Predictive diagnosis and microscopic image content information

Kayser, Klaus, Borkenfeld, Stephan, Kayser, Gian

Basic theoretical considerations

Diagnosis   Types & Genetic Information

Object – Structure – Texture

Gray Value Thresholds dependent Entropy

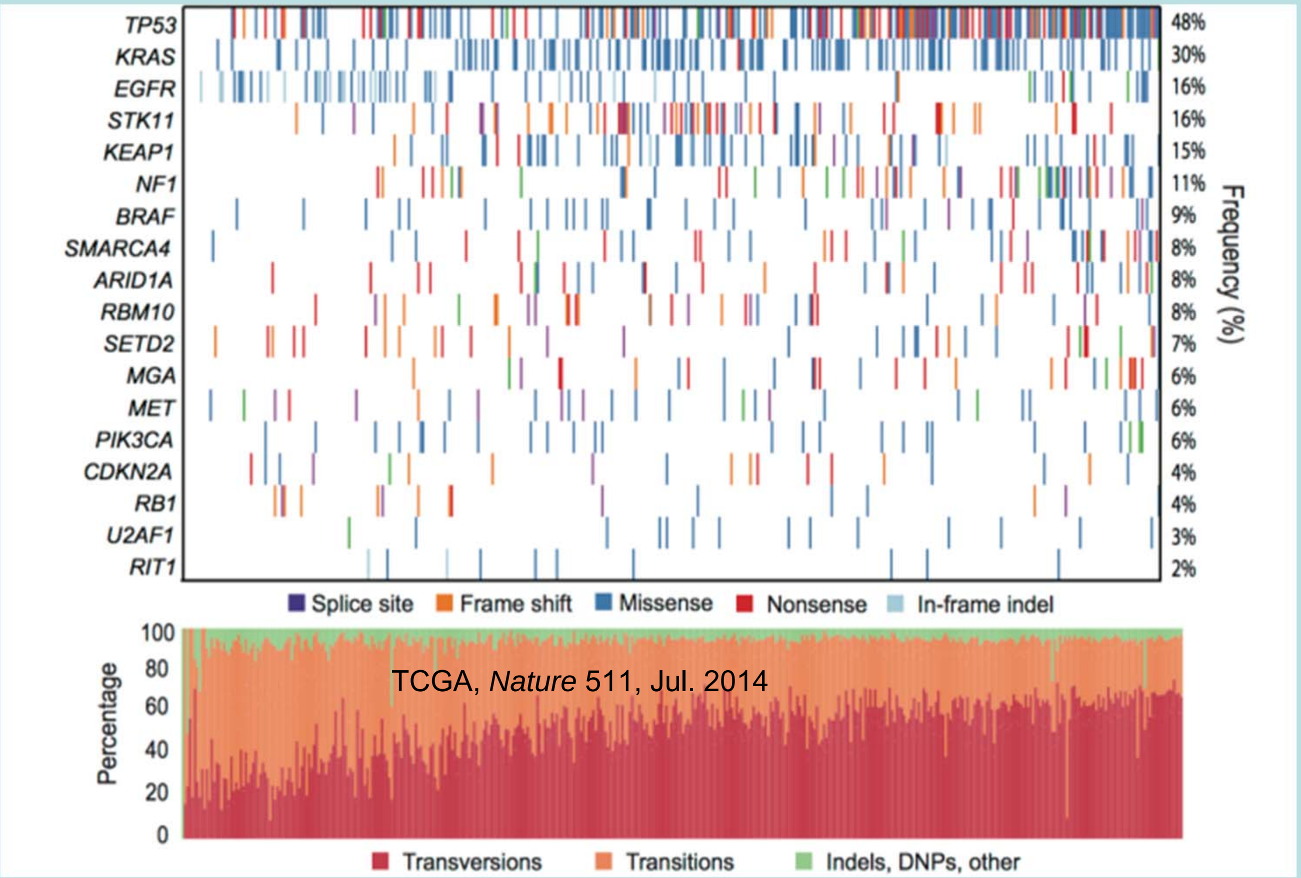
How to derive Function from Structure

Impact on Disease Classification & Prognosis.

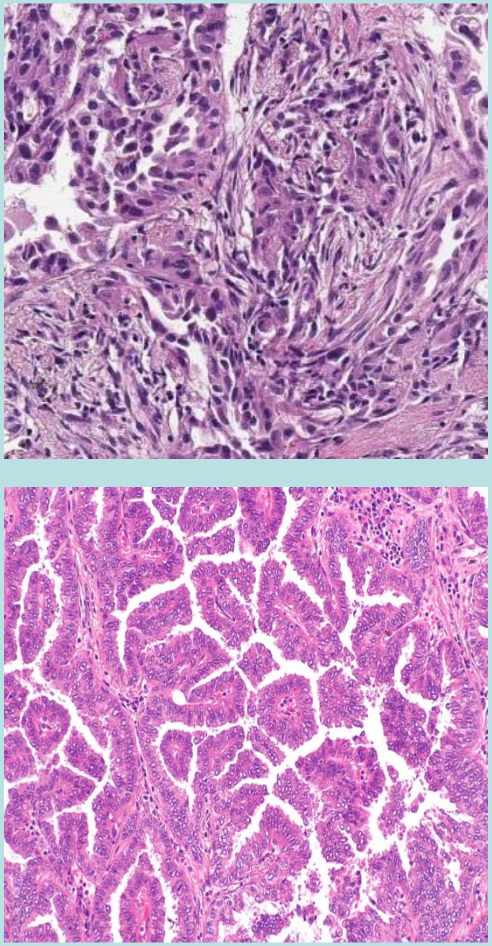
Genetic complexity - morphology in cancer

➔ **TCGA: lung cancer on ave. >800 genetic aberrations per tumor!**

Adenocarcinoma of the lung

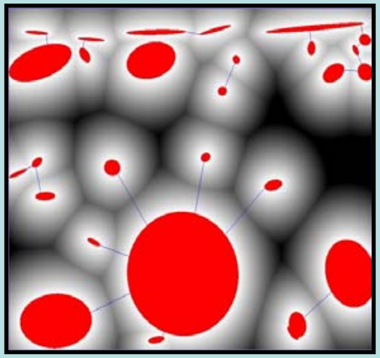


TCGA, *Nature* 511, Jul. 2014



How to describe life (a biological system)?

Biological systems (cells, cellular agglutinations, plants organs, animals, man) can be characterized by three fundamental aspects:



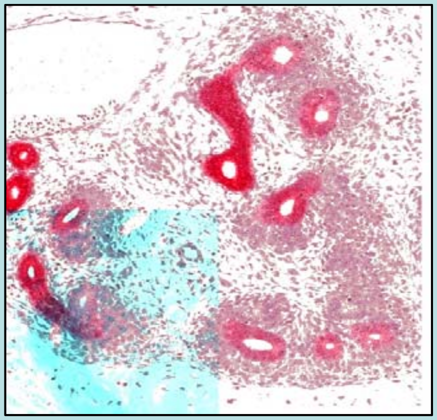
Exchange of energy and specific substances with the environment or specific basic units via their surface.

A discrete boundary separates the non - overlapping inner space from its surrounding (environment).

The inner space might be included in another (larger) space of a basic units (spatial hierarchical order).

Definition: The space within the boundaries is called an object. The (spatial) relationship between objects (for example computed by graph theory) is called a structure.

Definitions

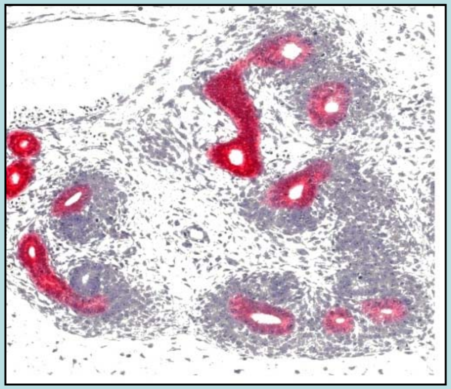


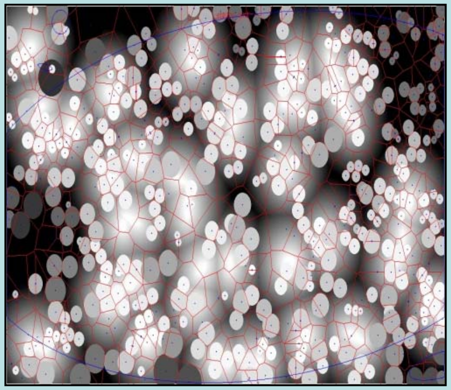
- Predictive diagnosis > forecasts probability and institutes preventive measures in order to prevent or to improve the outcome of a disease.
- Object > Circumscribed area (volume) separated from environment (background).
- Structure > Constellation of objects in the background.
- Texture > Constellation of smallest objects (pixels) in the background.
- Function > Changes of object features and location within the observation time
- Observation Time > Period of image information access

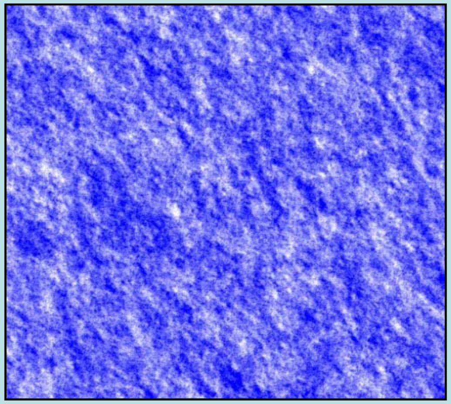
Remarks: A function might jump between objects of different orders, for example might change the structure only, or alter objects without changing the structure.

