

Reconocimiento de patrones (aprendizaje automático) aplicado a imágenes y video

Agenda

- Generalidades de los sistemas de reconocimiento de patrones (RP)
 - Categorías
 - Medidas de desempeño
 - Aplicaciones a imágenes.
- Algunos modelos
 - k-means
 - Support Vector Machine (SVM)
 - Decision Trees y Random Forests
- Más aplicaciones.

Reconocimiento de patrones/Aprendizaje automático



IN CS, IT CAN BE HARD TO EXPLAIN THE DIFFERENCE BETWEEN THE EASY AND THE VIRTUALLY IMPOSSIBLE.

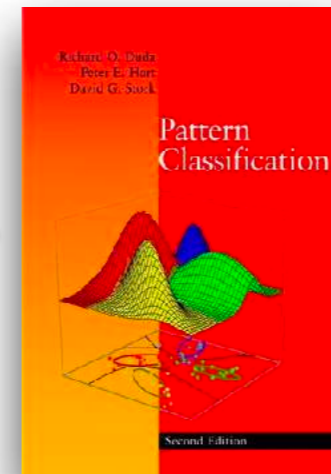


Reconocimiento de patrones/Aprendizaje automático

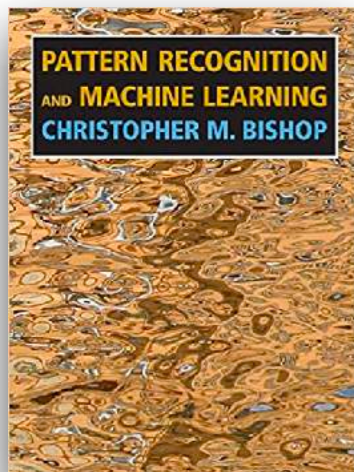
“El estudio de cómo las máquinas pueden, observando el ambiente, aprender a distinguir patrones de interés de un fondo y realizar decisiones razonables sobre las categorías de los mismos.”



“El acto de tomar datos crudos y hacer una acción basados en la categoría de los patrones.”



“Descubrir automáticamente regularidades en los datos usando algoritmos en computadora y a partir de esas regularidades tomar acciones.”

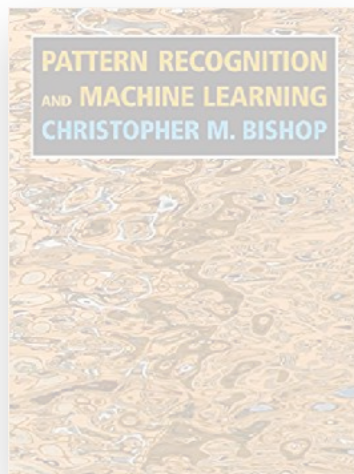


El Reconocimiento de Patrones es una disciplina que estudia métodos que permiten identificar objetos, tipos de objetos o relaciones a partir de señales o datos del mundo como usted hace en este momento con estas palabras.



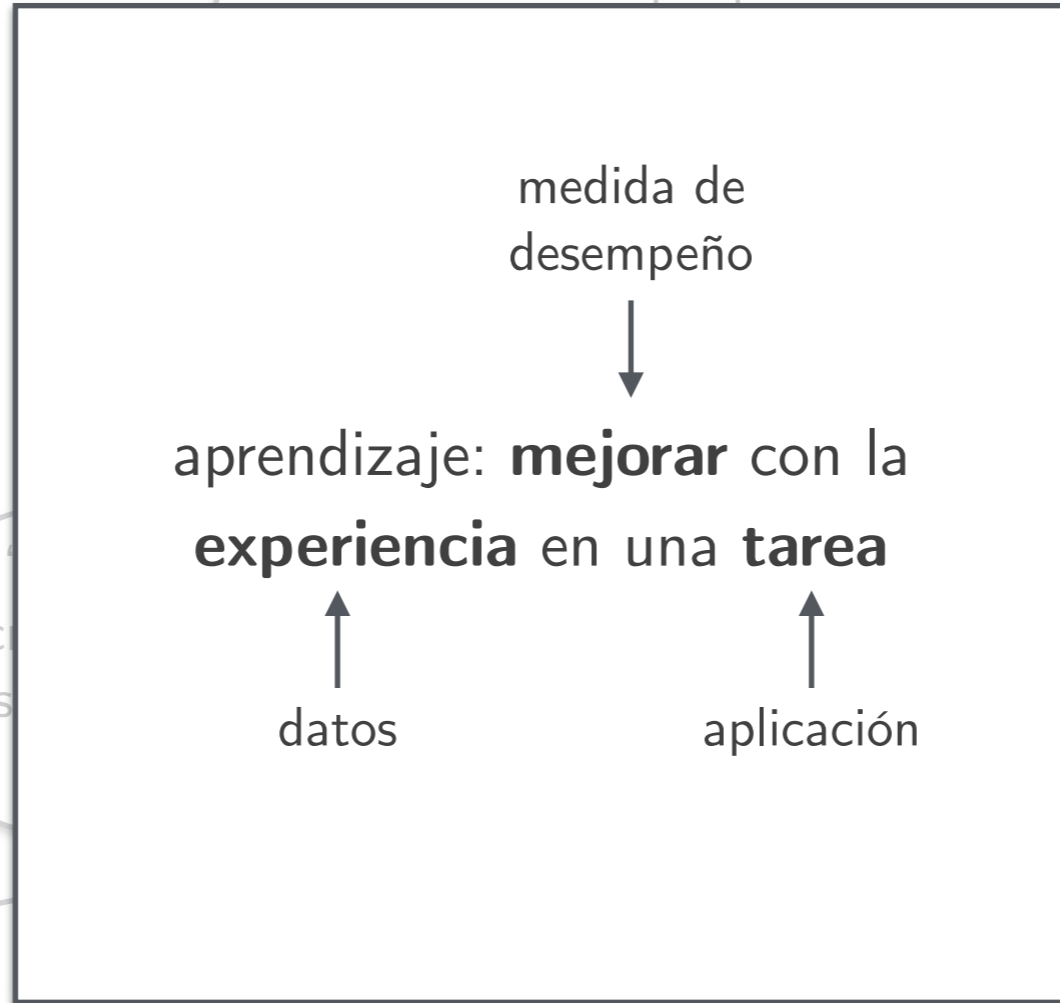
Reconocimiento de patrones/Aprendizaje automático

“El estudio de cómo las máquinas pueden, observando el ambiente, aprender a distinguir patrones de interés de un fondo y realizar decisiones razonables sobre las categorías de los mismos.”



“Descubrir automáticamente regularidades en los datos usando algoritmos en computadora y a partir de esas regularidades tomar acciones.”

El Reconocimiento de Patrones es una disciplina que estudia métodos que permiten identificar

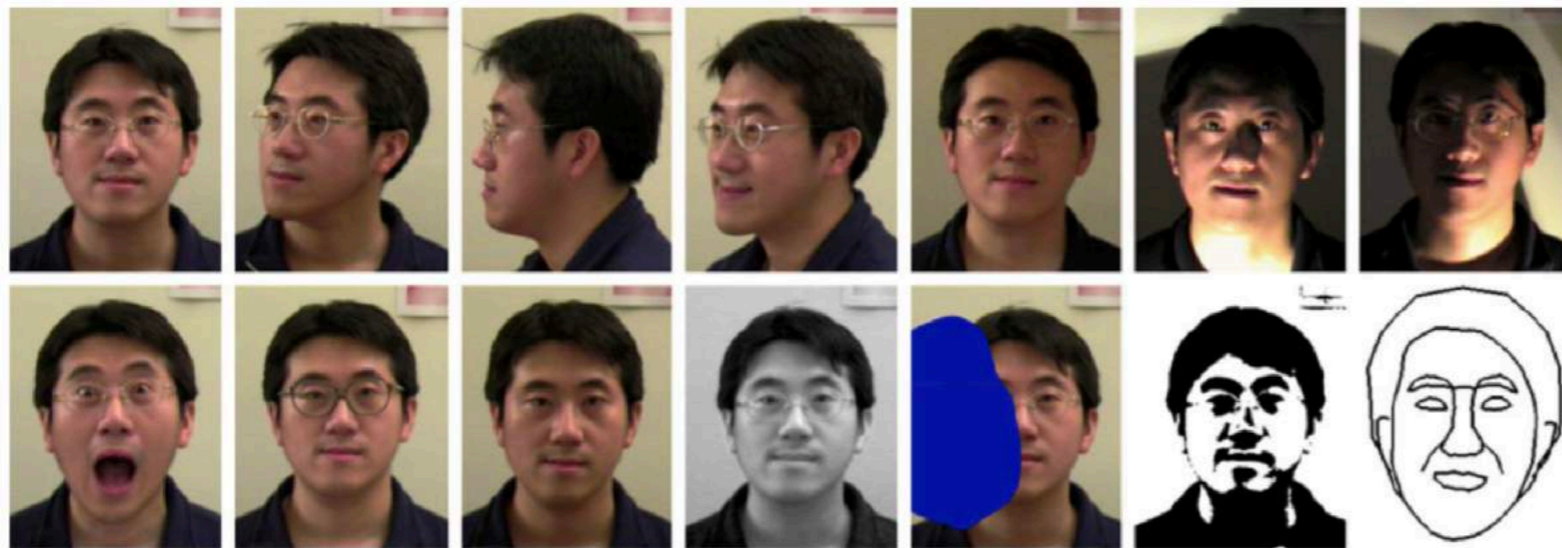


Detección y reconocimiento de caras



Detección y reconocimiento de caras

- Es muy simple para los humanos.
- Décadas de desarrollo en vision artificial.
- Desafíos: pose, iluminación, oclusiones, gestos, accesorios, envejecimiento, condiciones de captura, ...



Detección y reconocimiento de caras

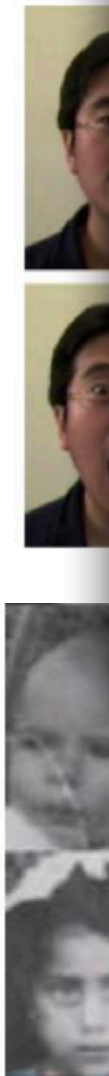
- Es muy simple para los humanos.
- Décadas de desarrollo en vision artificial.
- Desafíos: pose, iluminación, oclusiones, gestos, accesorios, envejecimiento, condiciones de captura, ...

ACCEPTED CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION 2001

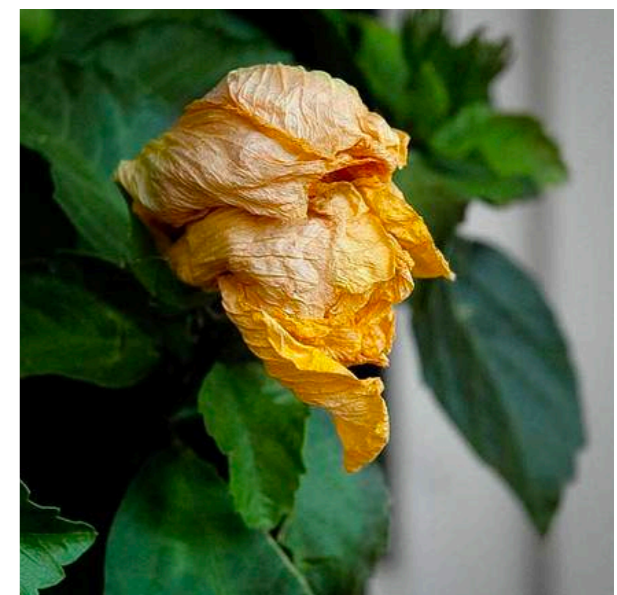
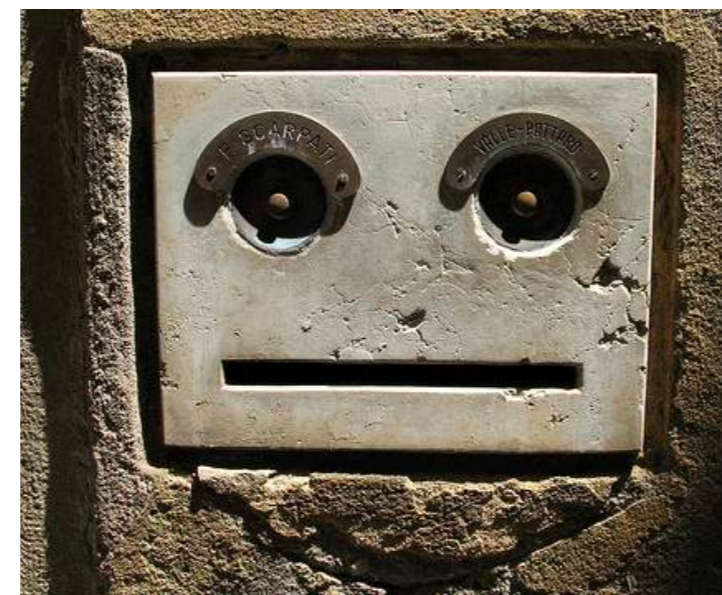
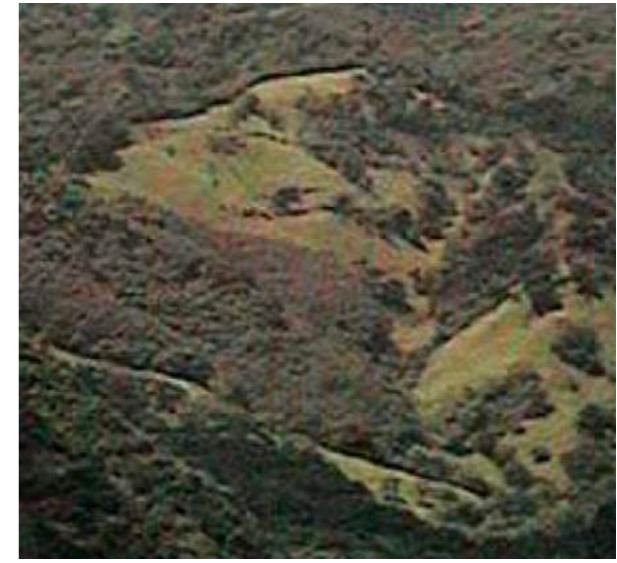
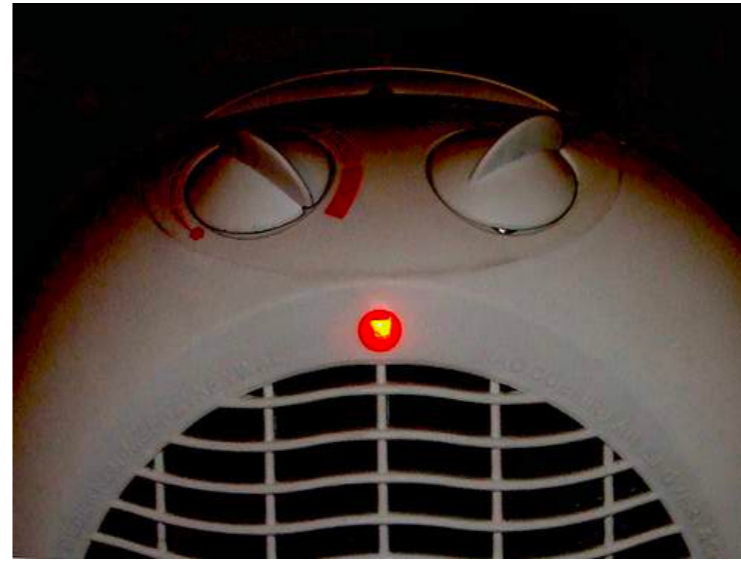
Rapid Object Detection using a Boosted Cascade of Simple Features

