

Computer vision in the new era of Artificial Intelligence and Deep Learning

Visión por computador en la nueva era de la Inteligencia Artificial y el Deep Learning

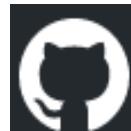
Rubén Usamentiaga*, **Alberto Fernández°**

***Universidad de Oviedo**

°TSK

Gijón

5 – 16 Abril 2021



<https://github.com/albertofernandezvillan/computer-vision-and-deep-learning-course>

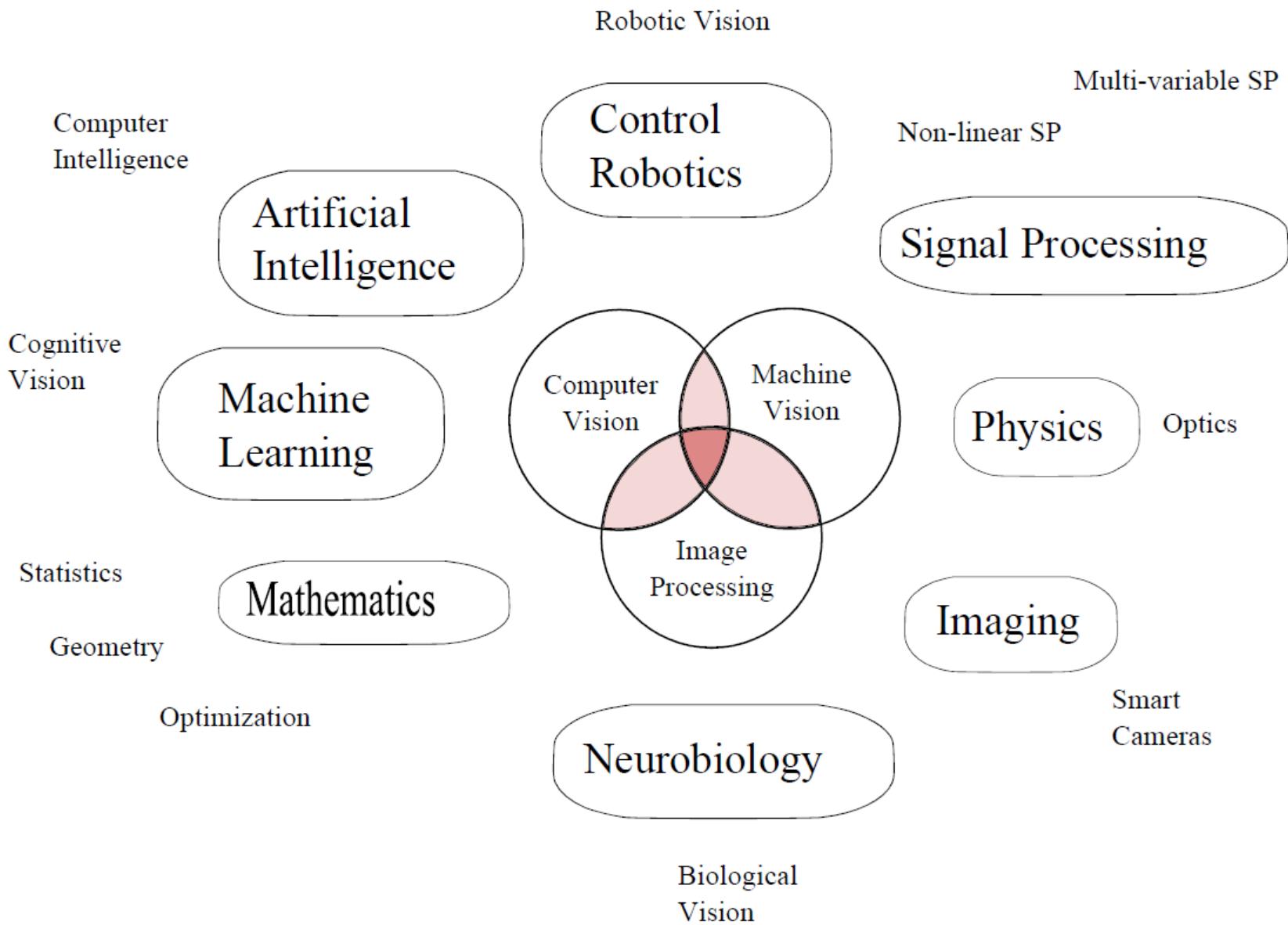
Professor of this course



- **Rubén Usamentiaga Fdez**
 - Profesor Titular de Universidad (acreditado a Catedrático)
 - Architecture and Computer Technology
 - >20 years of research in computer vision



- **Alberto Fernández Villán**
 - R&D Projects Engineer at TSK
 - PhD Universidad de Oviedo
 - Author of the book “Mastering OpenCV 4 with Python” and papers in conferences and journals





Computer vision

Computer vision

- Computer vision
 - Field of science that deals with the design of systems that are capable of understanding and interpreting the content of images
- Image
 - Representation that is formed in the human brain when receiving light produced or reflected by objects

Computer vision

□ Sampling



(a)



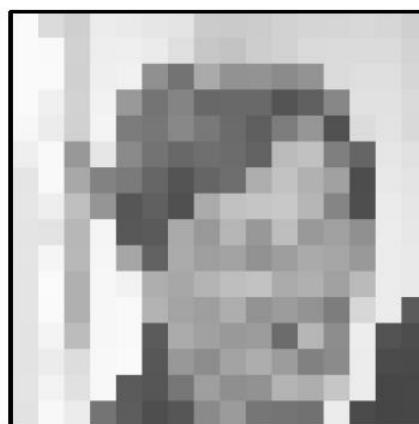
(b)



(c)



(d)



(e)



(f)

Computer vision

□ Quantization



(a)



(b)



(c)



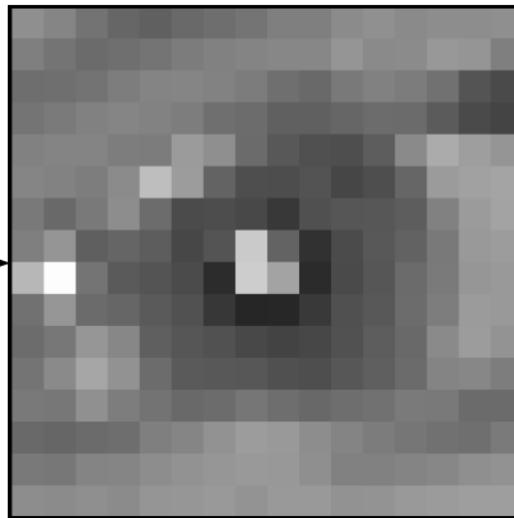
(d)



(e)



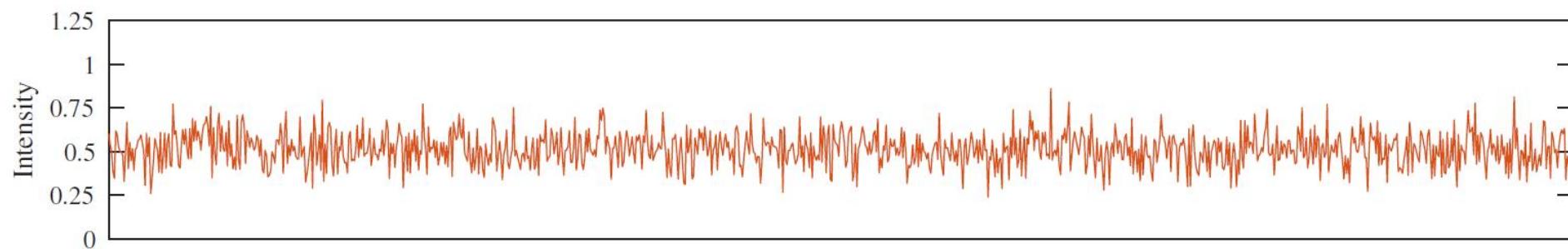
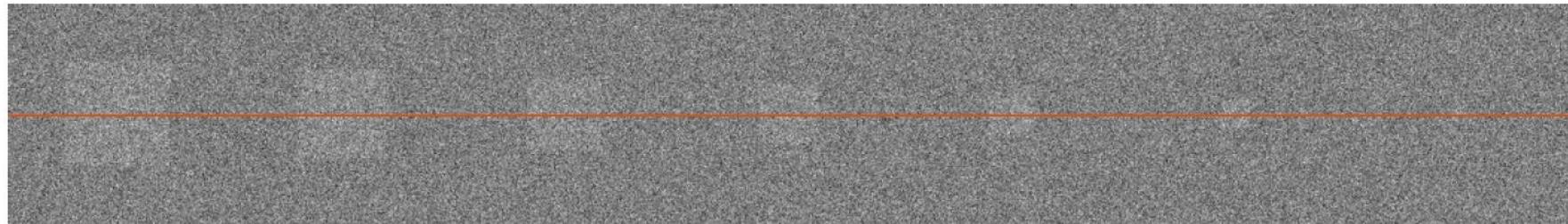
(f)



138	132	120	112	108	114	117	121	128	128	136	139	135	137	137	138
128	121	115	118	121	126	129	131	133	133	142	135	136	144	142	128
117	118	120	127	131	132	130	126	118	114	124	129	126	119	100	93
121	125	129	132	130	127	118	116	108	108	112	116	115	105	93	89
129	131	131	127	126	146	138	115	103	97	96	106	135	157	148	142
132	130	127	136	170	147	110	96	95	99	90	96	112	146	150	152
124	114	124	137	116	94	95	92	80	97	101	102	106	129	146	151
128	142	107	111	106	91	100	179	110	74	94	103	109	124	145	147
165	243	121	104	100	94	70	182	151	69	93	101	115	125	143	145
119	143	115	112	104	98	79	65	66	81	97	103	116	127	146	145
116	123	143	133	107	101	98	92	88	91	98	106	114	135	147	141
121	135	153	140	116	104	103	102	97	96	102	107	115	131	133	127
124	123	139	127	120	114	116	122	117	113	117	119	125	124	115	115
114	112	115	117	128	132	140	147	145	137	131	126	122	119	118	125
124	123	131	132	137	140	143	145	144	138	129	129	131	133	138	140
137	136	138	136	142	143	145	141	145	145	140	141	145	146	148	148

Computer vision

□ Difficulties: noise



Usamentiaga, R., Clemente Ibarra-Castanedo, and X. Maldague. "More than Fifty Shades of Grey: Quantitative Characterization of Defects and Interpretation Using SNR and CNR." *Journal of Nondestructive Evaluation* 37.2 (2018): 25.

Computer vision

- Difficulties: lighting



Source: Nacho Guzman

Computer vision

- Difficulties: differentiate object from background



Picture of Adele Bloch Bauer I (Gustav Klimt)

Computer vision

□ Difficulties: variation between classes



Brady Bunch House' Bought By HGTV – Variety
variety.com



House - Wikipedia
en.wikipedia.org



House Lights Turned on · Free Stock Photo
pexels.com



Sliding glass walls allow sea breezes ...
dezeen.com



Brady Bunch house is for sale for \$1.9M ...
la.curbed.com



Inheriting a House
aarp.org



Farmhouse Style House Plan - 3 Beds 2...
houseplans.com



Prefab smart home, Flex House ...
curbed.com



Simple House / Moon Hoon | ArchDaily
archdaily.com



Manmade gorge splits spectacular iron ...
newatlas.com



The Block 2017: Price range for Josh ...
domain.com.au



The Key Hole House | CIRCA Old House...
circaoldhouses.com



House VAP / Ney Lima | ArchDaily
archdaily.com



House Plans & Floor Plans | Porter Davis
porterdavis.com.au



WATCH: The Texas-Size Makeover ...
southernliving.com



Grass Field - Free Stock Photo
pexels.com



Why don't you own a house yet?
dailyexcelsior.com



Detached Houses For Sale in Pontefrac...
rightmove.co.uk

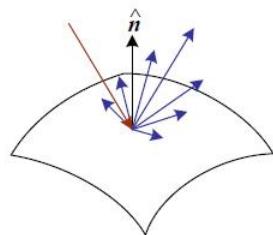
Source: google search

Visión por computador

□ Difficulties: occlusions



Source: LA times



2 Image formation

- Geometric primitives and transformations •
- Photometric image formation •
- The digital camera



3 Image processing

- Point operators • Linear filtering •
- More neighborhood operators • Fourier transforms •
- Pyramids and wavelets • Geometric transformations •
- Global optimization



4 Feature detection and matching

- Points and patches •
- Edges • Lines



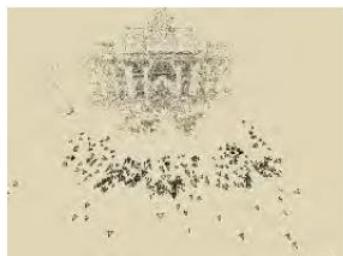
5 Segmentation

- Active contours • Split and merge •
- Mean shift and mode finding • Normalized cuts •
- Graph cuts and energy-based methods



