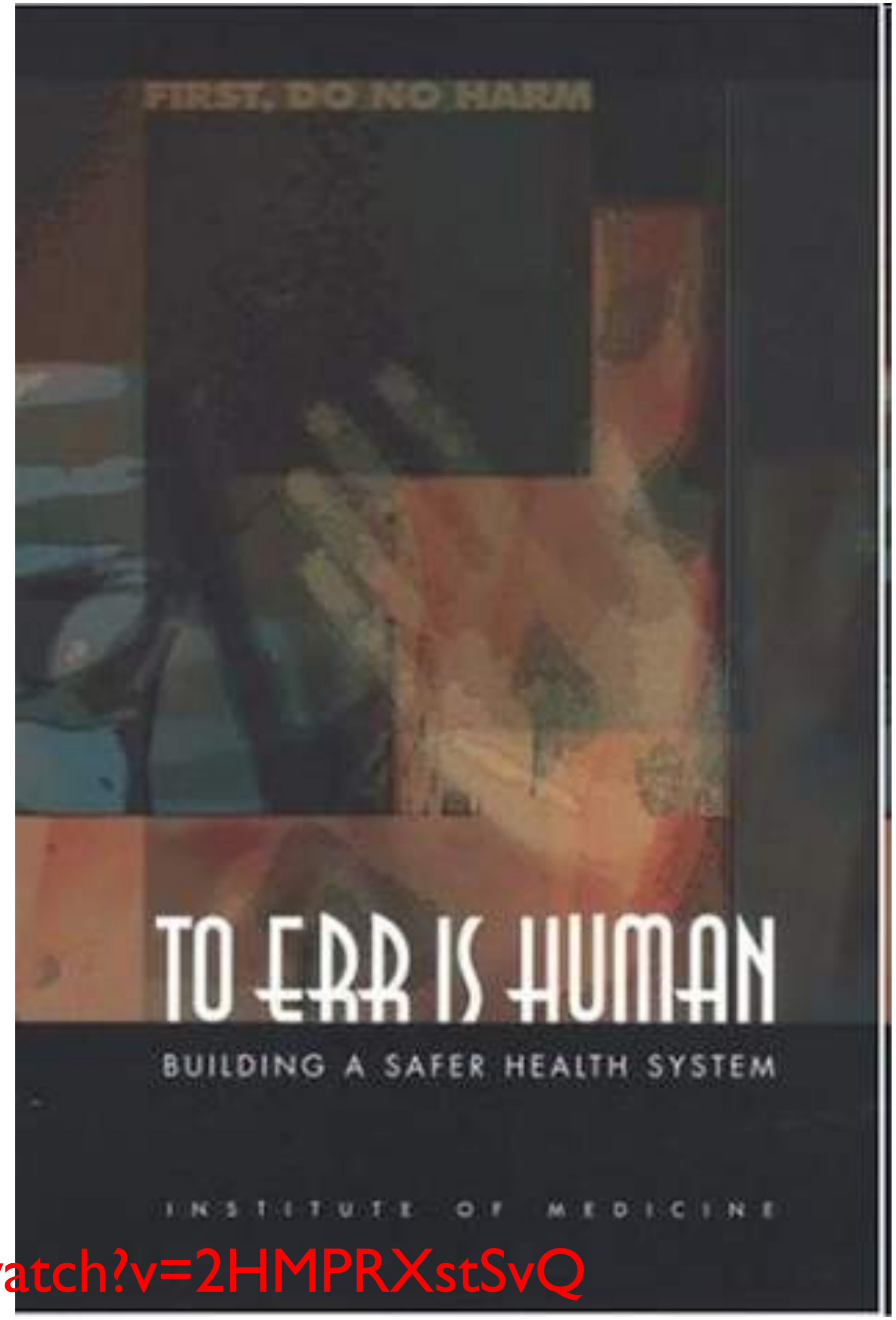

Query Image (1.000000)
Similarity: 0.332822
Similarity: 0.318930
Similarity: 0.314855
Similarity: 0.314855
top
Similarity: 0.313224
Similarity: 0.312622
Similarity: 0.312622
Similarity: 0.309210
Similarity: 0.306508
top
Similarity: 0.303397
Similarity: 0.302328
Similarity: 0.301460
Similarity: 0.301460
Similarity: 0.301460
top

Imaging applications: CADx, CADe

- Computer-Aided Diagnosis (CADx)
- Computer-Aided Detection (CADe)
  - Finding locations of lesions
- Computer-Aided Decision Support
- Many tools are in this area
  - Finding similar patients (retrieval)
  - Finding criteria for or against specific diseases (rules)
  - Prediction of findings such as tissue types
  - Predicting a diagnosis using machine learning

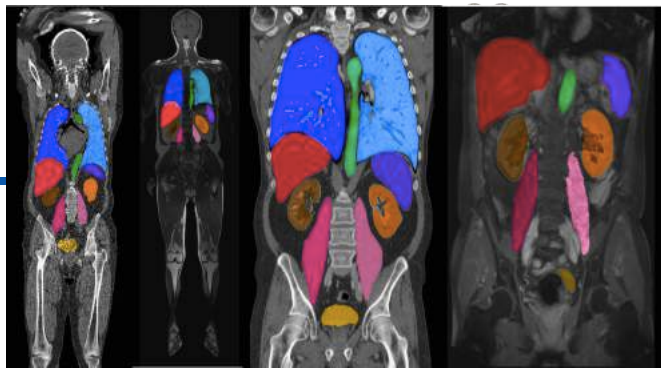


Why decision support?

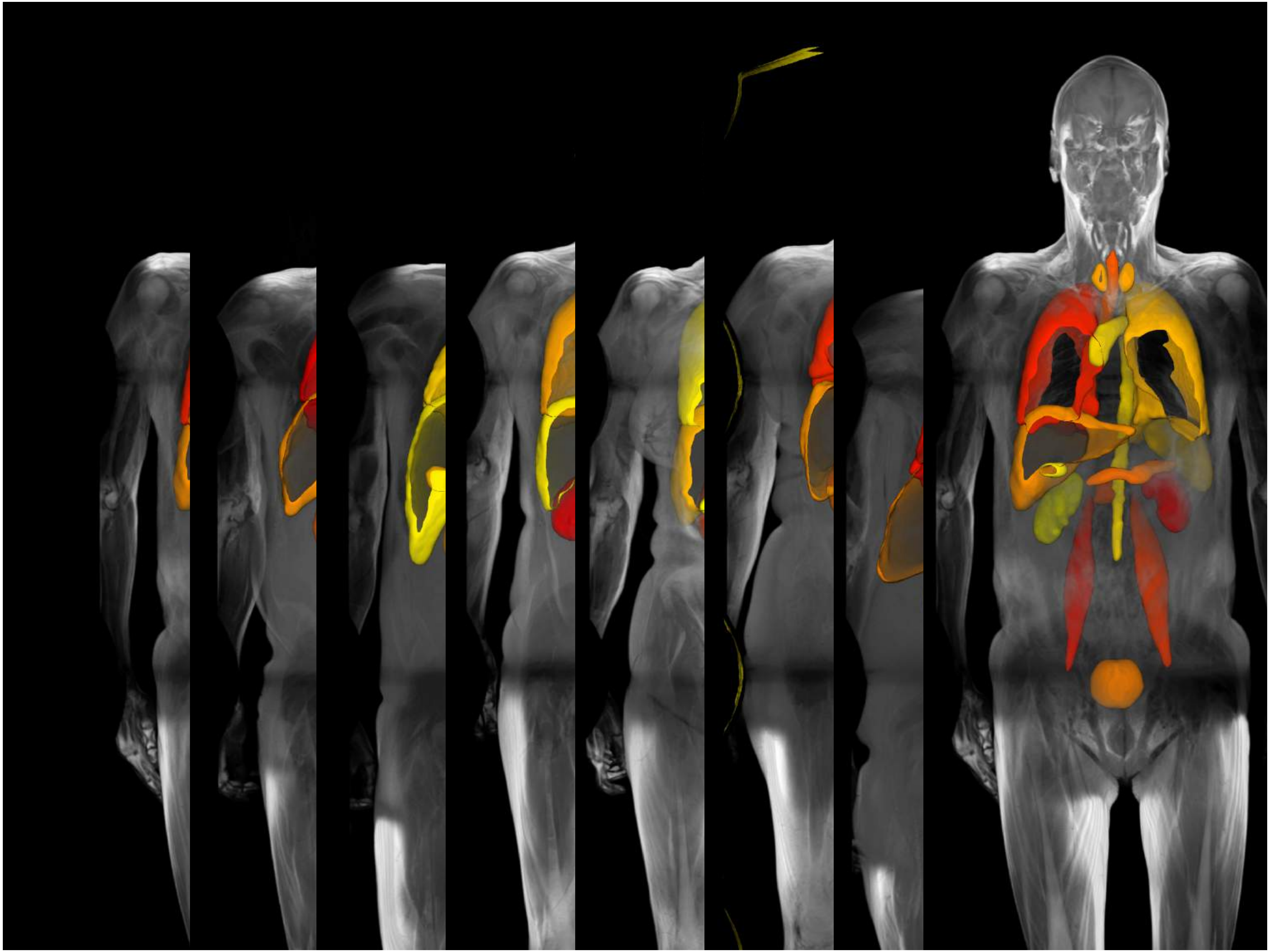


Geoff Hinton on radiology
https://www.youtube.com/watch?v=2HMPRXstSvQ

Annotating medical image regions



- Even "clear" annotations like entire organs are subjective
  - Even the same person at different moments annotates differently
  - Inter-annotator disagreement measures subjectivity
- Clear guidelines can create better annotations
  - Semi-automatic tools can harmonize but create a bias
- Automatic segmentations often try to model a human annotator closely
  - Data-driven vs. model-driven
  - Organs, lesions, landmarks, other structures

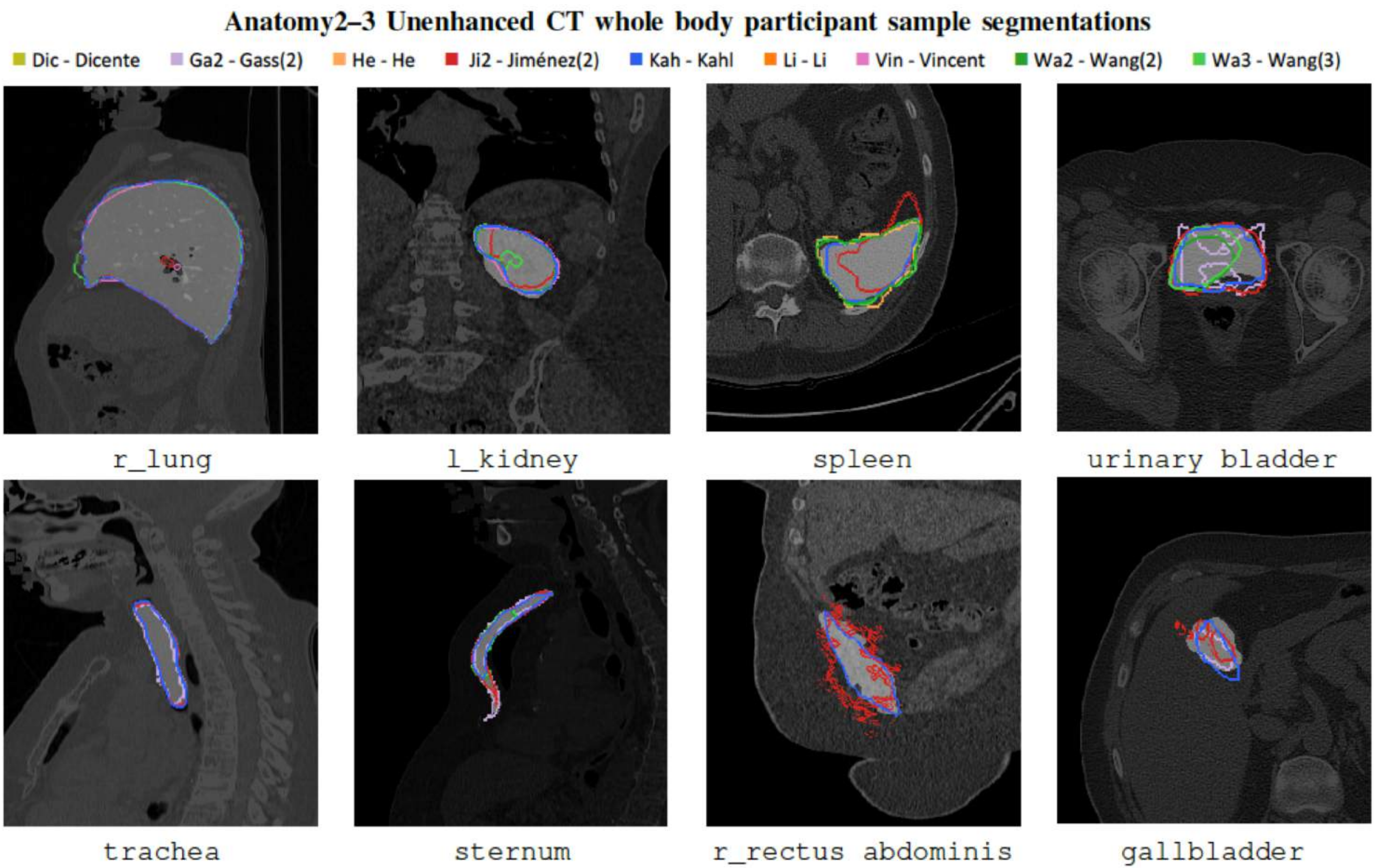



DOES IT HURT
WHEN I PRESS HERE?
PESSIN

Automatic segmentation techniques

- Based purely on grey levels/morphology (lung)
  - Requiring often some cleaning (on borders, holes)
- Simple region growing based on grey levels/texture
  - Based on a seed point in the structure
- Atlas-based annotations
  - Use manually annotated examples and register them to a new case
- Model-based segmentations using shape priors
- Deep learning-based segmentations

Comparing several organs/tools



Oscar Alfonso Jiménez del Toro et al., Cloud–based Evaluation of Organ Segmentation and Landmark Detection Algorithms: VISCERAL Anatomy Benchmarks, *IEEE Transactions on Medical Imaging*, 2016.

Detection of regions of interest

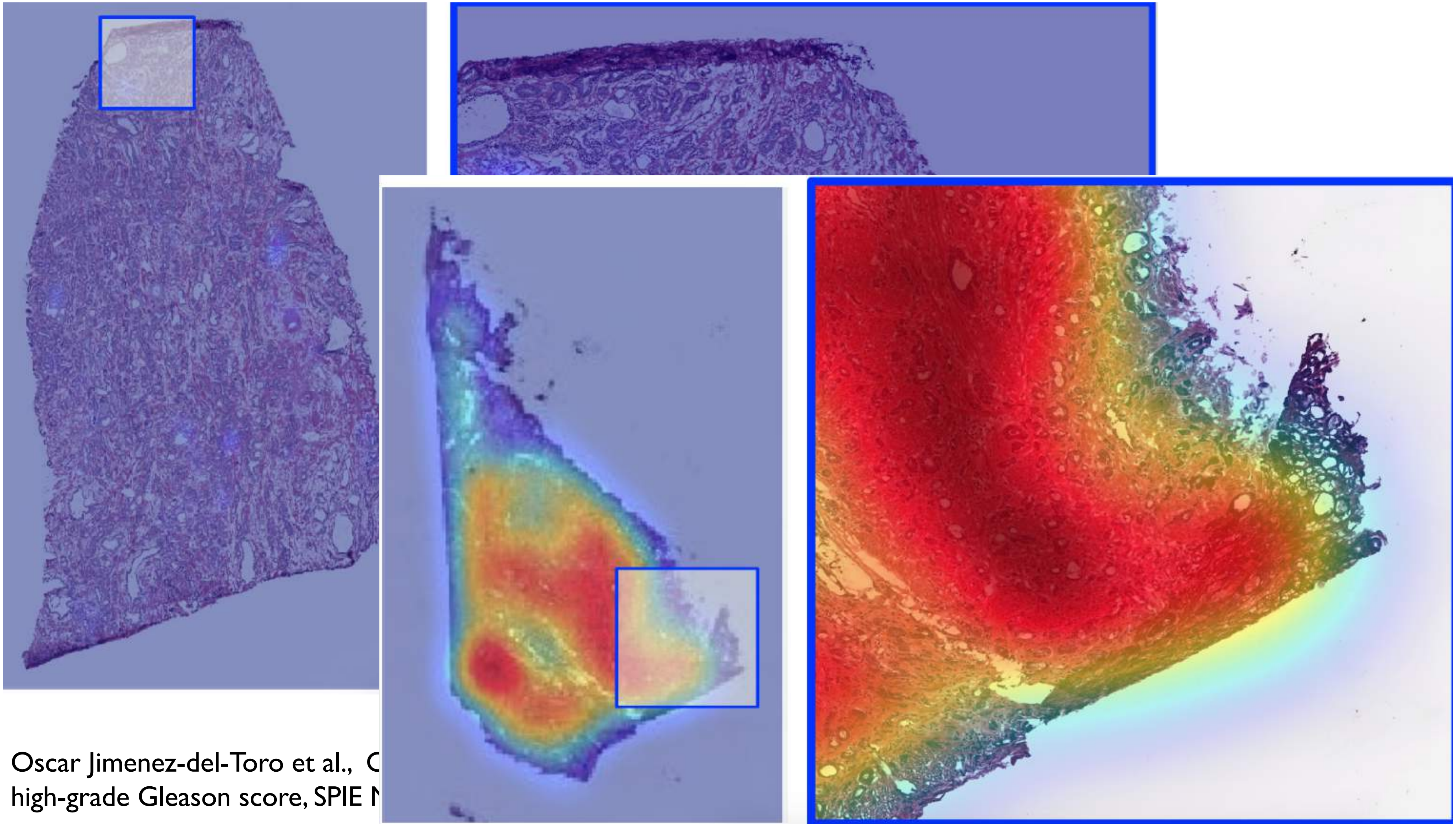
- Avoids having to read the entire image data (efficiency) or missing important findings
  - Finding a needle in a haystack
  - Incidental findings, how to deal with them
- Make results less observer-dependent/subjective
  - Extract quantitative criteria locally
- Input for a more detailed analysis of regions of interest, removing noise
  - Comparing apples with apples …
  - Automating the analysis workflow

Abnormality detection

- First question in a medical case is if there is a disease or if the person is healthy (normal)
- Most cases are (fortunately) normal
  - In screening of risk groups maybe 1% is positive, less in non risk groups (limit also false positives)
- Most images in a pathologic case are normal
  - A chest CT may contain 500 images and a tumor may be visible in 5-6 of them
- … and most pixels in an abnormal slice are still normal … only very few are usually part of the tumor

Examples: ROI detection (DL)

- Find areas with high vs. low Gleason grades



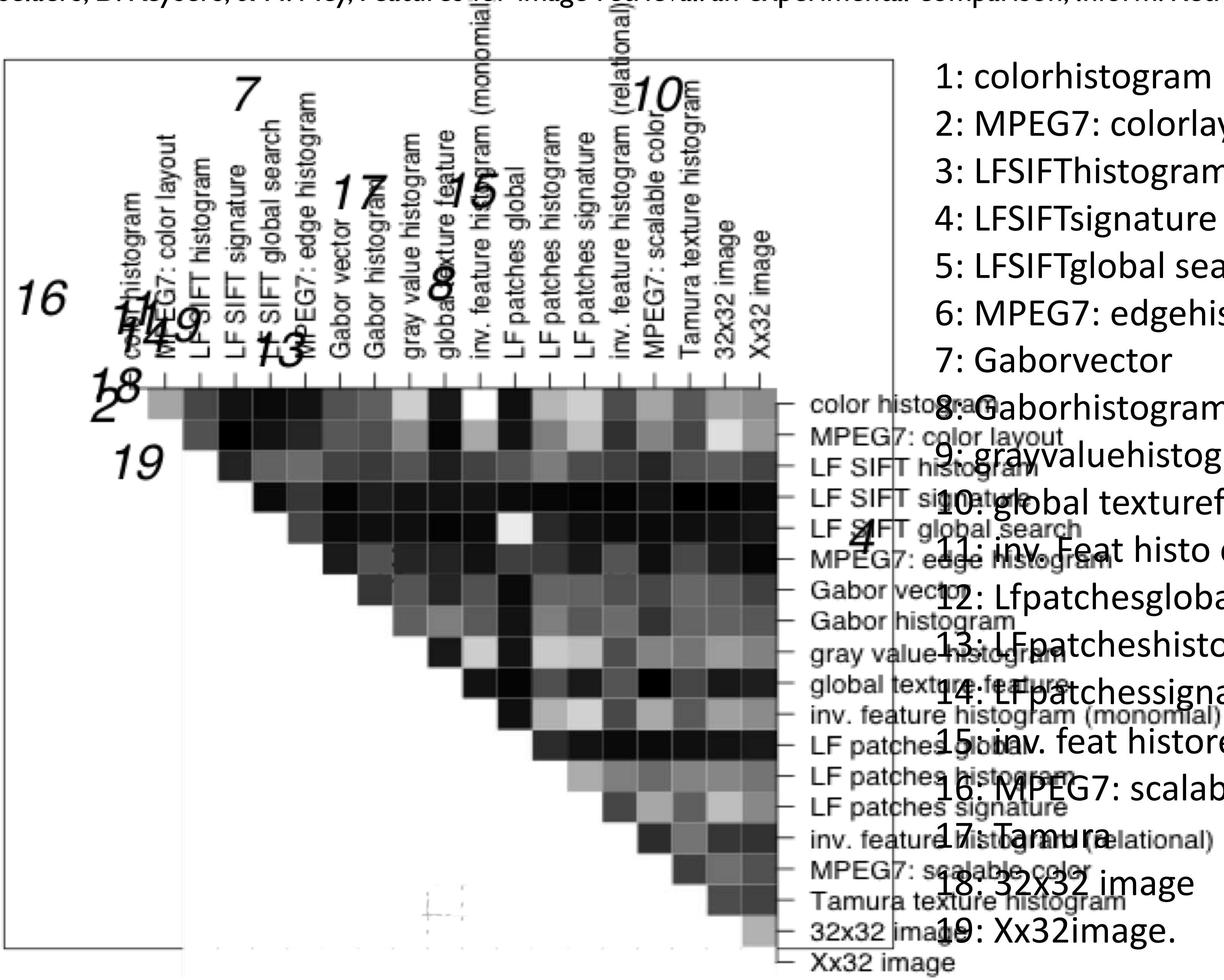
Oscar Jimenez-del-Toro et al., C
high-grade Gleason score, SPIE N