Paul Viola
viola@merl.com
Mitsubishi Electric Research Labs
201 Broadway, 8th FL
Cambridge, MA 02139

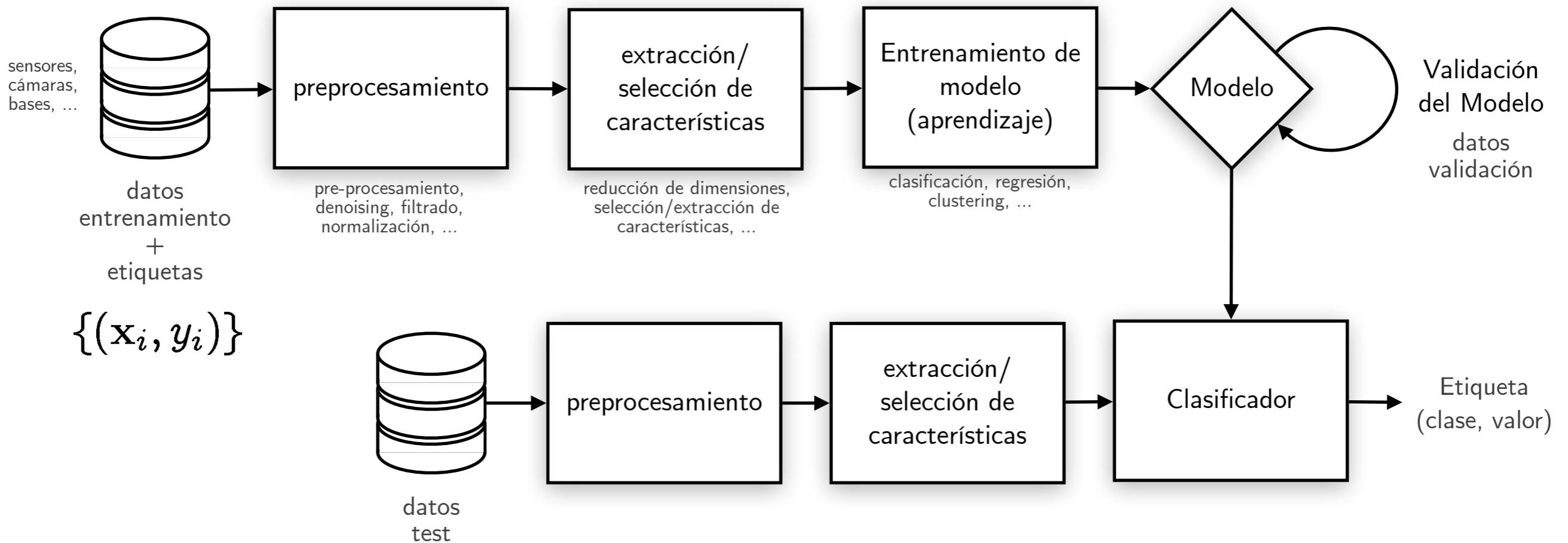
Michael Jones
mjones@crl.dec.com
Compaq CRL
One Cambridge Center
Cambridge, MA 02142



Detección y reconocimiento de caras



Sistema de reconocimiento de patrones



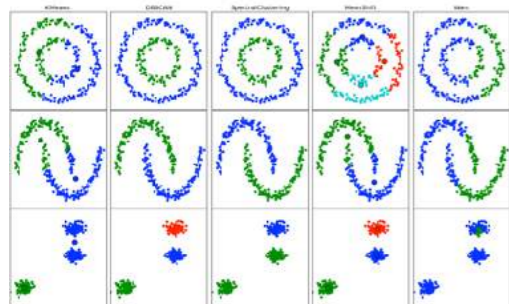
Categorías de los sistemas de RP

no supervisado

no se conocen las etiquetas (clases) de los datos, se organizan a partir de las características.

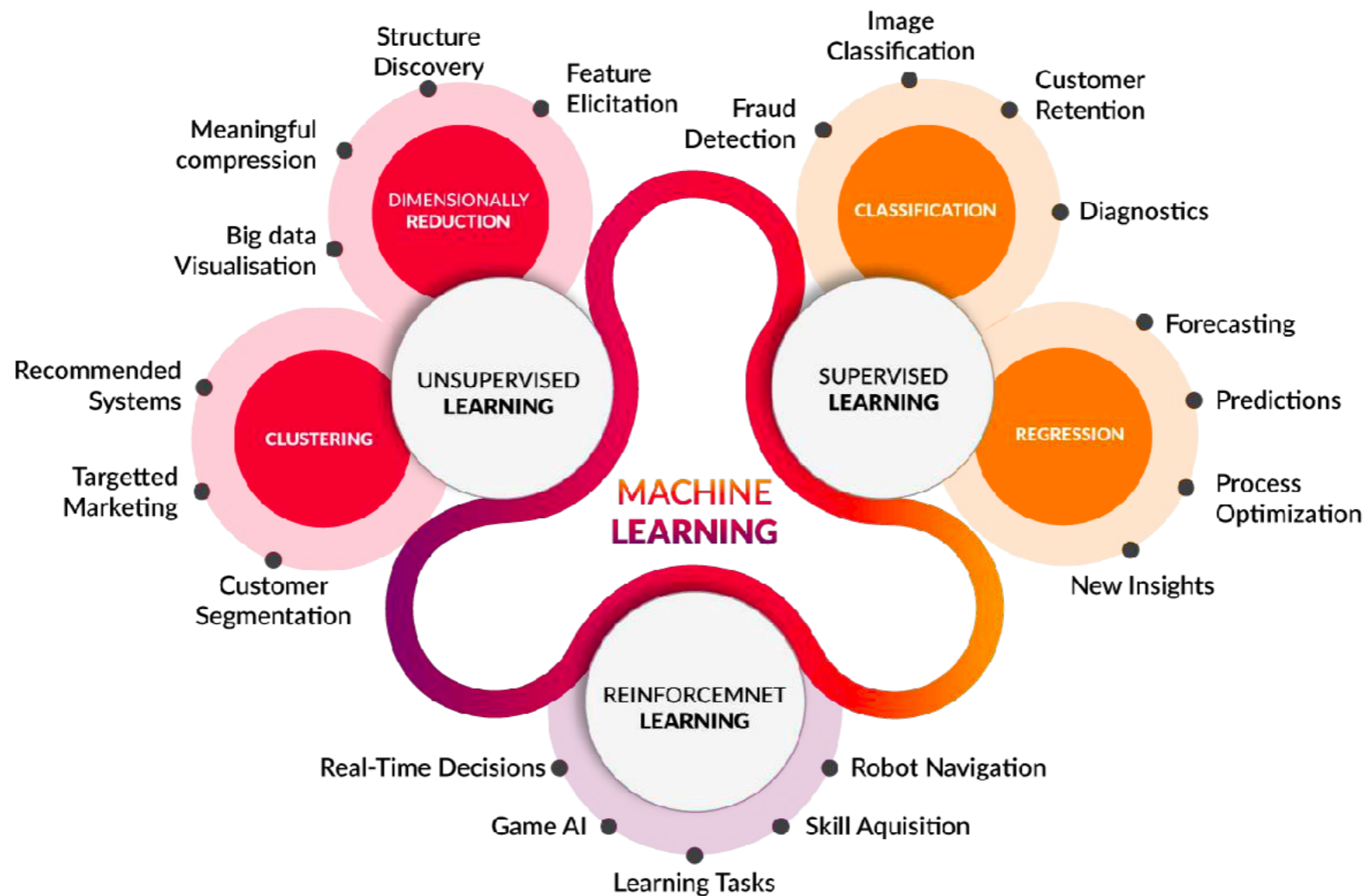
- identificación de patrones/ estructura
- evaluación indirecta o cualitativa
- organiza/agrupa clustering
- paramétrico o no
- # clases conocida o no

k-means, Fuzzy C-means, Hierarchical Clustering, Spectral Clustering, Gaussian Mixtures, Hidden Markov Model, Neural Networks, (Generalized/Robust) PCA, Isomap, MDS, Diffusion Maps, ...



semi-supervisado

se conocen las etiquetas (clases) de algunos datos



supervisado

datos de entrenamiento con etiquetas (clases) o valores de salida correctos para predecir datos nuevos

- clasificación y regresión
- aprendizaje explícito
- evaluación directa
- predicción clase/valor
- paramétrico o no

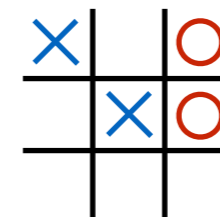
Nearest Neighbor, Support Vector Machines (SVM), Decision Trees, Random Forest, Discriminant Analysis, Naive Bayes, Neural Networks, Linear Regression, SVR, ...



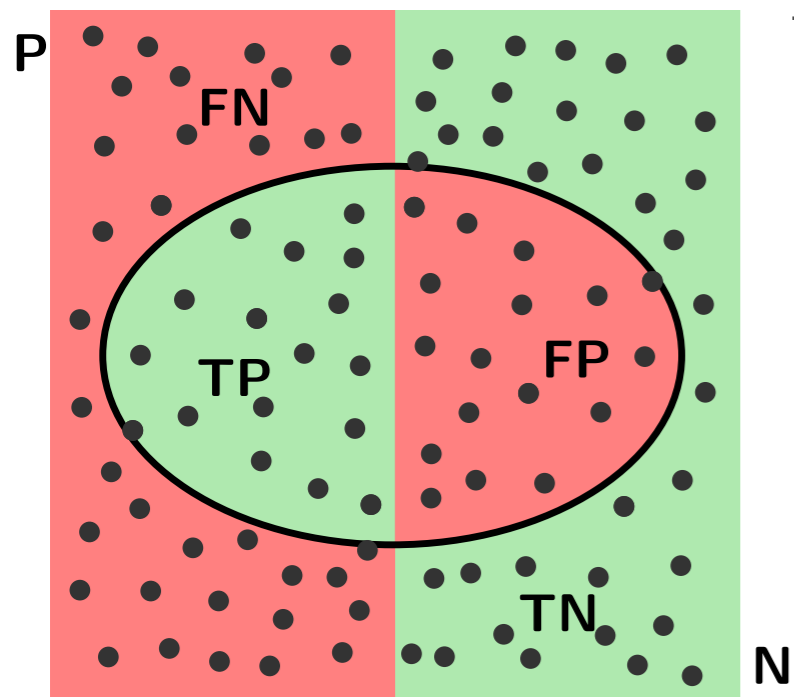
reinforcement learning

realimentación del resultado de la tarea: recompensa o penalización

- aproximación a IA
- definir estrategias ante eventos
- maximizar recompensa



Medidas de desempeño



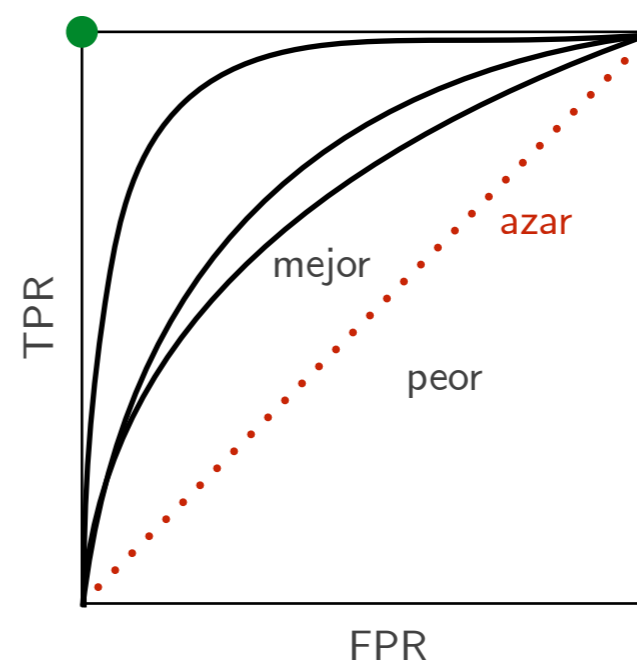
acertado rechazo correcto falsa alarma (error I) desacierto (error II)
 TP: true positives, TN: true negatives, FP: false positives, FN: false negatives.

- Accuracy (exactitud): ¿cuánto se acerca a los valores reales?
- Sensitivity/Recall (sensibilidad): ¿cuántos enfermos son correctamente detectados?
- Specificity (especificidad): ¿cuántos sanos no son seleccionados?
- Precision (presición): ¿cuántos son los enfermos de los seleccionados?
- False Positive Rate: ¿cuántos sanos son seleccionados?
- F1 score: Media armónica entre TPR y PPV.