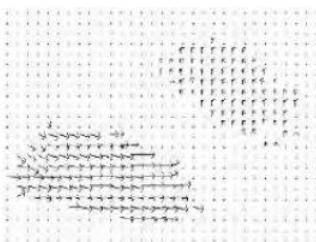
6 Feature-based alignment

- 2D and 3D feature-based alignment •
- Pose estimation •
- Geometric intrinsic calibration



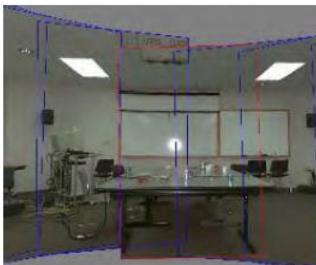
7 Structure from motion

- Triangulation • Two-frame structure from motion •
- Factorization • Bundle adjustment •
- Constrained structure and motion



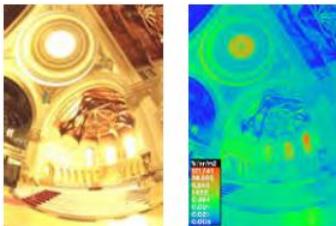
8 Dense motion estimation

- Translational alignment • Parametric motion •
- Spline-based motion • Optical flow •
- Layered motion



9 Image stitching

- Motion models • Global alignment •
- Compositing



10 Computational photography

- Photometric calibration • High dynamic range imaging •
- Super-resolution and blur removal •
- Image matting and compositing •
- Texture analysis and synthesis



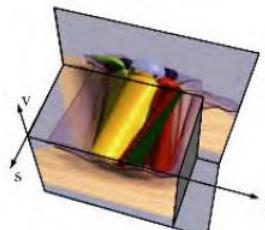
11 Stereo correspondence

- Epipolar geometry • Sparse correspondence •
- Dense correspondence • Local methods •
- Global optimization • Multi-view stereo



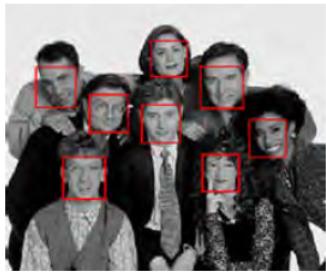
12 3D reconstruction

- Shape from X • Active rangefinding •
- Surface representations • Point-based representations •
- Volumetric representations • Model-based reconstruction •
- Recovering texture maps and albedos



13 Image-based rendering

- View interpolation • Layered depth images •
- Light fields and Lumigraphs • Environment mattes •
- Video-based rendering

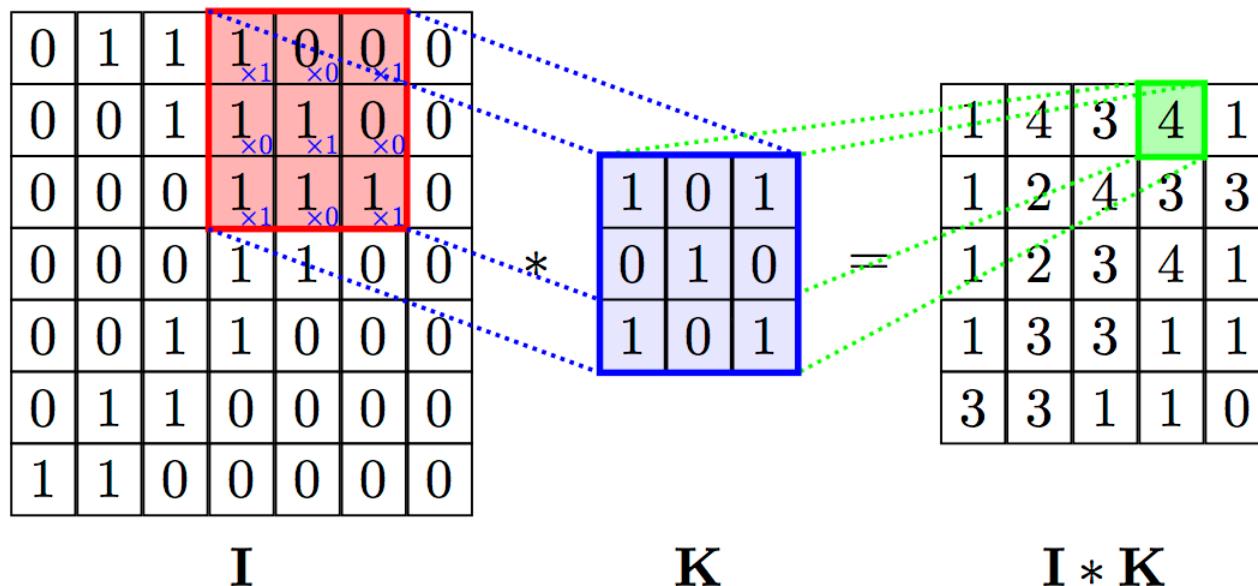


14 Recognition

Object detection • Face recognition •
Instance recognition • Category recognition •
Context and scene understanding •
Recognition databases and test sets