Object – Structure - Texture

Feature	Biological unit
• Object	nucleus, cell, vessel, etc. (observer defined)
• Structure*	clusters, dependent upon neighborhood condition (Voronoi, O'Callaghan)
• Texture*	pixel based, dependent upon applied transformation.







Remarks: Objects have to be predefined and can form regular arrangements, which can serve for objects at a different (lower) magnification (level of structures). *) Graph Theory applications

Some thermodynamic aspects of structure and function

The hierarchic order of structures assures an (unstable) equilibrium of its environment (object).
Violation of boundaries disturbs the unstable equilibrium.
Disturbances can only be repaired by external actions.
Functions describe different thermodynamic states.

Derivatives: Life is bound to limited spaces of (nearly) identical thermodynamic states, which are separated by boundaries, i.e. structures.
The term entropy is a general descriptor of a thermodynamic system and forecasts its future development.

Remarks: The commutative law (a * b) = (b* a) is not valid in biological systems (in contrast to virtual systems).

Definition of Entropy

Entropy is a concept applied in thermodynamics, statistical mechanics and information theory.

- ἐντροπία, εν- *en-* ‚in-‘ and τροπή = change.
- It is a state variable, a measure of the distance from a terminal ‘stable’ stage.
- Differential of transported heat divided by the absolute temperature $dS = d(Q)/T + d(W_{diss})/T$.
- $D(Q)$ = additive heat; $d(W_{diss})$ = inside created heat.
- Structure entropy $S(MST) = - k * \Sigma\{p(\sigma,\delta) * \ln p(\sigma,\delta)\}$.+)
- Entropy Flow $SF = dS * C^{-1}/dt$ (C = surface, t = time)

Examples: It is a measure of

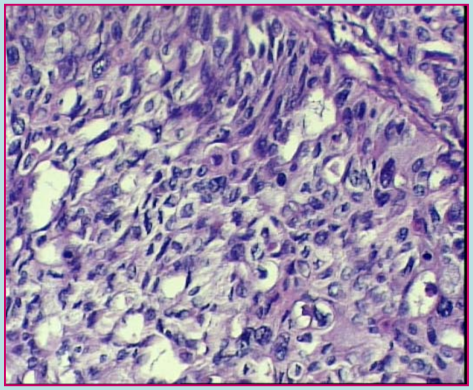
- direction to and distance from the final stage,
- amount of energy not available for work,
- disorder of a system.

+) $p(\sigma,\delta)$ = probability of structure features (object size, distance, etc.)

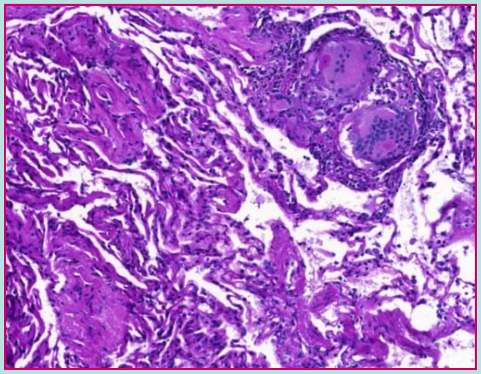
Image Information distribution

- Different image areas might display different expression of medical information.
- How to define those areas with significant information?
- How to estimate the most appropriate image size (for the pathologist, for the IT system)?
- How to detect?

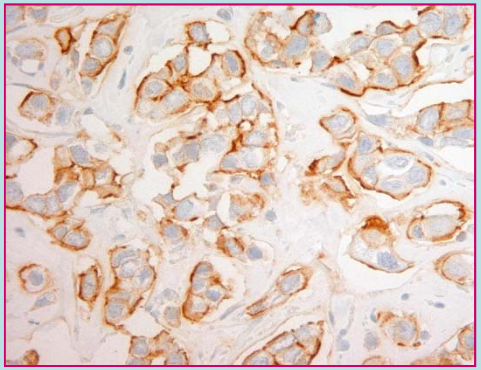
random



singular



multiple focal



How to detect region of interest?

Sliding frames (fixed size, >256 * 256 pixels):

- Applicable for texture, object, structure analysis
- Comparison of image distances

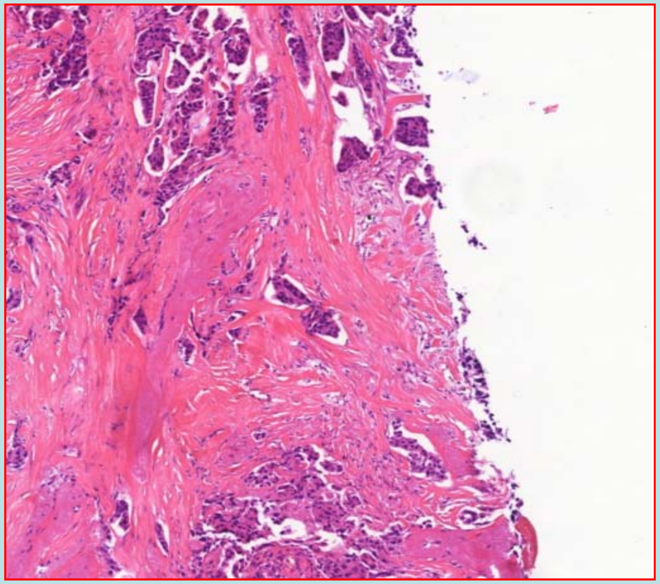
Variable frames, computed image information:

- Graph theory, neighborhood, connectivity, MST – entropy, gray value transformation.

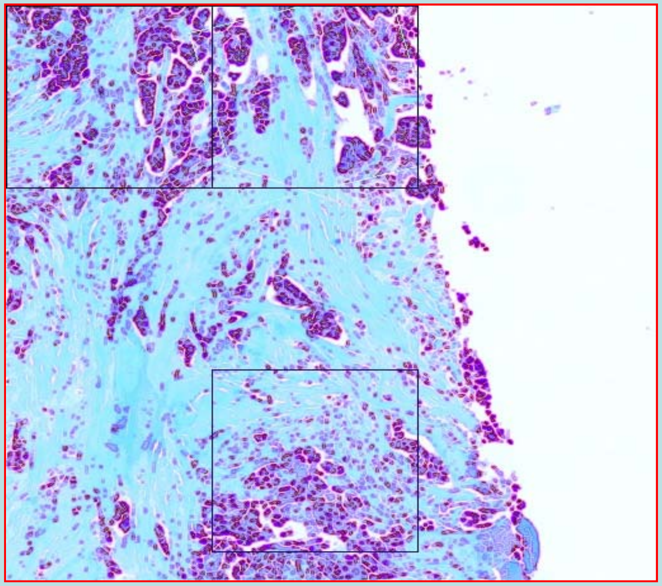
Frequency analysis (Oger et al, 2008)