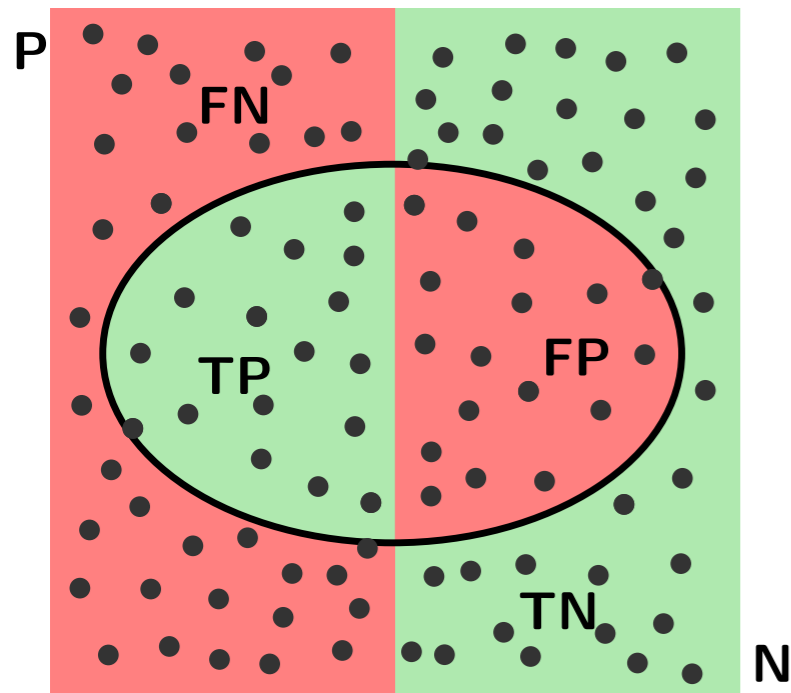
Matriz de confusión

		Valores reales	
		enfermos	sanos
Valores predichos	enfermos	TP	FP
	sanos	FN	TN

Receiver Operating Curve (ROC)



Medidas de desempeño



acertado rechazo correcto falsa alarma (error I) desacierto (error II)
 TP: true positives, TN: true negatives, FP: false positives, FN: false negatives.

login enfermedad

Accuracy (exactitud): $ACC = (TP + TN)/(T + N)$

Sensitivity/Recall (sensibilidad): $TPR = TP/P = TP/(TP + FN)$

Specificity (especificidad): $TNR = TN/N = TN/(TN + FP)$

Precision (presición): $PPV = TP/(TP + FP)$

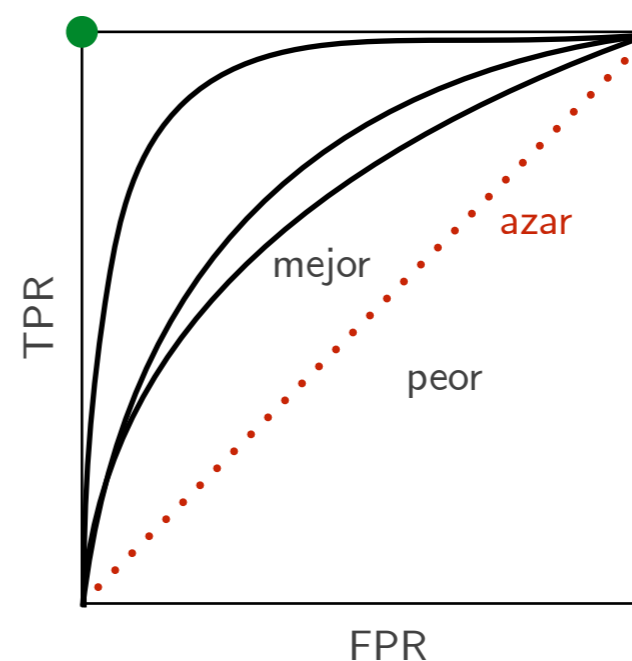
False Positive Rate: $FPR = FP/N$

F1 score: $F_1 = TPR \times PPV/(TPR + PPV)$

Matriz de confusión

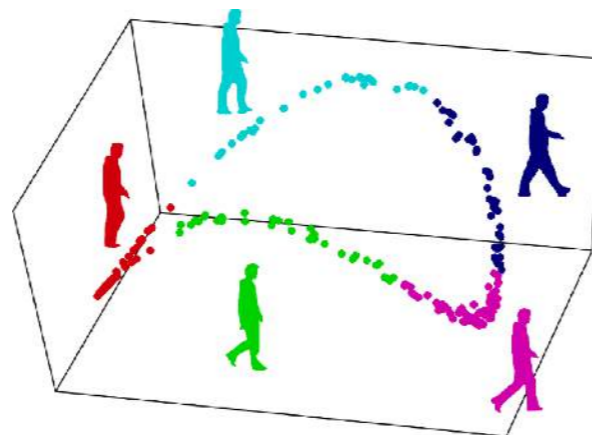
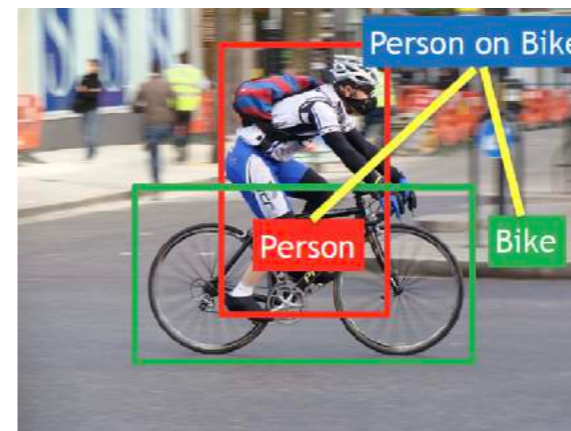
		Valores reales	
		enfermos	sanos
Valores predichos	enfermos	TP	FP
	sanos	FN	TN

Receiver Operating Curve (ROC)



Aplicaciones

- speech recognition
- biometría: huellas, iris, cara, firma, oreja, teclado, ...
- filtros de spam
- diagnóstico (desbalance de clases)
- detección anomalías
- recomendaciones
- identificación objetos
- predicción genómica
- visualización
- lenguaje natural
- mercados y stock
- ...



Selección y extracción de características

- Extraer la información que puede permitir la discriminación.
- Eliminar información redundante e irrelevante.
- Reducir la dimensionalidad del problema.
- **Extracción:** crear nuevas características combinación de las características originales. Determinar un conjunto de dimensionalidad *menor* en el espacio original.
- **Selección:** seleccionar las características (originales o transformadas en el proceso de extracción) con mayor poder de discriminación.
- Diseñadas por expertos en el tema o descubiertas por su capacidad de discriminación.

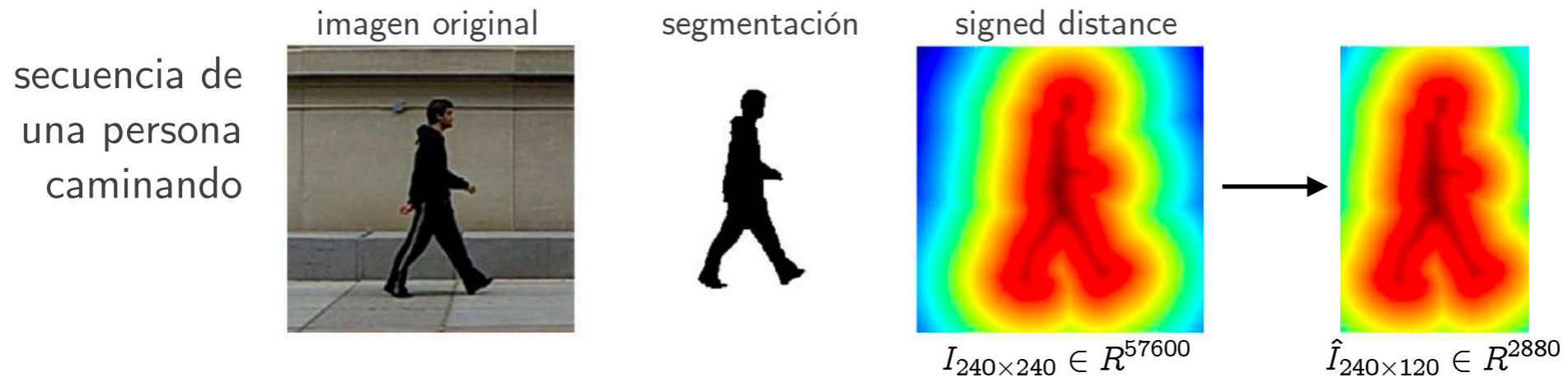


Botella de *vidrio* de altura h_b , ancho de base w_b , altura pico h_n y ancho de pico w_n

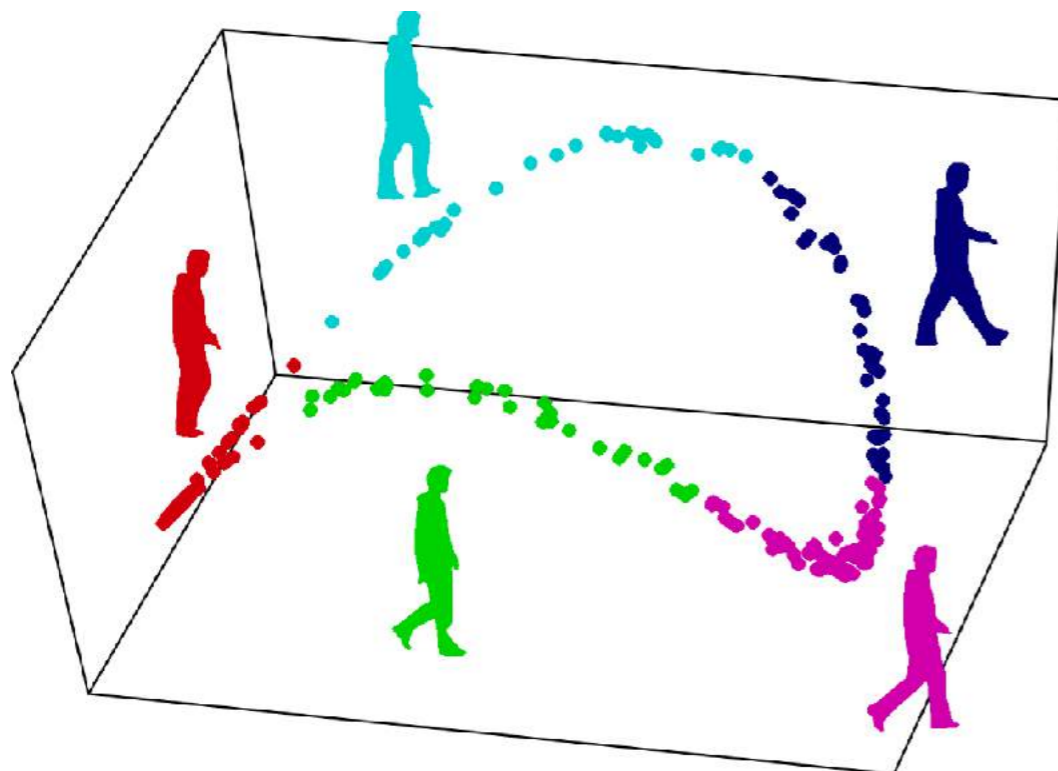


¿Descripción?

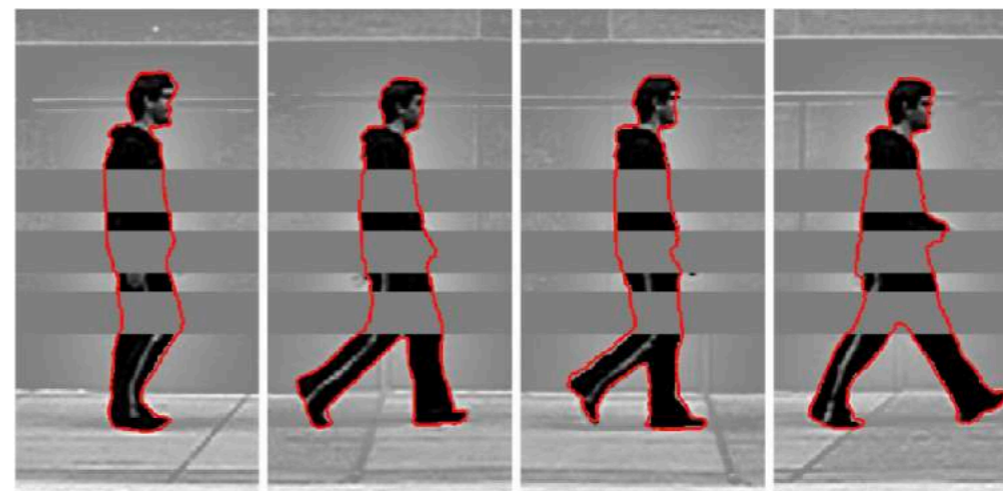
Reducción de dimensiones



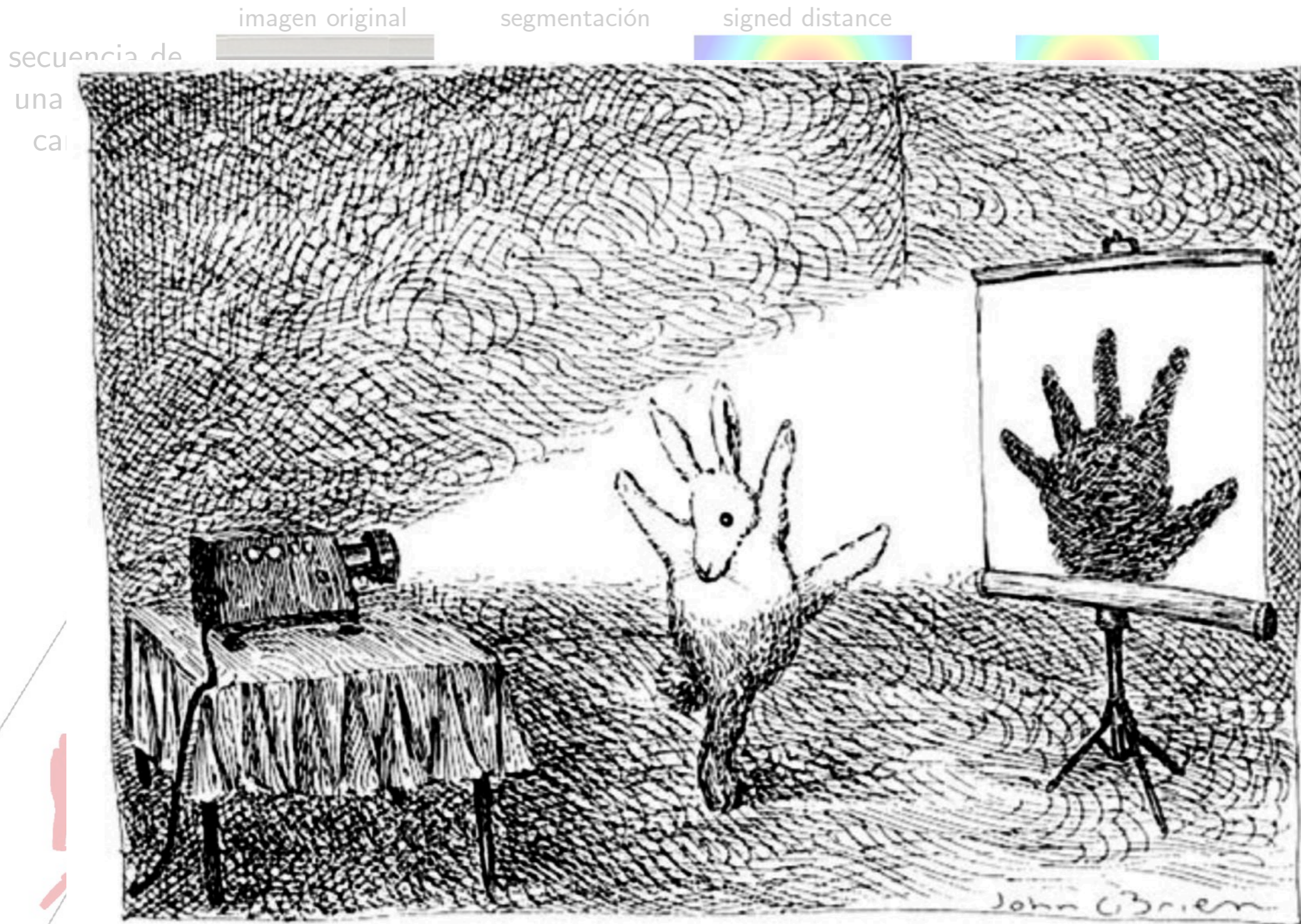
objetos segmentados, el color corresponde al cluster obtenido