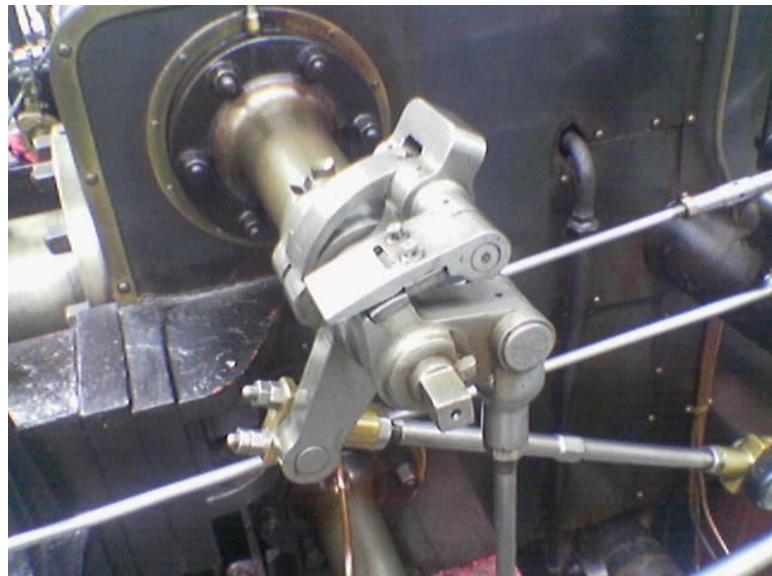
Computer vision

□ Convolution



Computer vision

□ Convolution



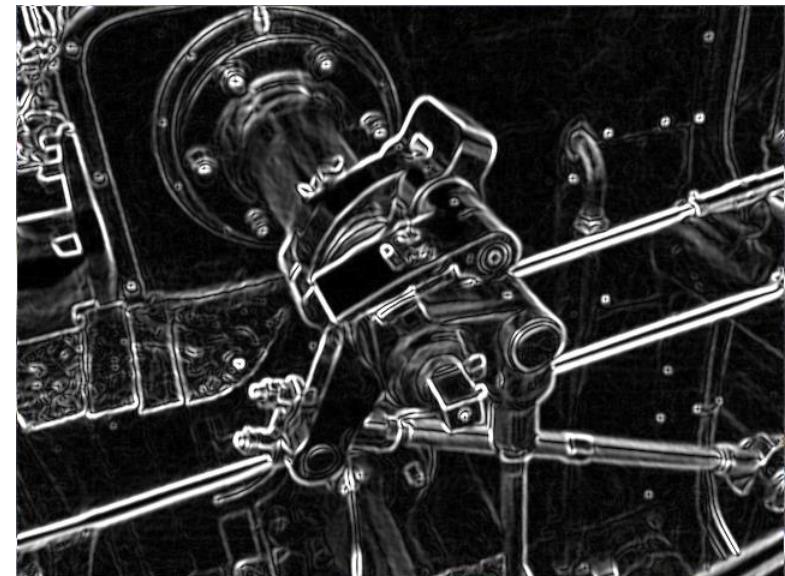
Sobel Kernel

-1	0	+1
-2	0	+2
-1	0	+1

Gx

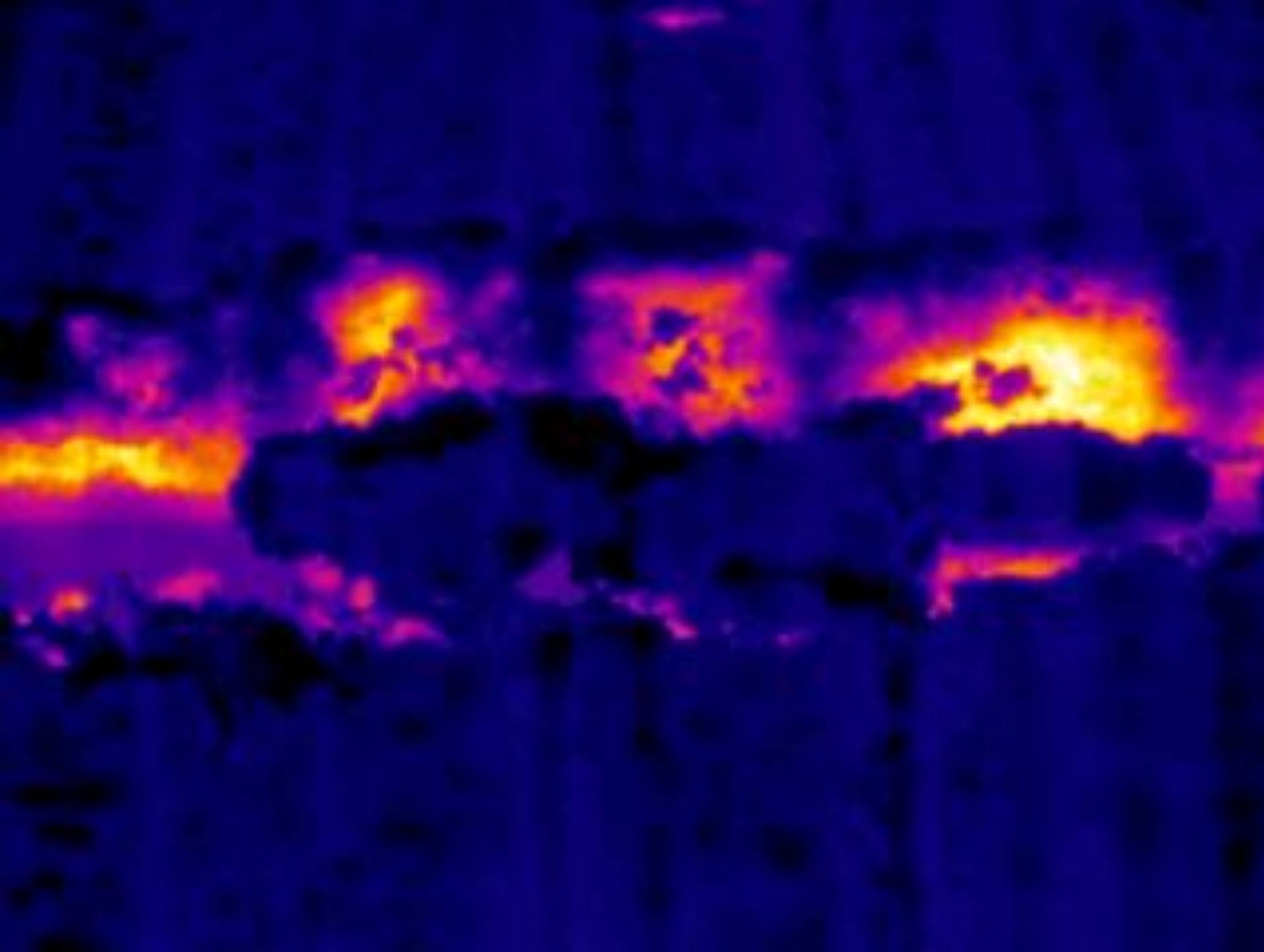
+1	+2	+1
0	0	0
-1	-2	-1

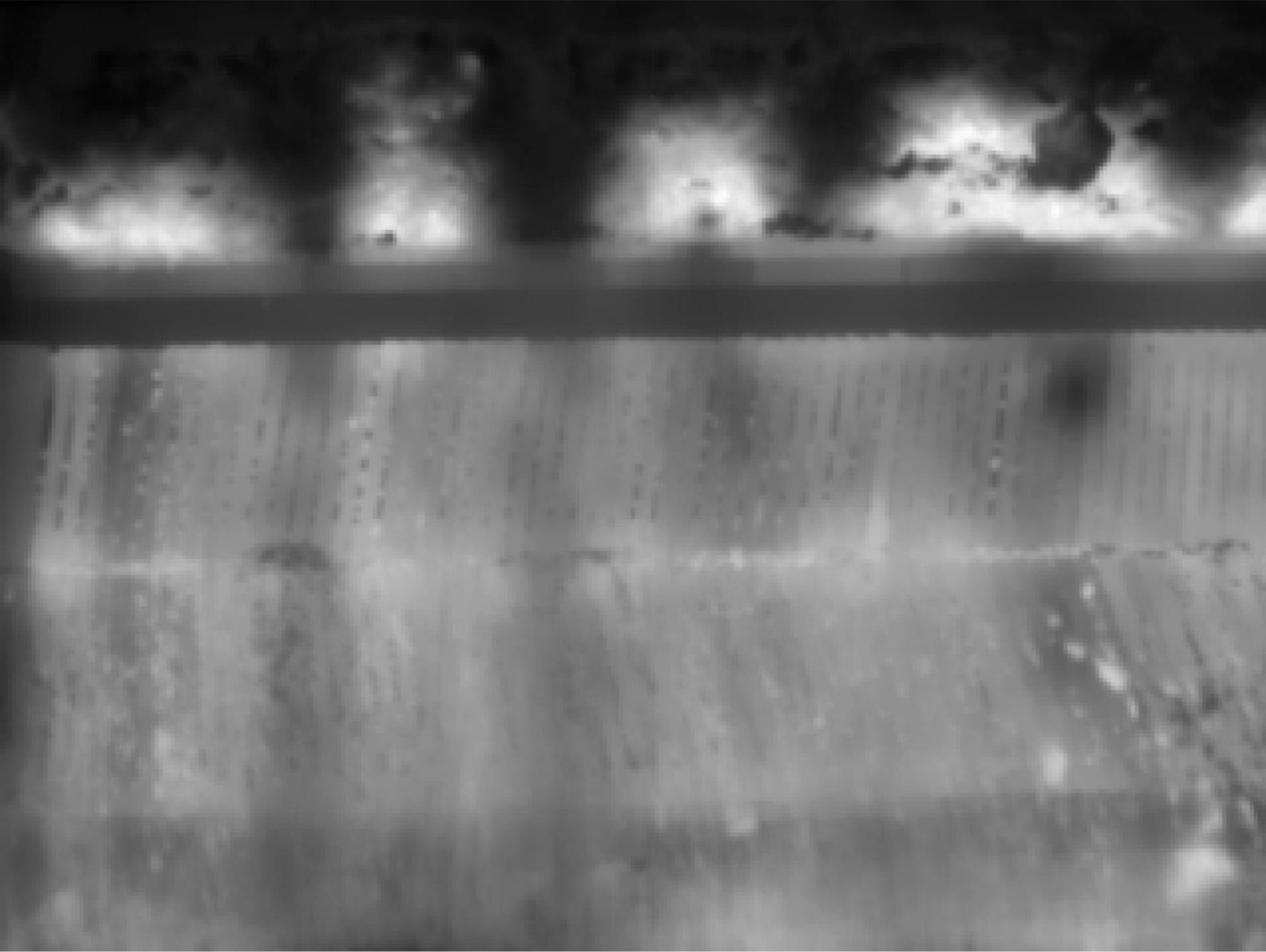
Gy

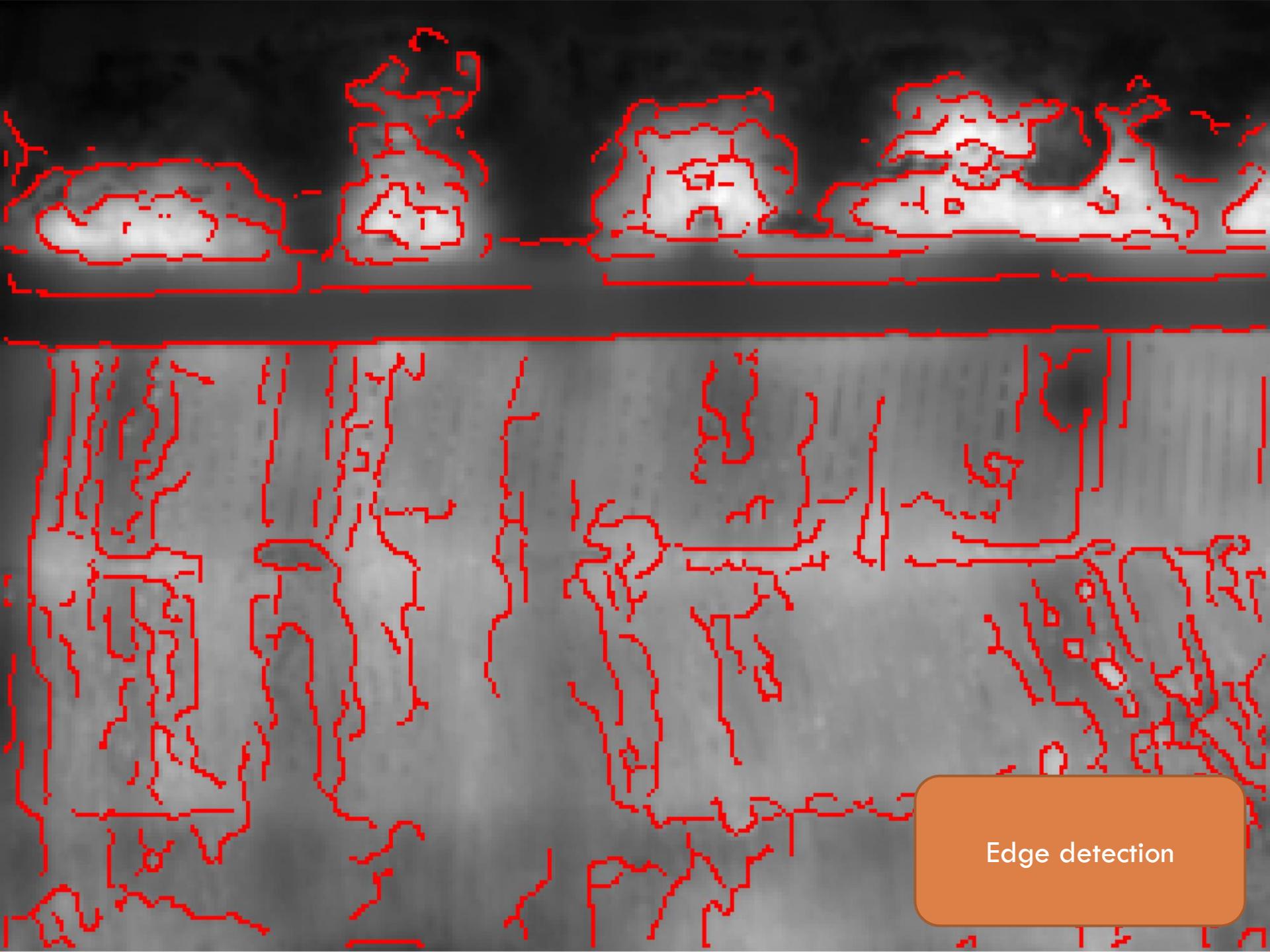