Types of visual features

- Handcrafted vs. partially learned vs. fully learned
  - Deep learning vs. traditional approaches
- Classifications of visual features
  - Low level vs. mid level vs. semantic/high-level
    - Higher levels via feature modeling (visual words) or latent semantic techniques, sometimes matching words and pictures
- Type of information that is modeled
  - Shape vs. grey level/color vs. texture
- Local vs. global features
  - Local based on segmentation or partitioning
- 2D vs. 3D vs. nD (3D +time, protocols)

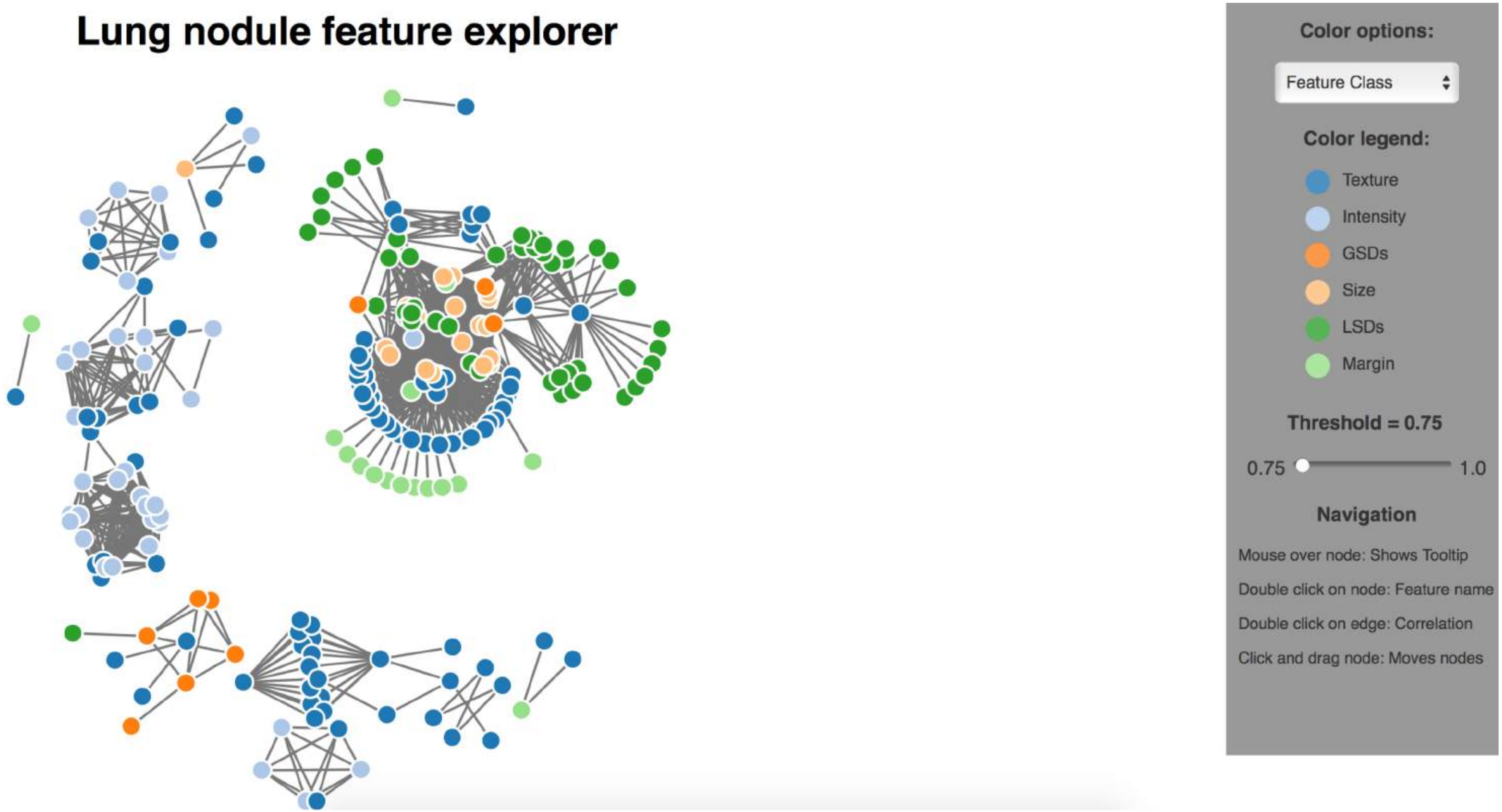
Correlation between Features

T. Deselaers, D. Keysers, & H. Ney, Features for image retrieval: an experimental comparison, Inform. Retrieval (2008) 11: 77.

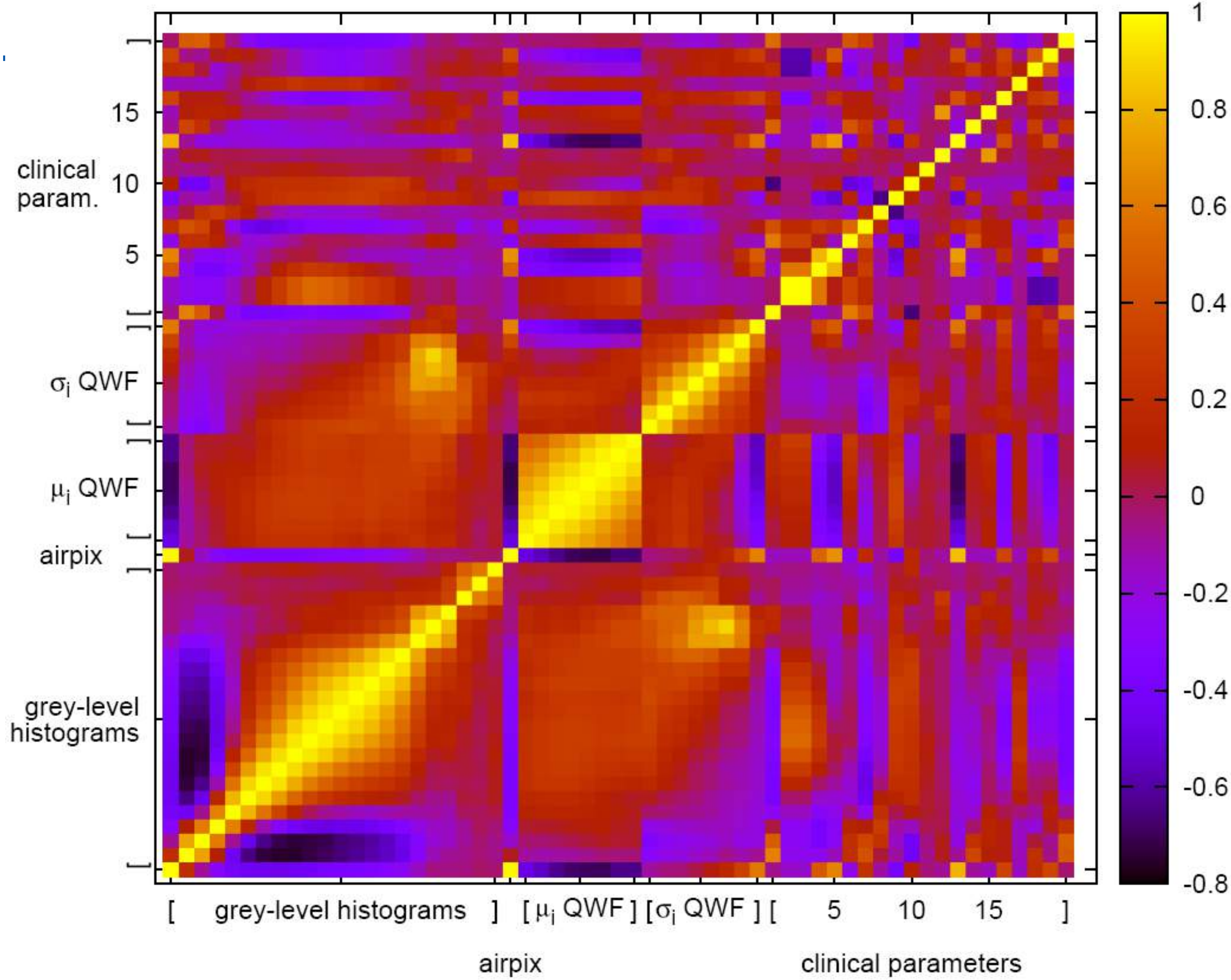


1: colorhistogram
2: MPEG7: colorlayout
3: LFSIFThistogram
4: LFSIFTsignature
5: LFSIFTglobal search
6: MPEG7: edgehistogram,
7: Gaborvector
8: Gaborhistograms
9: grayvaluehistogram
10: global texturefeature,
11: inv. Feat histo color
12: Lfpatchesglobal
13: LFpatcheshistogram
14: LFpatchessignature,
15: inv. feat historel
16: MPEG7: scalablecolor,
17: Tamura
18: 32x32 image
19: Xx32image.

Feature exploration



http://lung-nodule-feature-explorer.s3-website-us-east-1.amazonaws.com

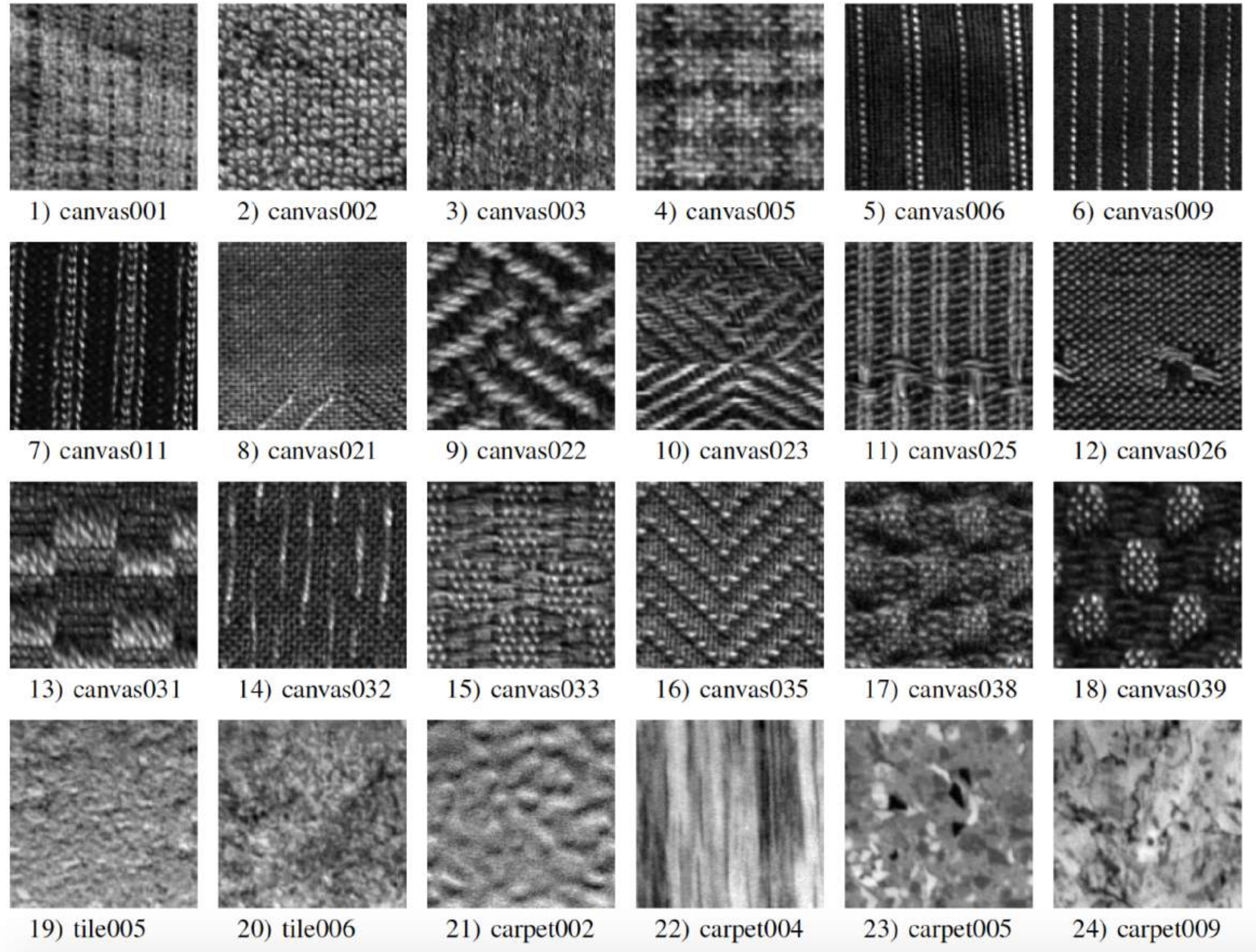

clinical param.
15
10
5
σi QWF
μi QWF
airpix
grey-level histograms
[grey-level histograms] [μi QWF] [σi QWF] [5 10 15]
airpix
clinical parameters
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8

Color/grey level features

- Resolution in the feature space is important for comparisons (particularly using histograms)
  - Level/window settings to concentrate on the interesting aspects (for example in the case of CT on the lung settings or abdomen settings)
- Different color spaces (RGB, HSV, CIE Lab/Luv) that correspond more or less to human perception can help
- Global counting or relationships of local colors
  - Often using a simple histogram intersection

Texture

- "the feel, appearance, or consistency of a surface or a substance"

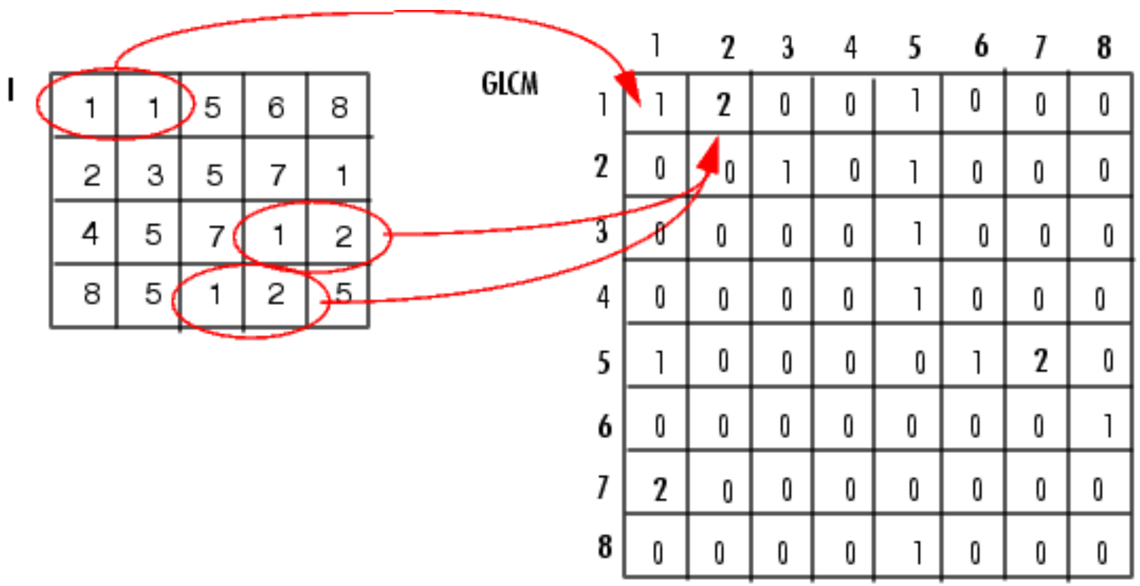


Texture features

- Tamura texture features (1978)
  - Coarseness – coarse vs. fine
  - Contrast – high vs. low
  - Directionality – directional vs. non-directional
  - Linelikeness – line-like vs. non-line-like
  - Regularity – regular vs. irregular
  - Roughness – rough vs. smooth
- Grey level co-occurrence matrices (Haralick, 1979)
  - Choice of scales, directions
- Gabor filters, Wavelets
  - Steerable wavelets allow for texture learning (Riesz)
- Local Binary Patterns (LBP)

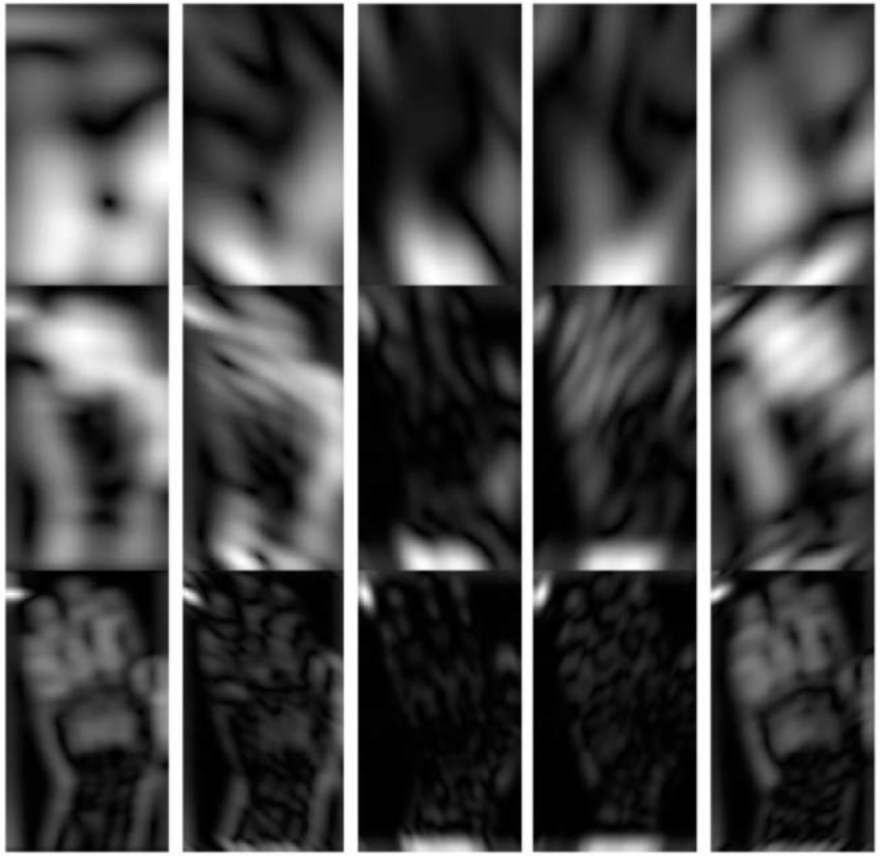
Gray-Level Co-Occurrence Matrices

- Statistical descriptor for texture properties of an image by comparing neighboring pixels
  - Direction and distance
    - Features extracted from several matrices in general
  - Extract features from matrix
    - Entropy
    - Contrast
    - Correlation
    - …



Gabor Features

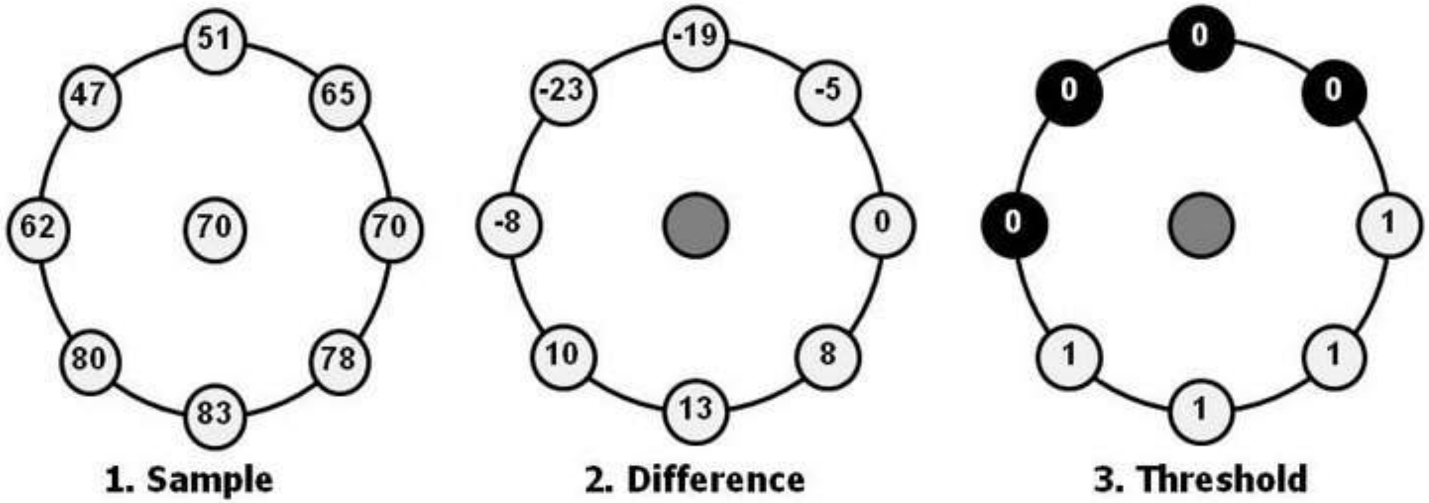
- Obtain several values per pixel denoting spatial frequencies and directions
  - Often four directions, three scales, similar to GLCM



Local binary patterns

The value of the LBP code of a pixel (x_c, y_c) is given by:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \qquad s(x) = \begin{cases} 1, if\ x \geq 0; \\ 0, otherwise. \end{cases}$$



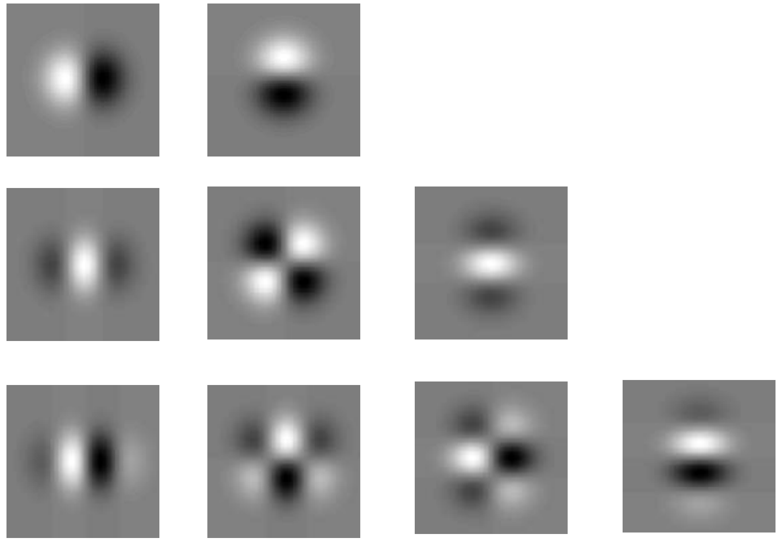
1*1 + 1*2 + 1*4 + 1*8 + 0*16 + 0*32 + 0*64 + 0*128 = **15**

4. Multiply by powers of two and sum

Ojala, T., Pietikäinen, M. and Harwood, D. (1996), A Comparative Study of Texture Measures with Classification Based on Feature Distributions. Pattern Recognition 19(3):51-59.