- Stochastic geometry
- Syntactic structure analysis of primitives

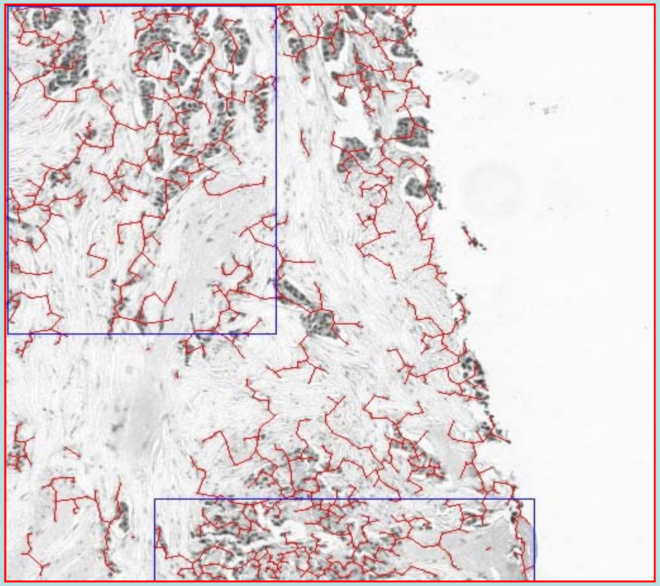
Automated selection of fields of view



Original



Selected areas each
10% of original



Selected areas by minimum
spanning tree, no relation
to original image size

Application: Automated selection of fields of view

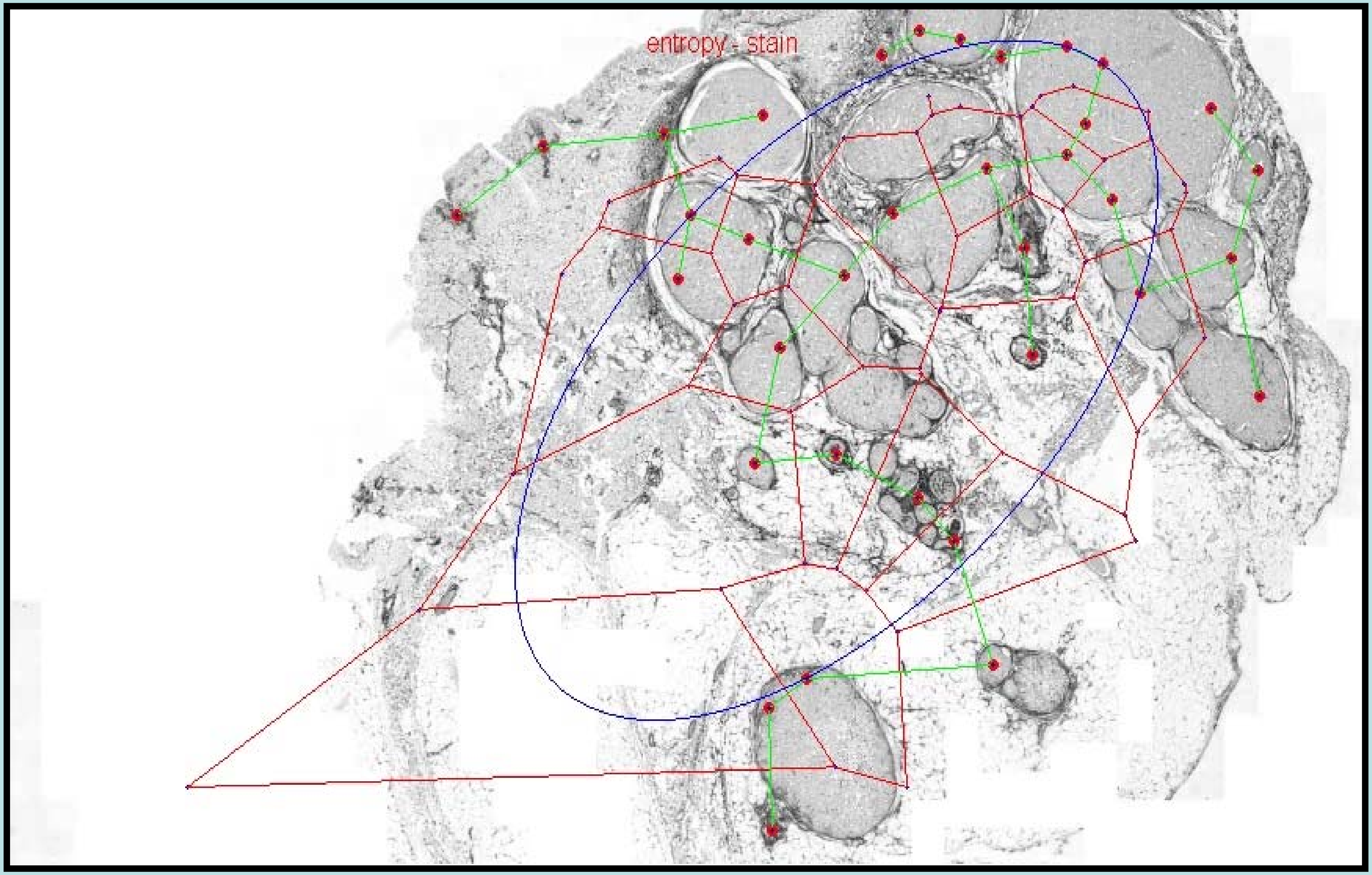
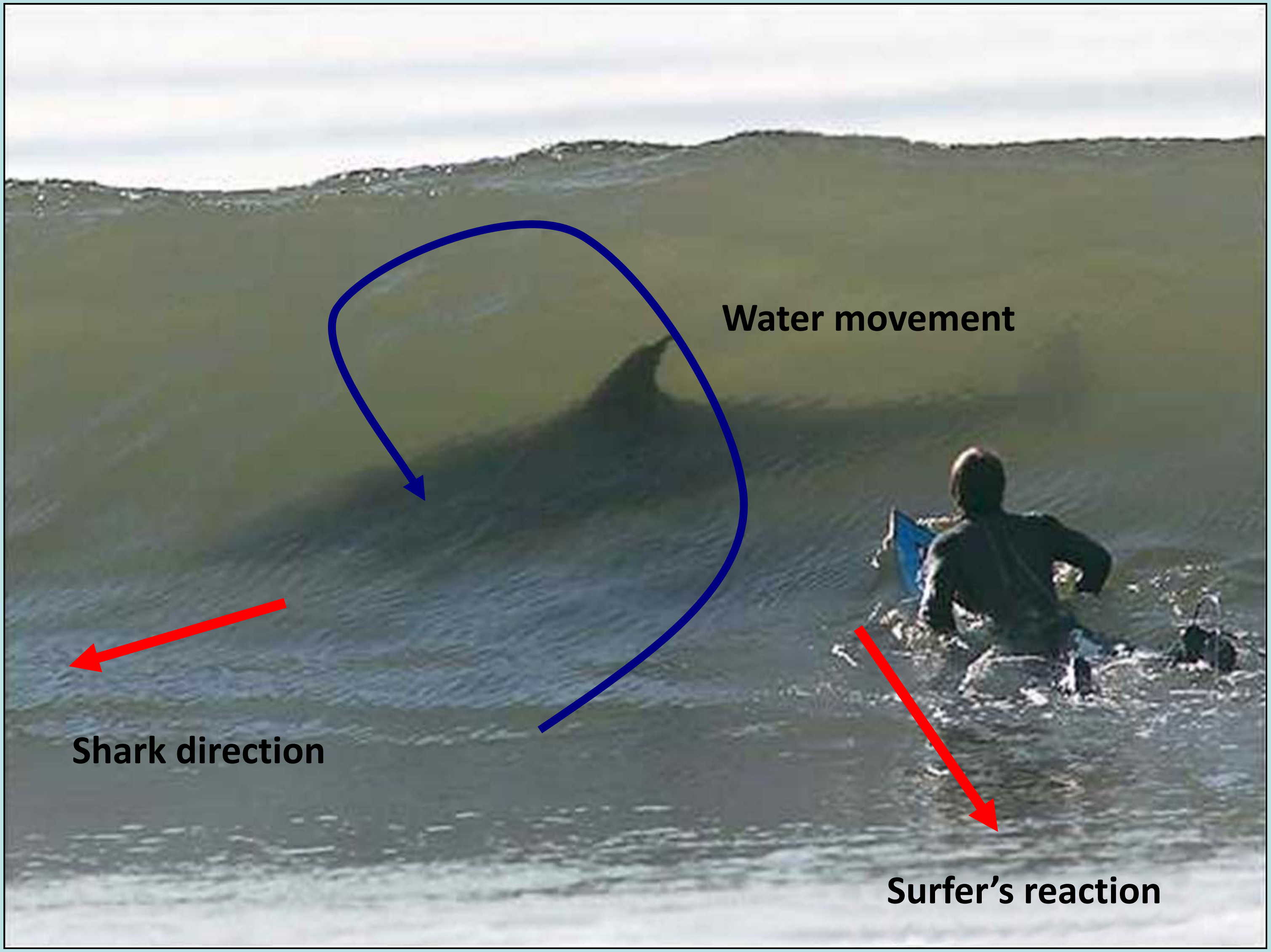


Image Content Information

Definitions:

- **Image content information** comprises the information that an external observer can extract solely from the image itself, i.e., without additional notifications (labels, classification, etc.).
- **Image analysis** is the procedure to extract meaningful information from the image.
- **Meaningful information** comprises the information which the observer can understand or which is transformed into an information related reaction.