- tres primeras dimensiones luego de un proceso de reducción de dimensiones (diffusion maps)
- los clusters se obtienen mediante k-means
- ajuste de modelo de la deformación de los objetos en cada cluster.



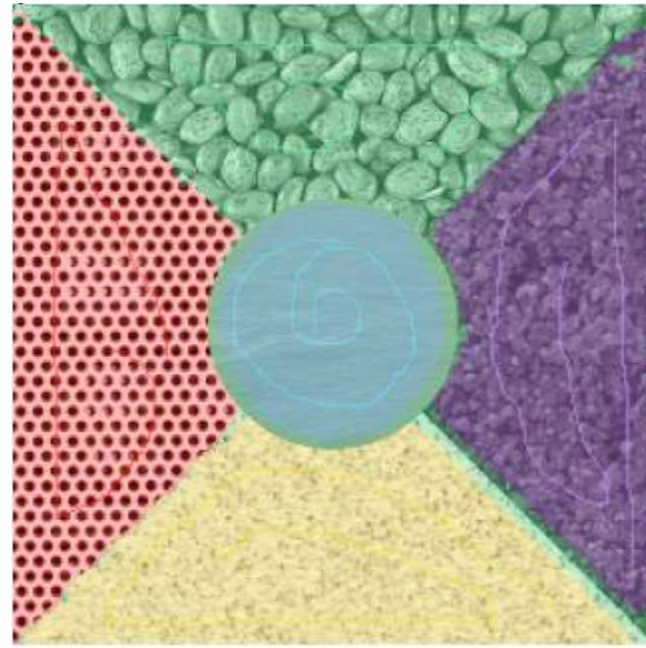
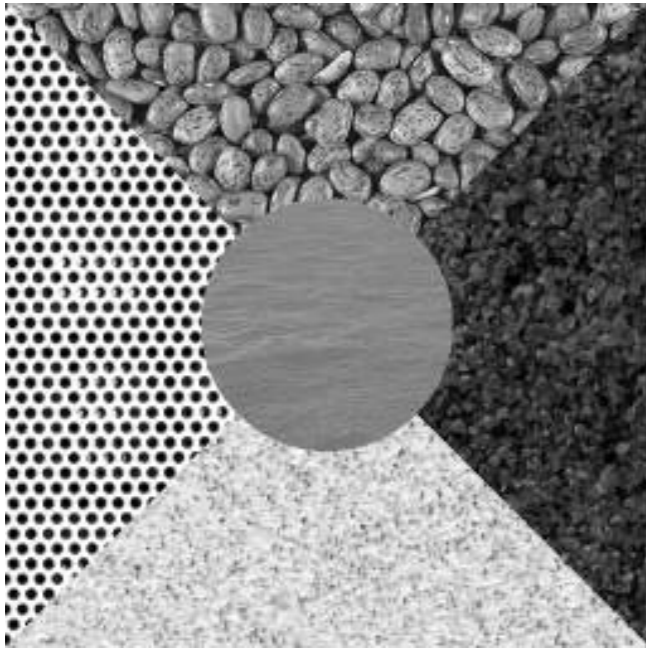
Reducción de dimensiones



Maldición de la dimensionalidad



RP en imágenes o video



clasificador (segmentación) de textura con Random Forest en ImageJ



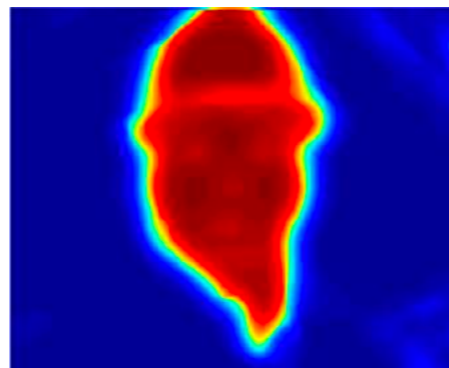
tracking objeto basado en probabilidad de color y posición



tracking objeto basado en probabilidad de color y posición



objeto de interés

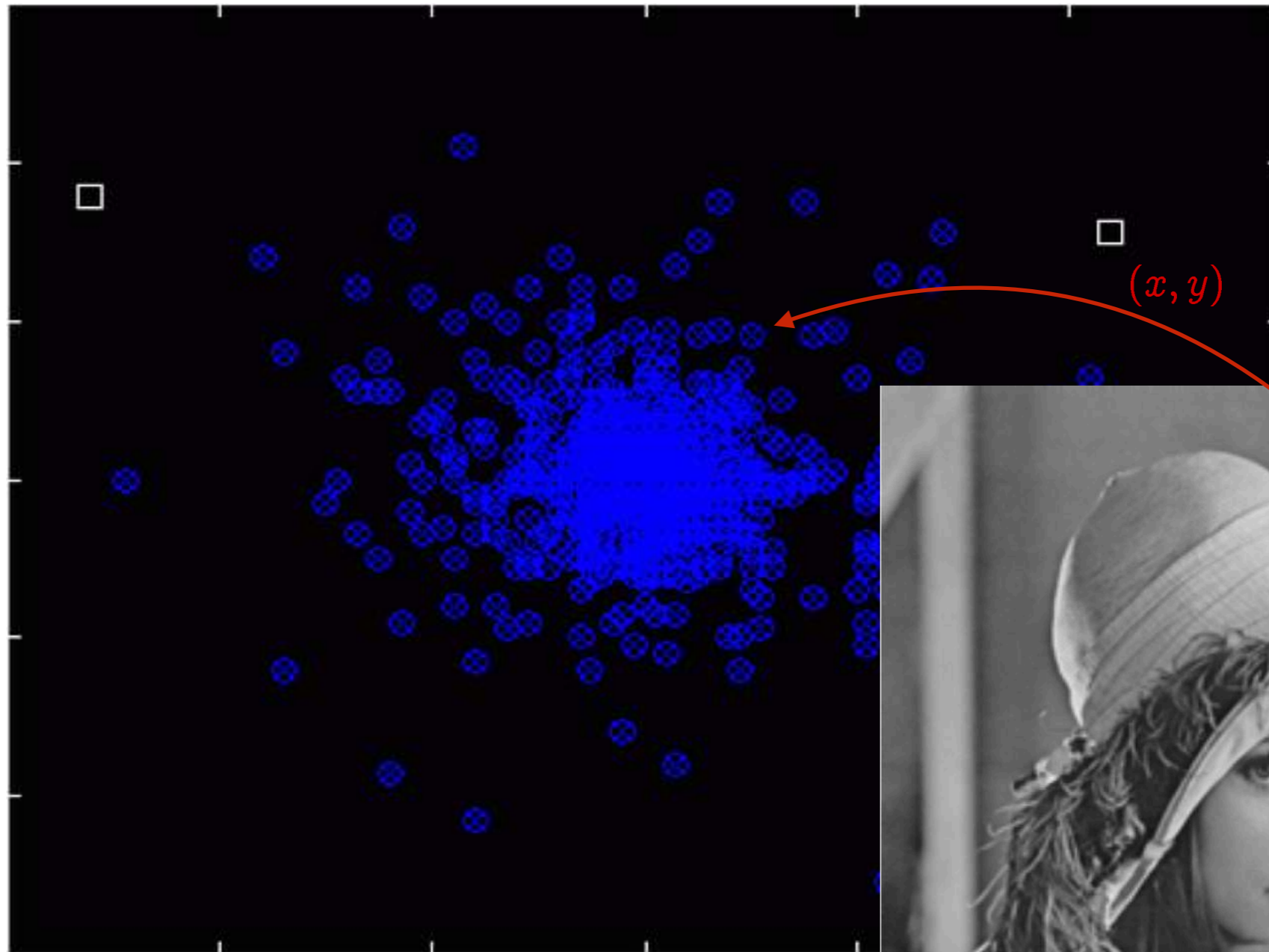


probabilidad del pixel de pertenecer al objeto



tracking objeto basado en probabilidad de color, posición y profundidad en la escena.

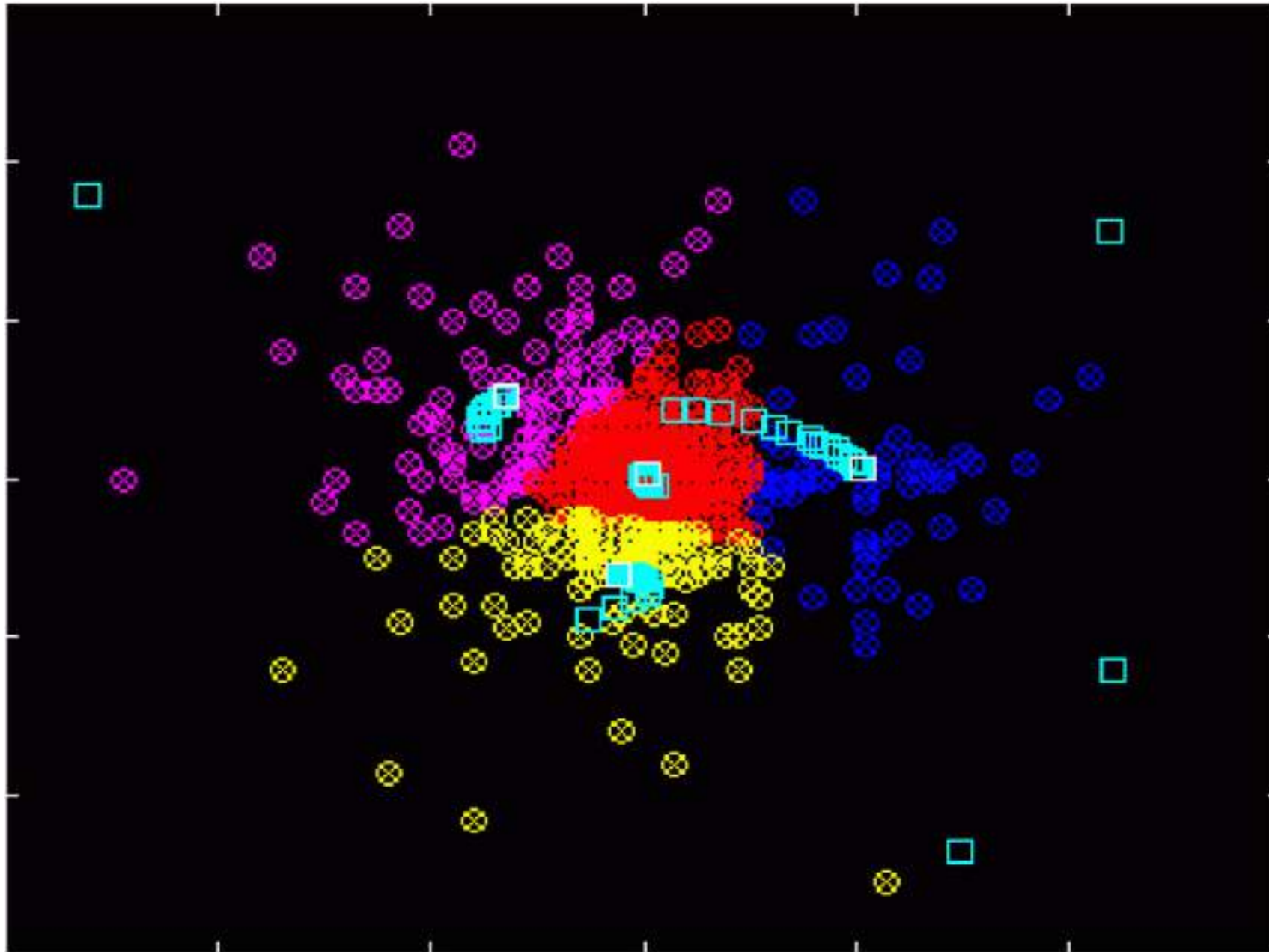
k-means



$k = 4$



k-means



iterar hasta cumplir alguna condición: error de representación (distorisión) o número de iteraciones

k-means



“Vectores” de dimensión 4 (bloques de 2×2), codebook de tamaño 32. Distorsión $D = 95.1$ y PSNR = 28.35dB.