Source: Simpsons

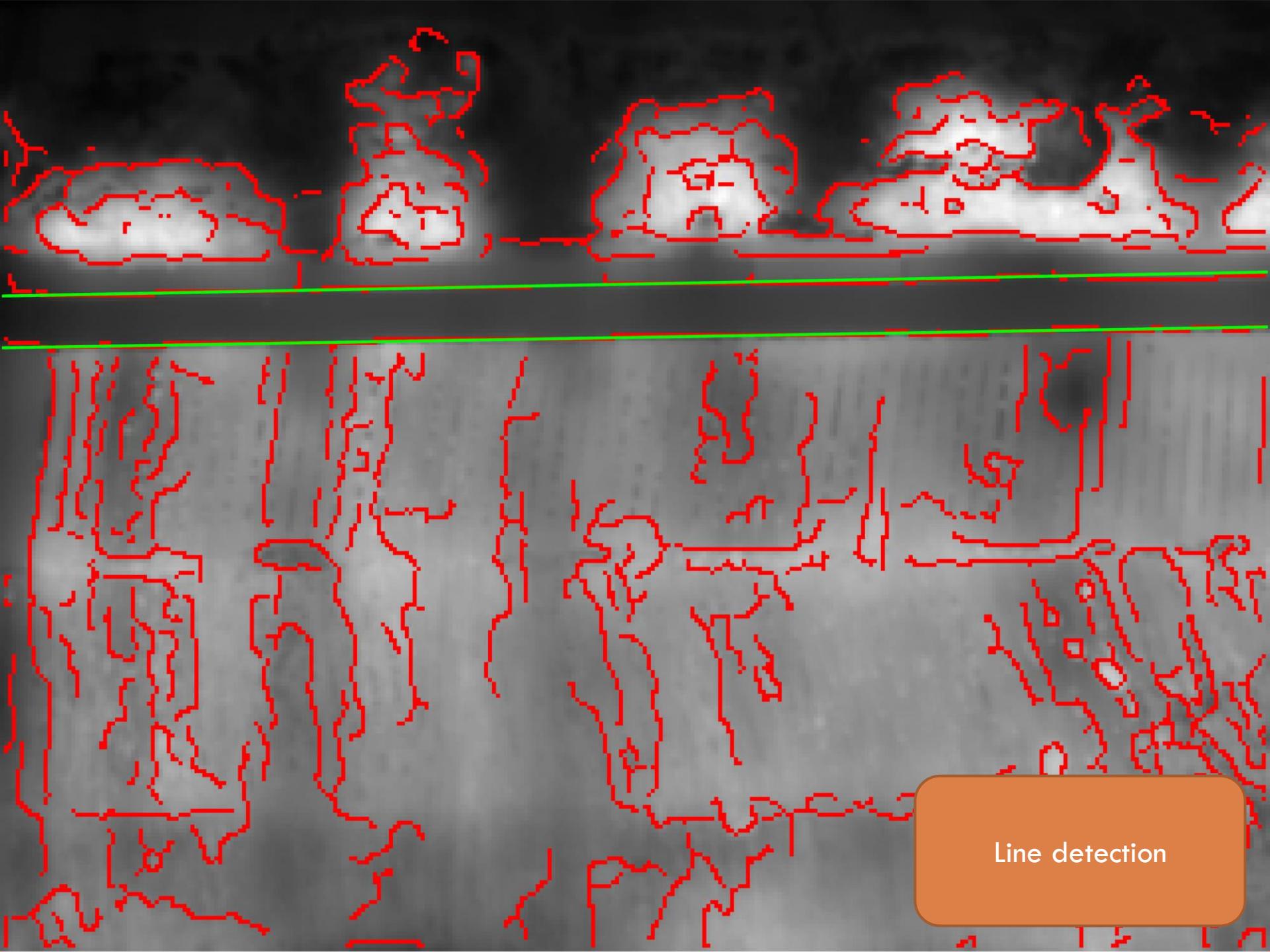
Traditional computer vision



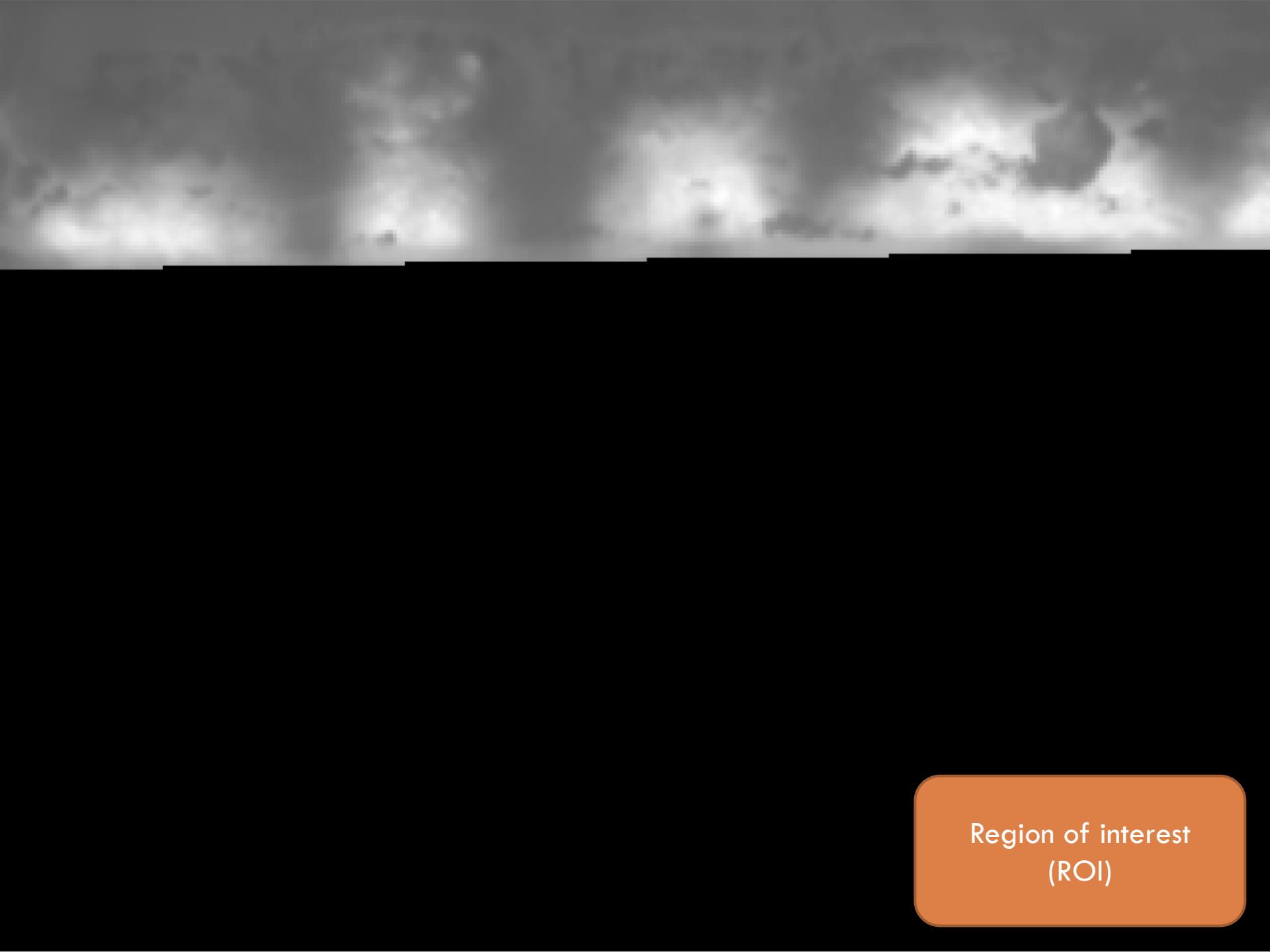




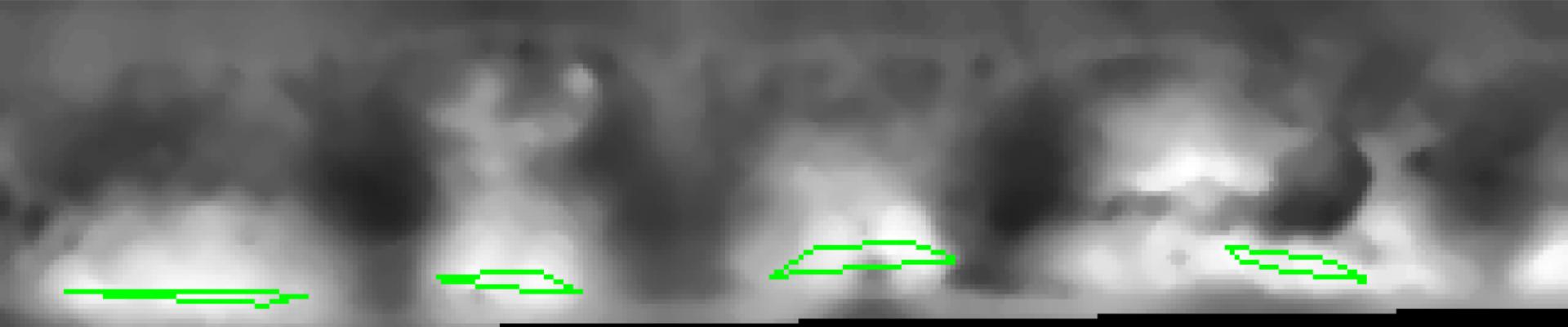
Edge detection



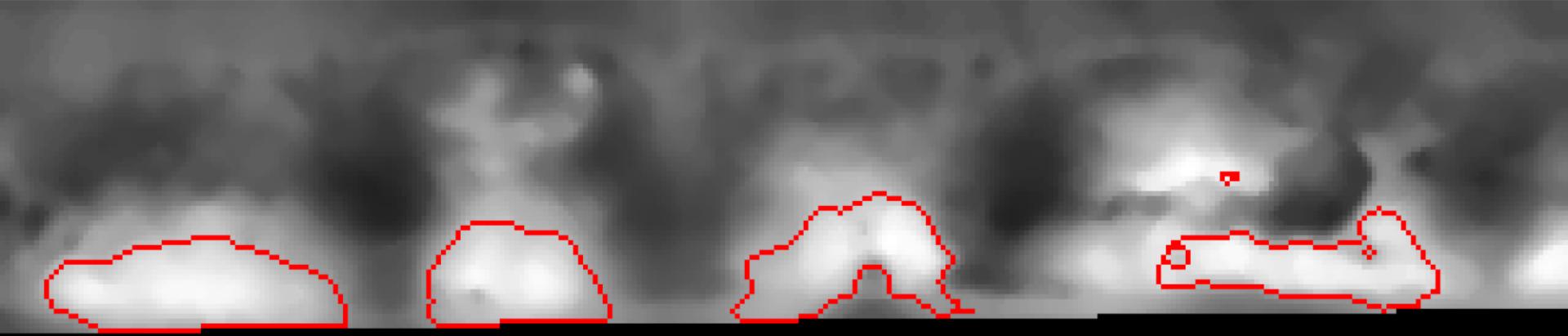
Line detection

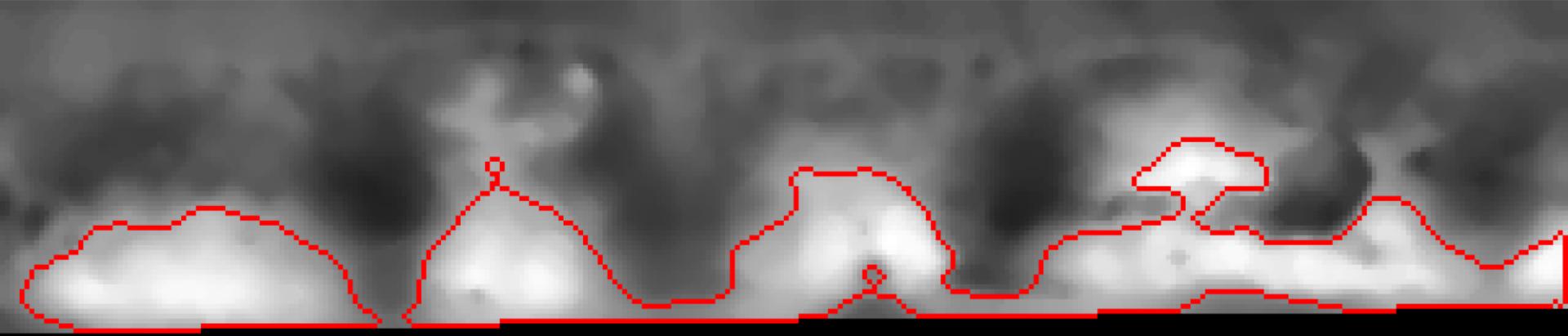


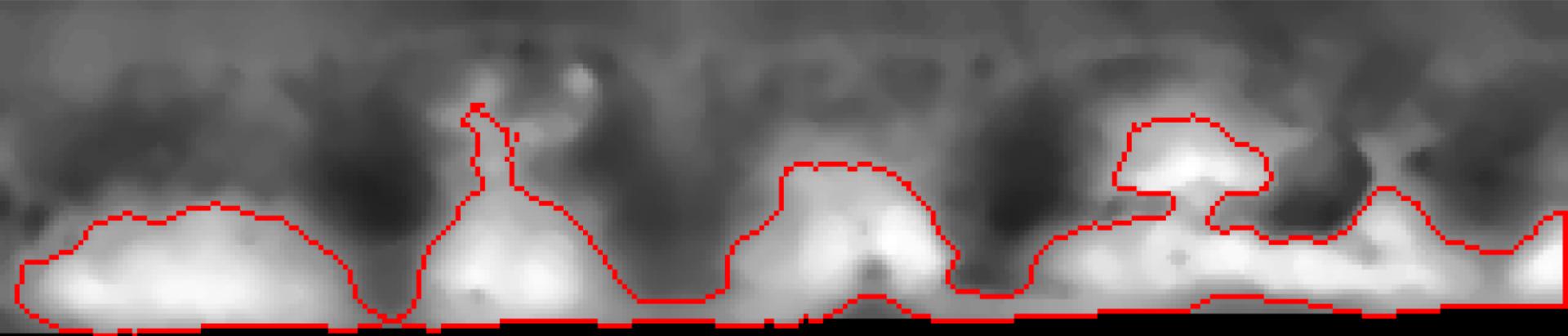
Region of interest
(ROI)