The Riesz transform

- The Riesz transform implements Nth-order directional derivatives:



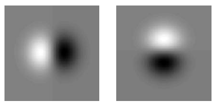
$N = 1$

$N = 2$

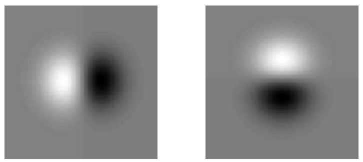
$N = 3$

- At multiple scales:

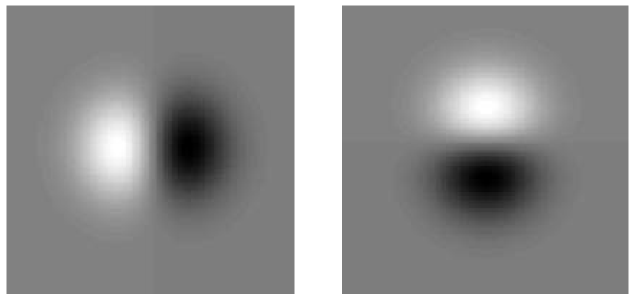
$N = 1$, 1st scale:



$N = 1$, 2nd scale:



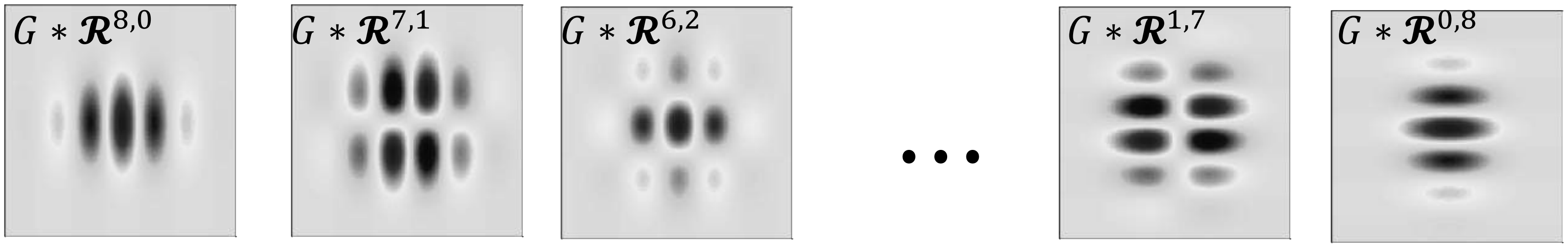
$N = 1$, 3rd scale:



The Riesz transform:

- A Riesz filterbank constitutes a dictionary of basic textures:

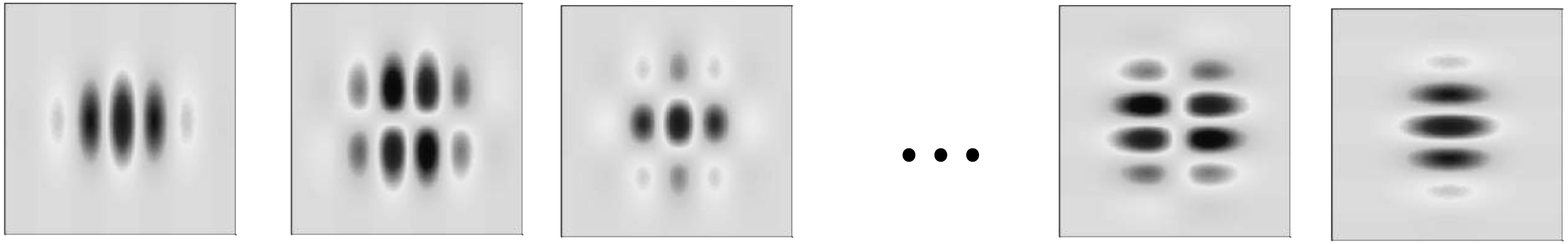
Riesz filterbank ($N = 8$)



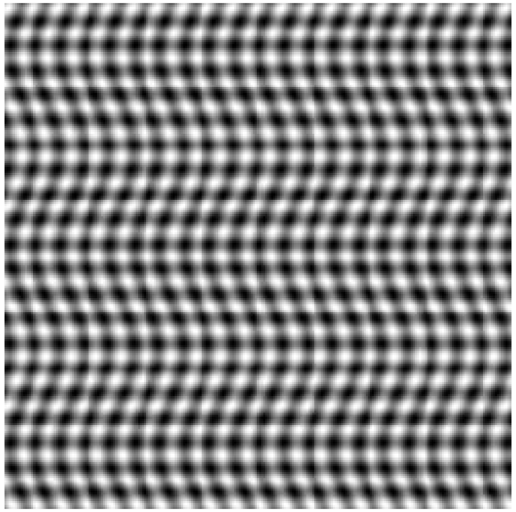
- Texture signatures are built using a linear combination of the Riesz templates

$$\Gamma_c{}^N = w_1 (G * \boldsymbol{\mathcal{R}}^{N,0})_{s_1} + w_2 (G * \boldsymbol{\mathcal{R}}^{N-1,1})_{s_1} + \cdots + w_{4N+4} (G * \boldsymbol{\mathcal{R}}^{0,N})_{s_4}$$

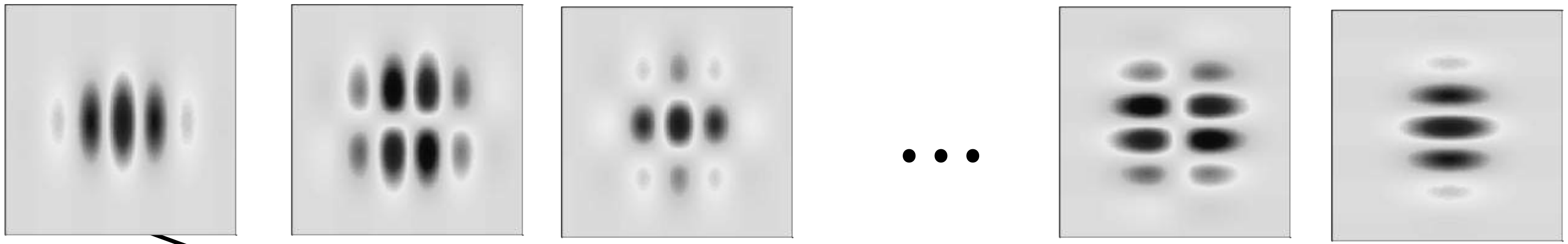
Riesz filterbank $(N = 8)$



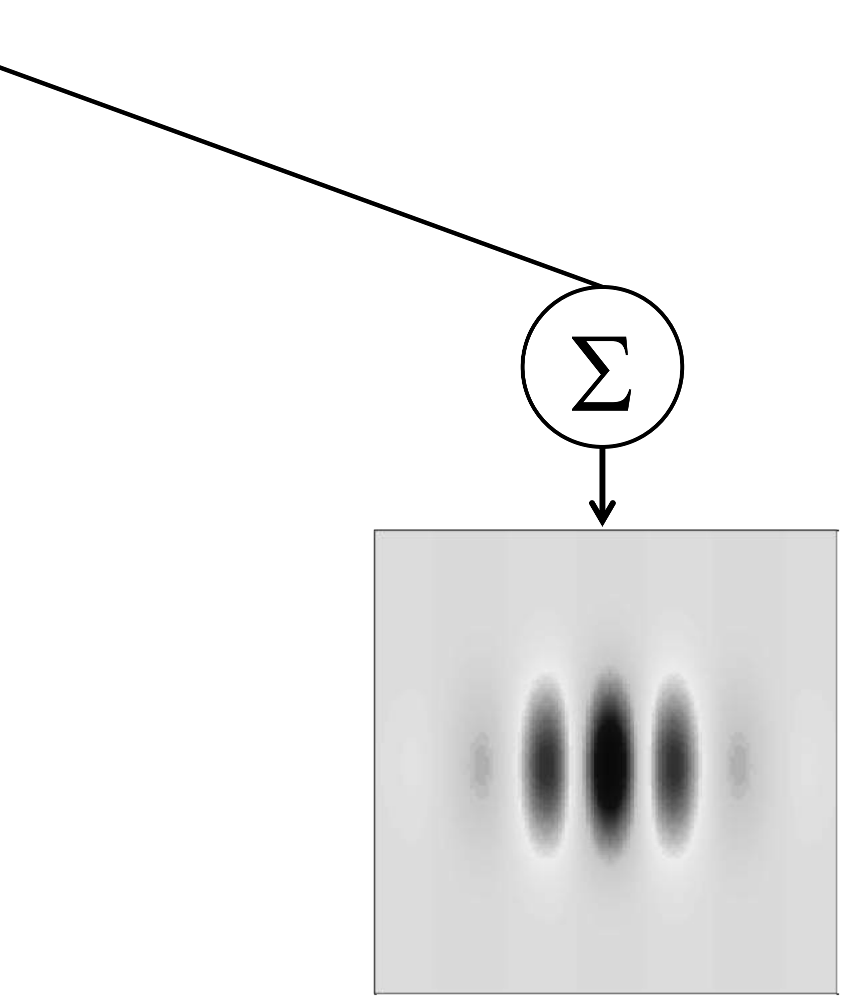
Texture to learn:



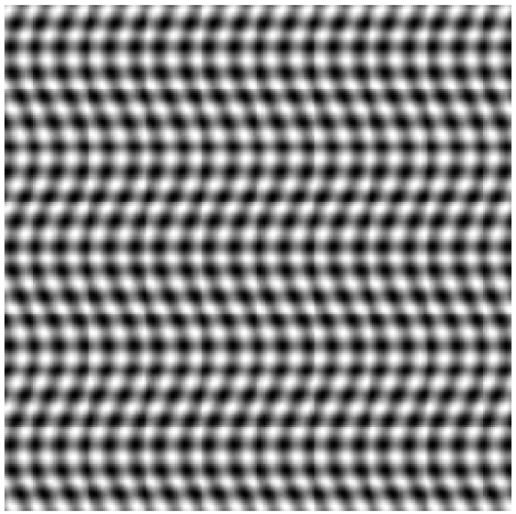
Riesz filterbank ($N = 8$)



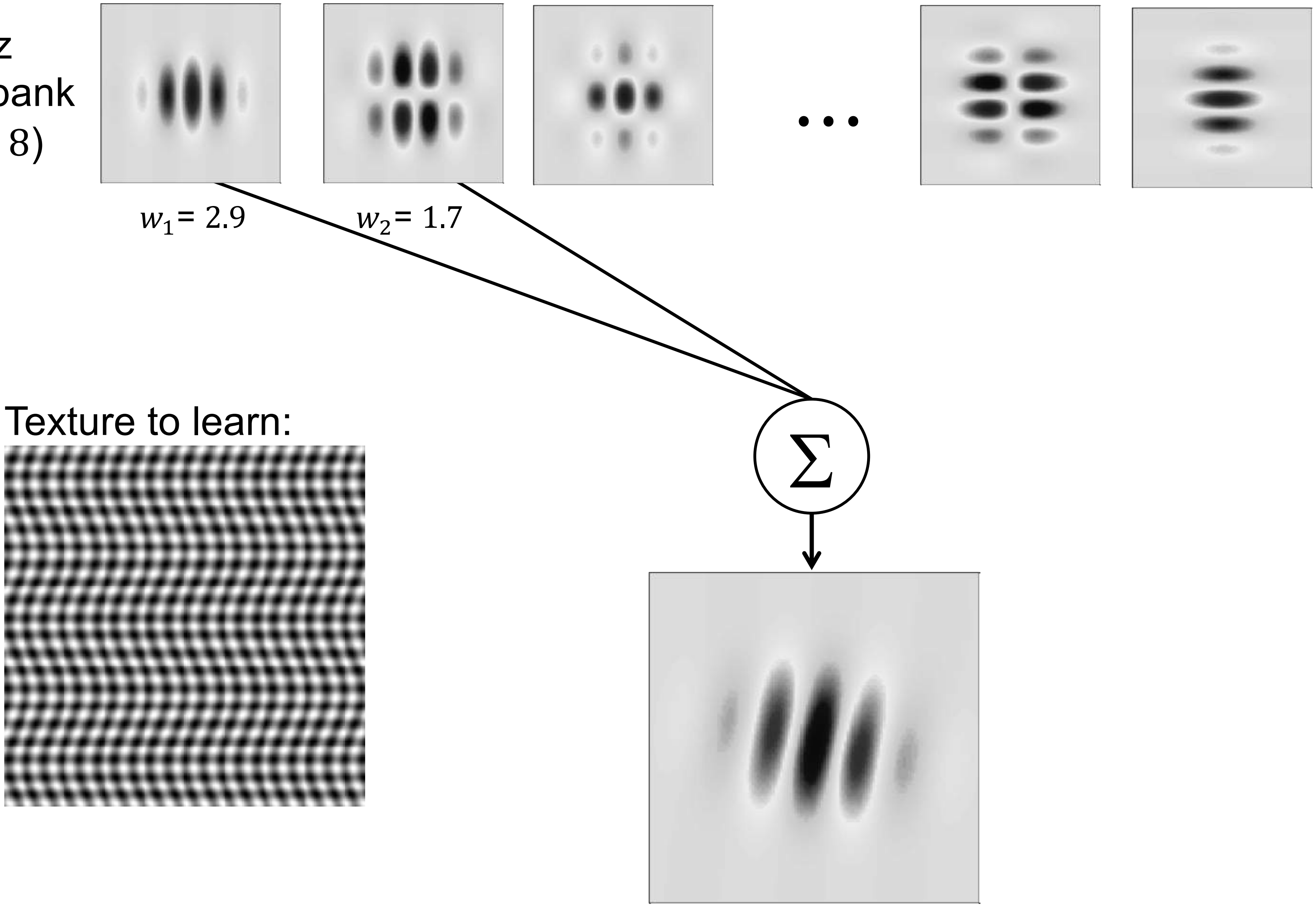
$w_1 = 2.9$



Texture to learn:

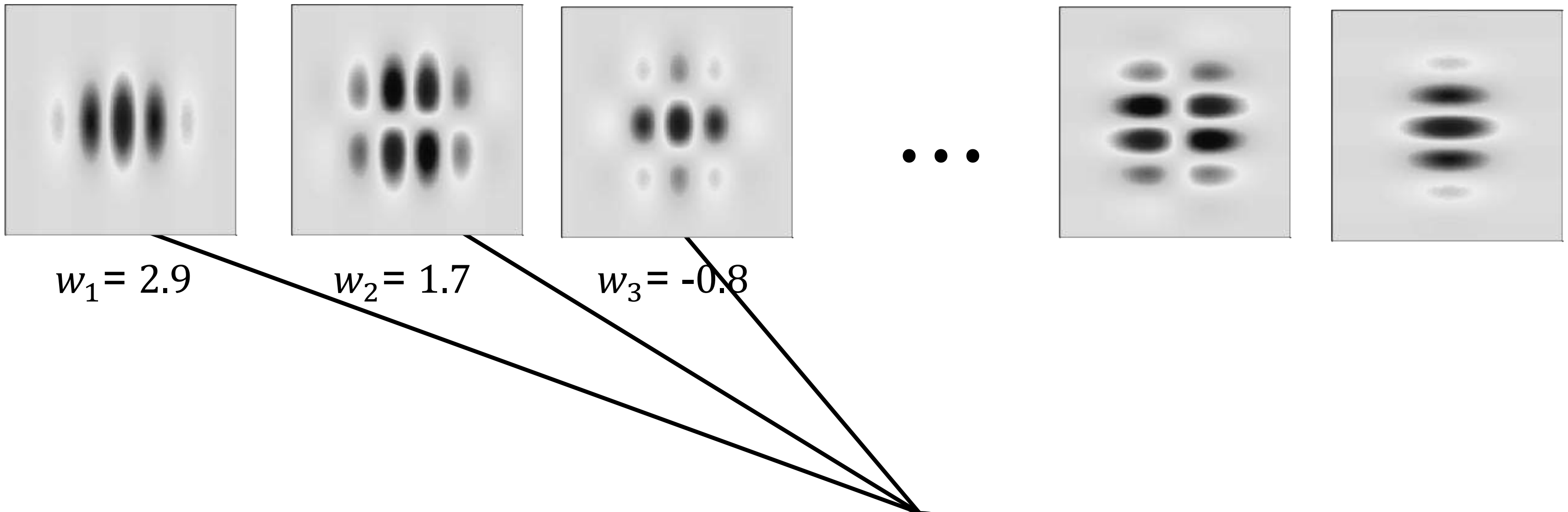


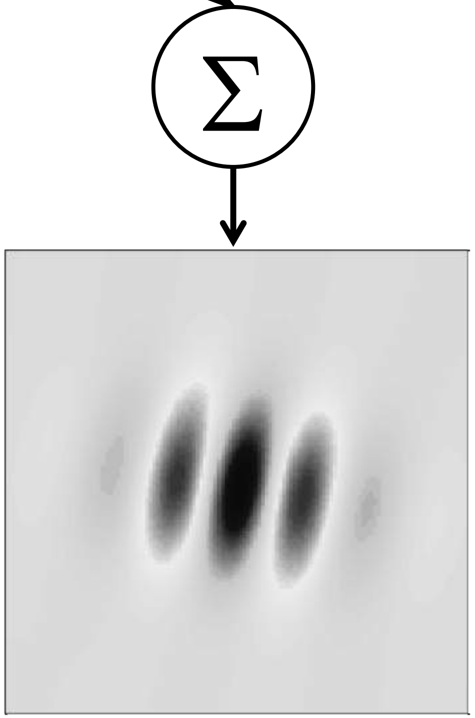
Riesz filterbank $(N = 8)$



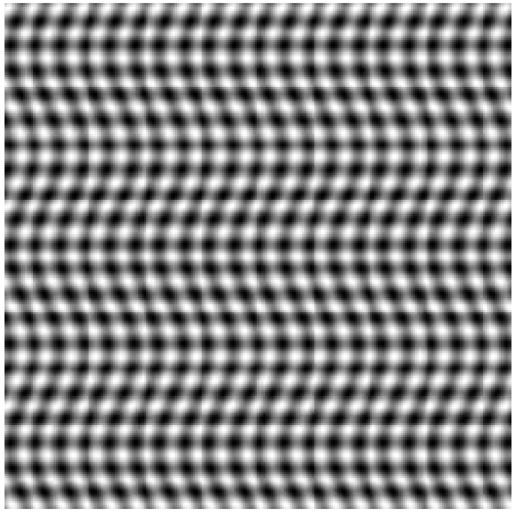
Texture to learn:

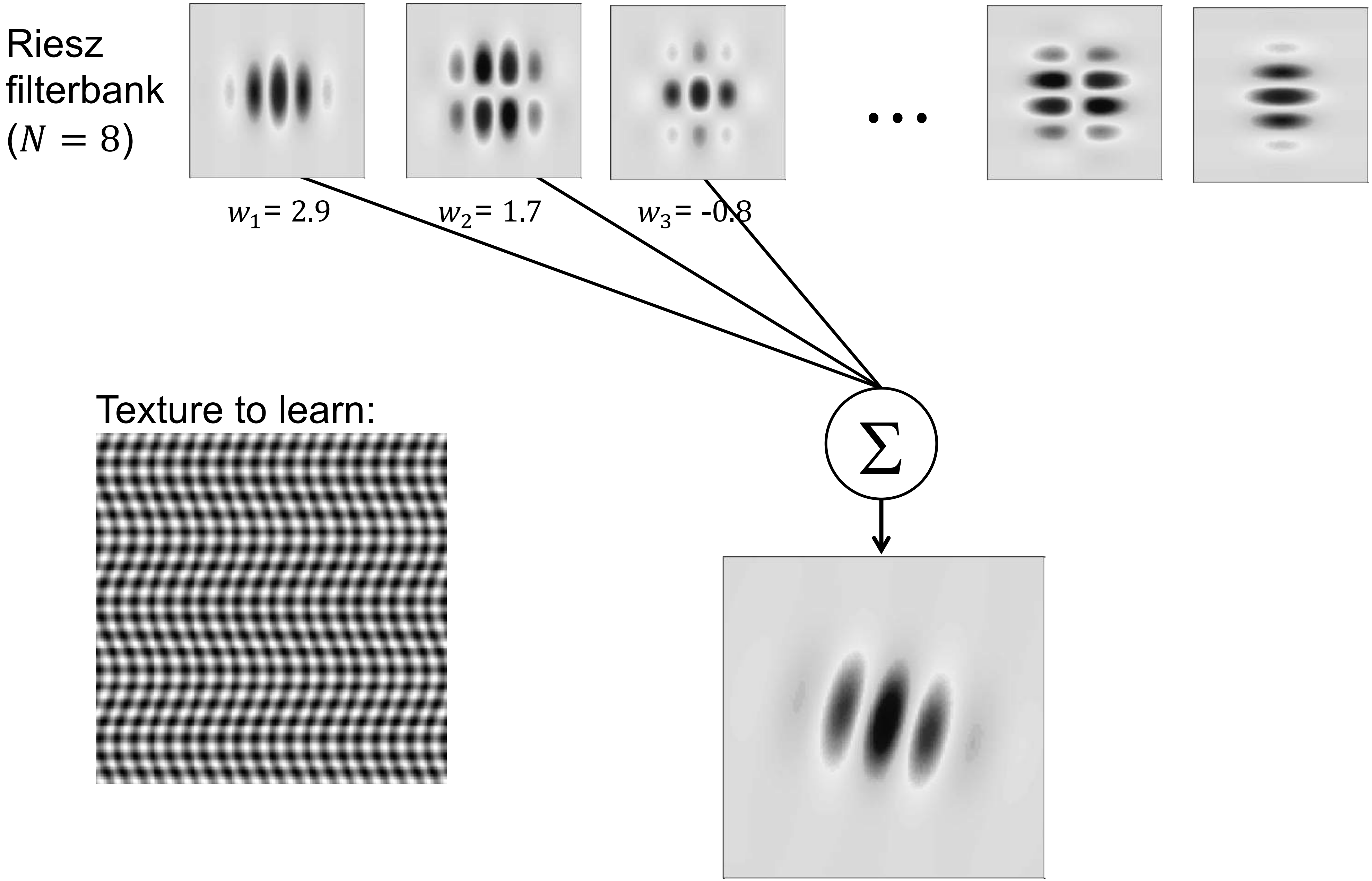
Riesz filterbank ($N = 8$)