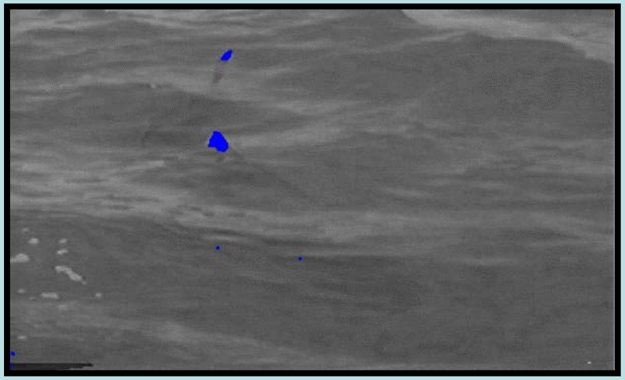
Hidden Information*



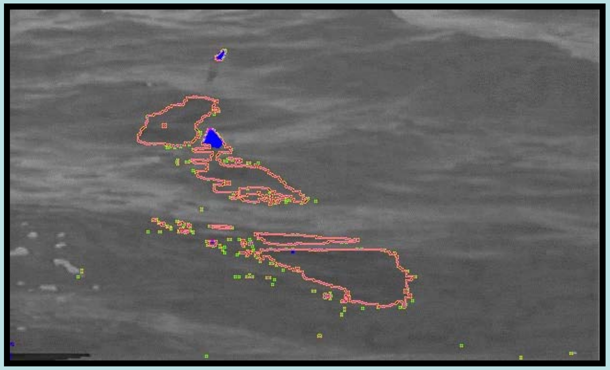
*) Frame (context) of Image Content Information

Application of advanced object identification

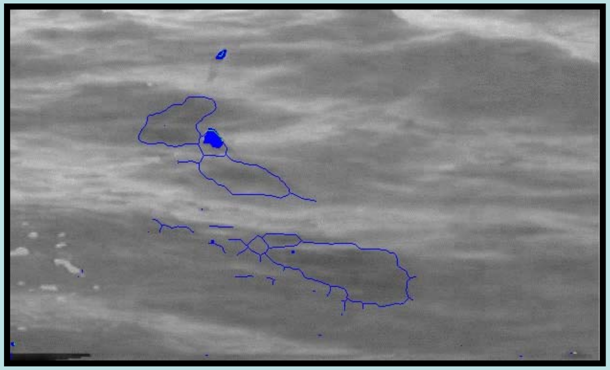
Original image



Fin segmentation



Color segmentation dependent on fin location



Gradient segmentation dependent on fin location

Evaluation of meaningful image information

Structure analysis

- Recognize objects
- Analyze features & spatial distribution
- Analyze spatial distribution (symmetries, clusters)
- Analyze gray value intensities (IHC, FTC)
- Classify results.

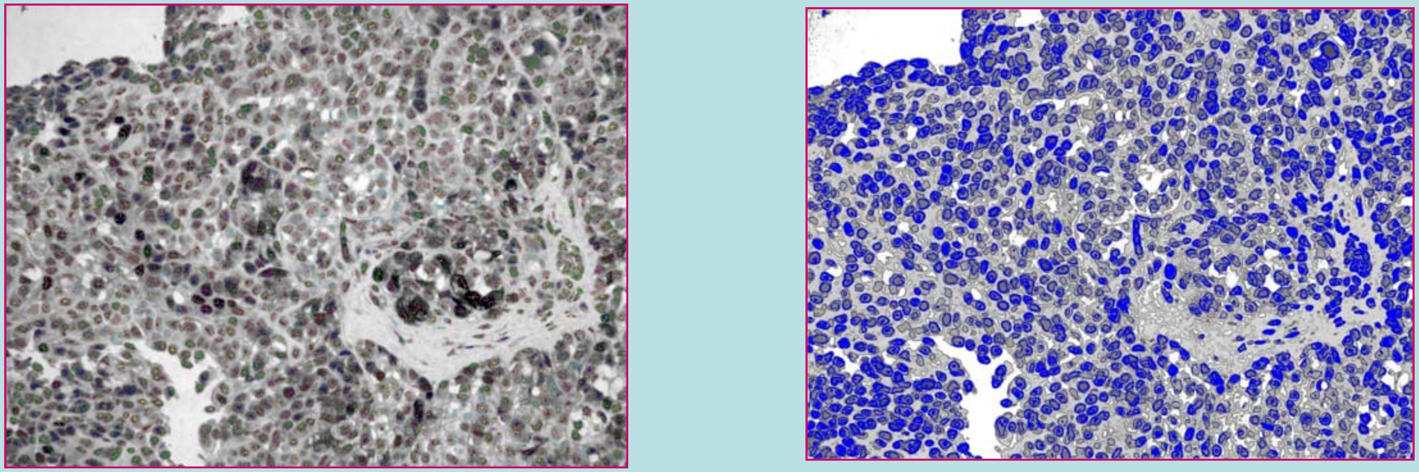
Function analysis

Analyze structure differences in ordered gray value planes, compute biological meaningful parameters (entropy, structural entropy).

Composition of objects by pixel related information

Objects consist of gray value compartments

1. **separated points p_{is} [$gr(p_{is}) >$ threshold, $n(p_{is})=0$]**
2. **fibers lp_{is} [$gr(lp_{is}) >$ threshold, connected p_{is}, $1 \leq n(p_{is}) < 3$, $p_i \# p_e$]**
3. **rings rp_{is} [connected p_{is} with $p_i=p_e$]**
4. **plateaus pp_{is} [$gr(pp_{is}) >$threshold, $n(pp_{is})>2$]**

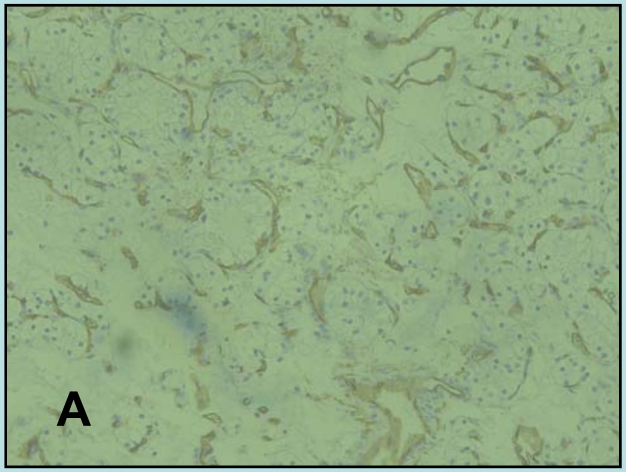


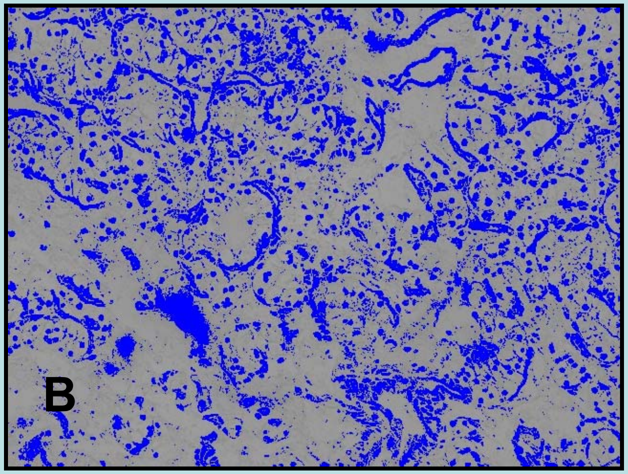
Any segmented object is composed of {p, l, r, pp} and their spatial relations.