Árboles de decisión y Bosques aleatorios

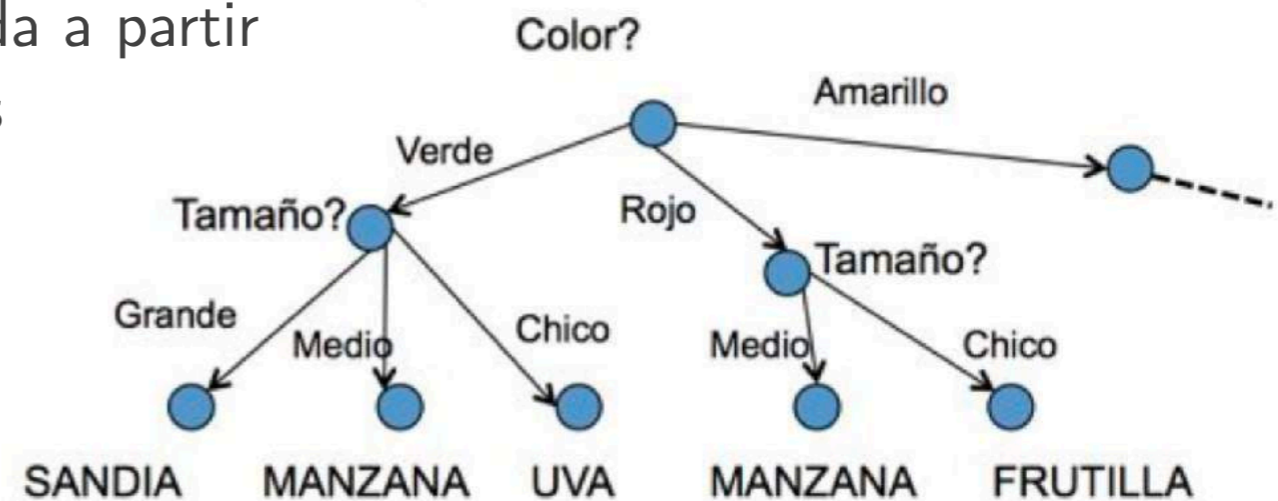
- Árboles de decisión (DTs): método no paramétrico supervisado para clasificación y regresión (CART). Busca crear un modelo para predecir el valor de salida a partir de *aprender* decisiones simples inferidas de las características.

- Ventajas

- Simple de entender, interpretar y visualizar.
- No requiere mucha preparación de los datos
- Puede manejar variables numéricas y categóricas.

- Desventajas:

- Sobreajuste (overfitting) a conjunto de entrenamiento.
- Inestables: pequeños cambios en datos dan árboles muy diferentes.
- Se aprende de heurísticas greedy que no garantizan optimalidad global
- Sesgo hacia una clase dominante.



color, tamaño, textura, sabor, ...

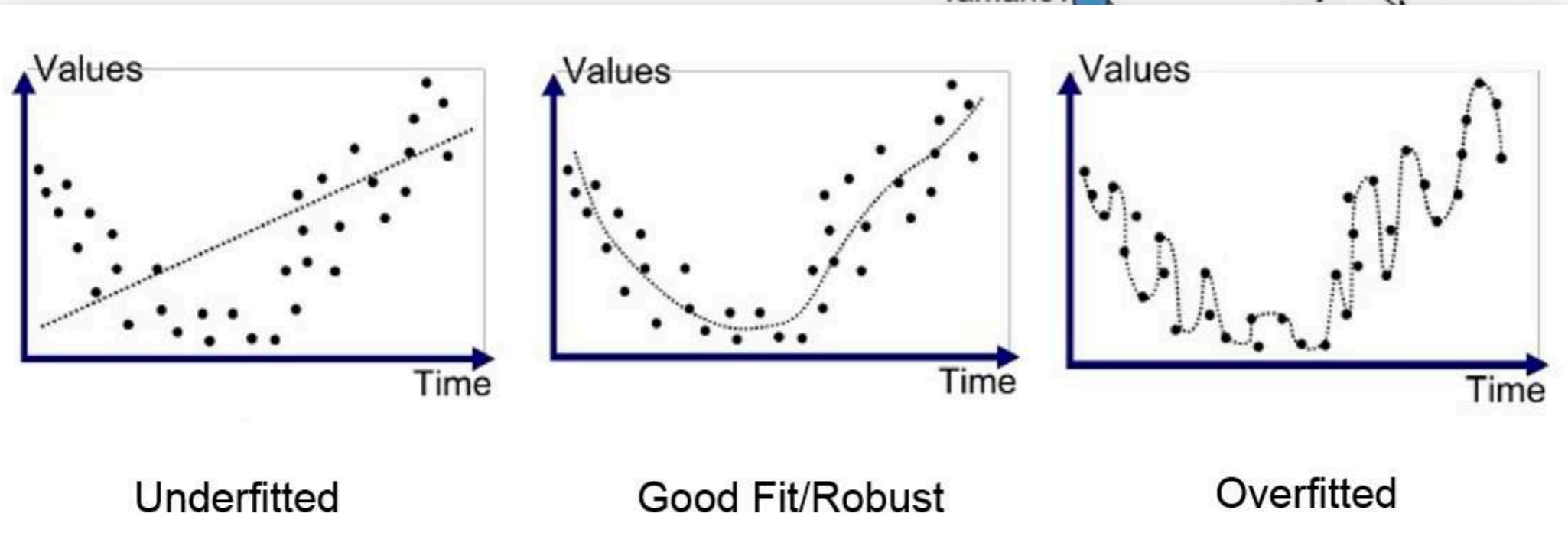
Combinación de DTs: Random Forest

Árboles de decisión y Bosques aleatorios

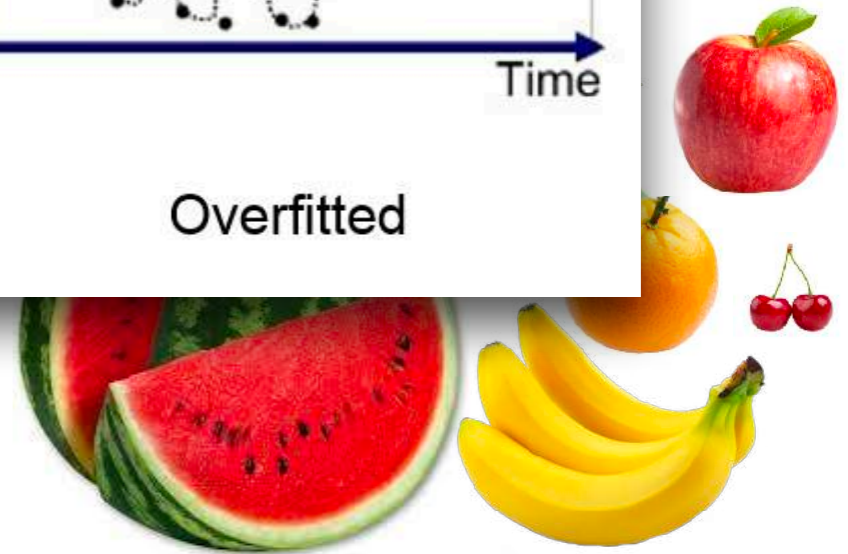
- Árboles de decisión (DTs): método no paramétrico supervisado para clasificación y regresión (CART). Busca crear un modelo para predecir el valor de salida a partir de *aprender* decisiones simples inferidas de las características.



- Venta
- Simp
- No r
- Puec
- Desve
- Sobr



- Inestables: pequeños cambios en datos dan árboles muy diferentes.
- Se aprende de heurísticas greedy que no garantizan optimalidad global
- Sesgo hacia una clase dominante.

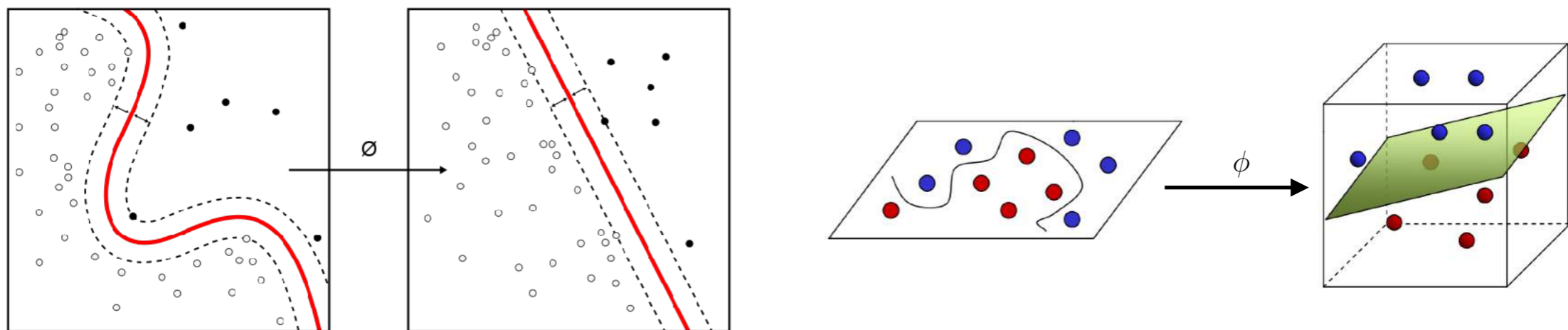


color, tamaño, textura, sabor, ...

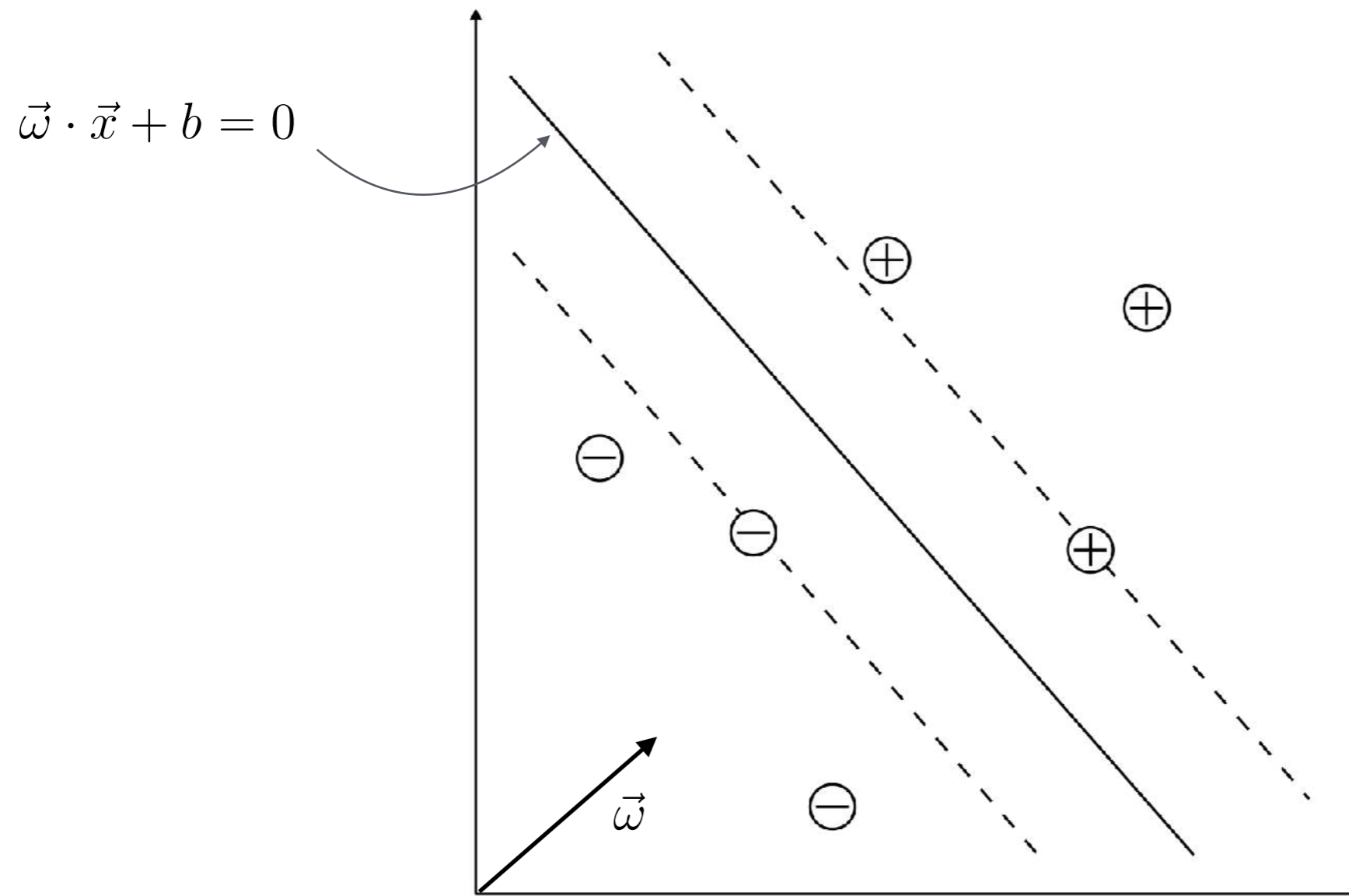
Combinación de DTs: Random Forest

Support Vector Machines (SVM)

- Las SVMs son métodos paramétricos supervisados para clasificación y regresión (SVR).
- Una SVM separa las clases en dos espacios lo más amplios posibles mediante un *hiperplano* de separación (lineal) definido como el vector entre los dos puntos, de las dos clases, más cercanos al que se llama **vectores soporte**.
- En el entrenamiento hay varios parámetros para ajustar, incluyendo la elección de posible núcleos (transformaciones).
- Hay variantes para semi-supervisado (Transductive SVM).