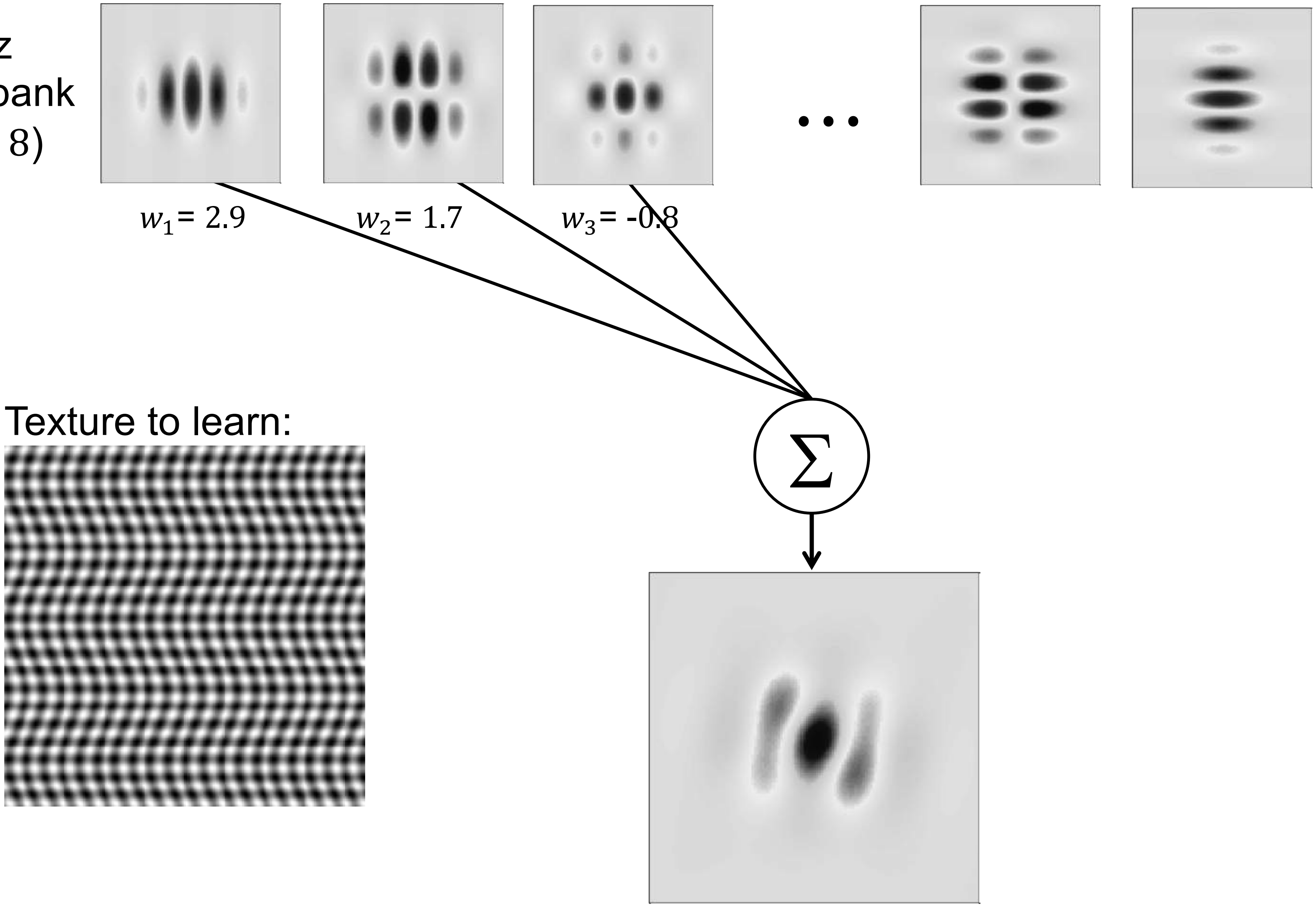


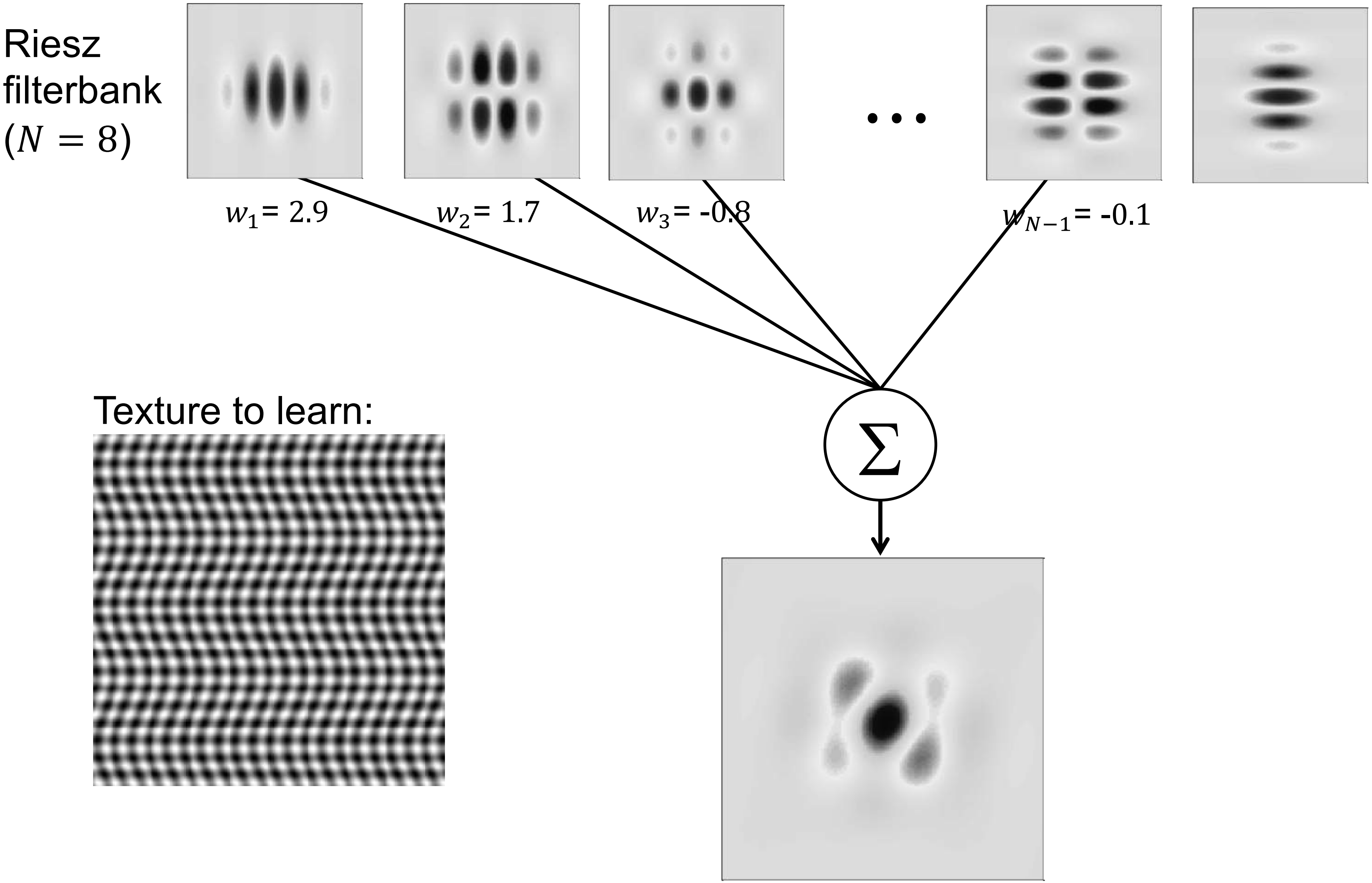
Texture to learn:



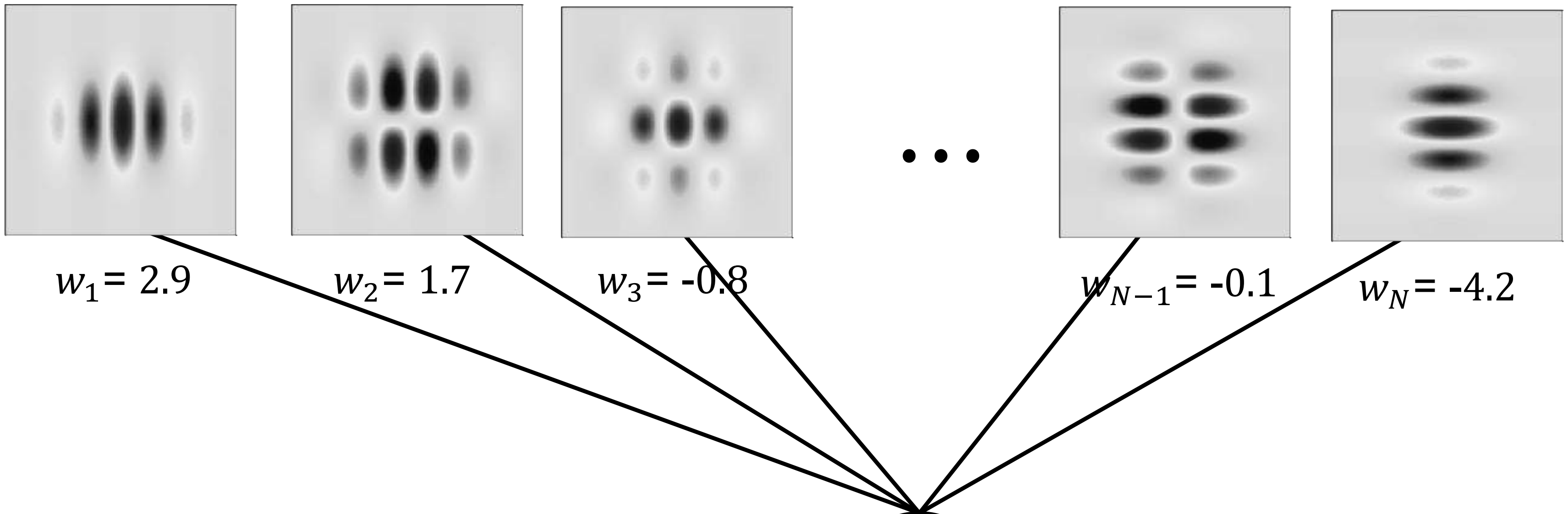

Riesz filterbank ($N = 8$)
w_1= 2.9
w_2= 1.7
w_3= -0.8
Texture to learn:
Σ

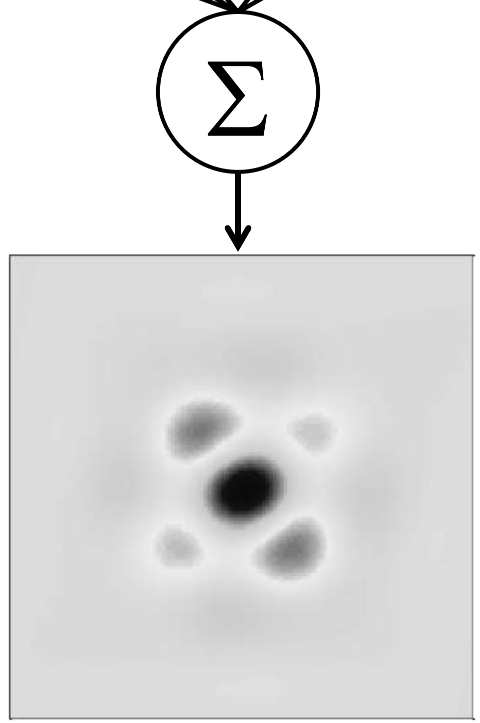
Riesz filterbank ($N = 8$)



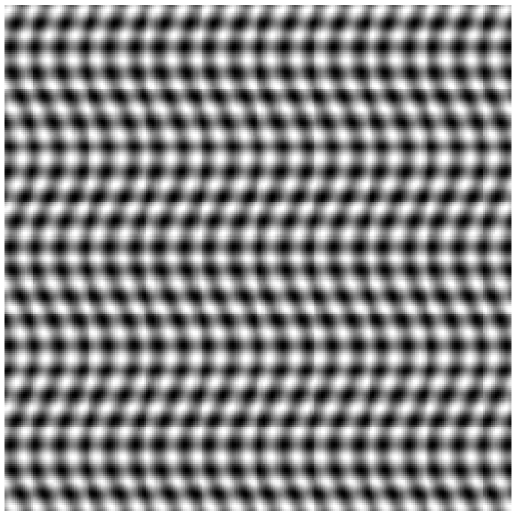

Riesz
filterbank
$(N = 8)$
$w_1 = 2.9$
$w_2 = 1.7$
$w_3 = -0.8$
$w_{N-1} = -0.1$
Σ
Texture to learn:

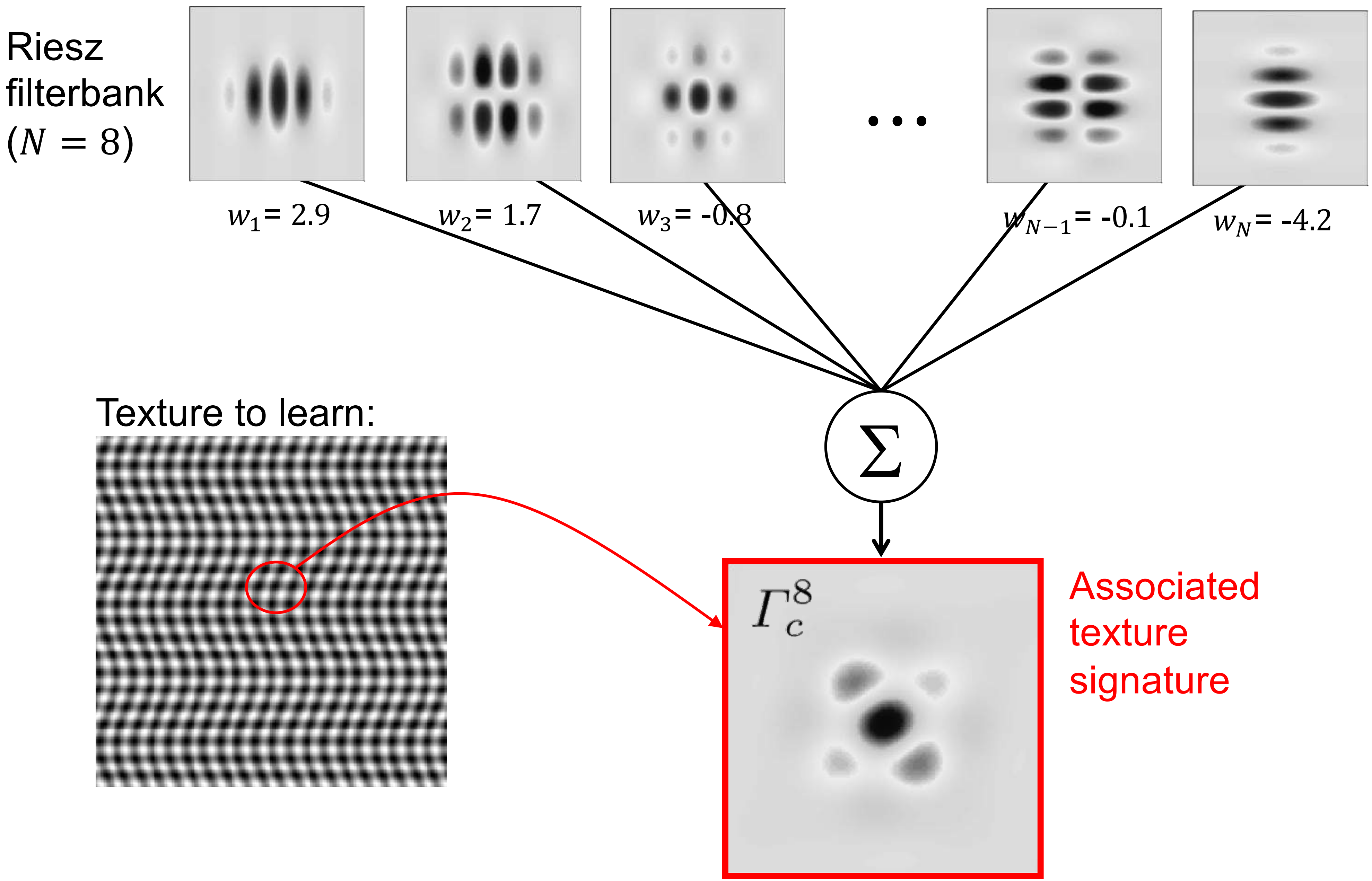
Riesz filterbank ($N = 8$)





Texture to learn:

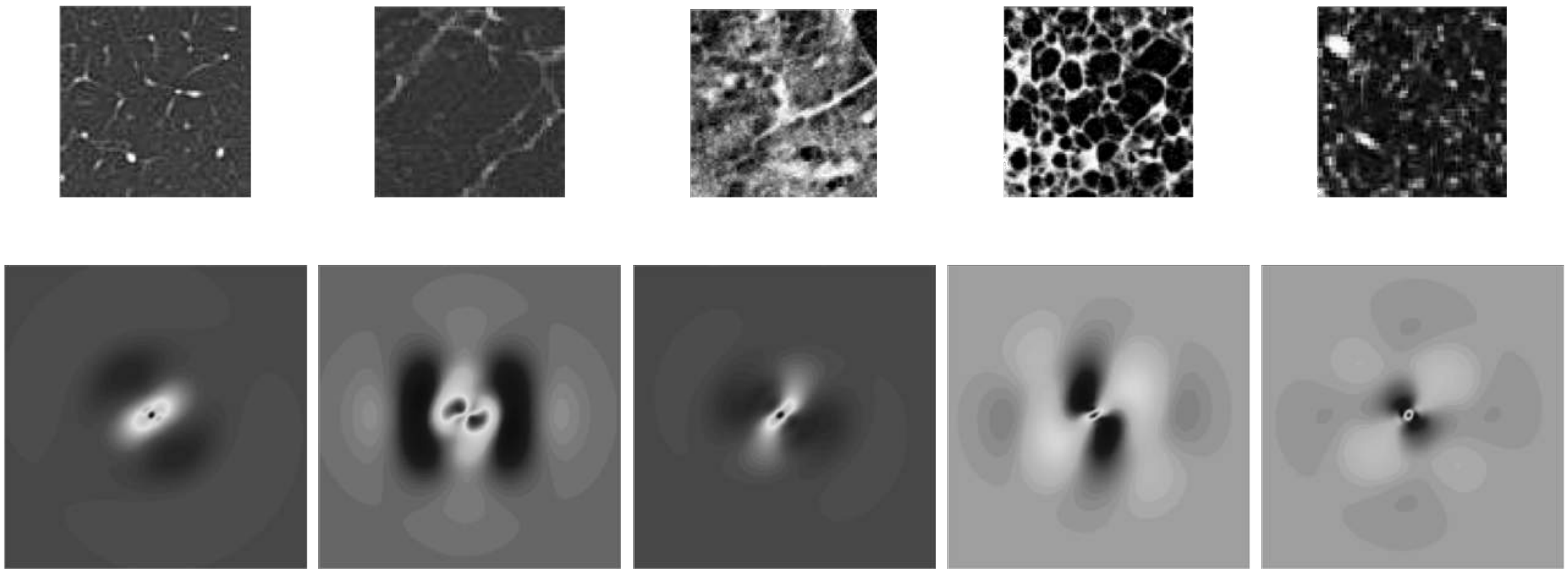




Matlab code available from: http://publications.hevs.ch/index.php/publications/show/1373

Learned signatures

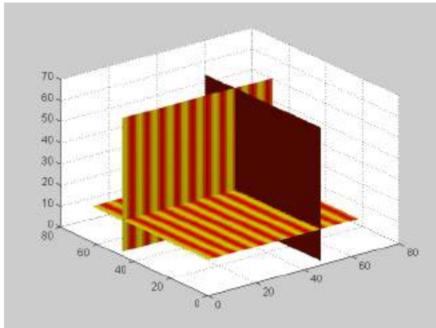
- Learn combinations of Riesz wavelets as digital signatures using SVMs
  - Create signatures to detect small local lesions and visualize them

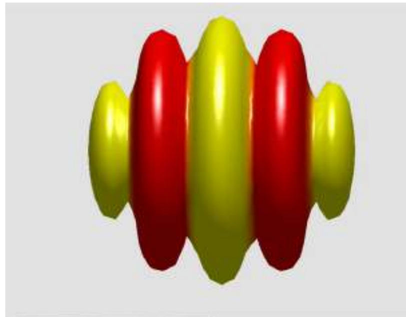


Adrien Depeursinge, Antonio Foncubierta–Rodriguez, Dimitri Van de Ville, and Henning Müller, Rotation–covariant feature learning using steerable Riesz wavelets, IEEE Transactions on Image Processing, volume 23, number 2, page 898-908, 2014.