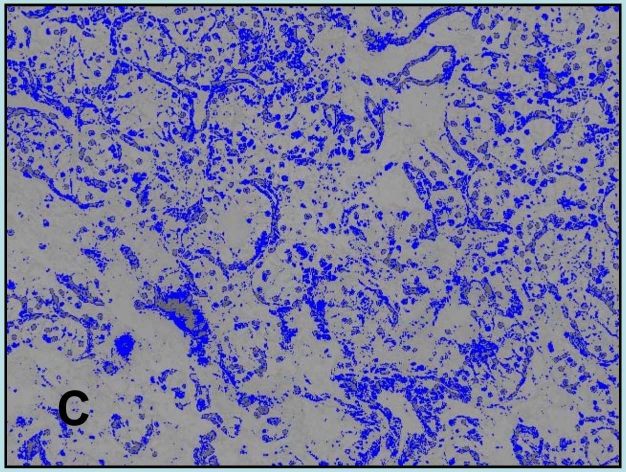
Image information <– mapped structure & function of elements

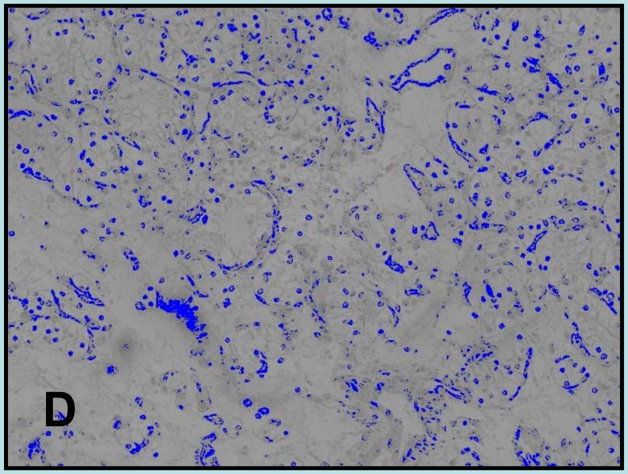
- Structure = {elements which do not change during the observation (measurement) period (f(t)},
- Function = {changes of elements in relation to the total environment at different observation times}, i.e. changes & movement of structures within a certain period, i.e. the spatial gradient of elements.
- Thus: The visualization of structures & functions depends on the spatial resolution & observation time.
- In life, structures form closed spaces that define specific environments for new structures (order of structures).
- Functions usually involve neighboring structures at the same and higher order structures (from the bottom up).
- Simulation by gray value adjusted segmentation (0 < tr < 8 (12) bits).

Impact of observation period on structure recognition









Simulation of observation period *t* by gray level thresholds:

A: Original

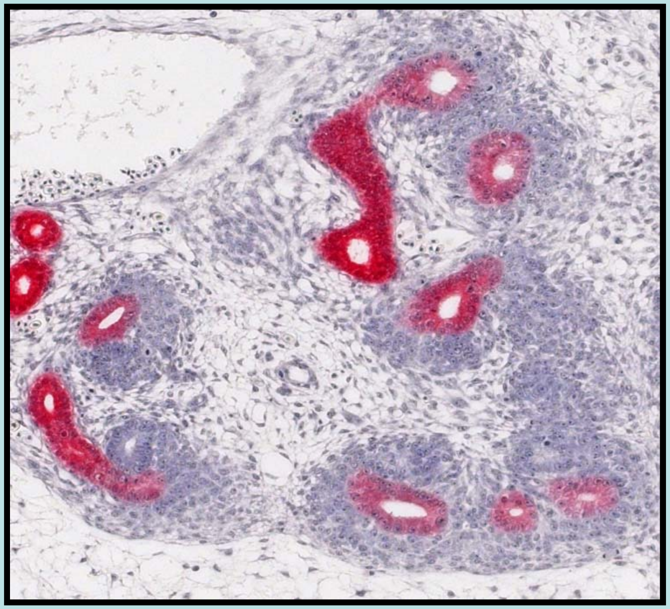
B: *t* >> s(t)*

C: t << s(t)

D: derived space related function f(s)

*) s(t) = period of fixed structures

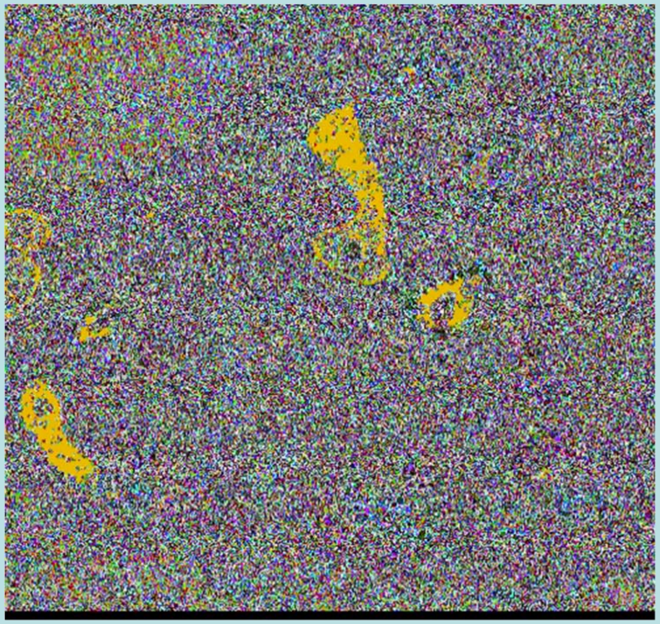
Virtual Simulations (*Day 13, CG-3*)



Original

Segmented
Gradient

Functions



Segmented
Original

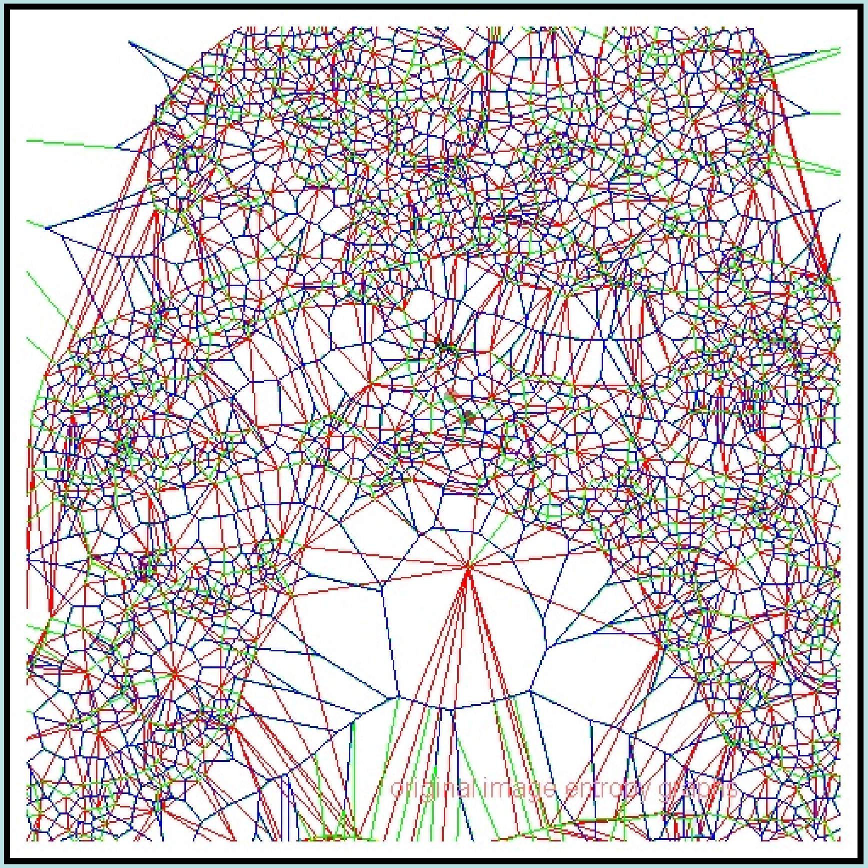
Algorithms of predictive diagnosis

Surgery, Radiology H&E, IHC	Conventional diagnosis	Tumor extension
IHC, FISH	Membrane receptors (*)	Cellular communication
IHC, PCR, FISH, TMA	Proteins, RNA, miRNA (**)	Intracellular pathways
PCR, RNA, DNA assays	Cancer-related target genes (***)	Gene anomalities

(*) EGFR, VGFR, etc.; (**) (*BAG1, BRCA1, CDC6, CDK2AP1, ERBB3, FUT3, IL11, LCK, RND3, SH3BGR, WNT3A*); (***) raf, myc, EML4,…

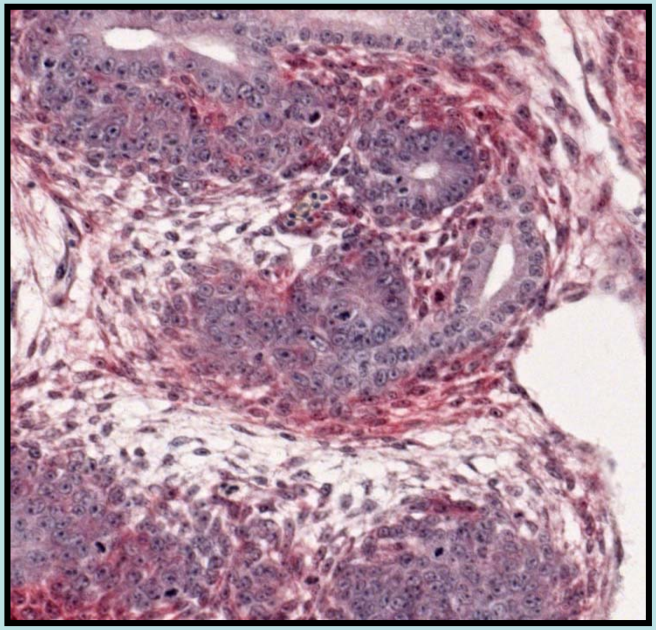
Example: development of chicken kidneys embryos+)

- Study aim: To analyze the network of potentially involved galectins (CG-1A, CG-2, CG-3) in the embryonic stages of chicken kidneys.
- Chicken kidney at day 9, 11, 16, 19 21.
- Number of cases: 8 per cohort.
- Applied markers: Galectin 1A, Galectin 2, Galectin 3.