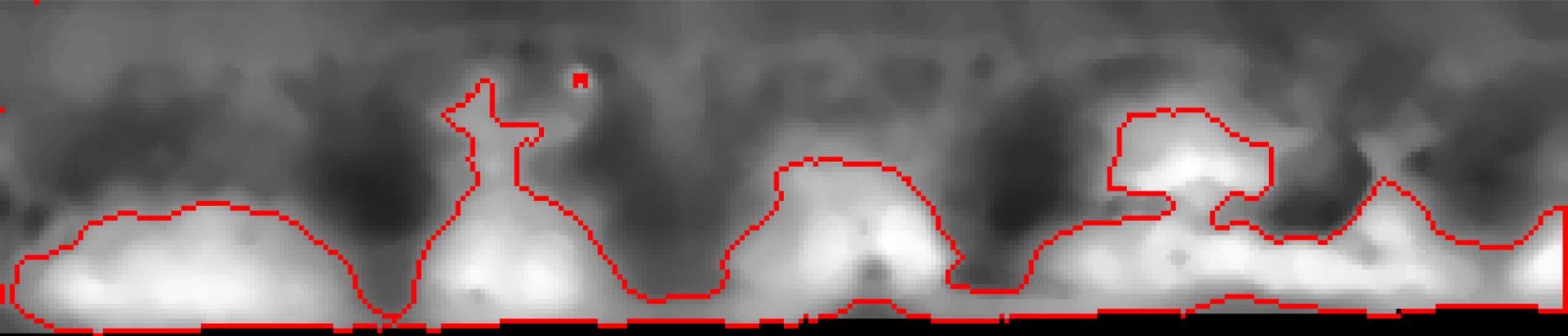


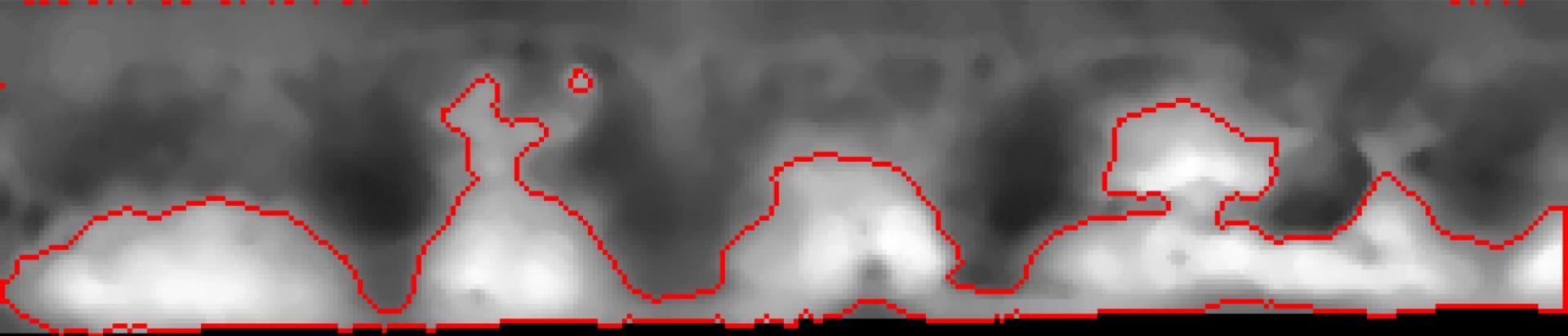
Segmentation using
active contours

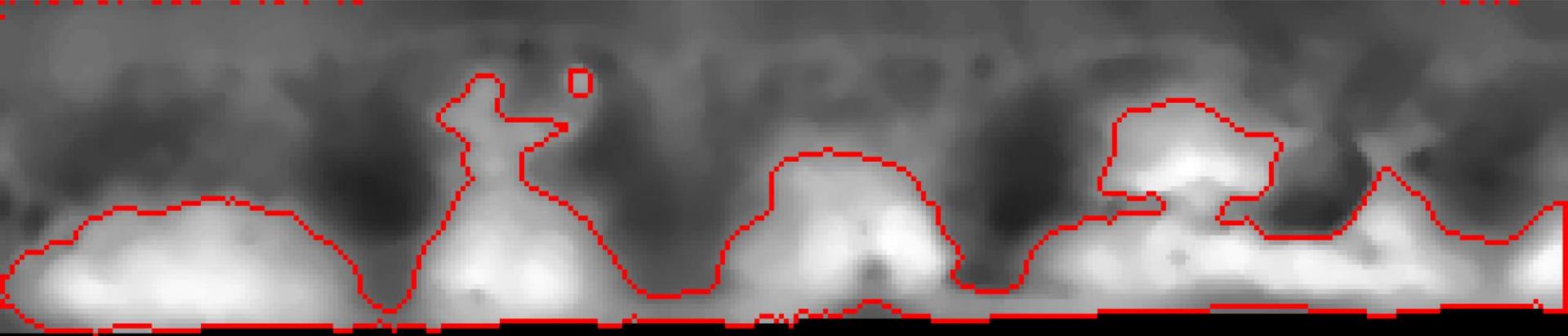




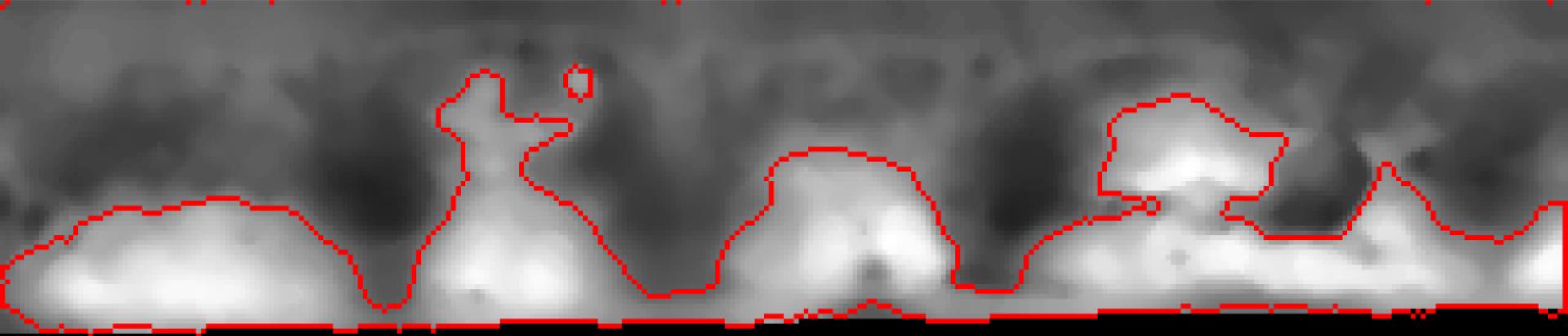


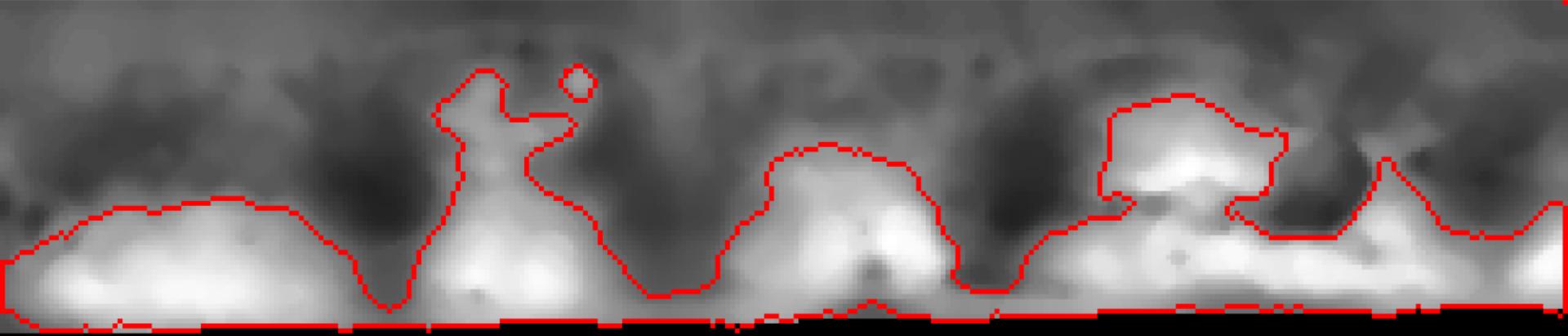


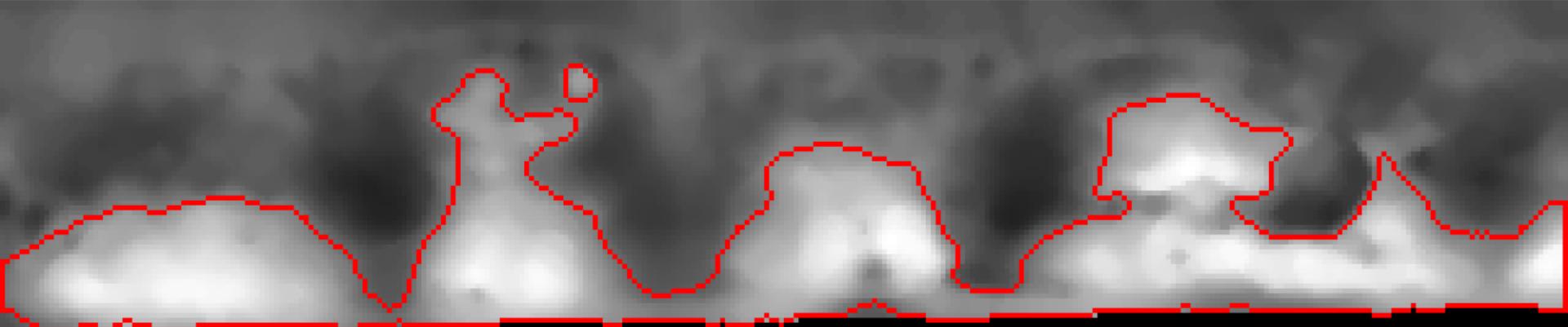


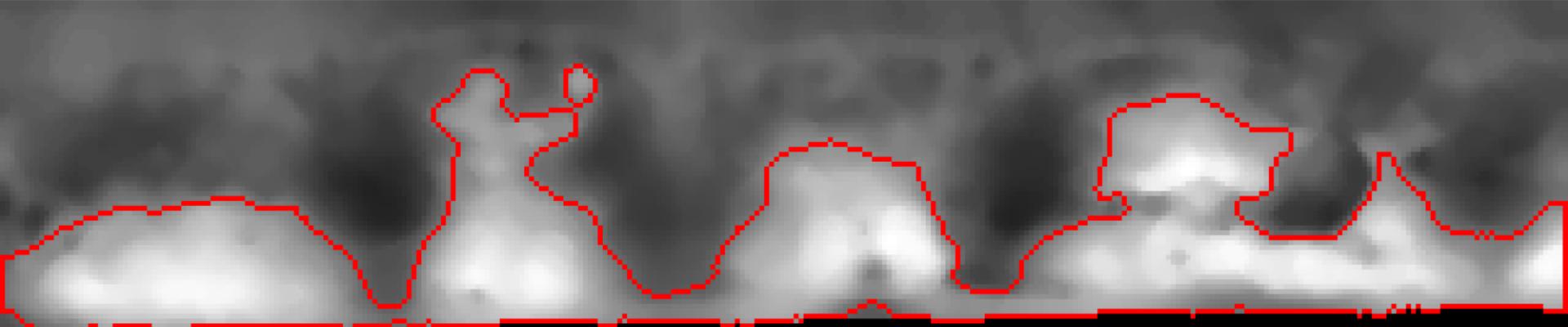


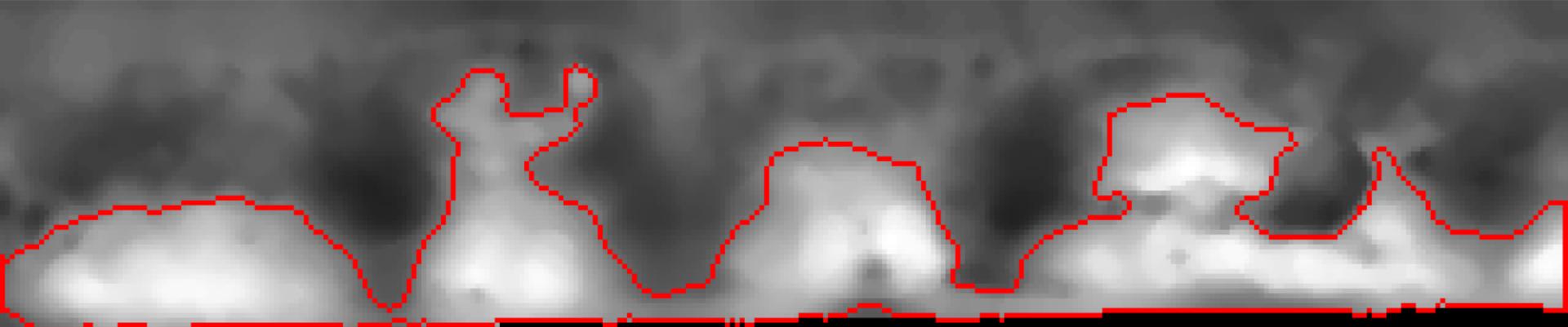
C

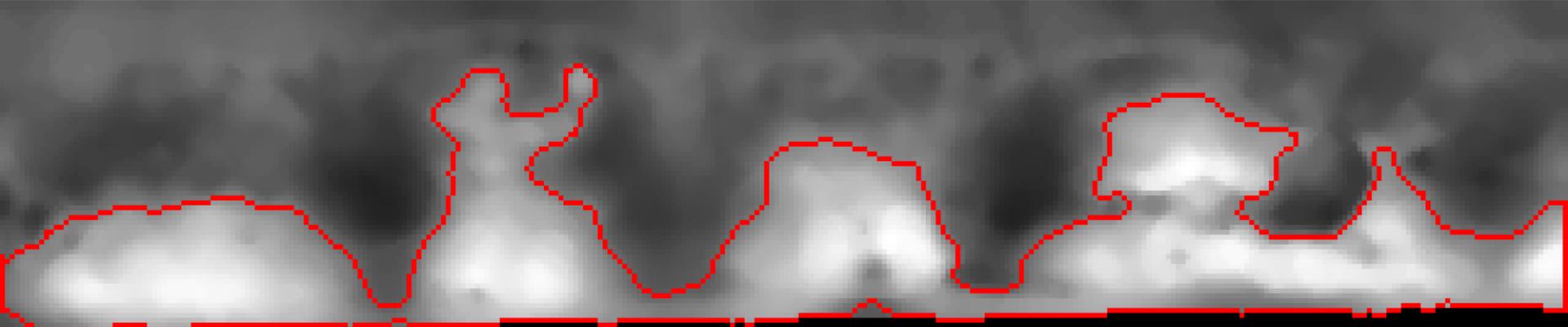


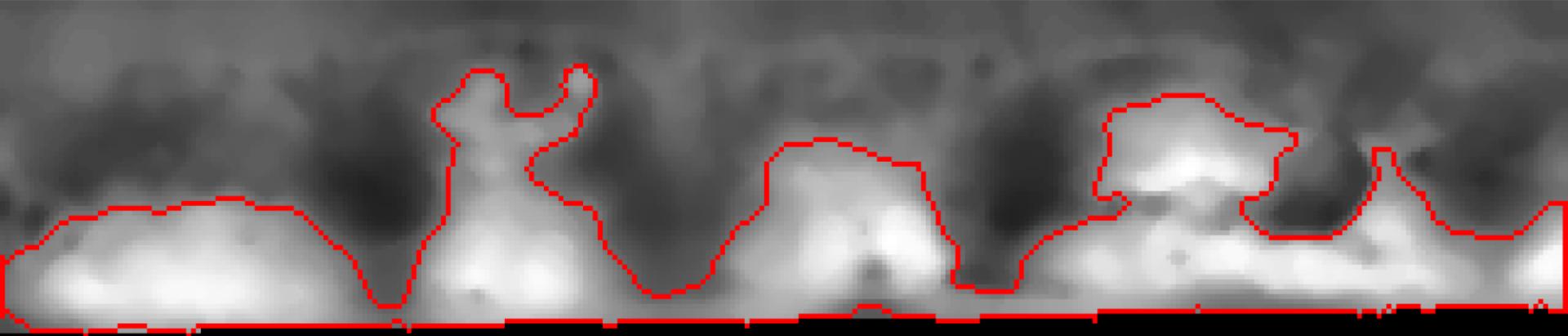


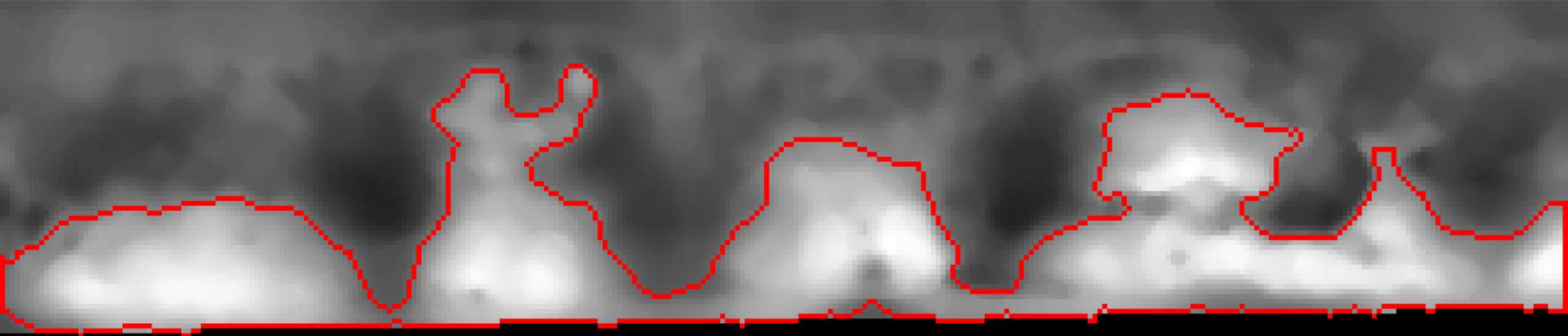


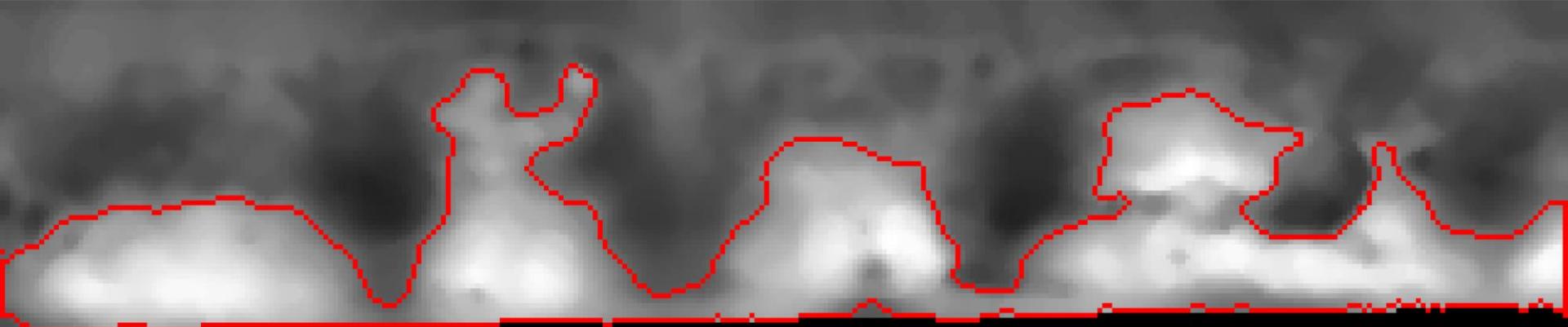




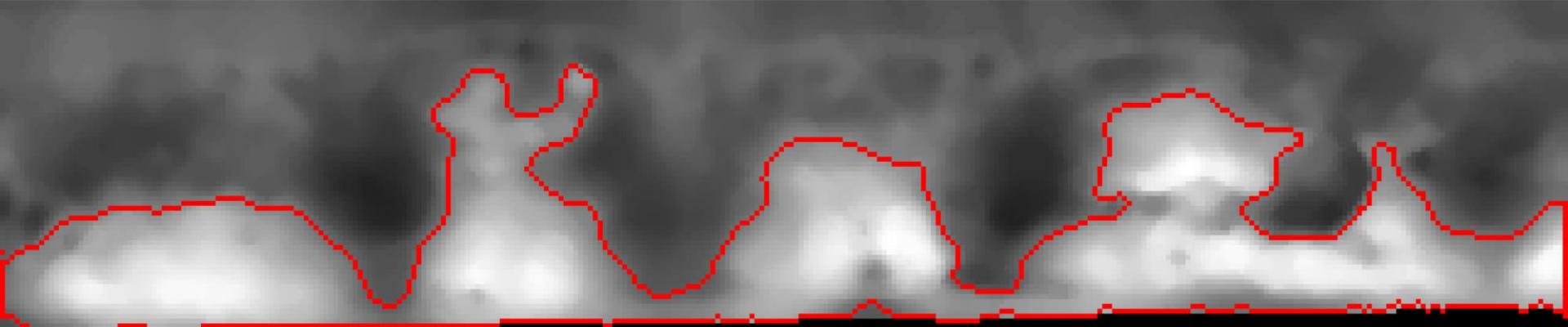








End of the
segmentation

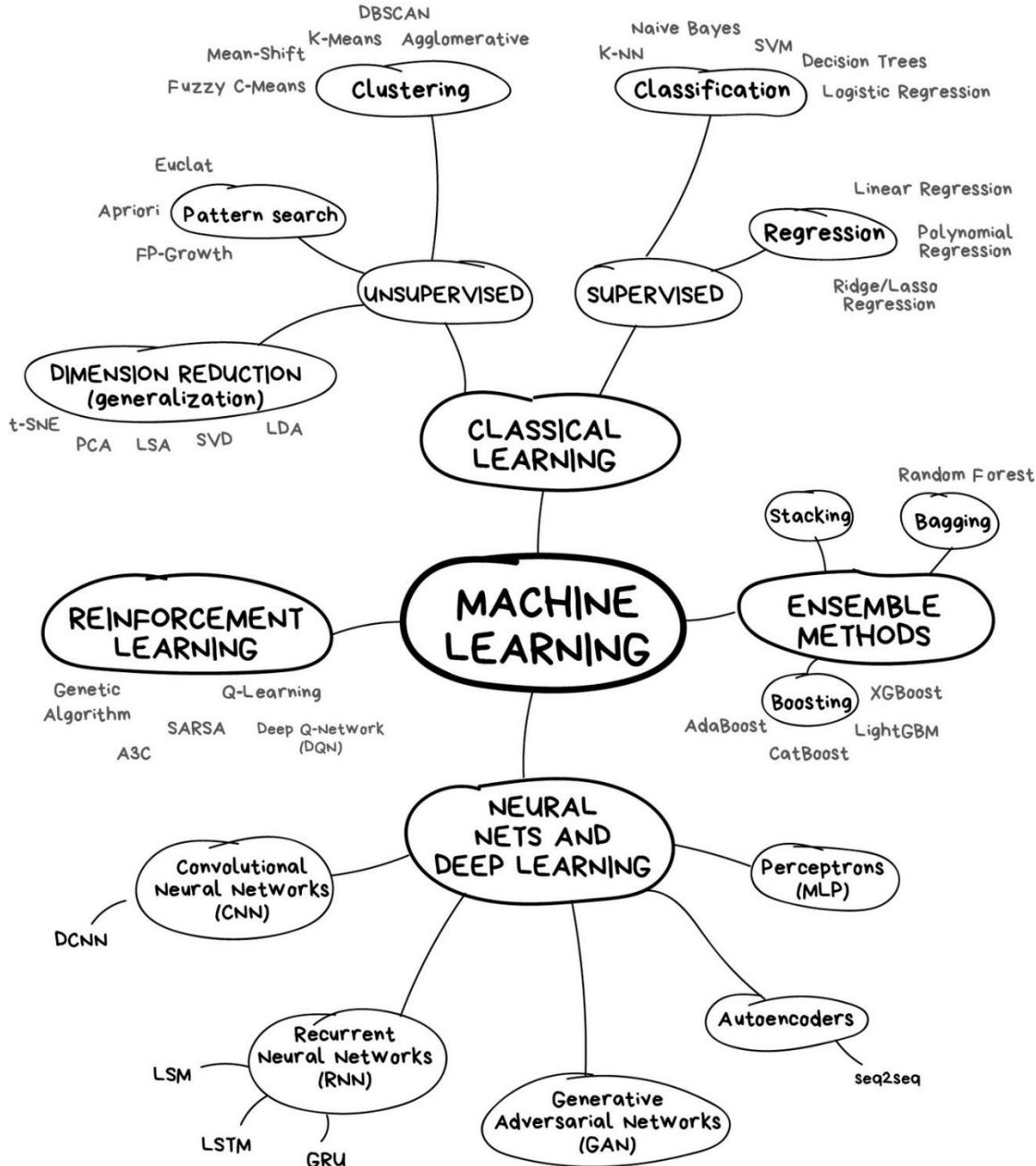


R. Usamentiaga, et al., "Temperature tracking system for sinter material in a rotatory cooler based on infrared thermography", In IEEE Transactions on Industry Applications, IEEE, vol. 50, no. 5, pp. 3095-3102, doi: 10.1109/TIA.2014.2306984, 2014.

Feature extraction and classification