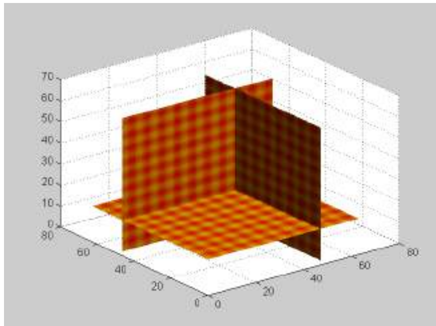
Learning Riesz in 3D

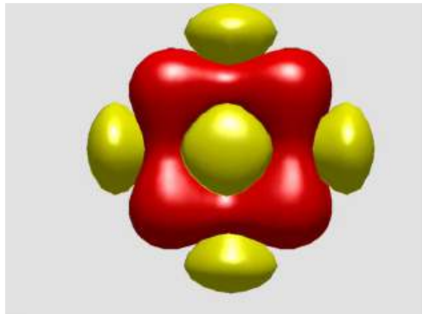
- Most medical tissues are naturally 3D
- But modeling gets more complex
  - Vertical planes



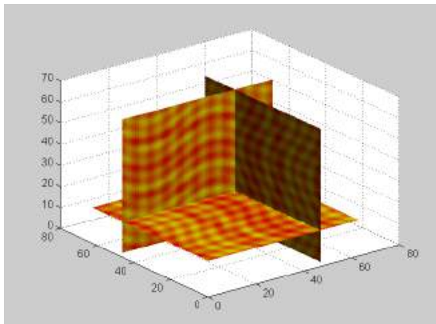


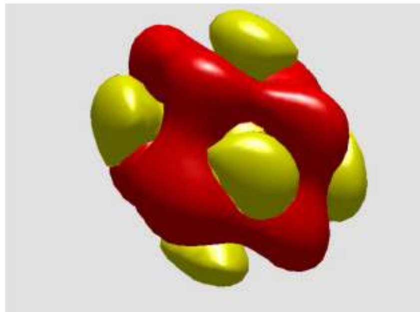
  - 3-D checkerboard





  - 3-D wiggled checkerboard



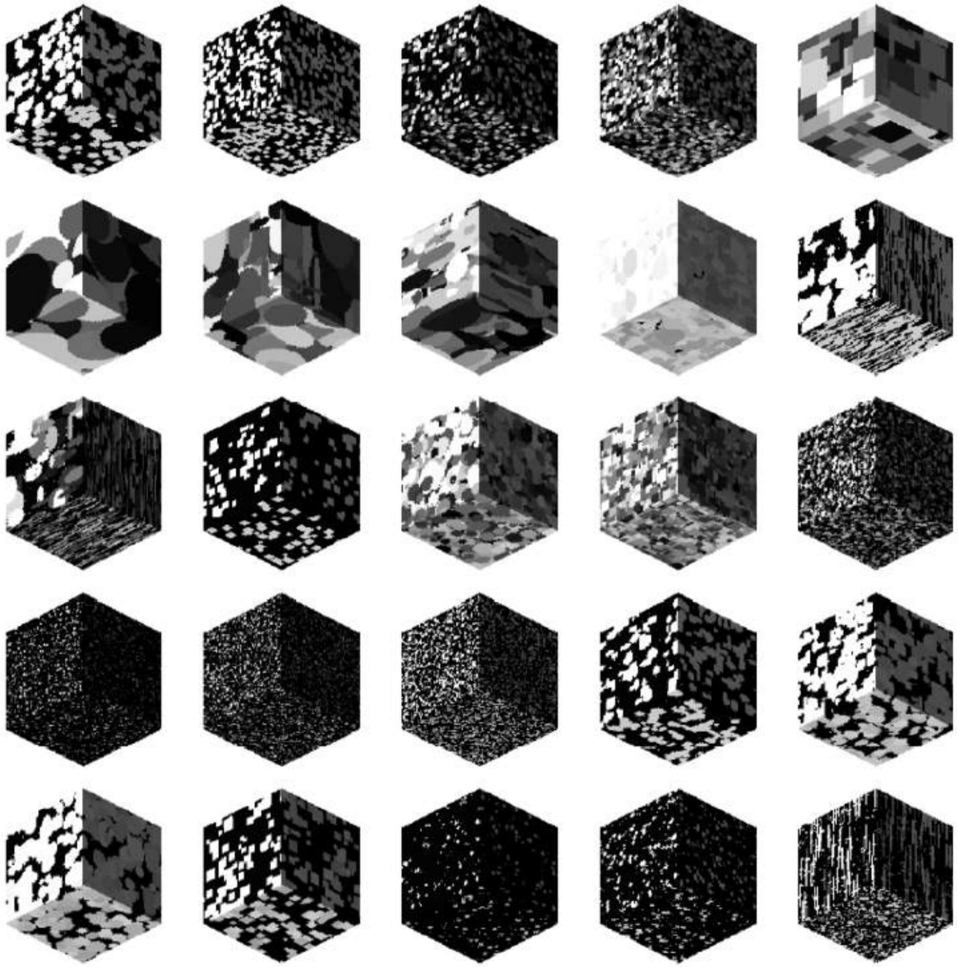


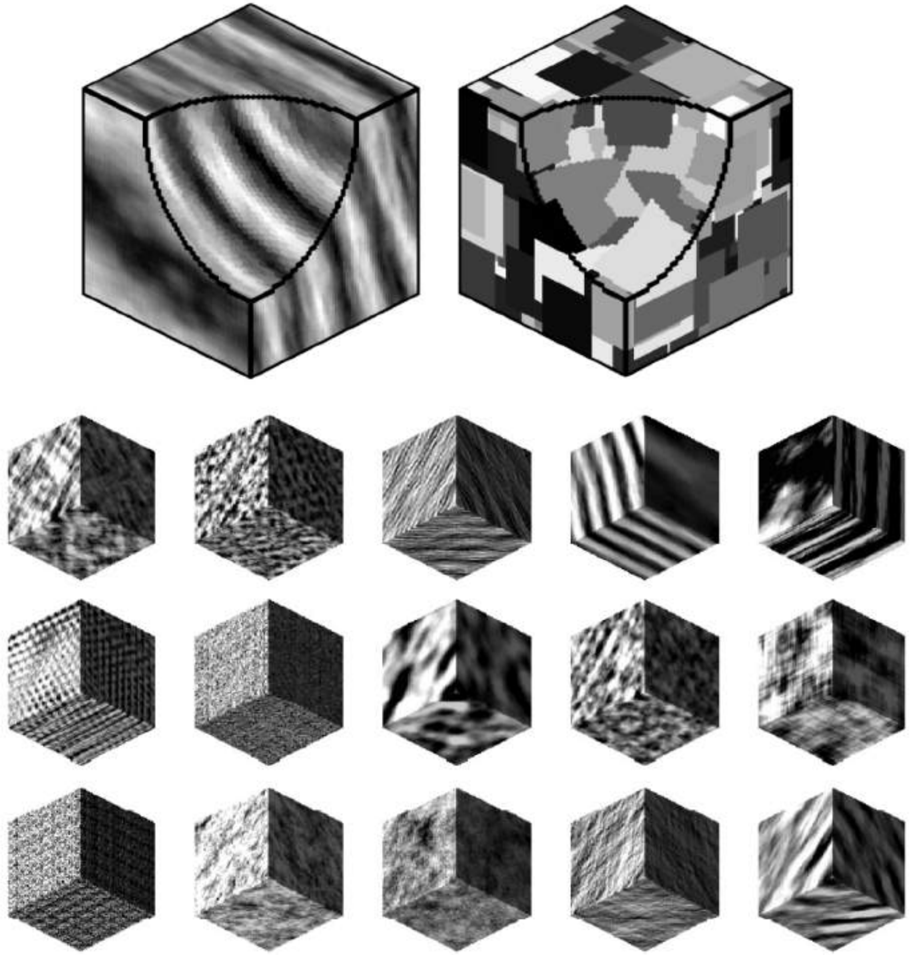
Solid 3D texture

- Hard to visualize

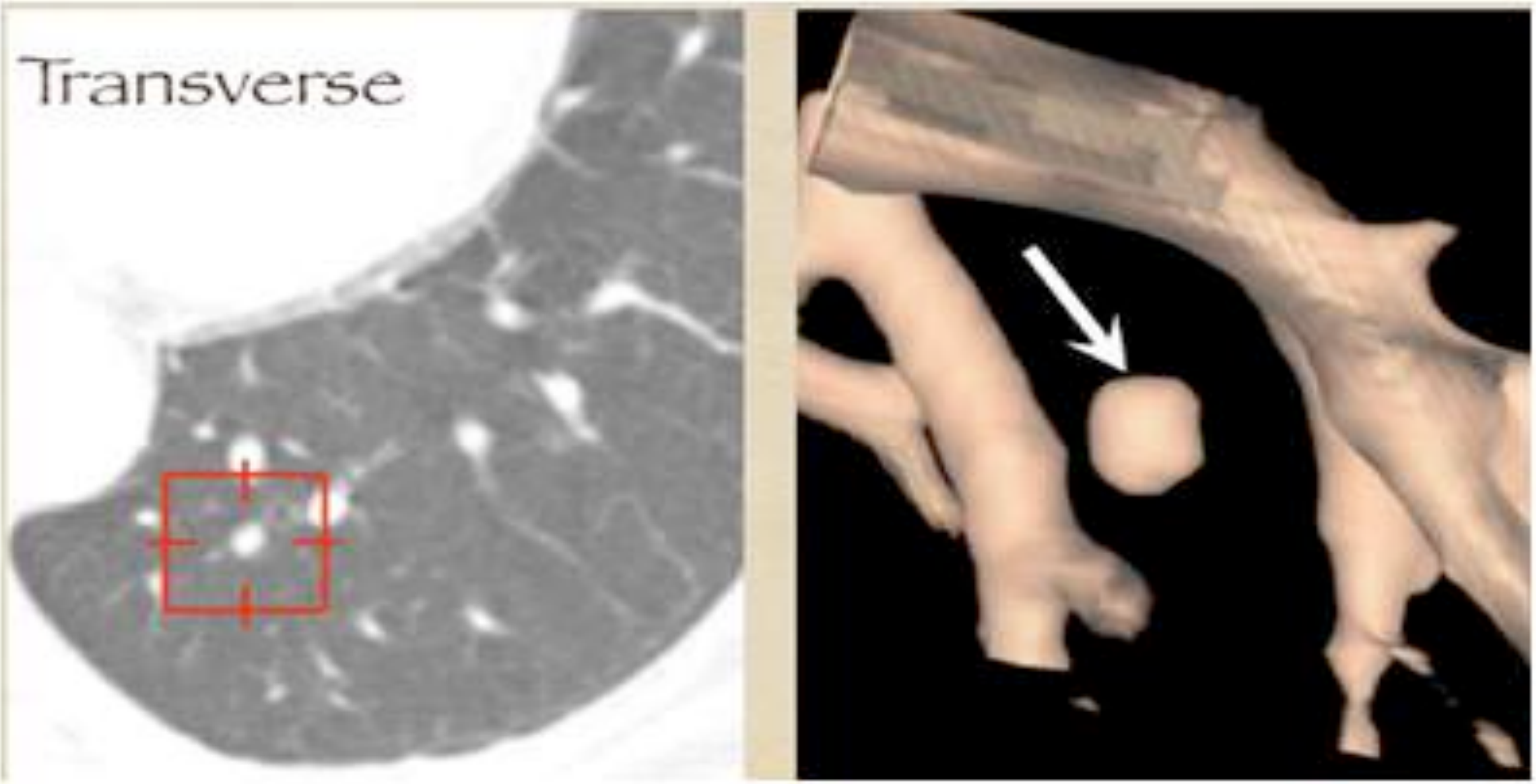
Y. Dicente Cid, H. Müller, A. Platon, PA Poletti, A. Depeursinge, Locally–Oriented Wavelet Transforms for 3–D Solid Texture Classification, *IEEE Transactions on Image Processing*, 2017.

- Most texture measures are translated to 3D



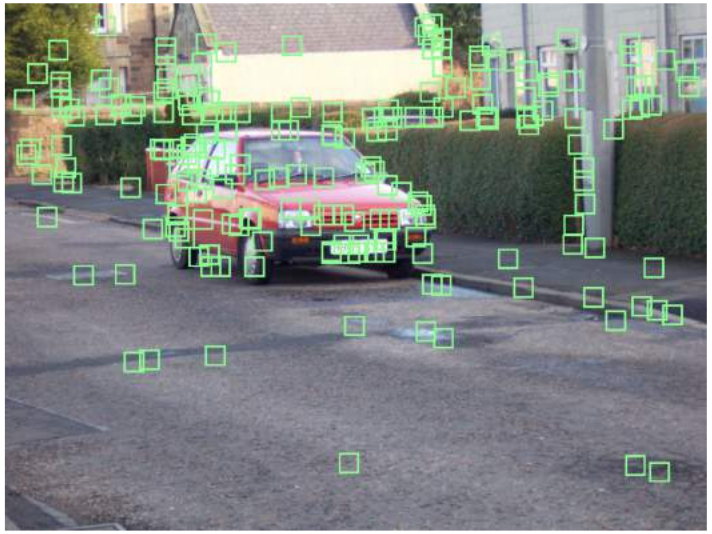


3D can be really important

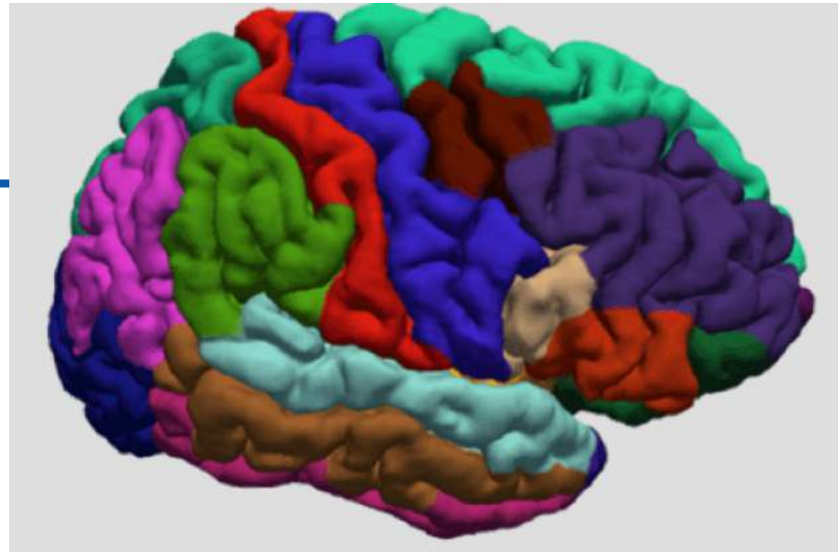


CT finding (left) has the appearance of an adjacent vessel in transverse-section reconstruction and was not called by any of the four LIDC readers. After viewing transverse, coronal, sagittal, and volume-rendered reconstructions (right), all four university readers called the finding a lung nodule.

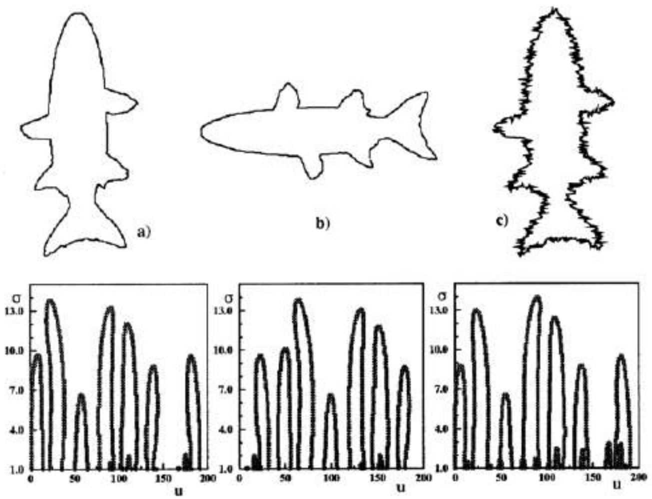
Interest Points – local information

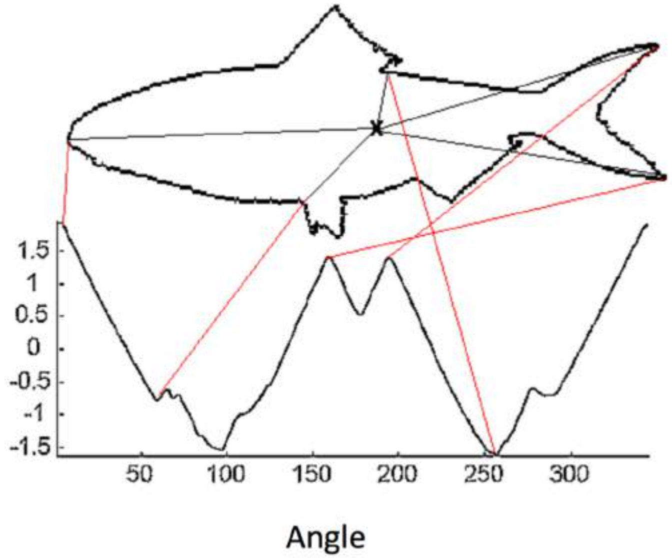


Shape features



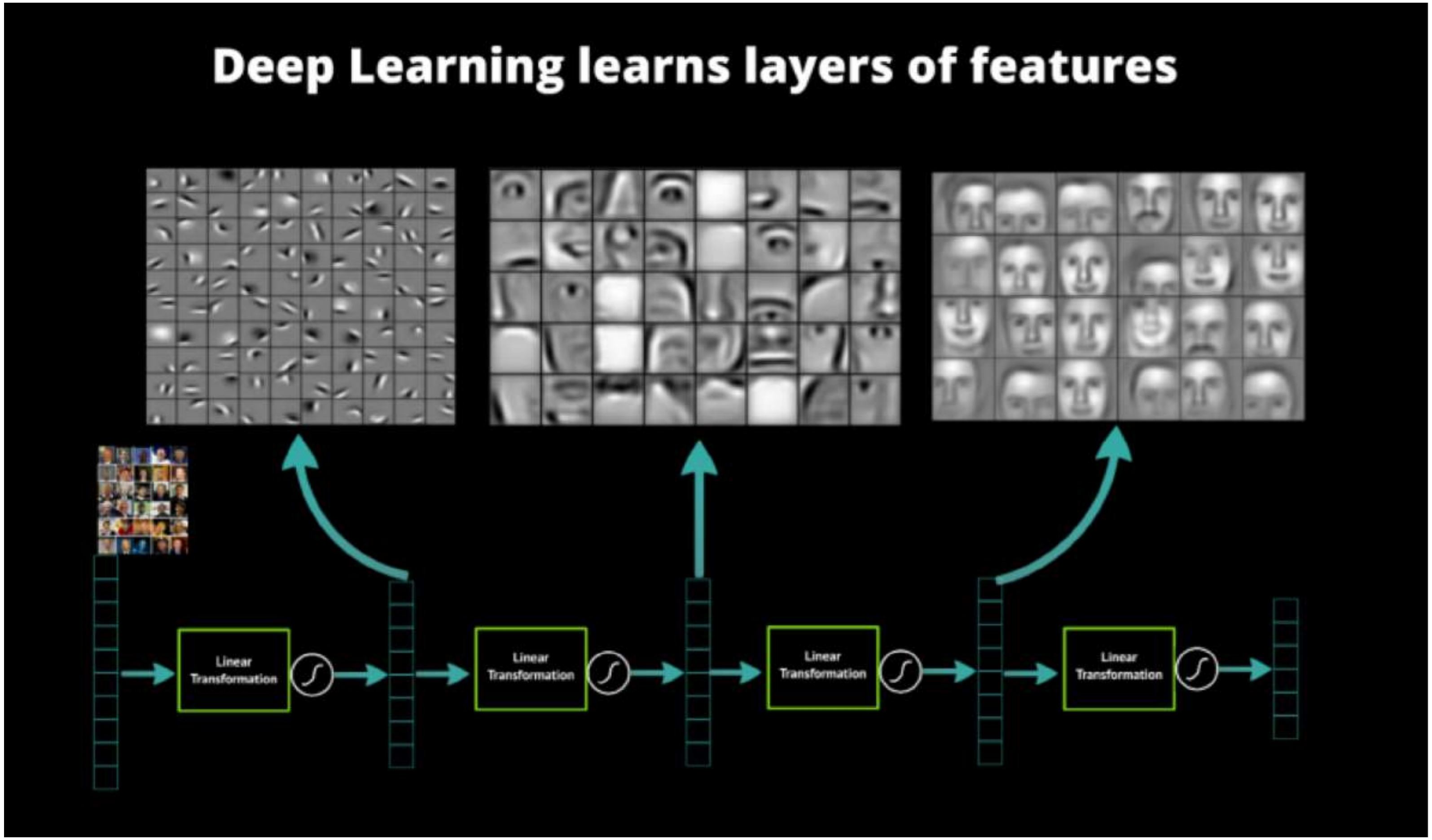
- 2D vs. 3D
- Invariances (rotations, size)
- Bounding box, aspect ratio, circularity, centroid, chain code, …
- Invariant statistical moments (Zernicke moments)
- Curvature scale space





Deep learning-based features

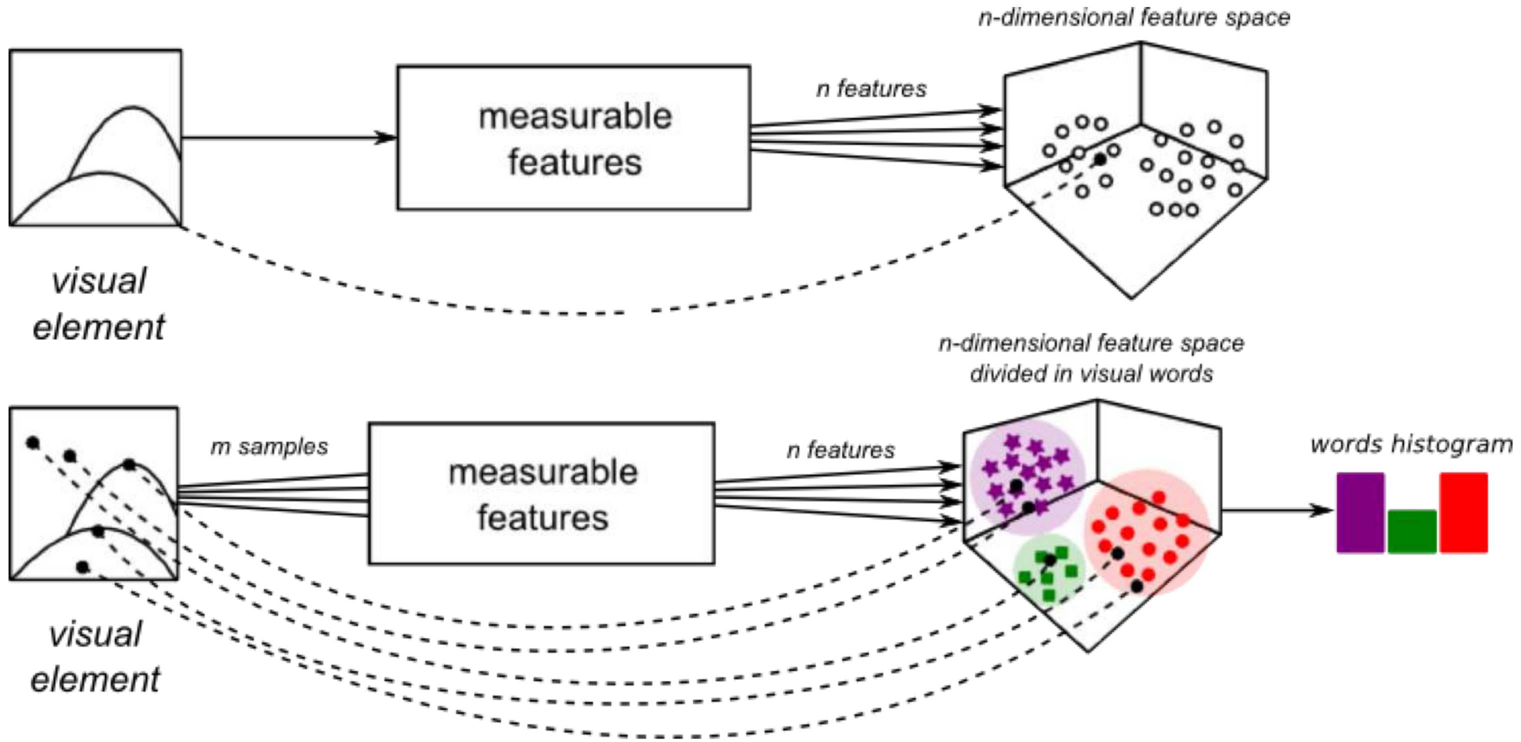
- Deep learning has several layers of neurons
- Usually "features" are not used and all is learned
  - Specific output layers can be used as features



Datarobot.com, a primer on deep learning

Visual feature modeling

- Visual words instead of raw visual features
  - Reducing the "curse of dimensionality"
- Use features based on the data actually present in a database



A Foncubierta, AG Seco de Herrera, H Müller, Medical Image Retrieval using a Bag of Meaningful Visual Words, ACM MM workshop on medical multimedia retrieval, Barcelona, Spain, 2013.