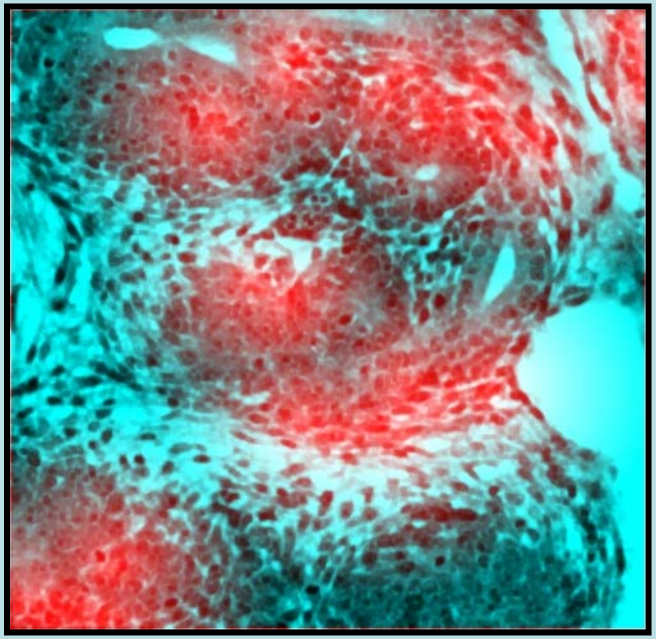


+) In collaboration with H.-J Gabius, H. Kaltner, J. Manning, Munich.

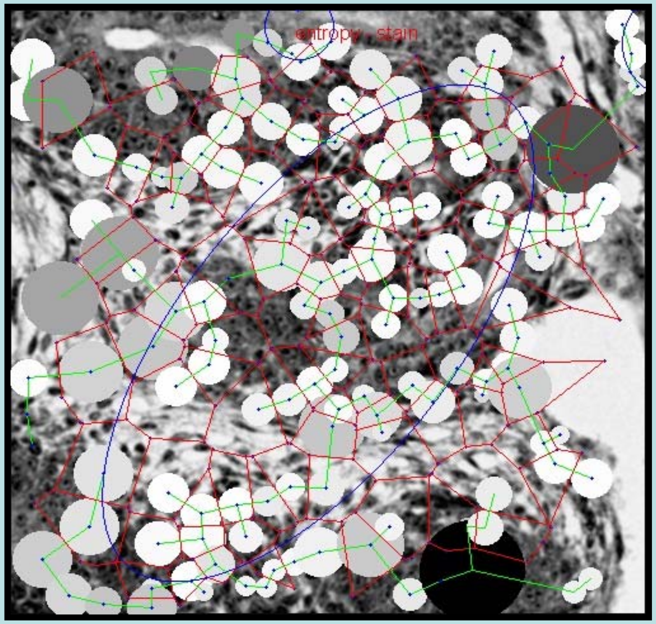
Morphology of CG1A applied at day 11, objective *20



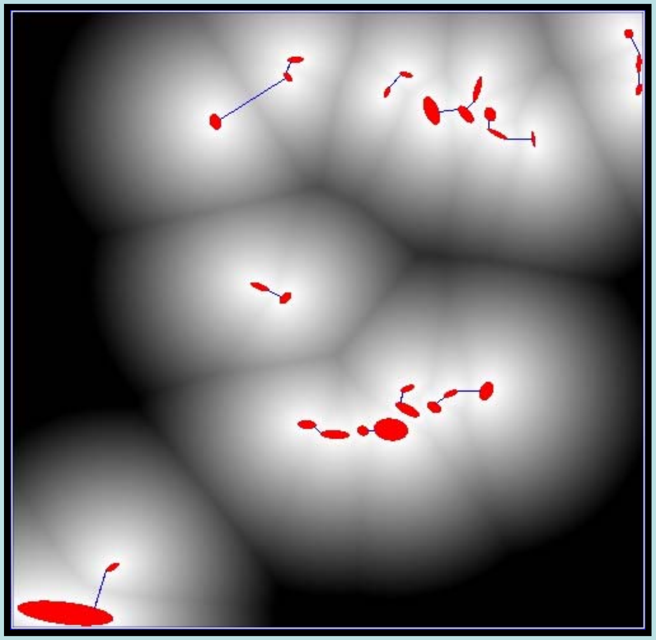
Original



CG1A regions

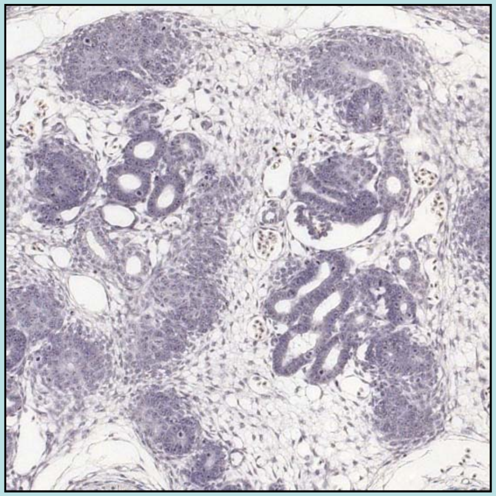


Entropy patterns

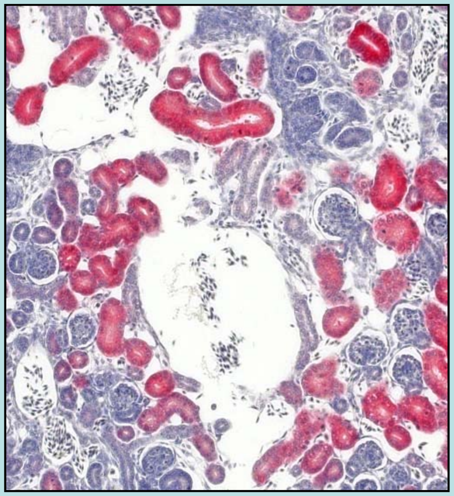


Intensity patterns

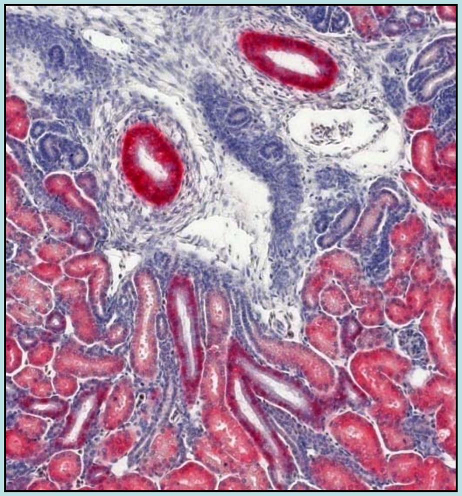
Entropy gradients of grey value thresholds
chicken kidneys, embryonic period, day 11 - 21