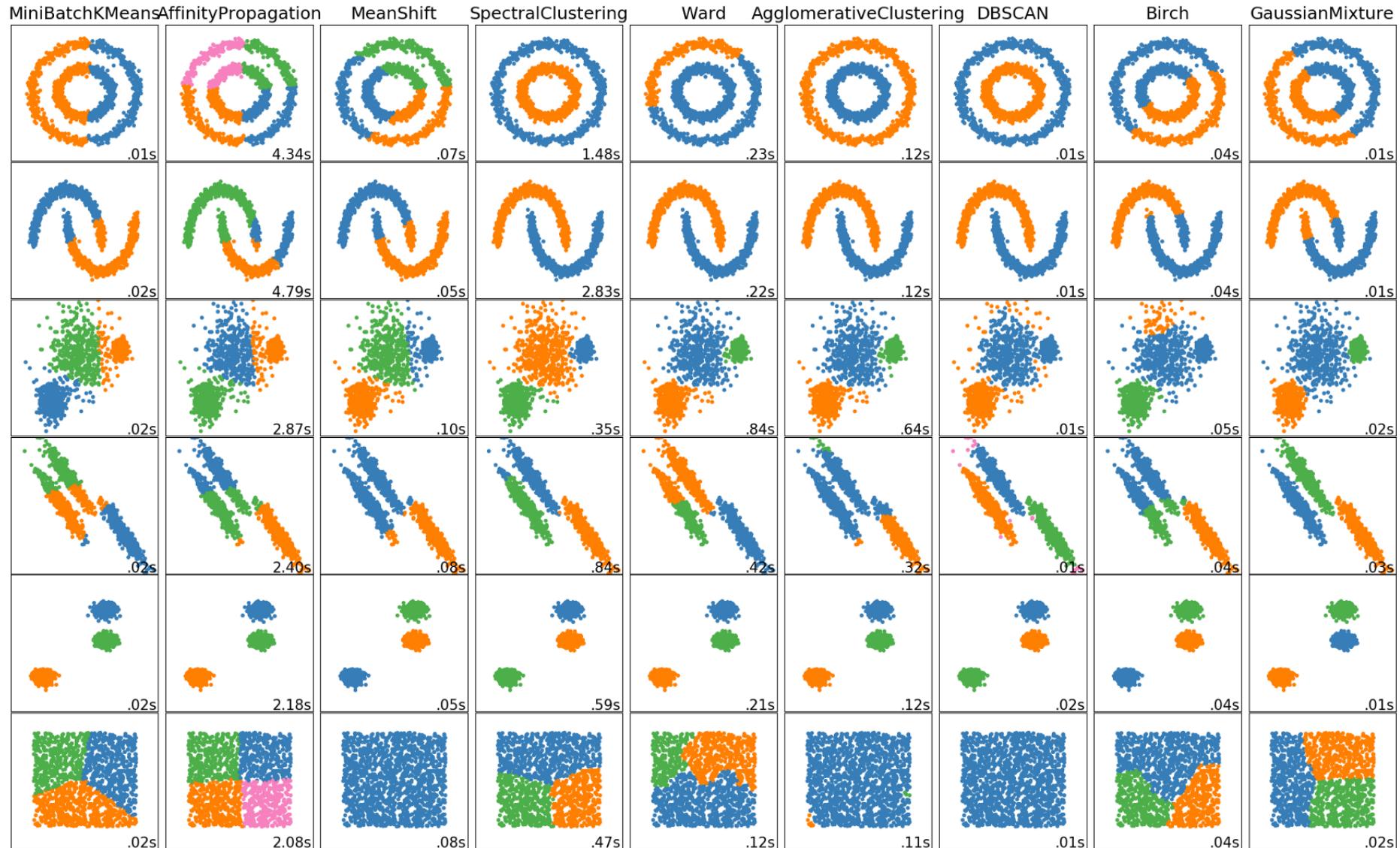
Machine learning



Machine learning

- Machine learning
 - Field of science that deals with the design of systems that learn
 - Machine learning in computer vision
 - Recognize or characterize regions or objects in images
 - Classify regions
 - Supervised
 - Find patterns and structures
 - Non-supervised

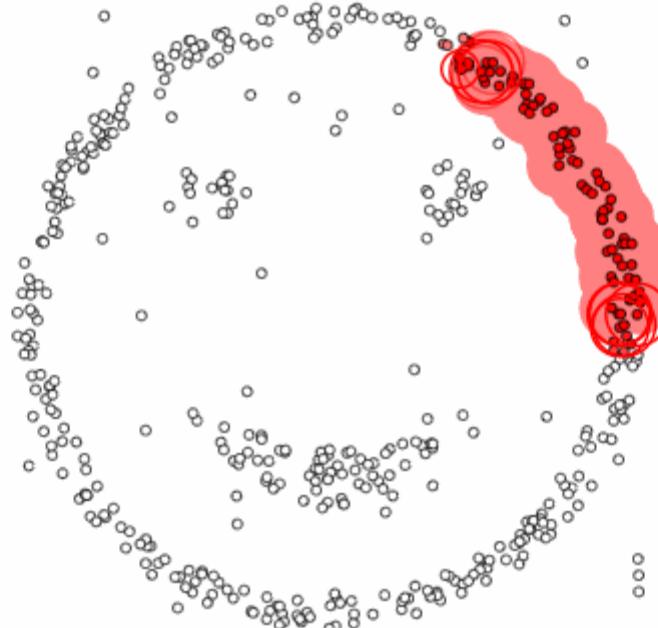
Clustering



Source: scikit-learn

DBSCAN

epsilon = 1.00
minPoints = 4



Restart



Pause

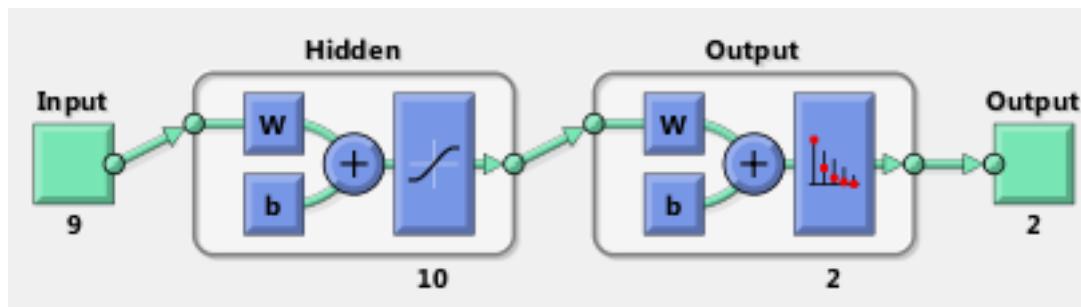
Machine learning

- Bayesian classifier
 - No previous information
 - Based on the probability of the objects for the considered classes and make the decision
 - With previous information
 - Determine the probability that a feature takes a certain value for a certain class: $P(x/w)$
 - Determine the probability that an object belong to that class: $P(w)$
 - Determine the probability that this feature takes that value: $P(x)$
 - $P(w/x) = P(x/w)P(w)/P(x)$