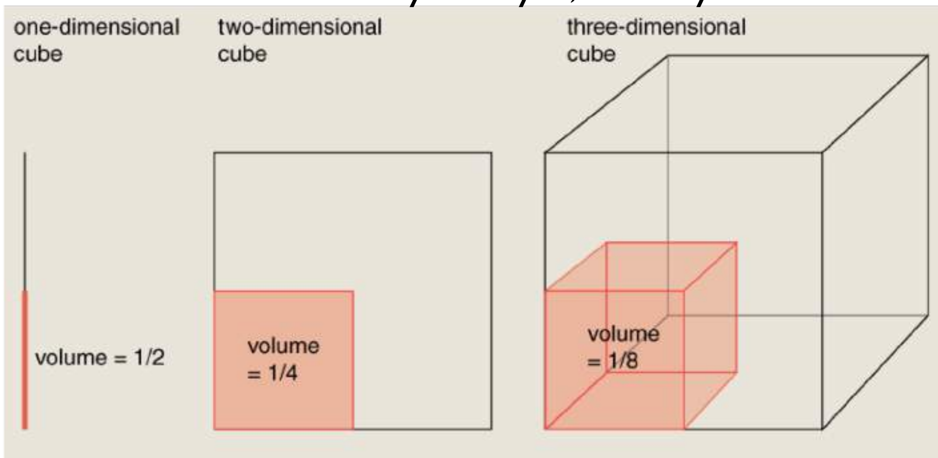
Curse of dimensionality

- "The curse of dimensionality refers to various phenomena that arise when analyzing and organizing data in high-dimensional spaces (often with hundreds or thousands of dimensions) that do not occur in low-dimensional settings such as the three-dimensional physical space of everyday experience." Wikipedia

- Increasing numbers of features mean that generalization requires exponential data amounts

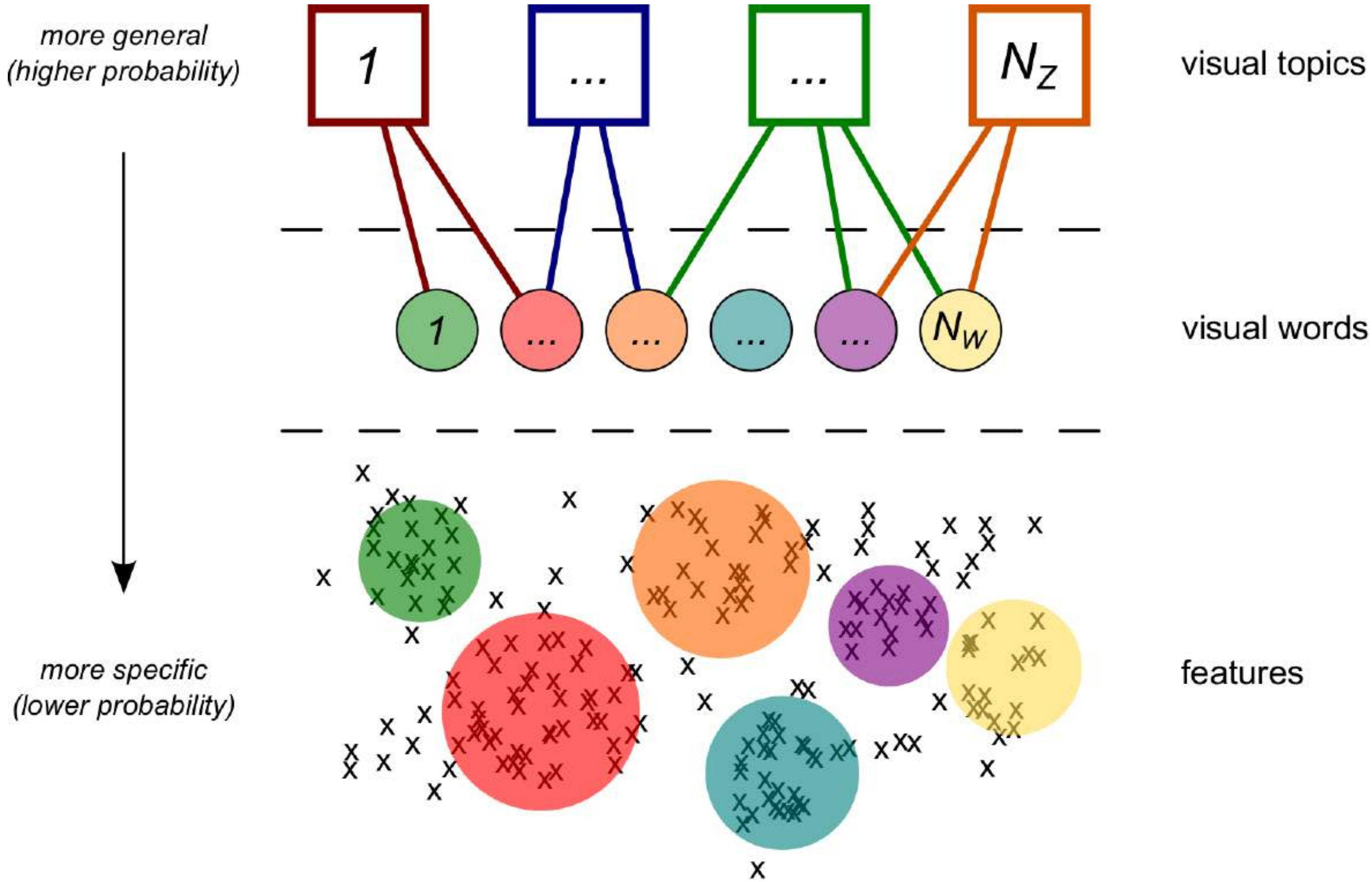
- Volume of a space increase with more dimensions, so data gets increasingly sparse
  - Distance between all items becomes similar
  - Automatic classification is then hard (using kNN)

Bryan Hayes, Orderly Randomness



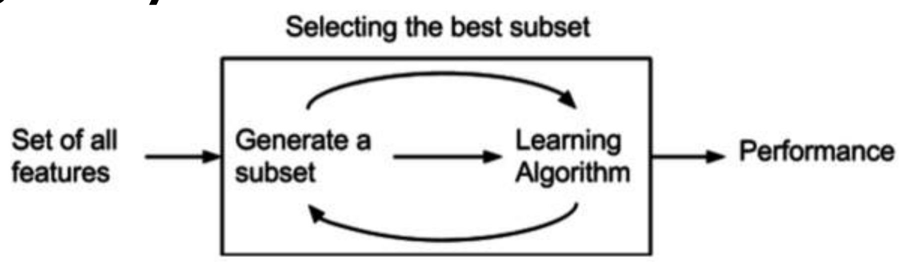
Latent semantic analysis

- Find underlying links between (visual) words
  - Find synonyms, homonyms, so ambiguous words



Feature selection strategies

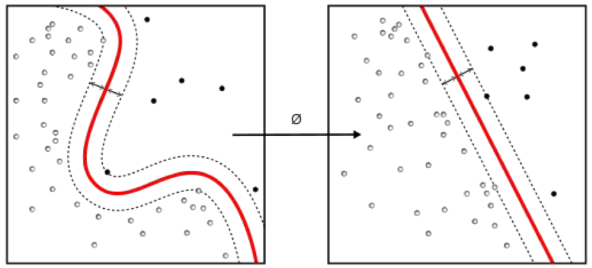
- Variable selection, attribute selection, …
- Goal: simpler models, lower storage, faster training, avoid overfitting, remove redundancy
- Many approaches
  - Exhaustive search, best first, greedy forward and backward, ...
  - Filter, wrapper, embedded



- Criteria to select features
  - Correlation between features, mutual information, accuracy of the created models

Dimensionality reduction

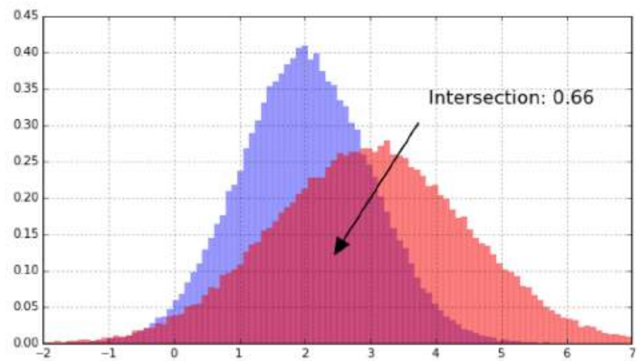
- Feature selection
- Principal component analysis (PCA)
  - Linear mapping of data onto fewer dimensions
  - Mapping to 2D, 3D allows to visualize data
- Kernel PCA
  - Nonlinear space, maximizing variance

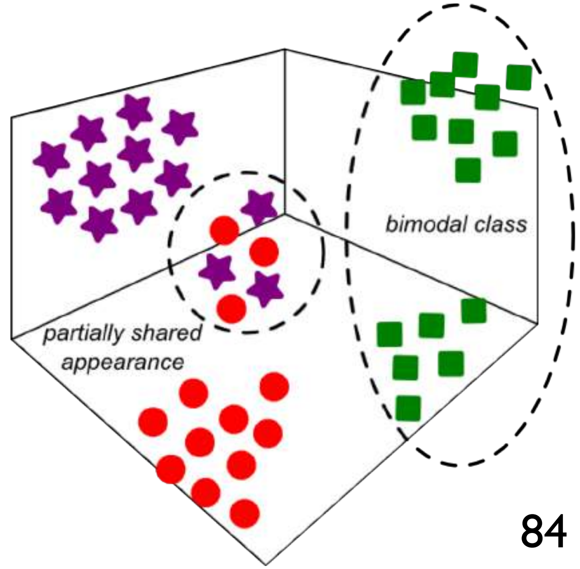


- Linear Discriminant Analysis (LDA)
  - Finding a linear combination to best separate classes
- …

Distance measures

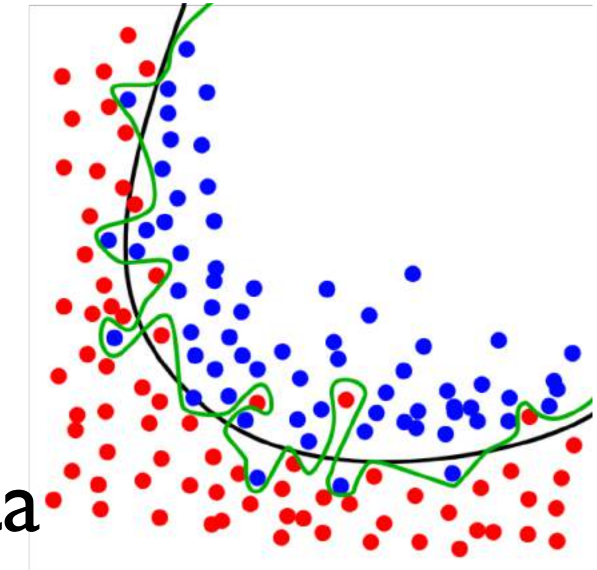
- Visual features represent structures in an n-dimensional space
  - Hopefully our visual features separate the items well
- Many distance metrics exist
  - Histogram intersection
  - City block, Manhatten distance
  - Euclidean distance
  - Earth Movers distance
  - Mahalanobis, Bhattacharyya, …





Overfitting

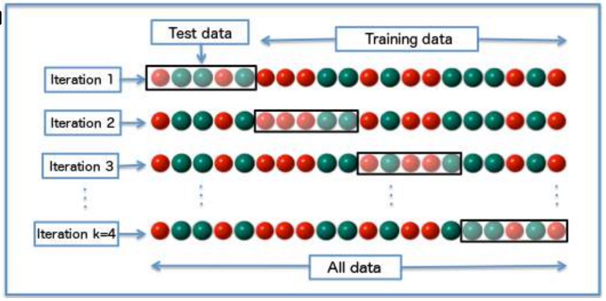
- "In overfitting, a statistical model describes random error or noise instead of the underlying relationship. Overfitting occurs when a model is excessively complex, such as having too many parameters relative to the number of observations." Wikipedia



- Over fit models do not generalize
  - Not good on new or unseen data
- Real risk in learning with many parameters or training on test data
  - Manual tuning to get good results
- A model should perform well on unseen data!
  - Methods such as testing on unseen data can help

Evaluation methodologies for ML

- Clear split of training and test data
  - Sometimes validation data as part of training data
  - All parameters set only on training data
- Leave-one-out cross validation
  - When few items are in the data, maximize training data
  - Leave-one-patient-out, limits overfitting
- N-fold cross validation
  - Separate data in N folds, then use each time one for testing and all others for training
  - Random splits can make this non-reproducible

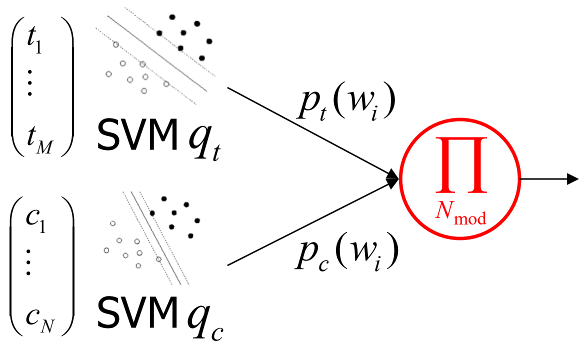


Information fusion

COMPUTERTOMOGRAPHIE THORAX / ABDOMEN Indikation: Z.n. Semicastratio rechts. Staging erbeten. Metastasestase? Der Patient wurde ueber die moeglichen Risiken und Nebenwirkungen im Rahmen der Kontrastmittel-Applikation informiert und bestaetigt sein Einverstaendnis. Der Patient hat keine weiteren Fragen. Untersuchungstechnik: Brilliance 64, Philips Medical Systems, Kollimation 1x64x0.625mm; Zwerchfell bis Leberunterrand Arteriell , Applikation 300 Jopamiro , 120 + 40 ml 4 ml/sec, BT+ 16 sec, Obere Thoraxapertur bis Symphyse Portalvenoes , Delay 50 sec, Rekonstruktionen: MPR axial und coronal 3/2mm Weichteilfenster, MPR axial und Wirbelsaeule sagittal 3/2mm Knochenfenster, Thorax MPR axial und coronal 3/2mm und MIP axial 15/2mm Lungenfenster. Thorax: Kleinstes, offenbar verkalktes Ganulom mit einem DM von etwa 1mm im apicalen Oberlappen rechts. Ansonsten kein Nachweis intrapulmonaler *Rundherde*. Keine Pleuraerguesse, kein *Pericarderguss*. Das zentrale Bronchialsystem frei, kein Nachweis eines bronchobstruktiven Prozesses. Kein Nachweis pathologisch vergroesserter mediastinaler oder hilaerer Lymphknoten. Die Pulmonalarterien und die supraaortalen Ã„ste homogen perfundiert. Abdomen: Die Leber von normaler Groesse und homogenem Enhancement. 2mm haltende Hypodensitaet subkapsulaer im Lebersegment VII. Dieses bei der geringen Groesse nicht naeher charakterisierbar. Die Gallenblase normal gross, kein Nachweis roentgendichter Konkremente. Keine *intra- oder extrahepatische Cholangiektasie*. Das Pankreas von regulaerer Parenchymsaumbreite und homogenem Enhancement. Die Milz normgross. Die Nebennieren bds. schlank. Die Nieren bds. von regulaerer Lage, Form und Groesse. Das Nierenhohlraumsystem bds. normal weit. Einzelne bis 3mm im QDM haltende Lymphknoten paraaortal bds. Die Harnblase mit minimaler Fuellung, soweit beurteilbar unauffaellig. Bei St.p. Semicastratio rechts geringes Weichteiloedem inguinal rechts. Im Knochenfenster kein Nachweis *Metastasestase*-suspekter *Laesion*.

__Legend:
POSPATHOLOGY
NEGPATHOLOGY
IDX_KEIN
NEGATEDSTRING
ANATOMY

- Combine information from several sources (i.e. text or structured data with visual)
- Text data can be mapped to semantics to understand links
  - Also language-independent
- Early fusion
  - Combing all features using a single classifier in the end
- Late fusion
  - Using classifier output
  - Rank-based vs. score-based



Many fusion techniques

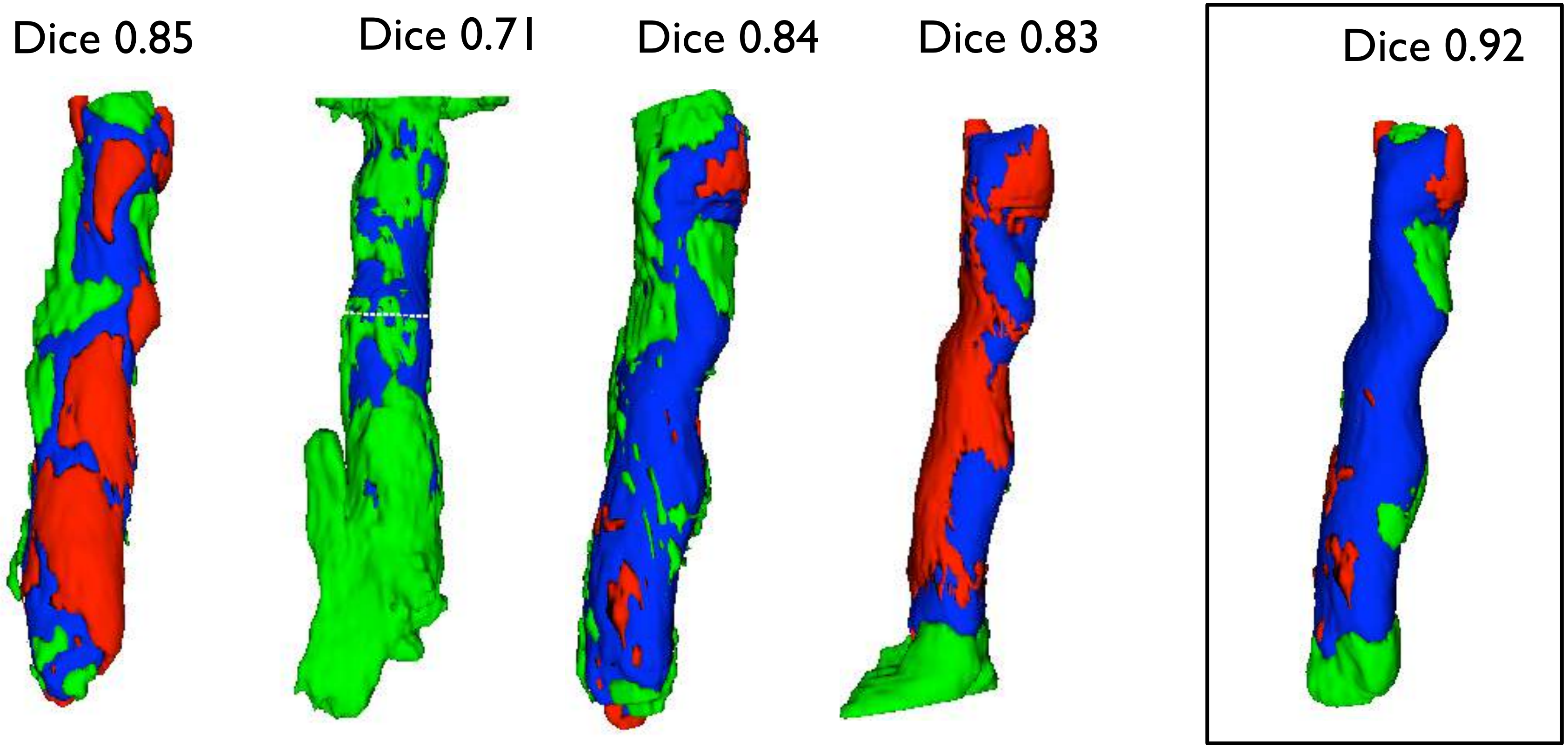
- Early fusion can work better but often requires dimensionality reduction
- Rank-based fusion (late)
  - Ranked list of items or classes are combined
  - If score distributions are not the same this can be better
- Score-based fusion (late)
  - Using scores of items for combined decision
    - Weighted linear combination
    - Sum of scores or simply minimum/maximum value
- Borda count, …

Silver corpus (example trachea)

- Executable code of the VISCERAL scientific challenge for segmentation
  - Run it on new data, do label fusion of the results
    - Silver corpus that can be used for training

Participant segmentations

Silver Corpus

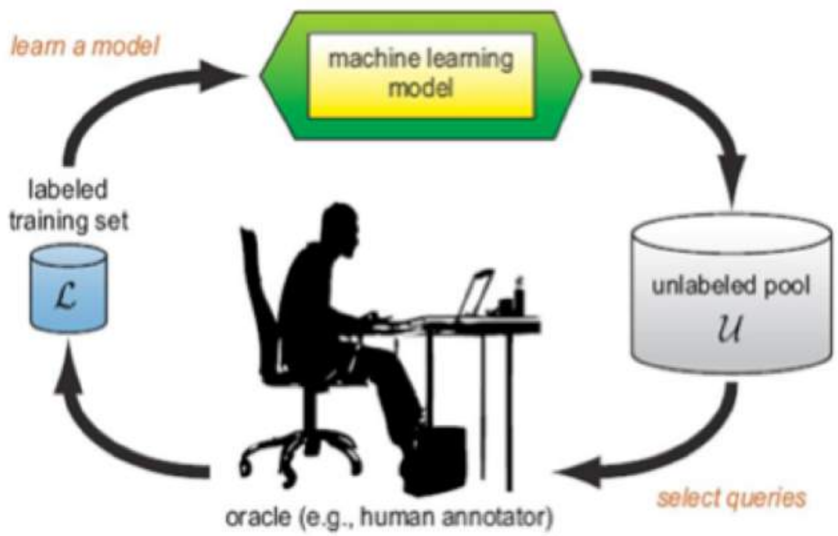


How to create labels for learning?