Clasificación de clases linealmente separables



Decision rule: $\vec{\omega} \cdot \vec{u} + b \geq 0 \Rightarrow \vec{u} \in \oplus$

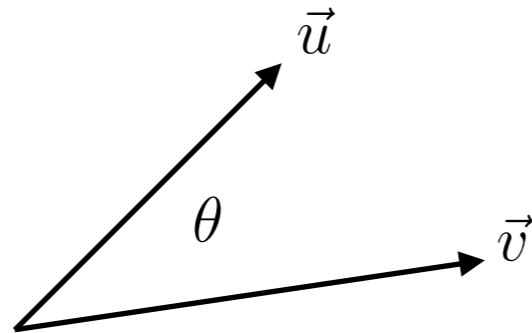
(Recordatorio)

- Inner product (dot product) between two vectors

$$\vec{u} \cdot \vec{v} = \langle \vec{u}, \vec{v} \rangle = \vec{u}^T \vec{v} = \vec{v}^T \vec{u} = \sum_i u_i v_i = u_1 v_1 + u_2 v_2 + \cdots + u_m v_m$$

- Geometric interpretation:

- Angle between vectors
- Projection



$$\vec{u} \cdot \vec{v} = \|\vec{u}\| \|\vec{v}\| \cos \theta$$

Clasificación de clases linealmente separables

- The regions are defined by

$$\vec{\omega} \cdot \vec{x}_+ + b \geq 0$$

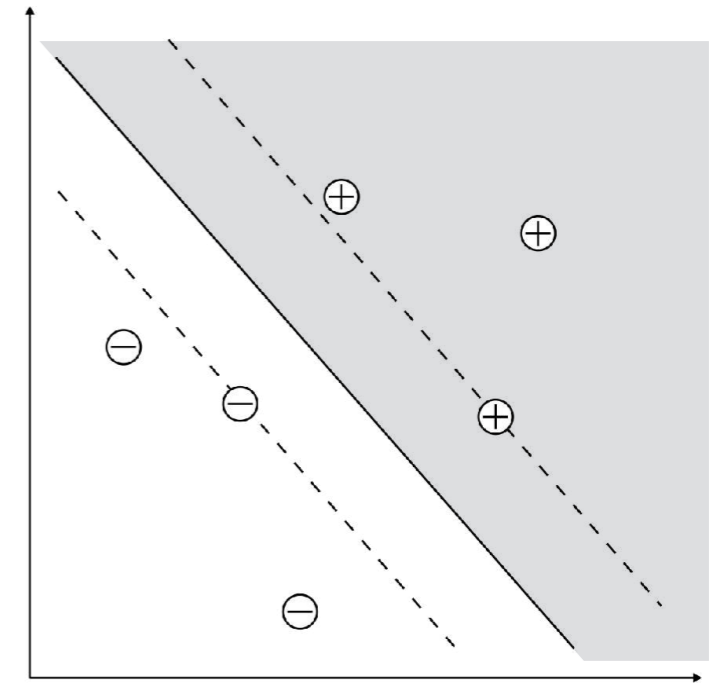
$$\vec{\omega} \cdot \vec{x}_- + b \leq 0$$

- Increase the gap to the decision boundary

$$\vec{\omega} \cdot \vec{x}_+ + b \geq 1$$

$$\vec{\omega} \cdot \vec{x}_- + b \leq -1$$

- Using
$$y_i = \begin{cases} +1 & \vec{x}_+ \in \oplus \\ -1 & \vec{x}_- \in \ominus \end{cases} \quad y_i(\vec{\omega} \cdot \vec{x} + b) - 1 \geq 0$$



and the equality holds for the samples in the “gutter” of this “street”.

Clasificación de clases linealmente separables

- For these two samples

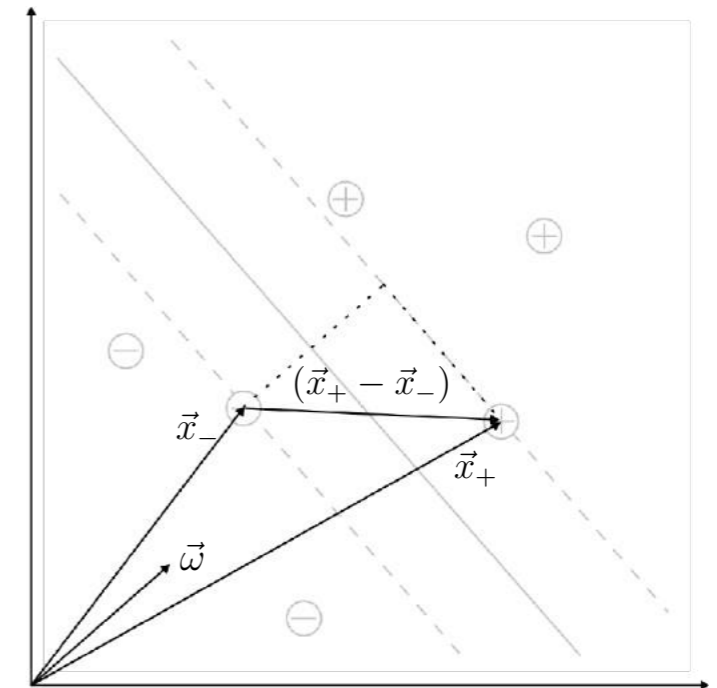
$$y_i(\vec{\omega} \cdot \vec{x} + b) - 1 = 0 \quad \Rightarrow \quad \begin{aligned} \vec{\omega} \cdot \vec{x}_+ &= 1 - b \\ \vec{\omega} \cdot \vec{x}_- &= -(1 + b) \end{aligned}$$

- The width of the gap can be computed by

$$\begin{aligned} (\vec{x}_+ - \vec{x}_-) \cdot \frac{\vec{\omega}}{\|\vec{\omega}\|} &= ((1 - b) + (1 + b)) \frac{1}{\|\vec{\omega}\|} \\ &= \frac{2}{\|\vec{\omega}\|} \end{aligned}$$

...and we want it as big as possible.

$$\max \frac{1}{\|\vec{\omega}\|}$$



Clasificación de clases linealmente separables

$$\min \frac{1}{2} \|\vec{\omega}\|^2$$

With some constraints $y_i(\vec{\omega} \cdot \vec{x} + b) - 1 = 0$

- Classical constrained optimisation: Lagrange multipliers!

$$\mathcal{L} = \frac{1}{2} \|\vec{\omega}\|^2 - \sum_i \lambda_i [y_i(\vec{\omega} \cdot \vec{x} + b) - 1]$$

$$\frac{\partial \mathcal{L}}{\partial \vec{\omega}} = \vec{\omega} - \sum_i \lambda_i y_i \vec{x}_i = 0 \Rightarrow \vec{\omega}^* = \sum_i \lambda_i y_i \vec{x}_i$$

$$\frac{\partial \mathcal{L}}{\partial b} = \sum_i \lambda_i y_i = 0$$

- The optimal $\vec{\omega}^*$ (hyperplane) is a linear combination of some ($\lambda_i \neq 0$) of the sample vectors (the *support vectors*).

Clasificación de clases linealmente separables

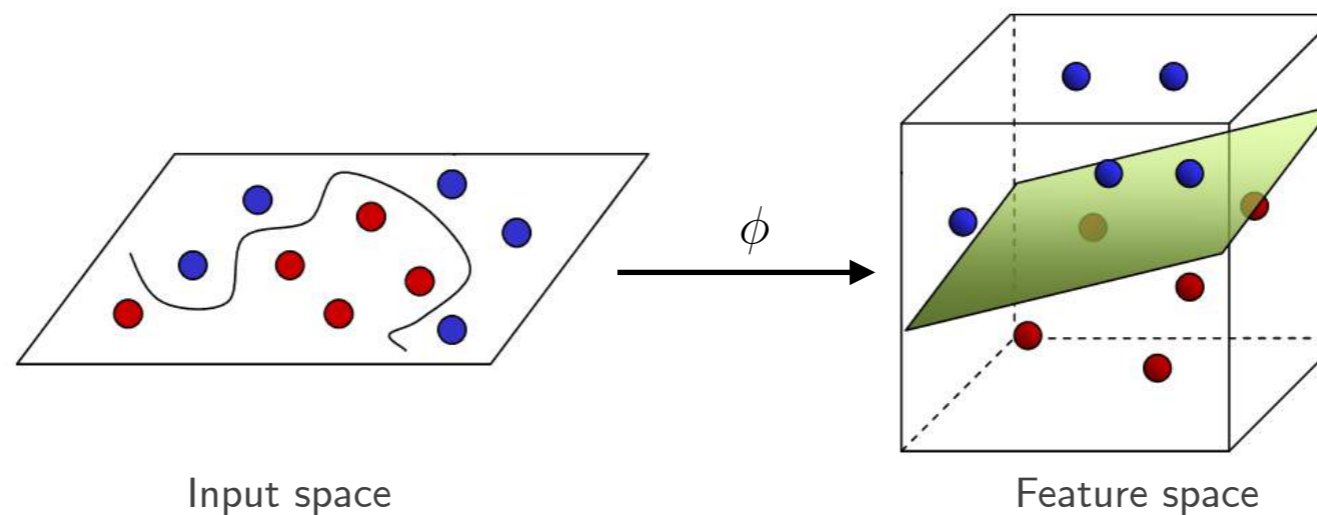
- The **optimal decision rule** also depends only in the **inner product** between some samples and the input vector

$$\vec{\omega} \cdot \vec{u} + b \geq 0 \Rightarrow \vec{u} \in \oplus$$

$$\left(\sum_i \lambda_i y_i \vec{x}_i \right) \cdot \vec{u} + b = \sum_i \lambda_i y_i (\vec{x}_i \cdot \vec{u}) + b \geq 0 \Rightarrow \vec{u} \in \oplus$$

Técnicas no-lineales

- Previous deduction is for linear decision boundaries.
- What can we do with non-linearly separable dataset?
 - Transform the *input space* by some mapping function ϕ into a new (higher dimensionality) *feature space*.
 - Use linear techniques in the *feature space*.



- The deduction is the same replacing \vec{x}_i by $\phi(\vec{x}_i)$.
- The decision rule will be

$$\sum_i \lambda_i y_i (\phi(\vec{x}_i) \cdot \phi(\vec{u})) + b \geq 0$$

Kernel Trick

- Mapping into high dimensional spaces has some practical disadvantages
 - Memory usage, computing time
 - Find the transformation
- But, there is no need of explicitly compute the transformation!

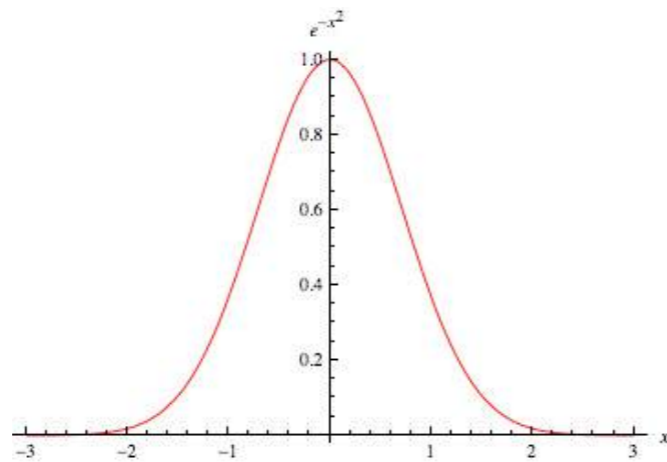
$$\sum_i \lambda_i y_i (\phi(\vec{x}_i) \cdot \phi(\vec{u})) + b = \sum_i \lambda_i y_i K(\vec{x}_i, \vec{u}) + b \geq 0$$

$$K(\vec{v}, \vec{u}) = \phi(\vec{v}) \cdot \phi(\vec{u}) \quad K : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$$

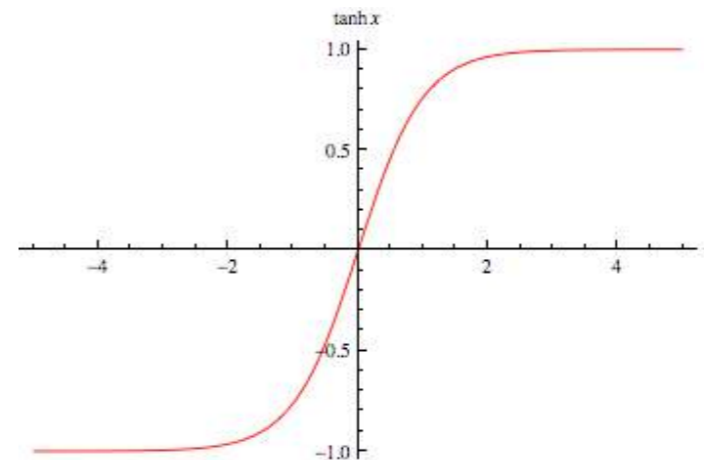
- Kernel functions
 - Implicitly transform samples and compute their inner product.
 - Efficiently construct non-linear versions of linear techniques.
 - They should be based in the inner product.

Popular Kernels

- Linear kernel: $\vec{u} \cdot \vec{v}$.
- Polynomial kernels: $(\gamma \vec{u} \cdot \vec{v} + \gamma_0)^d$
- Radial Basis Function (RBF) kernels: $\exp(-\gamma \|\vec{u} \cdot \vec{v}\|)$ $\gamma > 0$
 $\exp(-\gamma \|\vec{u} - \vec{v}\|^2)$ $\gamma > 0$
- Sigmoid kernel: $\tanh(\gamma \vec{u} \cdot \vec{v} + \gamma_0)$



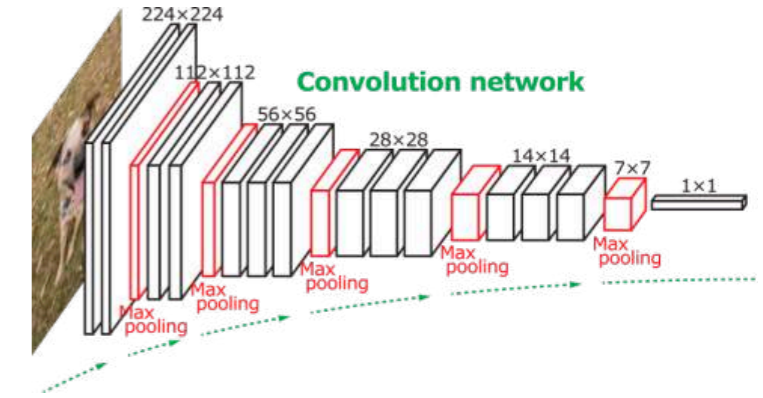
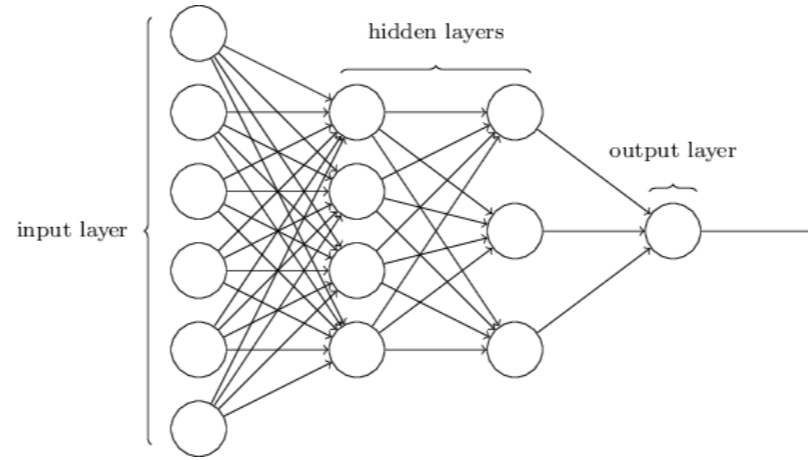
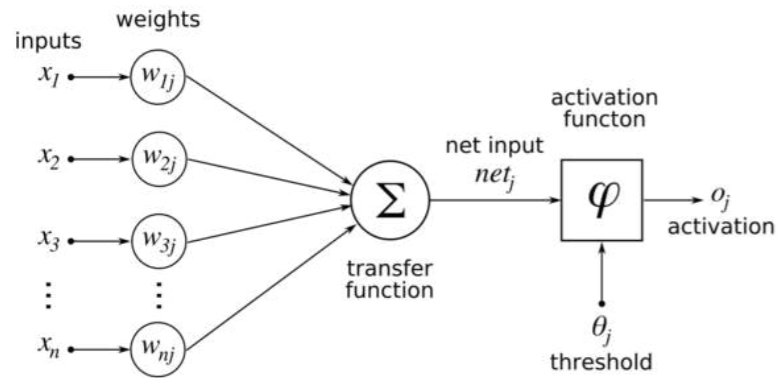
RBF kernel



Sigmoid kernel

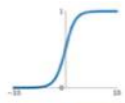
- More kernels: Bessel function, ANOVA RBF, linear splines, ...