Machine learning

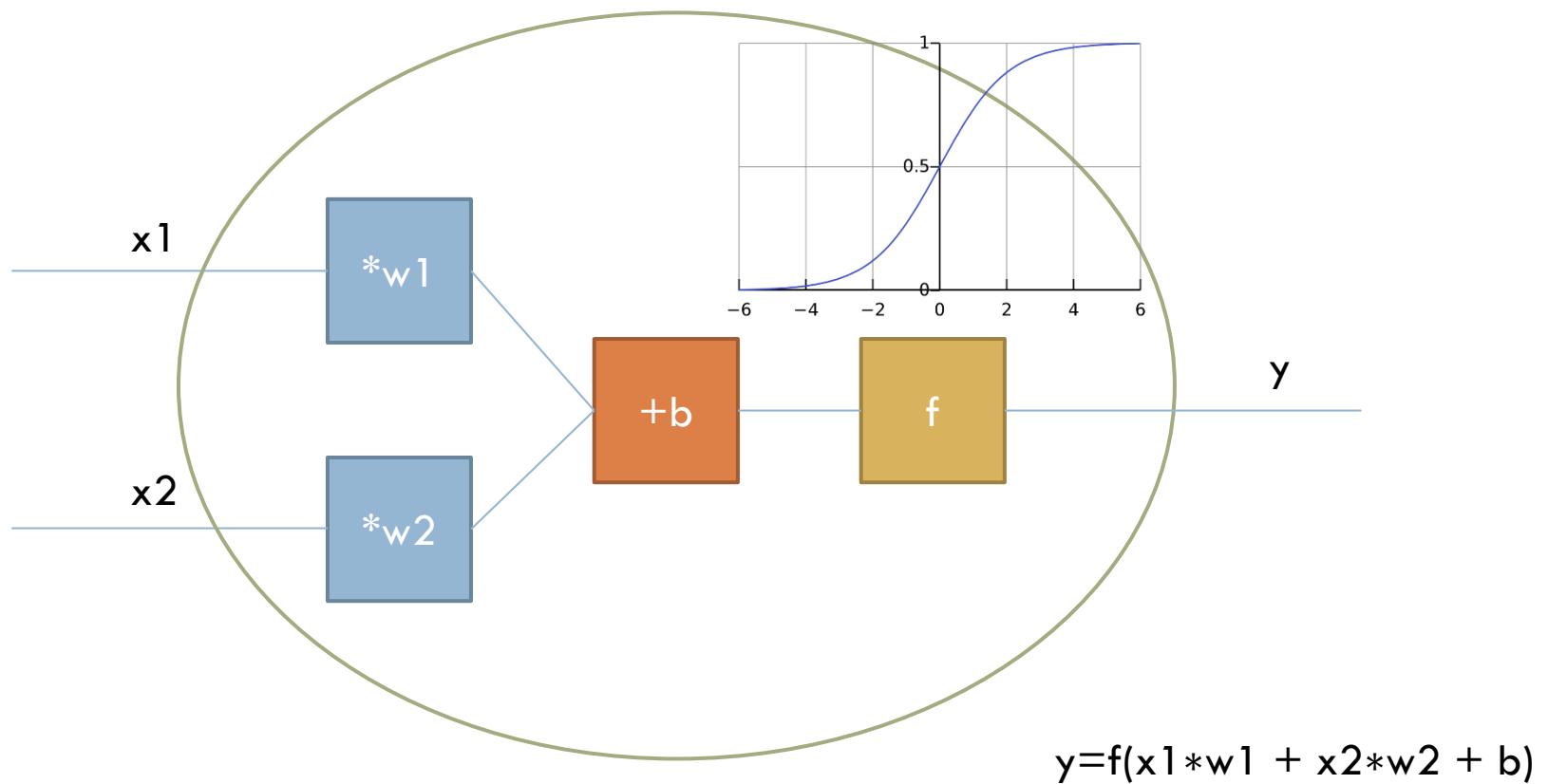
□ Neural networks

- Procedure used to estimate and fit models with a large number of parameters
- They can model complex relations between inputs and outputs
- They are based on artificial neurons connected in different layers (Ej: multilayer perceptron)



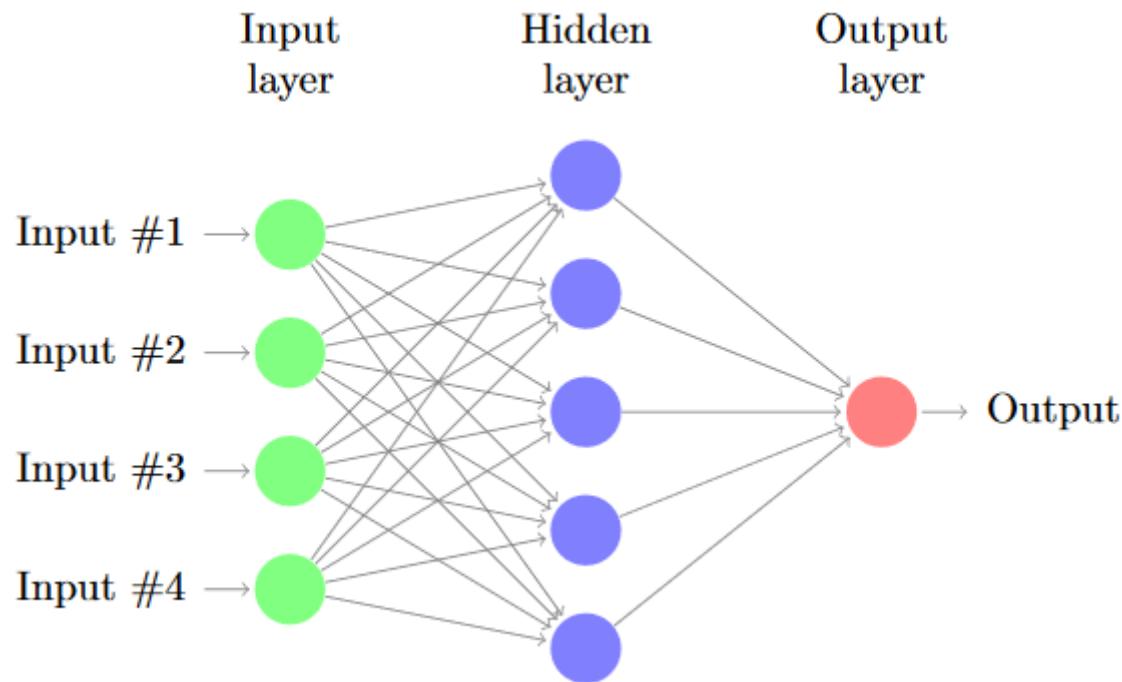
Machine learning

□ Neural networks



Machine learning

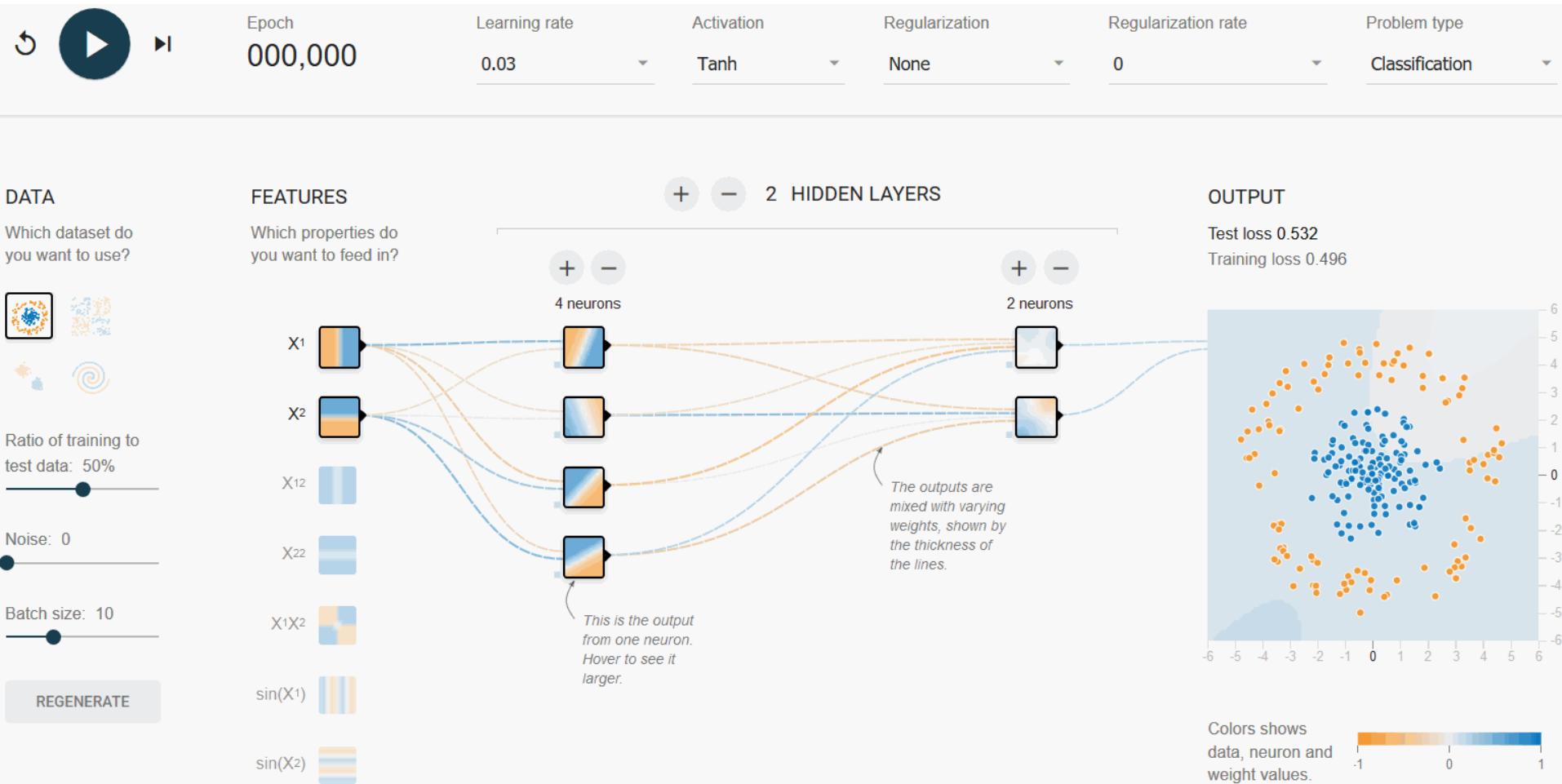
□ Neural networks



$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_{true} - y_{pred})^2$$

Machine learning

□ Neural networks



<https://playground.tensorflow.org>

Show test data Discretize output

Machine learning

□ Neural networks

Epoch 000,000 Learning rate 0.03 Activation ReLU Regularization L2 Regularization rate 0.001 Problem type Classification

DATA Which dataset do you want to use?

FEATURES Which properties do you want to feed in?