- Manual expert annotators
  - Expensive, time-consuming, does not scale well, quality
- Use associated documents
  - Radiology report, anamnesis, ICD codes
  - Quality not guaranteed
- Crowdsourcing
  - Non-expert annotations
  - Strict quality control is needed and focused task definition
- Leverage outcomes of automatic tools

Active learning

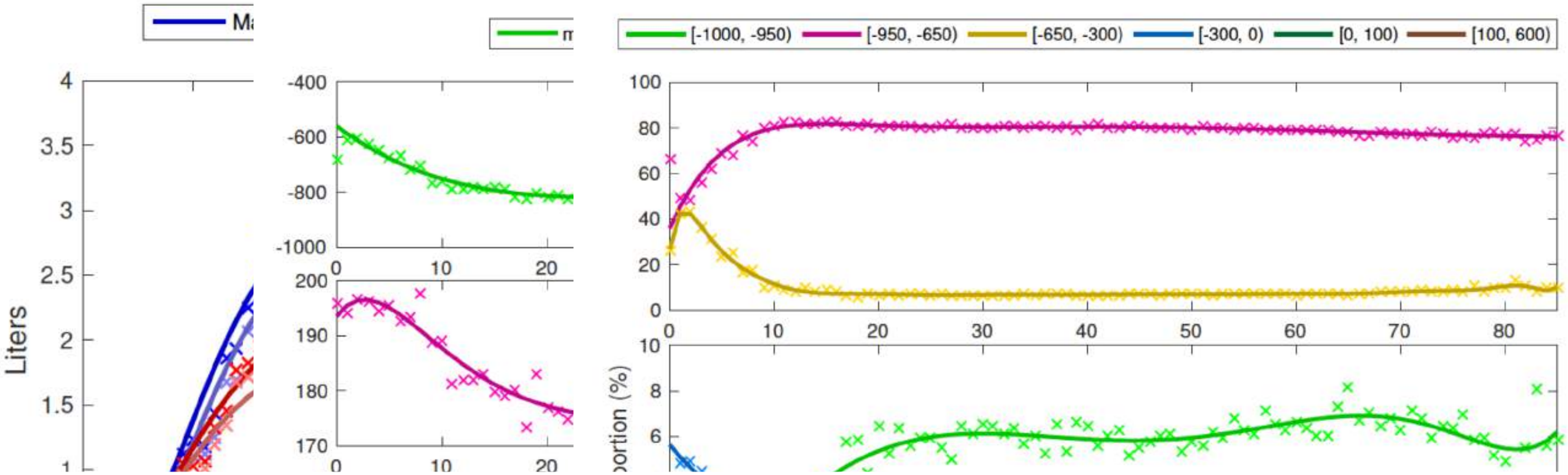
- Let the algorithm decide which non-labelled instances are best to be annotated
  - Usually iteratively to maximize information gain
- Interactive way to limit the amount of annotated data needed
- Visualization can make things easier

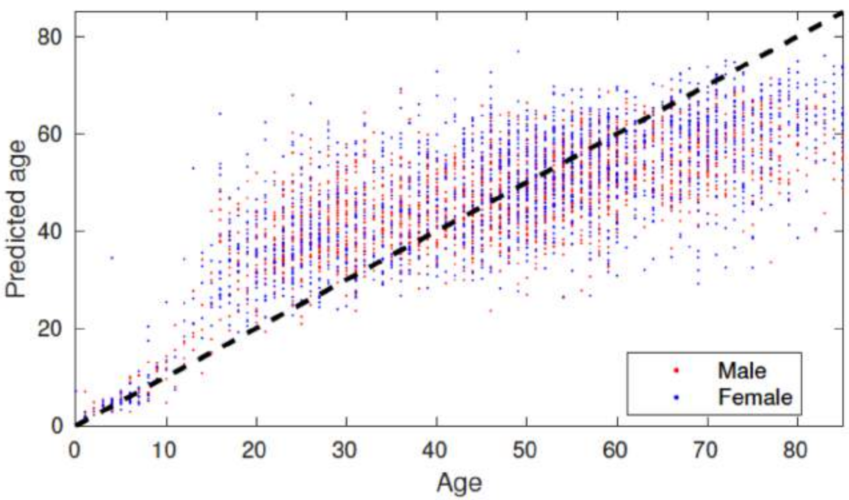


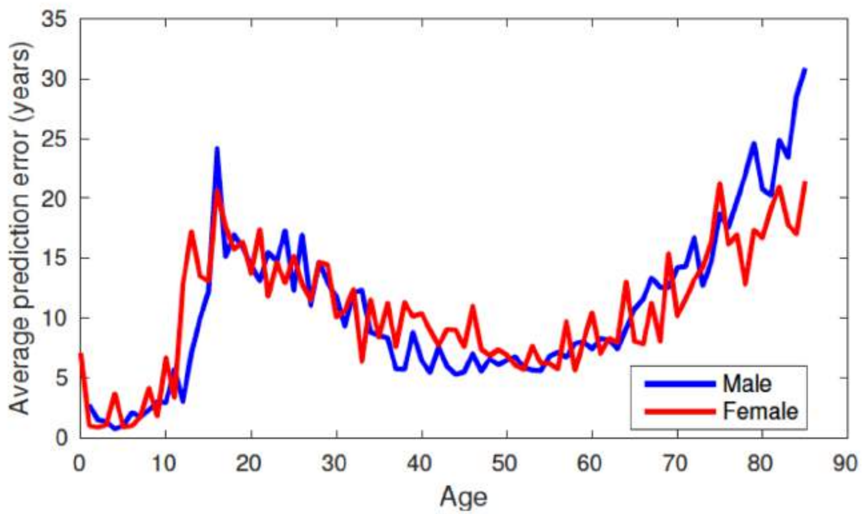
Pressforward.org

Analysis with (almost) no labels

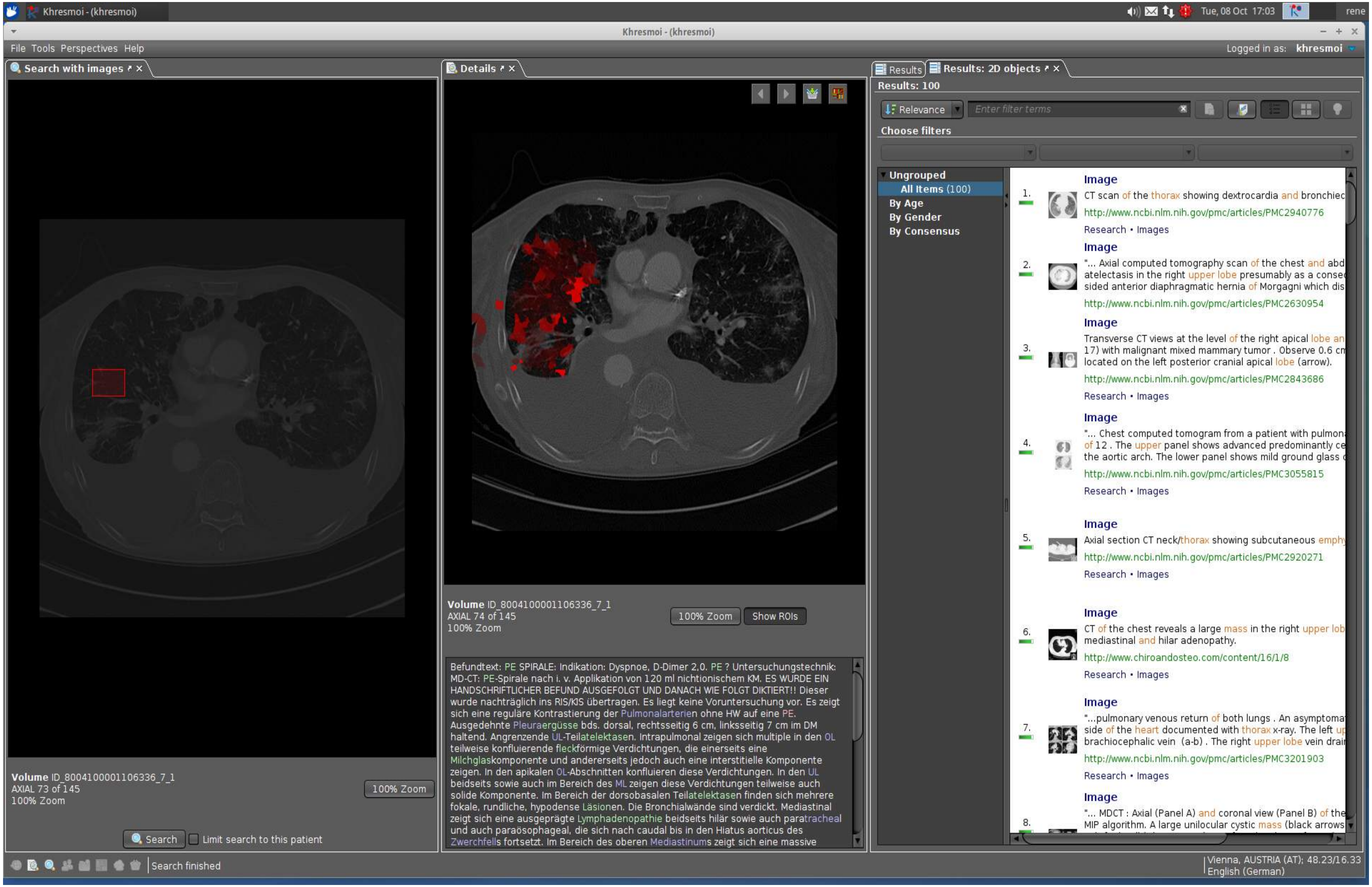
- Age prediction from lung CTs







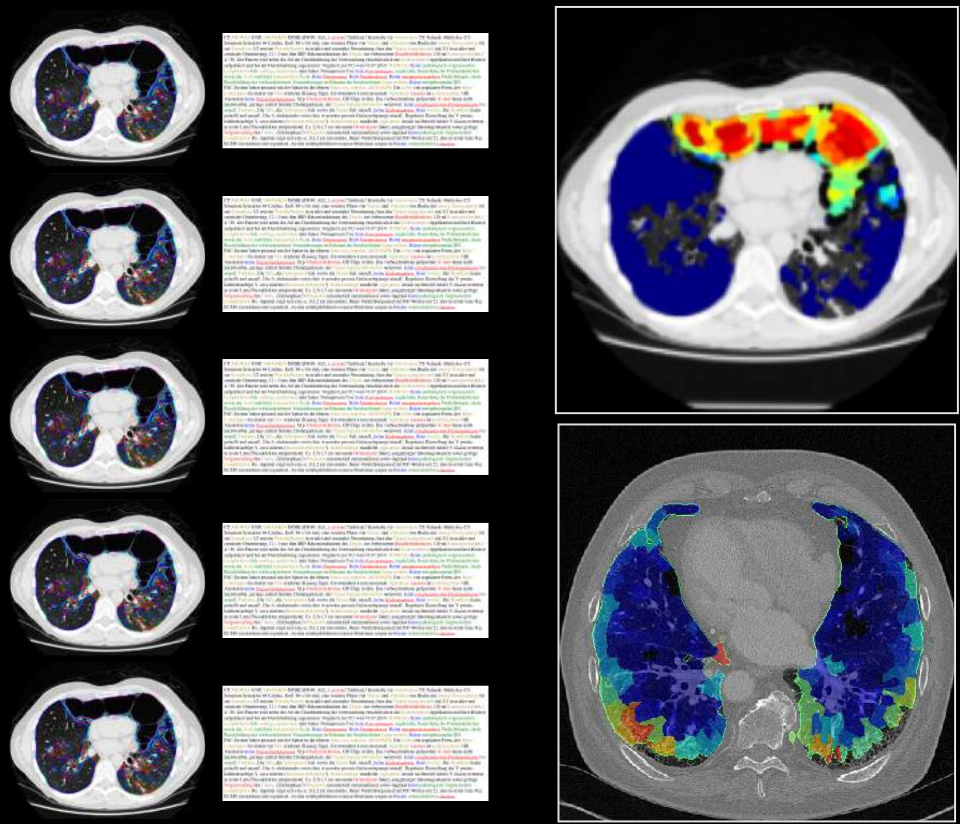
Khresmoi4radiology interface



KHRESMOI: towards retrieval in clinical data

Hofmanninger CVPR 2015

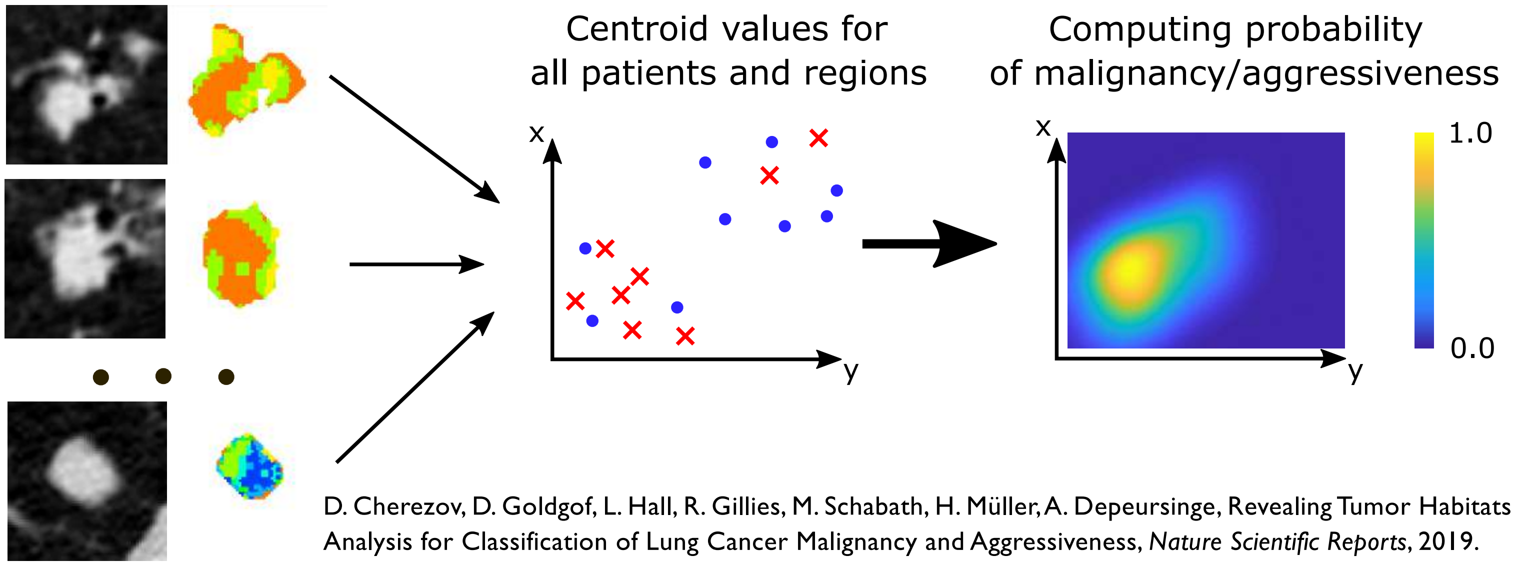
- Combine imaging data and radiology report information
- Extract semantic information from radiologe reports
- Identify visual signatures of findings from this combined data



Markers for disease patterns, learned based on images + reports

Tumor heterogeneity

- Using the NLST lung screening data set & Moffitt data
  - Both are publicly available



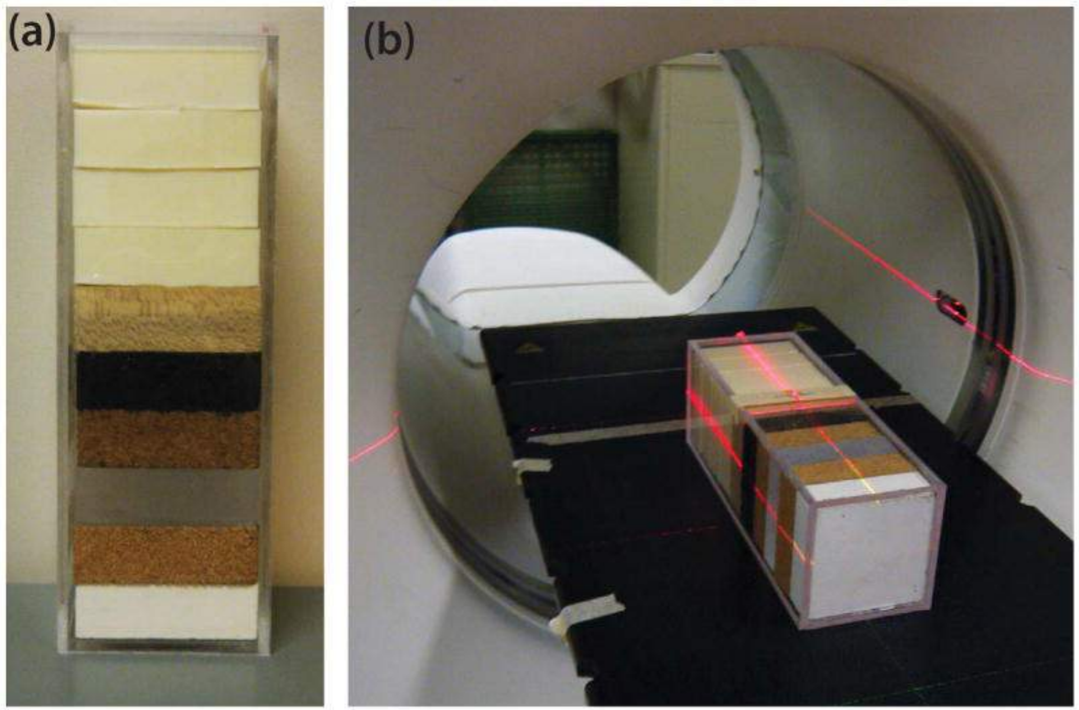
D. Cherezov, D. Goldgof, L. Hall, R. Gillies, M. Schabath, H. Müller, A. Depeursinge, Revealing Tumor Habitats from Texture Heterogeneity Analysis for Classification of Lung Cancer Malignancy and Aggressiveness, *Nature Scientific Reports*, 2019.

- For larger tumors the results are very good and complementary to radiomics features
  - For small tumors only a small increase

Image harmonization for radiomics

- Different scanners produce different images
  - Many protocols, construction kernels, producers, voxel sizes, …
    - Strong influence on features extracted
- How can we harmonize this?
  - Deep learning!
- Phantom study with 17 scanners
  - 10 solid textures
  - Features should be invariant to scanner and discriminative

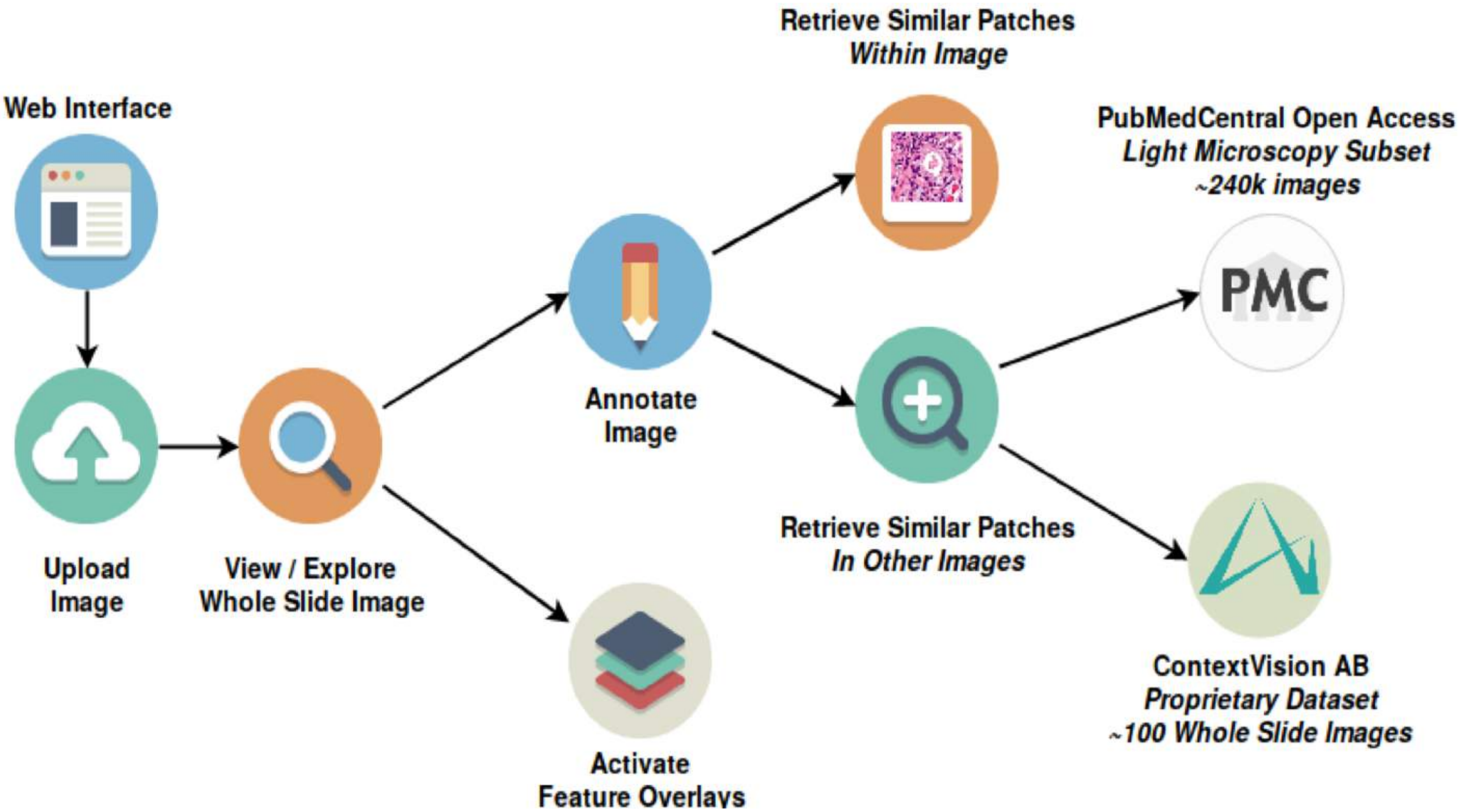
V. Andrearczyk, A. Depeursinge, H. Müller, Neural Network Training for Cross-Protocol Radiomic Feature Standardization in Computed Tomography, *Journal of Medical Imaging*, 2019.