Deep Learning



Activation functions

Sigmoid
 $\sigma(x) = \frac{1}{1+e^{-x}}$



Leaky ReLU
 $\max(0.1x, x)$



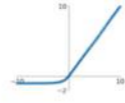
tanh
 $\tanh(x)$



Maxout
 $\max(w_1^T x + b_1, w_2^T x + b_2)$

ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



ReLU
 $\max(0, x)$



Curso de Actualización y Posgrado, 2020 (Se dictará en modo remoto)



Aprendizaje Profundo para Visión Artificial

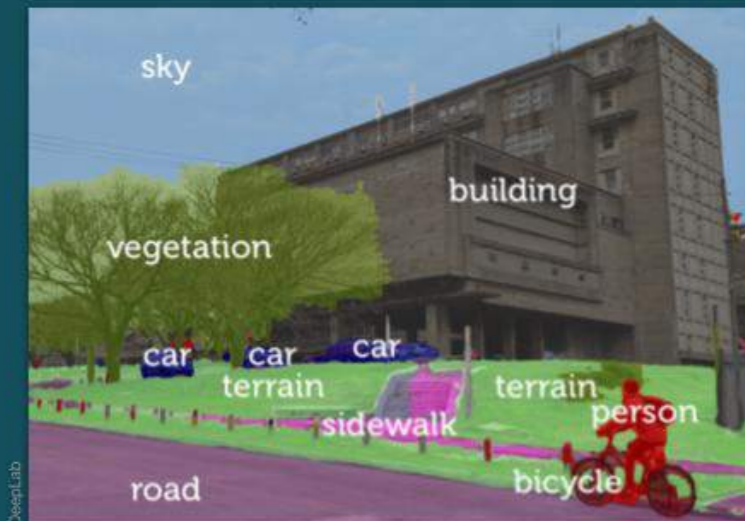


DLVIS2020

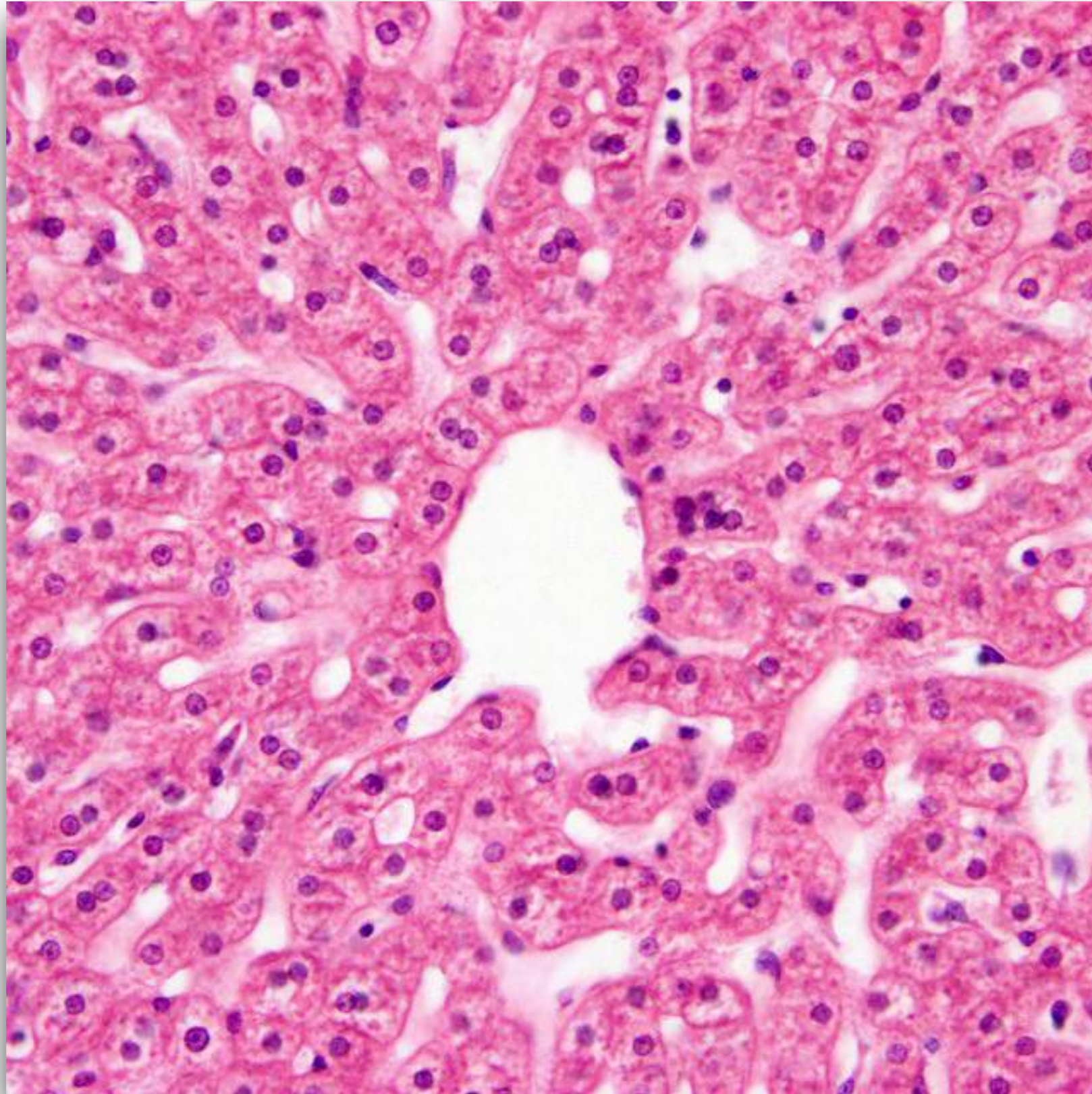
- Aprendizaje supervisado y no supervisado
- Redes neuronales de convolución (CNN)
- Reconocimiento-Detección-Segmentación
- Problemas inversos en procesamiento de imágenes
- Modelos secuenciales (RNN-LSTM-Transformer)
- Modelos generativos (GAN-VAE-Normalizing Flows)
- Aprendizaje profundo en la práctica (bibliotecas)
- Generalización - Aprendizaje profundo Bayesiano



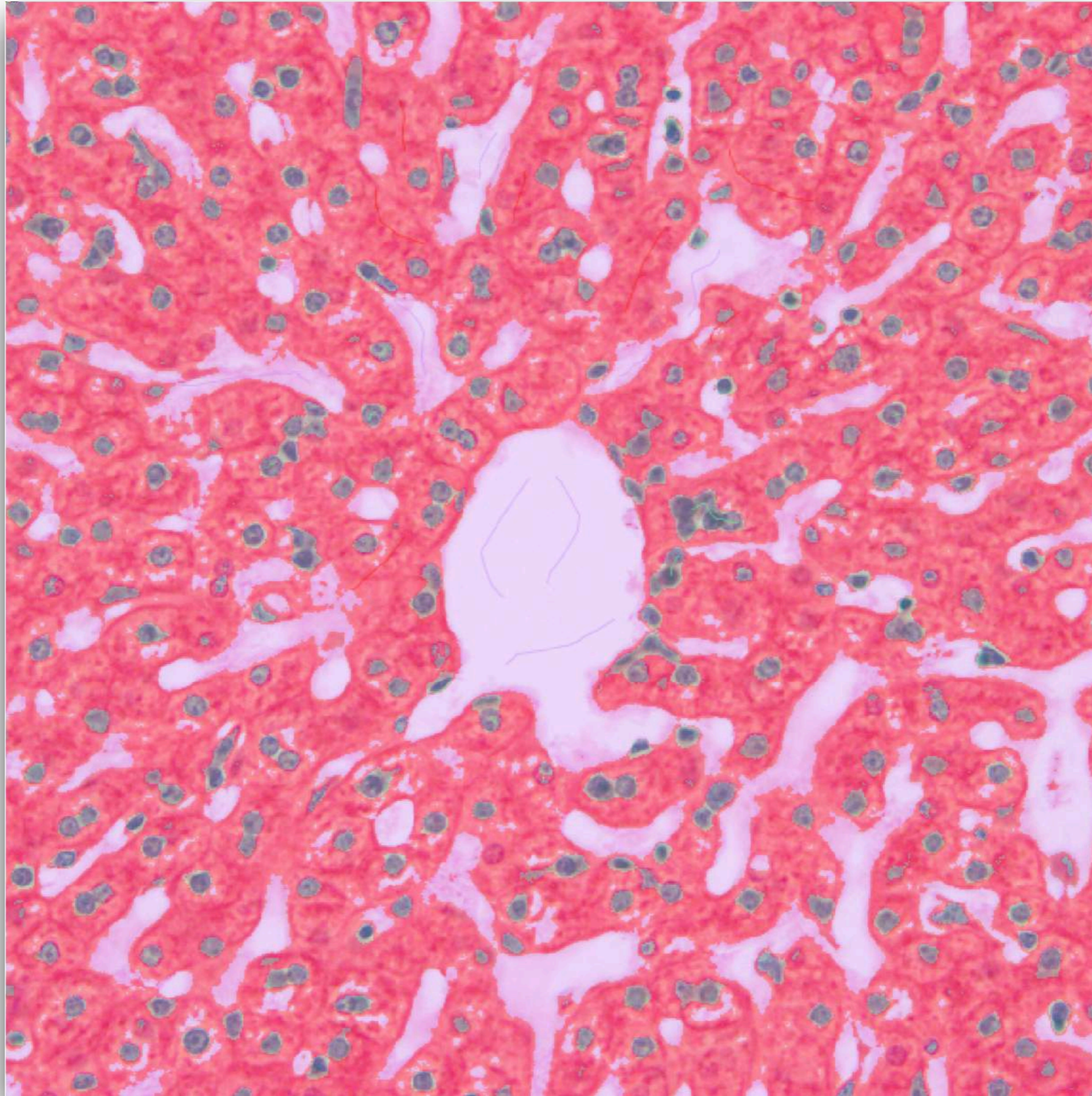
Inicio: 29 de Setiembre
 Contacto: jlezama@fing.edu.uy
<https://eva.fing.edu.uy> (dlvis)



RP para imágenes en FIJI



RP para imágenes en FIJI



RP para imágenes en FIJI

Trainable Weka Segmentation v2.1.0-SNAPSHOT

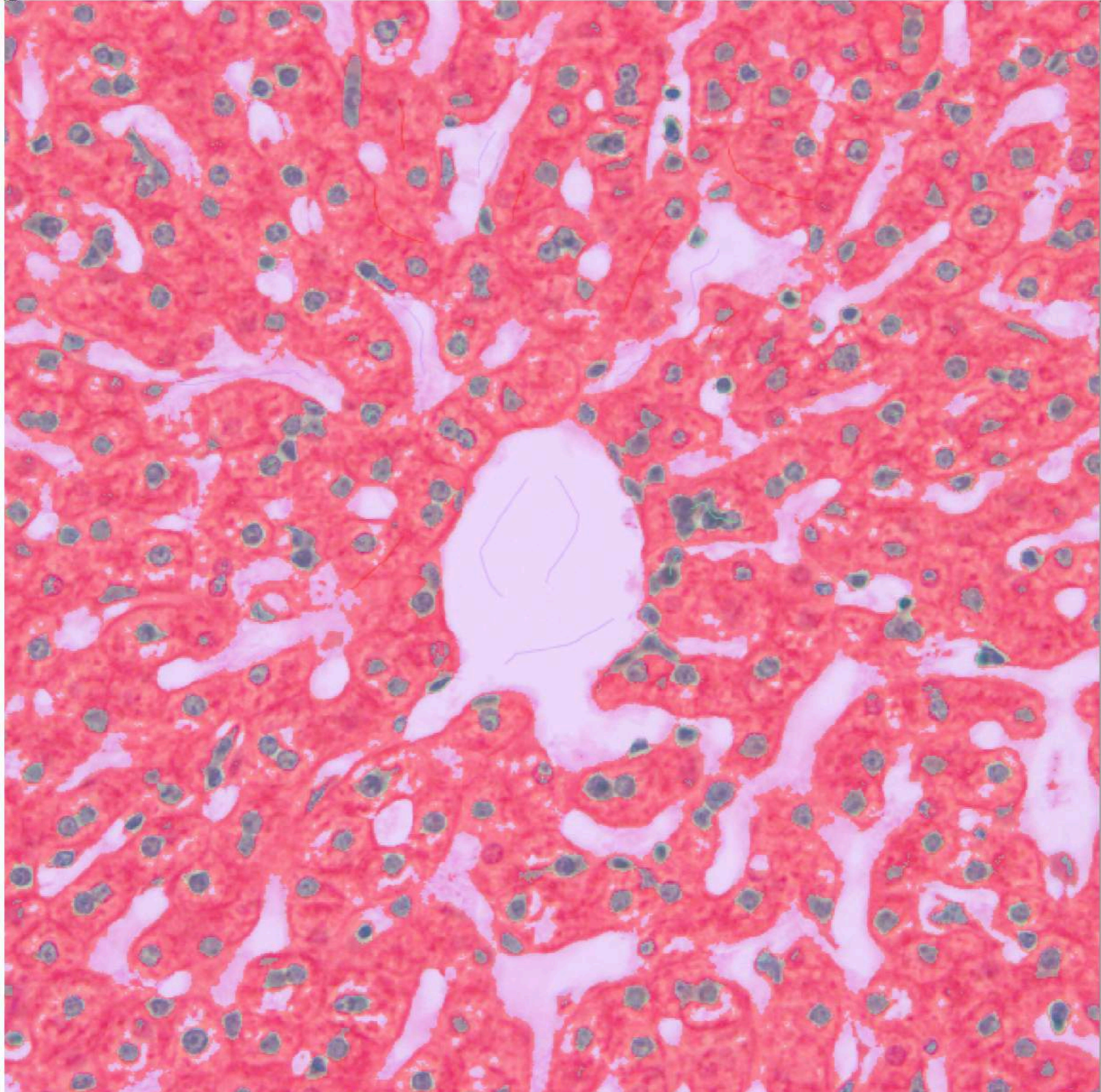
640x640 pixels: RGB: 1.6MB

Training

- Train classifier
- Toggle overlay
- Create result
- Get probability
- Plot result