6 HIDDEN LAYERS

OUTPUT Test loss 0.506
Training loss 0.516

X₁

X₂

X₁₂

X₂₂

X_{1X2}

$\sin(X_1)$

$\sin(X_2)$

Ratio of training to test data: 50%

Noise: 0

Batch size: 10

REGENERATE

This is the output from one neuron

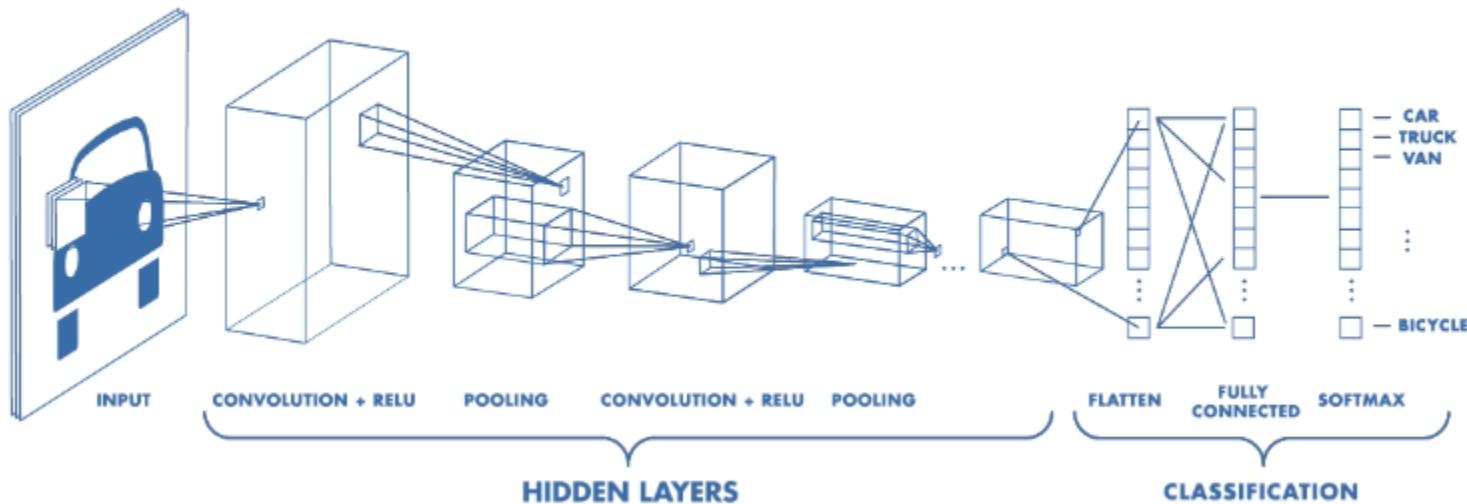
The outputs are mixed with varying

Colors shows data, neuron and weight values.

Show test data Discretize output

Machine learning

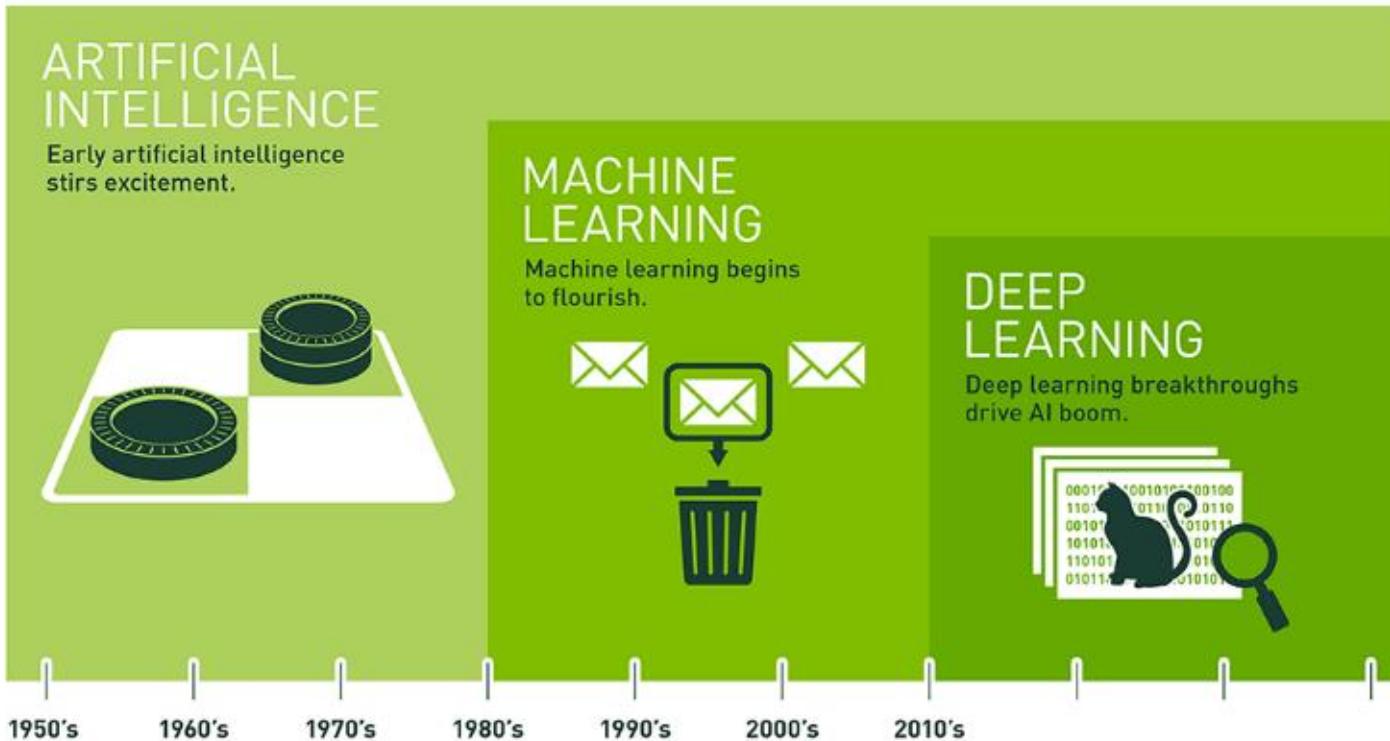
□ Neural networks



Source: matlab

Machine learning

□ Neural networks

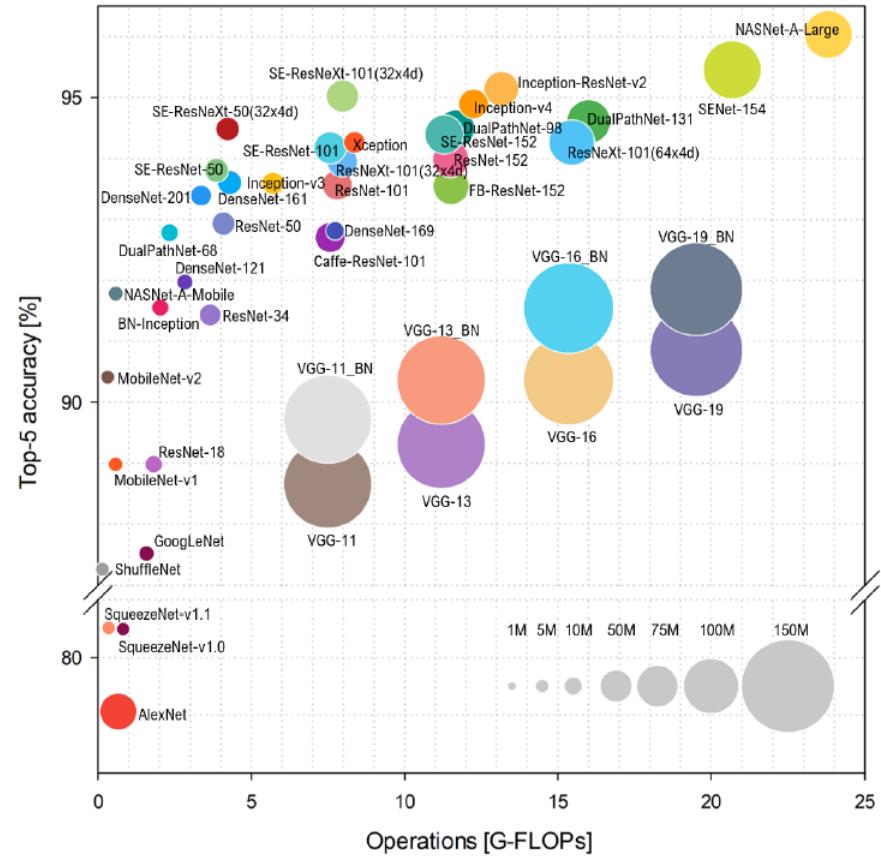
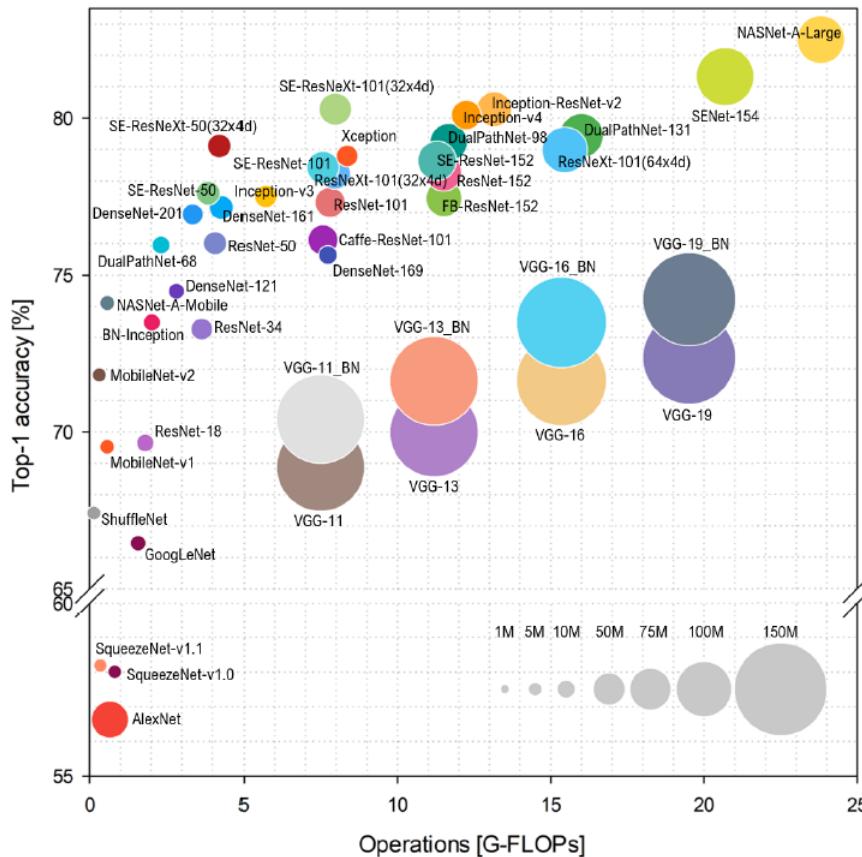


Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

Source: nvidia

Machine learning

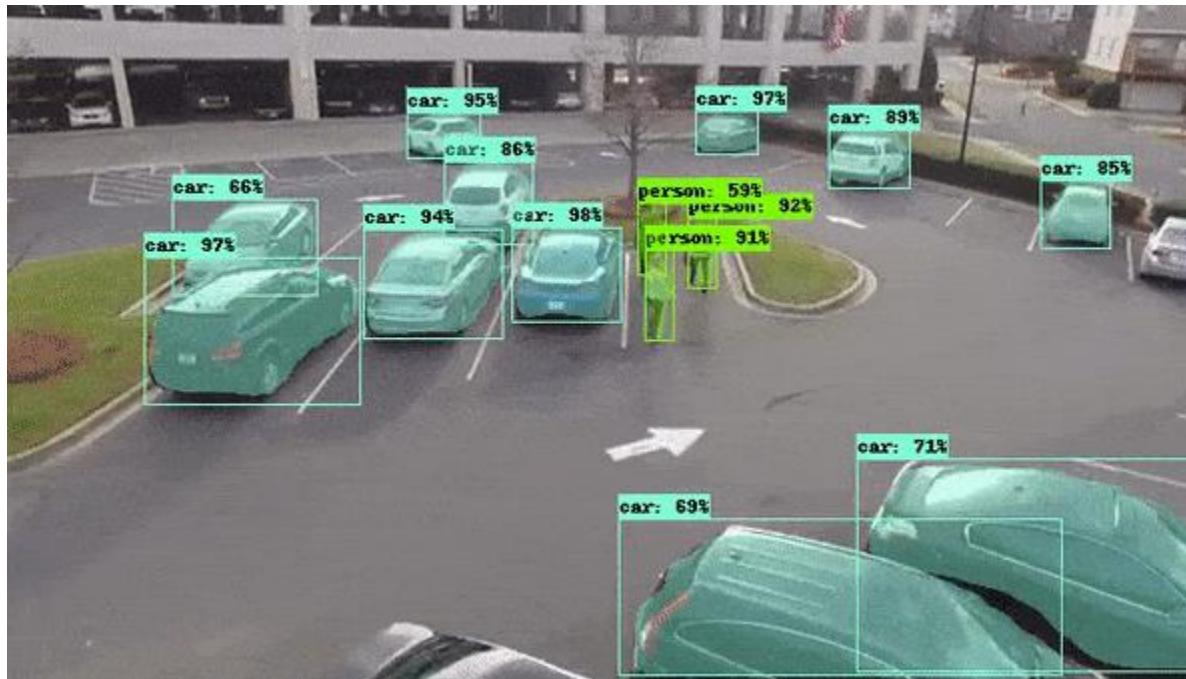
□ Neural networks



Source: Bianco, Simone, et al. "Benchmark Analysis of Representative Deep Neural Network Architectures." *IEEE Access* 6 (2018): 64270-64277.

Machine learning

□ Neural networks



Source: towardsdatascience