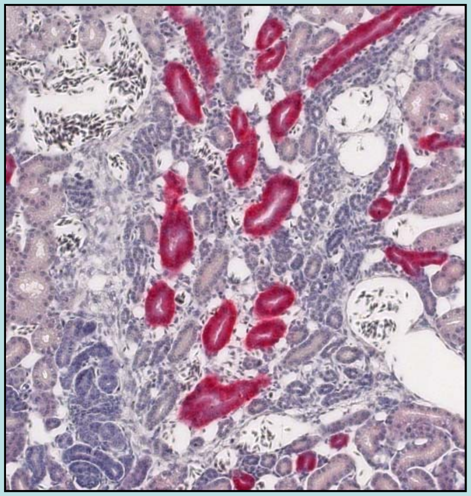
35_20 CG-2 day 11



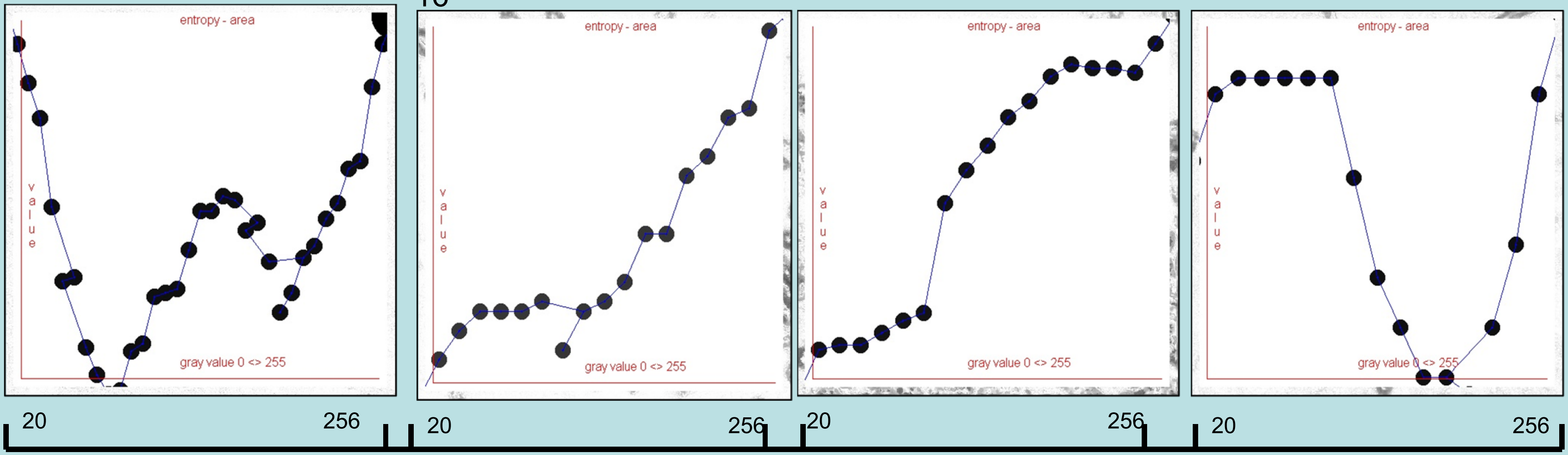
69_20 CG-2 day 16



104_20 CG-2 day 19



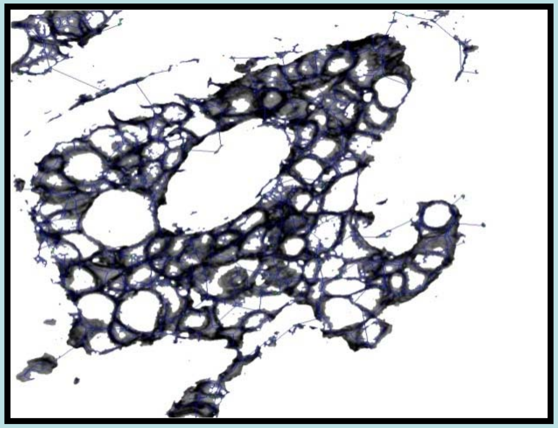
130_20 CG2 day 21

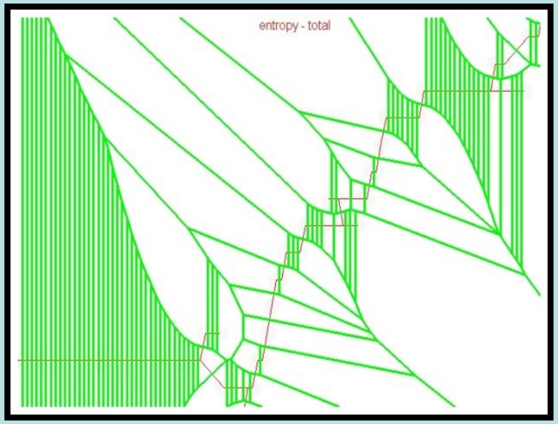


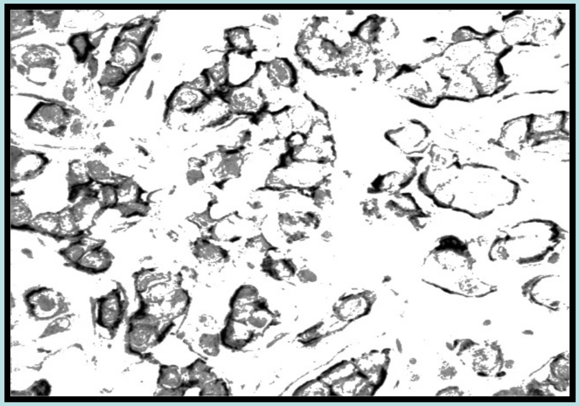
Grey value thresholds

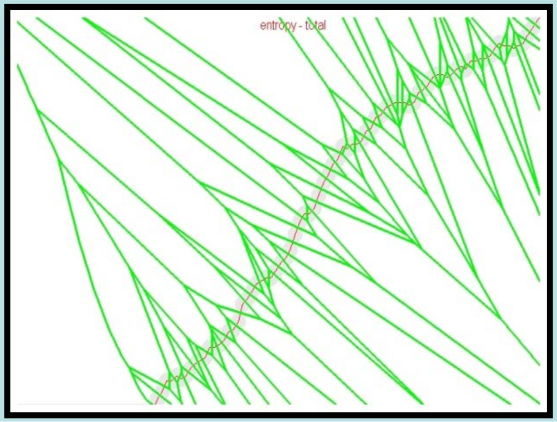
Entropy approach via gray value levels

- Take the rgb (or his) gray value levels
- Segment stepwise at all gray value
- Compute entropies
- Create entropy – gray value curves
- Transform the image (Fourier, Hough, recursive, gradient, etc.)
- Repeat algorithm.



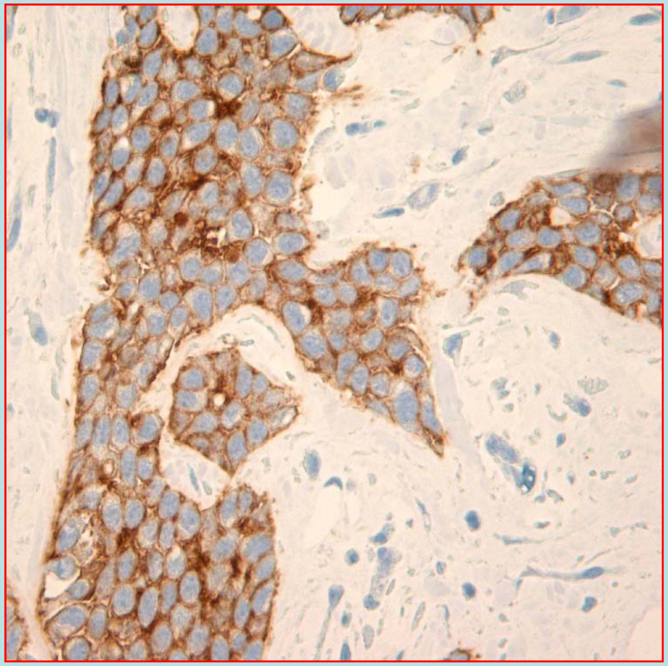






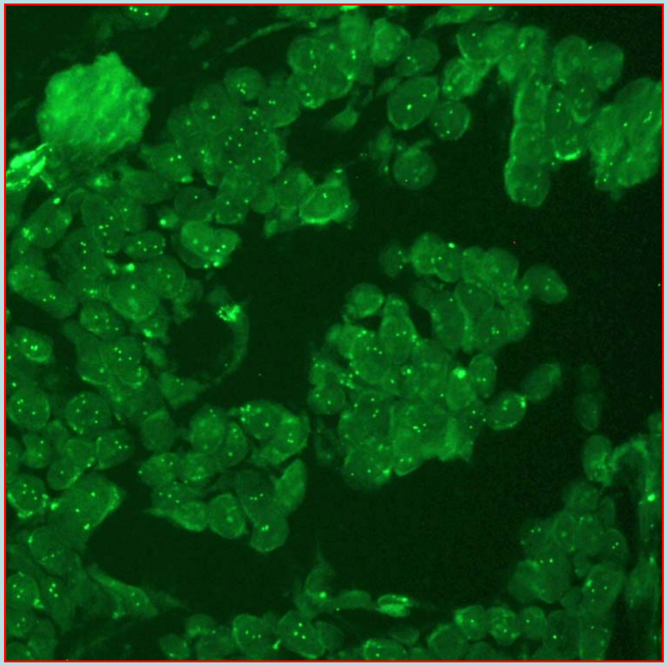
Her2_new Membrane stain:
Up: score 3
Down: score 2

Her2_neu breast carcinoma / entropy



Score 2+
Trastuzumab
therapy ?