Options


- Apply classifier
- Load classifier
- Save classifier
- Load data
- Save data
- Create new class
- Settings



The central image shows a histological section of tissue, likely kidney, with various cellular structures. A large, irregularly shaped region in the center is highlighted in light purple, representing a segmented area. The rest of the image is predominantly red and pink, with scattered dark blue/purple spots representing nuclei. The segmentation is overlaid on the original image.

Labels

- Add to class 1
- trace 0 (Z=1)
- trace 1 (Z=1)
- trace 2 (Z=1)
- trace 3 (Z=1)
- trace 4 (Z=1)
- Add to class 2
- trace 0 (Z=1)
- trace 1 (Z=1)
- trace 2 (Z=1)
- trace 3 (Z=1)
- trace 4 (Z=1)
- Add to class 3
- trace 0 (Z=1)
- trace 1 (Z=1)
- trace 2 (Z=1)
- trace 3 (Z=1)
- trace 4 (Z=1)



WEKA
The University of Waikato

RP para imágenes en FIJI

Trainable Weka Segmentation v2.1.0-SNAPSHOT

640x640 pixels: RGB: 1.6MB

Training

Train classifier

Toggle overlay

Create result

Get probability

Plot result

Options

Apply classifier

Load classifier


Save classifier

Load data

Save data

Create new class

Settings



Segmentation settings

Training features:

<input checked="" type="checkbox"/> Gaussian blur	<input checked="" type="checkbox"/> Sobel filter
<input checked="" type="checkbox"/> Hessian	<input checked="" type="checkbox"/> Difference of gaussians
<input checked="" type="checkbox"/> Membrane projections	<input type="checkbox"/> Variance
<input type="checkbox"/> Mean	<input type="checkbox"/> Minimum
<input type="checkbox"/> Maximum	<input type="checkbox"/> Median
<input type="checkbox"/> Anisotropic diffusion	<input type="checkbox"/> Bilateral
<input type="checkbox"/> Lipschitz	<input type="checkbox"/> Kuwahara
<input type="checkbox"/> Gabor	<input type="checkbox"/> Derivatives
<input type="checkbox"/> Laplacian	<input type="checkbox"/> Structure
<input type="checkbox"/> Entropy	<input type="checkbox"/> Neighbors

Membrane thickness:

Membrane patch size:

Minimum sigma:

Maximum sigma:

Classifier options:

Choose

Class names:

Class 1

Class 2

Advanced options:

Balance classes

Save feature stack

Result overlay opacity 33

Help Cancel OK

Labels

Add to class 1

trace 0 (Z=1)

trace 1 (Z=1)

trace 2 (Z=1)

trace 3 (Z=1)

trace 4 (Z=1)

Add to class 2

trace 0 (Z=1)

trace 1 (Z=1)

trace 2 (Z=1)

trace 3 (Z=1)

trace 4 (Z=1)

Add to class 3

trace 0 (Z=1)

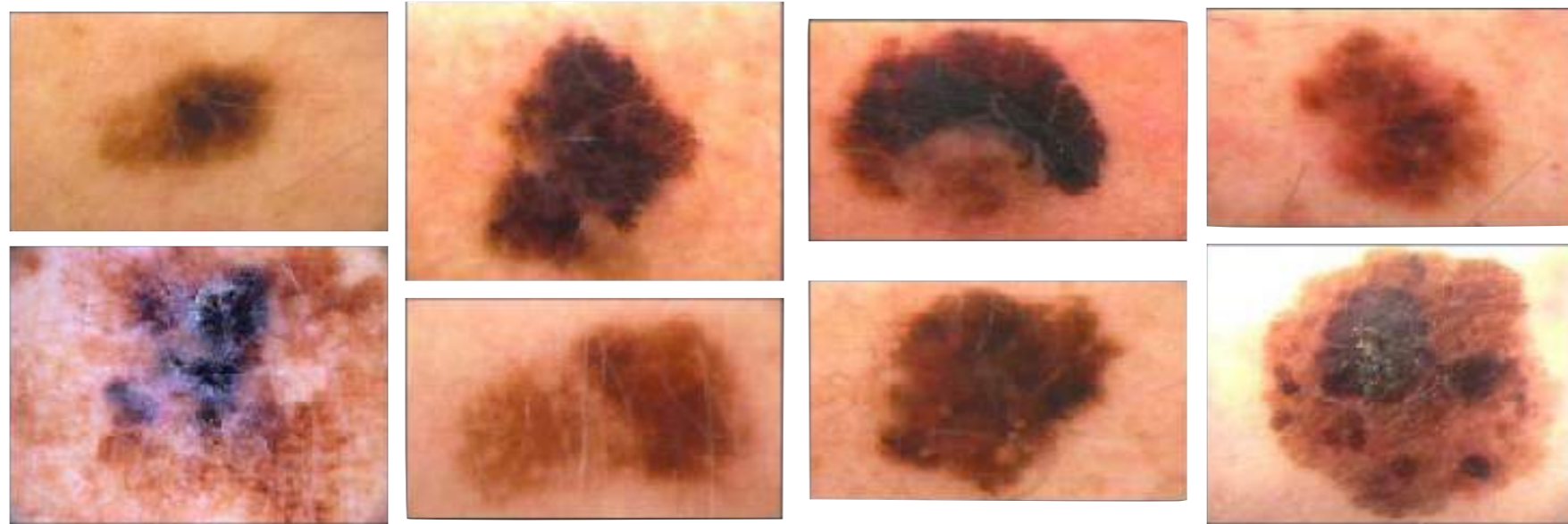
trace 1 (Z=1)

trace 2 (Z=1)

trace 3 (Z=1)

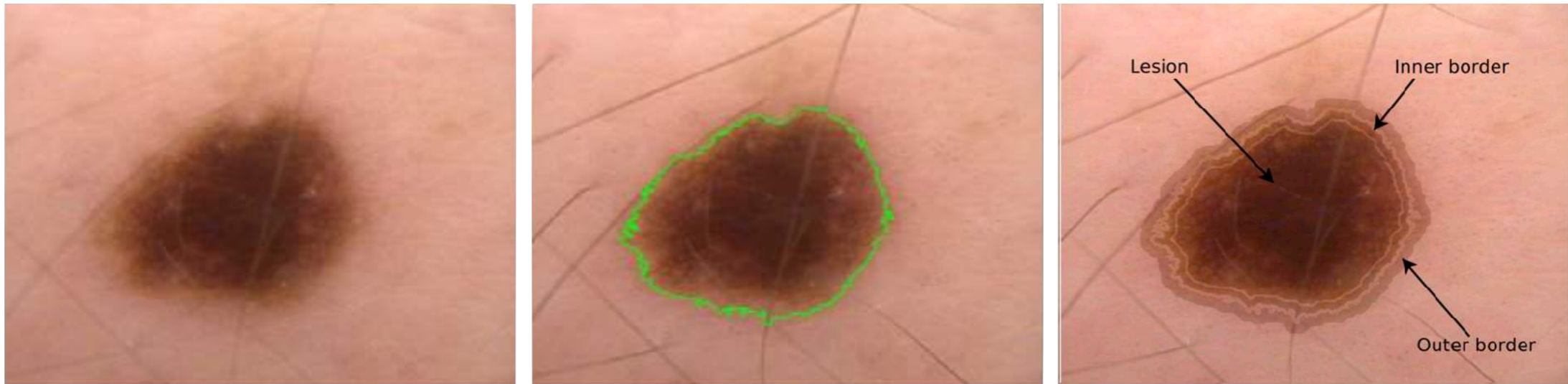
trace 4 (Z=1)

Más aplicaciones: diagnóstico de melanomas

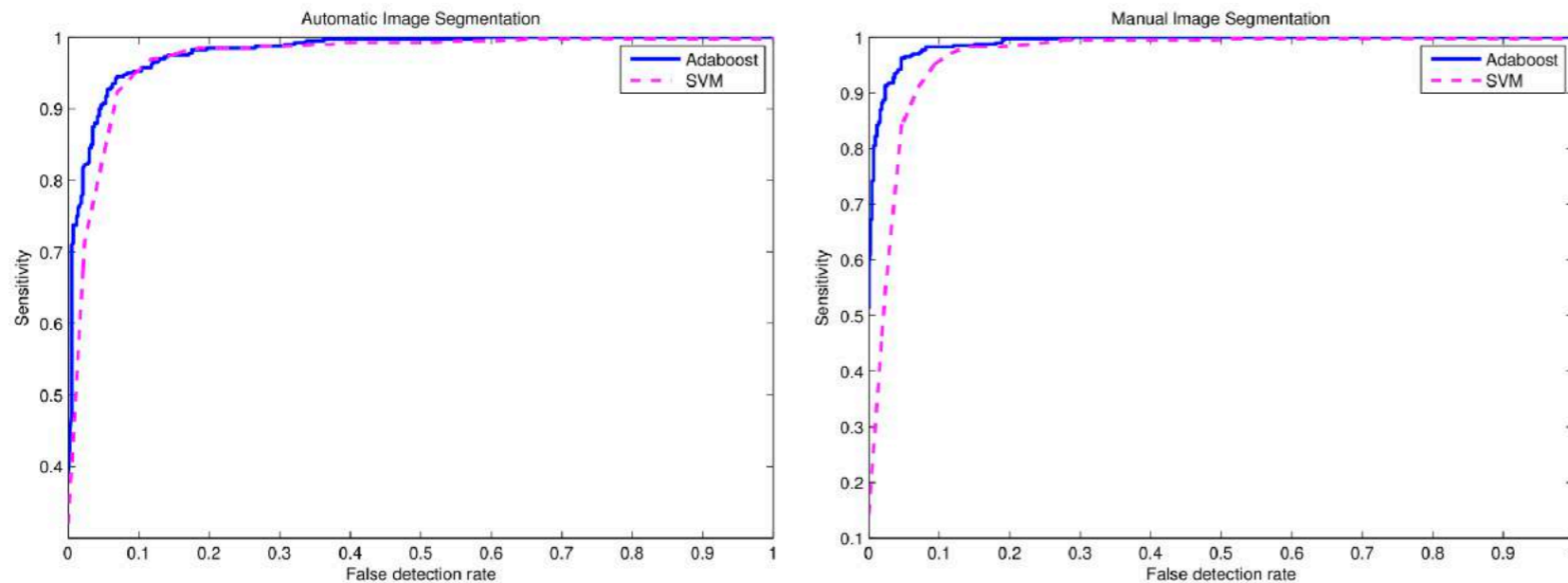


- Desarrollo algoritmo detección de melanomas usando imágenes dermoscópicas.
- Proceso semi-automático, sistemático y objetivo.
- Reproduciendo el enfoque del especialista... 57 características basadas en el diagnóstico clínico forma, bordes, color, textura

Más aplicaciones: diagnóstico de melanomas



- Desarrollo algoritmo detección de melanomas usando imágenes dermoscópicas.
- Proceso semi-automático, sistemático y objetivo.
- Reproduciendo el enfoque del especialista... 57 características basadas en el diagnóstico clínico forma, bordes, color, textura



Method	FPR for 95 % sensitivity	Area under ROC
Automatic segmentation, AdaBoost - C4.5	8.75 %	0.981
Automatic segmentation, SVM	9.52 %	0.963
Manual segmentation, AdaBoost - C4.5	4.62 %	0.991
Manual segmentation, SVM	9.23 %	0.966

Performance indicators for the ROC curves.

ROC curves for AdaBoost/C4.5 and SVM for automatic (L) and manual (R) segmentation.

Más aplicaciones

Predicción de grasa intramuscular por ultrasonido.



Nunes, J. L., Piquerez, M., Pujadas, L., Armstrong, E., Fernández, A., & Lecumberry, F. (2015). Beef quality parameters estimation using ultrasound and color images. *BMC bioinformatics*, 16(4), S6.

Detección de endometritis subclínica en imágenes de ultrasonido en bovinos de leche



Tailanián, M., Lecumberry, F., Fernández, A., Gnemmi, G., Meikle, A., Pereira, I., & Randall, G. (2014, November). Dairy cattle sub-clinical uterine disease diagnosis using pattern recognition and image processing techniques. In *Iberoamerican Congress on Pattern Recognition* (pp. 690-697). Springer, Cham.

Bibliografía

- Szeliski, R. *Computer Vision: Algorithms and Applications*. Springer, 2010 (<http://szeliski.org/Book>)
- Bishop, C. *Pattern Recognition and Machine Learning*. Springer, 2006.