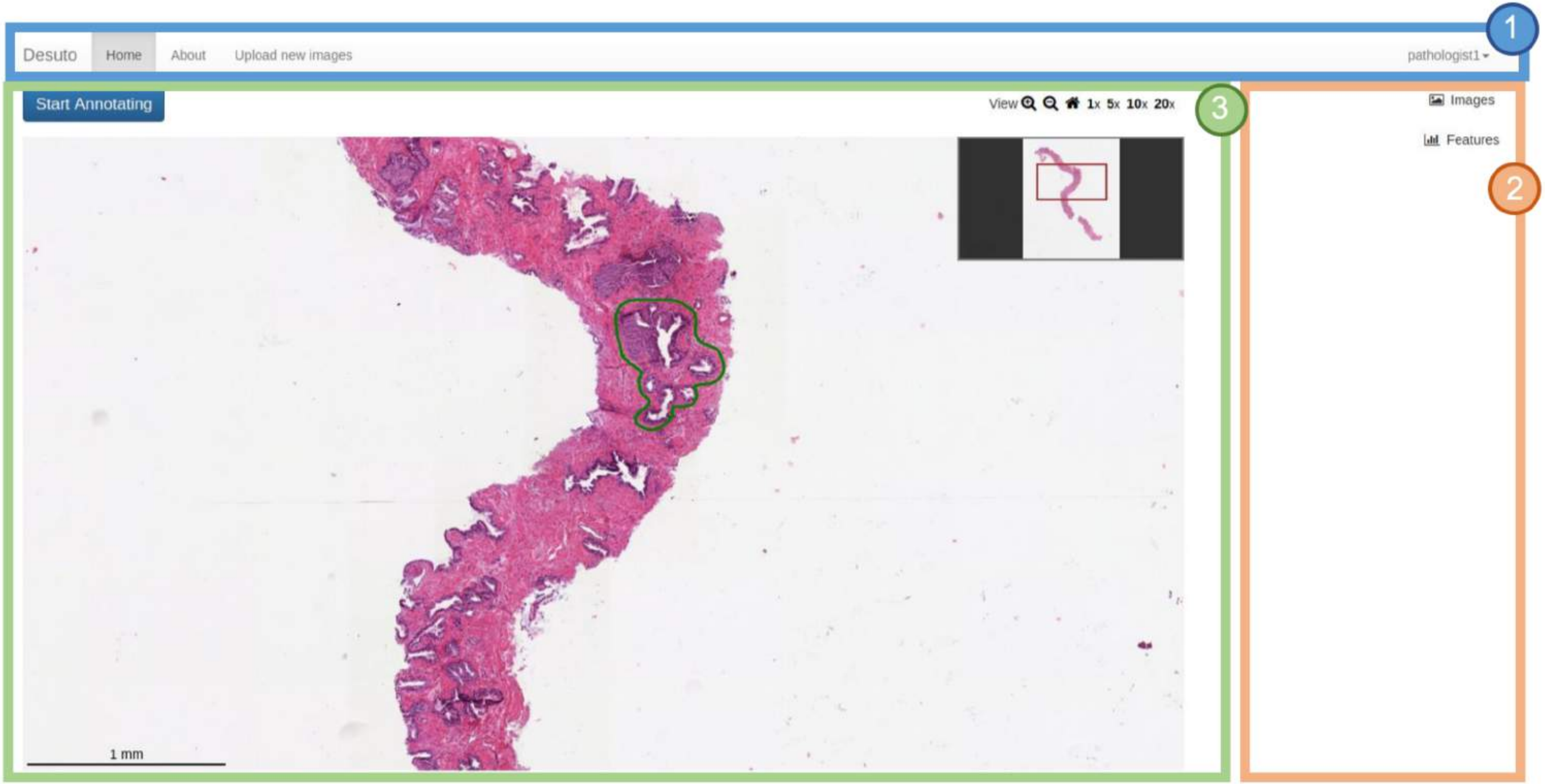
Histopathology image retrieval

- Slides are very large (100,000x100,000 pixels)
  - A case contains several slides
- The magnification level varies for structures to be observed
  - So images need to be compared on the same magnification
- Few annotated data sets of whole slide images exist but many on-annotated resources
  - Images from the biomedical literature (PubMed Central)
  - The Cancer Imaging Archive, The Cancer Genome Atlas

Main functions

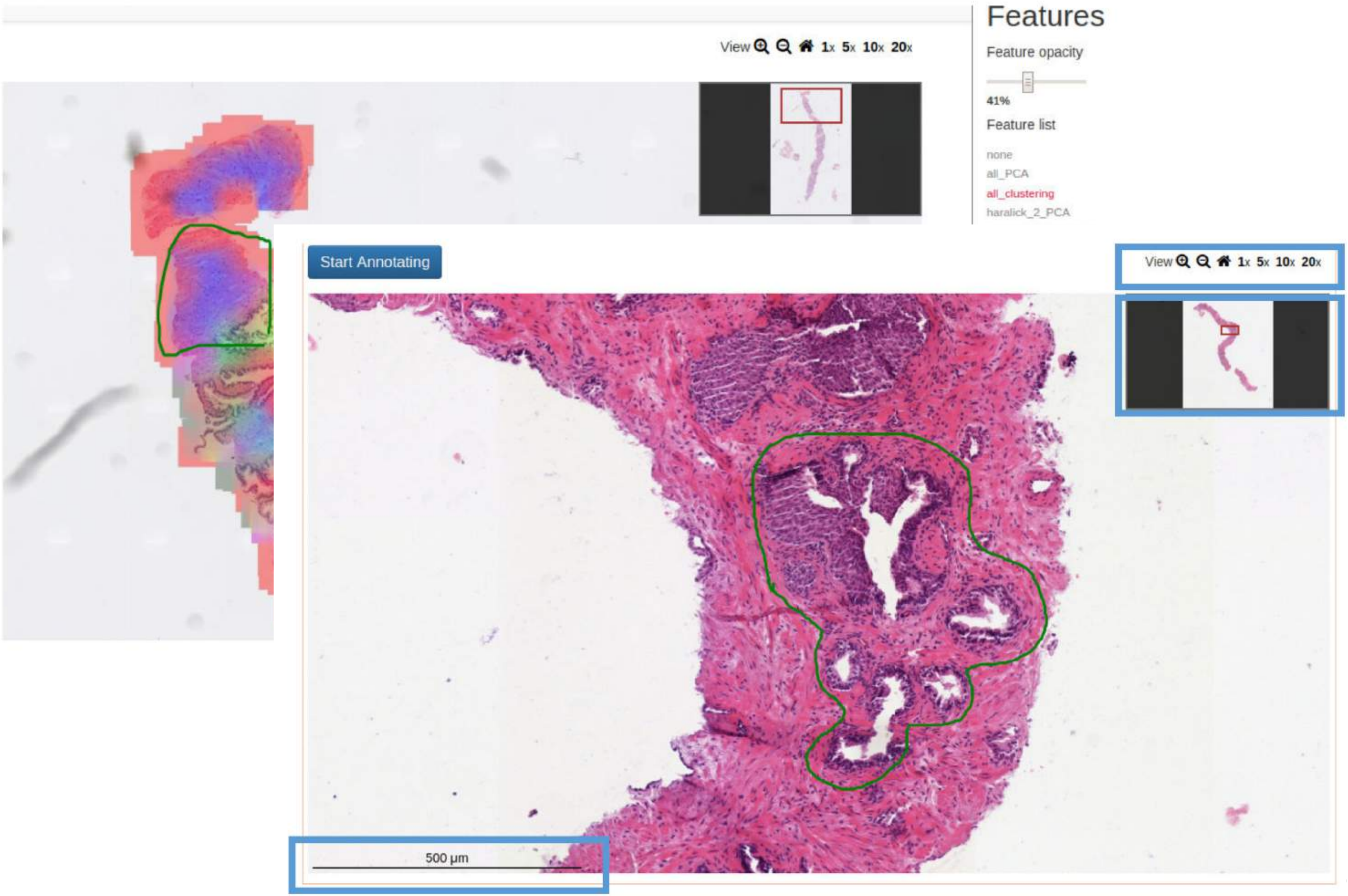


Interface overview

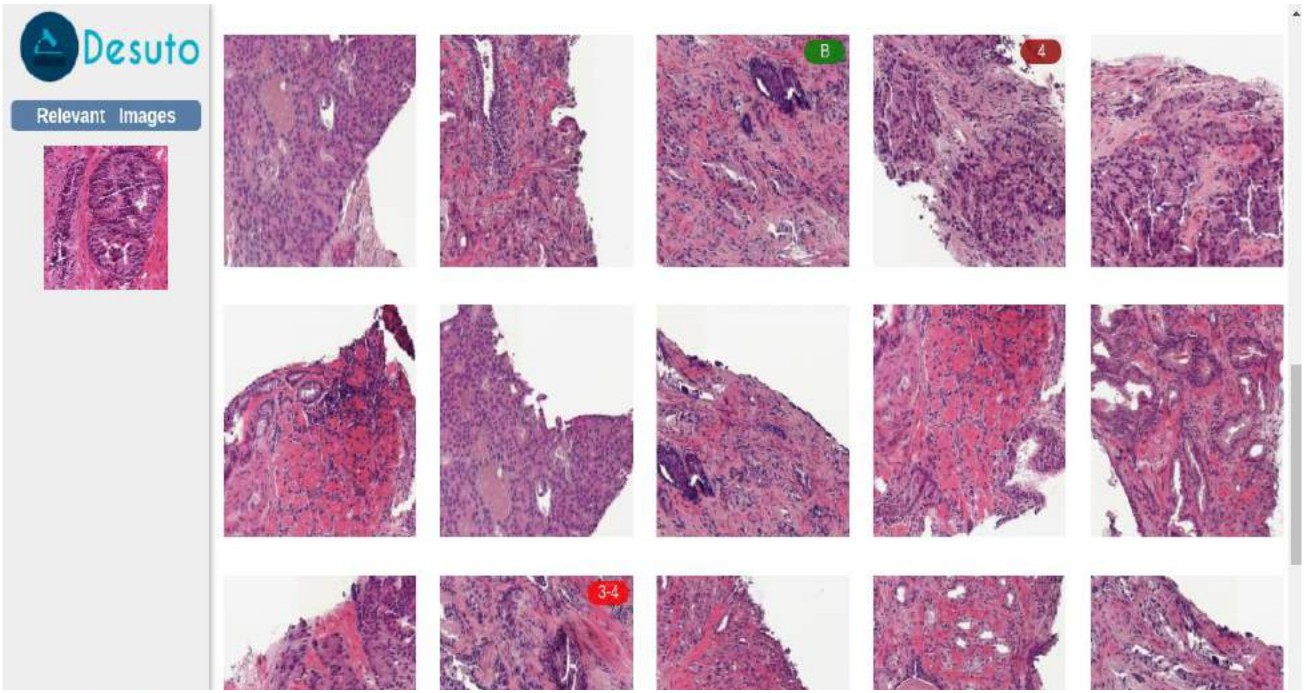


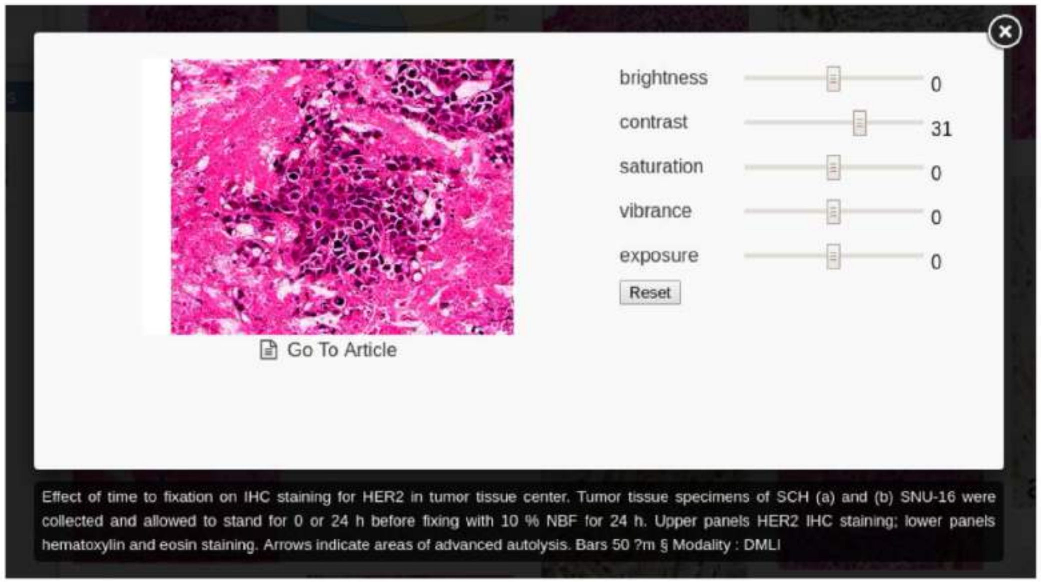
R. Schaer, S. Otalora, O. Jimenez-del-Toro, M. Atzori, H. Müller, Deep learning based retrieval system for gigapixel histopathology cases and open access literature, Pathology Informatics, 2019.

Overlays and annotations

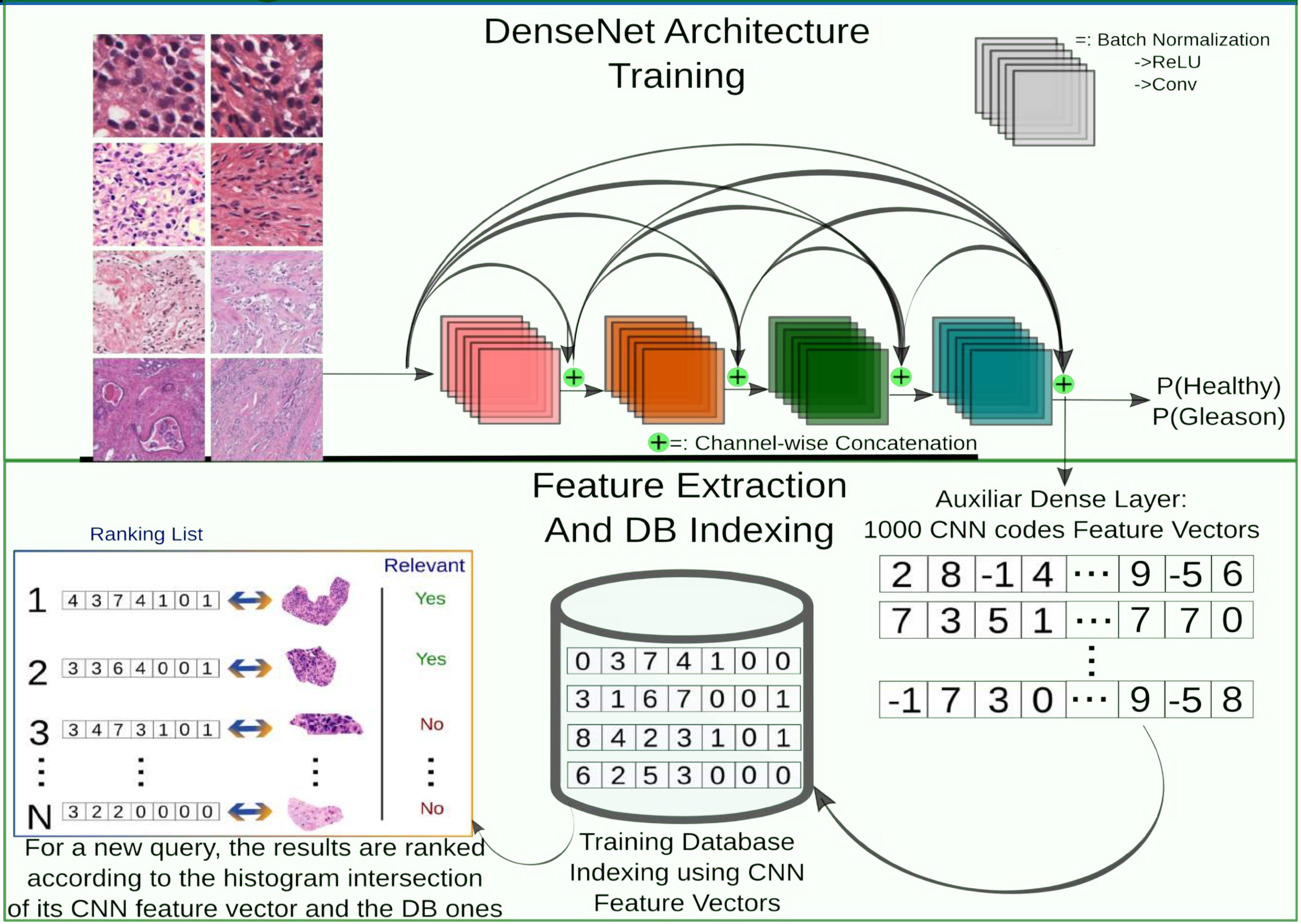


Visual retrieval





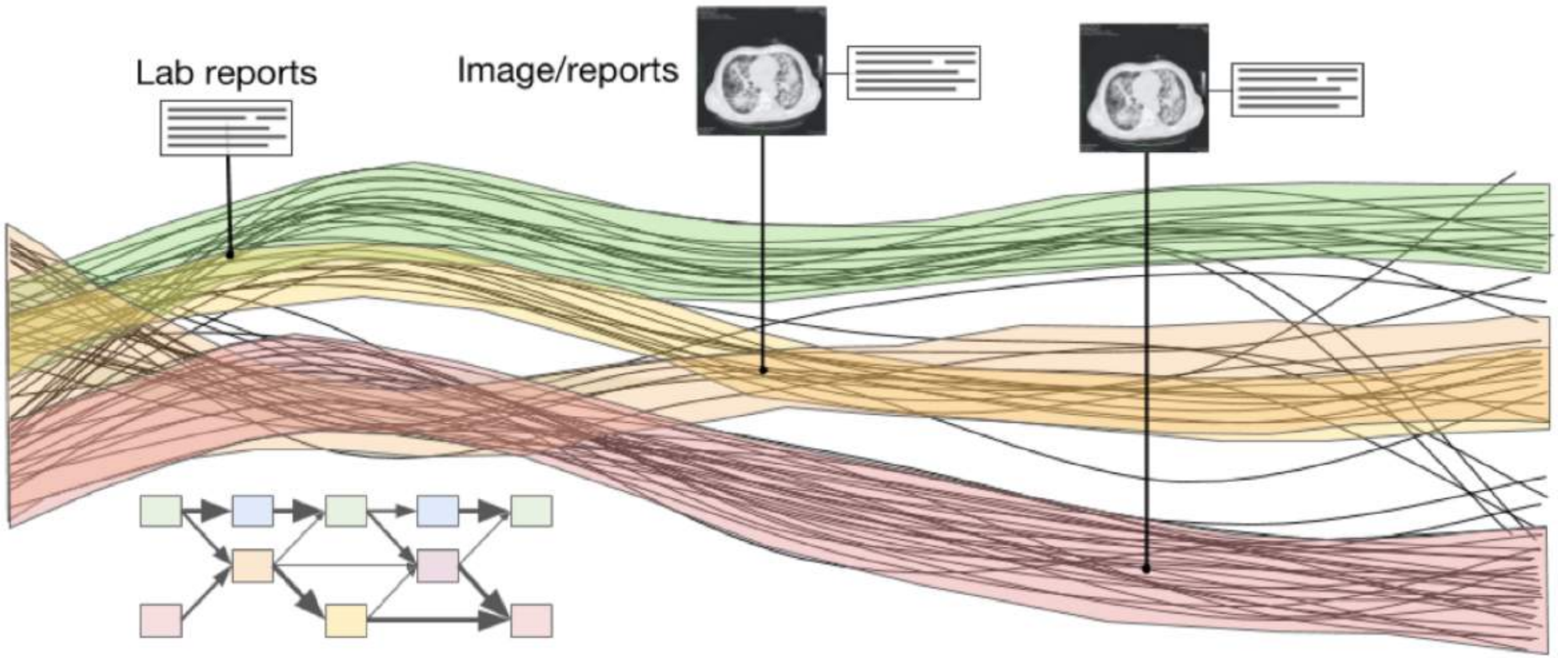
Deep Learning retrieval



S. Otalora, V. Andrearczyk, M. Atzori, H. Müller, Histopathology Nuclei Area Regression for Exploiting Open Access Content Using Deep Learning, *MICCAI workshop on Computational Pathology Imaging (COMPAY) 2018*, Granada, Spain, 2018.

Trajectories of patients

- Changes often matter more than absolute values
  - Which direction is a patient coming from
  - Longitudinal data are needed but hard to obtain
- Makes computation even more complex …
  - Much possible noise in the data



Conclusions

- Intelligent information access will likely replace manual analysis in routine tasks in medicine (i.e. sorting cases by importance)
  - Particularly with deep learning
- Data and data quality are the key to success
  - Amounts and quality matter
- Unsupervised or weakly supervised techniques are key to success to limit manual work
  - As manual annotation does not scale
- How to deal with changing images
  - New protocols, techniques, modalities

Contact

- More information can be found at
  - http://medgift.hevs.ch/
  - http://publications.hevs.ch/
- Contact:
  - Henning.mueller@hevs.ch