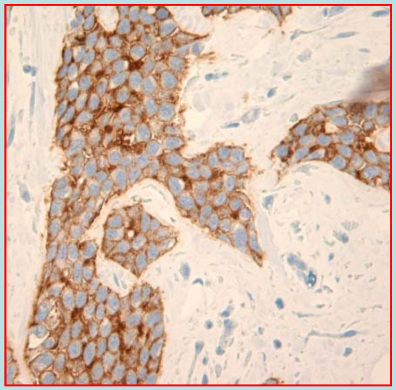
FISH
amplification:
Her2/neu gene



Microstages: tumor cells
IHC macrostages:
Tumor cells {3+, 2+, 1+, 0}
Connecting membranes
{n,M+}, (n=1-4, M = 0 - 3)
$\Sigma = 16$

Microstages: tumor cells
FISH macrostages:
Nuclear signals {n, F+}
n number of neighbors
(n=1- 4, F = 0 – 3)
$\Sigma = 16$

Results: IHC



Microstages: Σ tumor cells: 320

M(0+) = 22, M(1+)=66, M(2+)=214,

M(3+)=38, N = No neighbors

Entropy: 1.03 Macrostages: 16

	N	1	2	3	4	Entropy
M	*0*	9	13	0	0	0.58
M	*1*	28	20	8	10	1.19
M	*2*	16	68	82	47	1.19
M	*3*	19	16	3	0	1.09

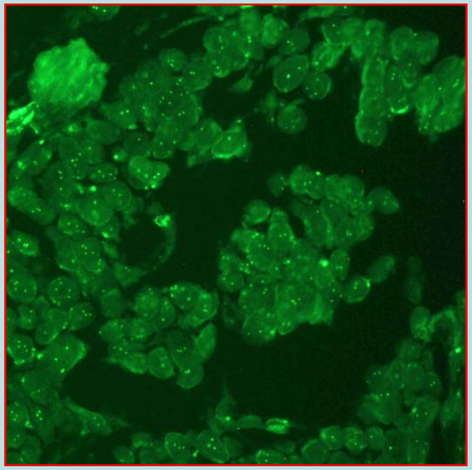
Results: FISH

Microstages: Σ tumor cells: 120

FISH(0+)= 8, FISH(1+)=16,

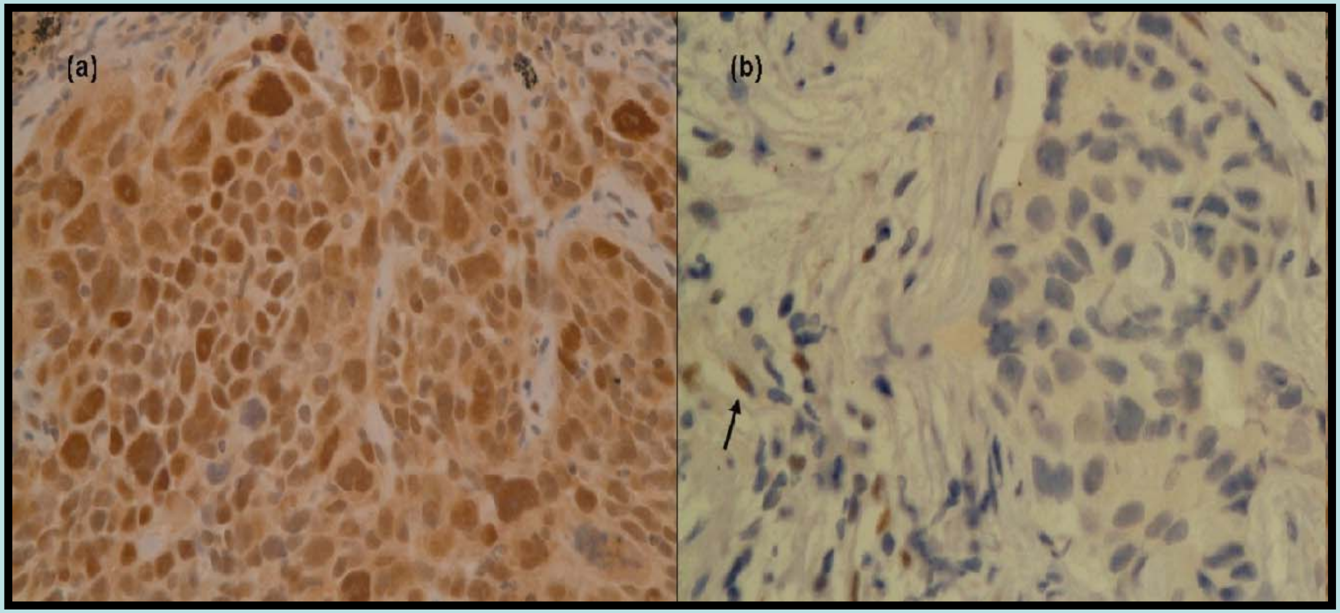
FISH(2+)=46, FISH(3+)=50

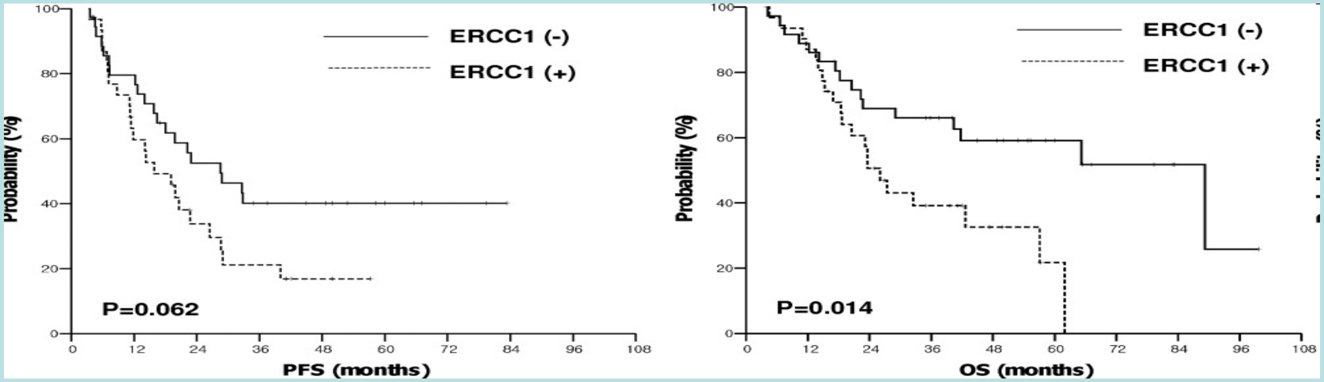
Entropy: 1.18 Macrostages: 16



	N	1	2	3	4	Entropy
FISH	*0*	5	3	0	0	0.35
FISH	*1*	6	4	6	0	1.66
FISH	*2*	6	9	17	14	1.13
FISH	*3*	9	23	13	5	1.23

Molecular marker – ERCC1, entropy micro- and macrostages

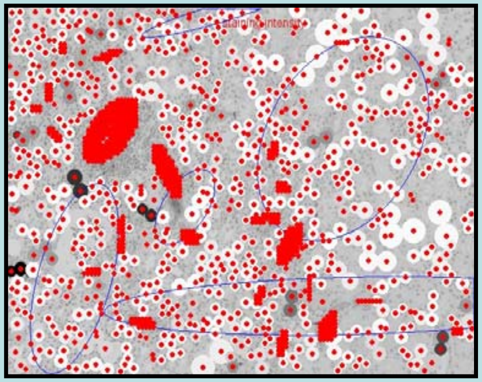




Hwang Cancer 2008 113:1379-1386

320	**Σ No cells**	**160**
{0, 30, 210, 80}	**{0, 1+, 2+, 3+}**	**{148, 12, 0, 0}**
0.84	**Entropy**	**0.28**
1.67 ± 0.3	**MST Entropy**	**1.43 ± 0.2**
2.43	**Σ Entropy macrostages**	**0.92**

Conclusions



- The development of a disease (predictive diagnosis) can be successfully predicted be measuring entropy (flows, functions) between different structural levels.
- These functions depend upon the observation time.
- Mapping entropy and entropy flow of correspondingly stained slides are appropriate tools to automate predictive diagnosis measures.
- The procedure might be applied to different levels of structures, including genes and macromolecules.

thank you for your attention