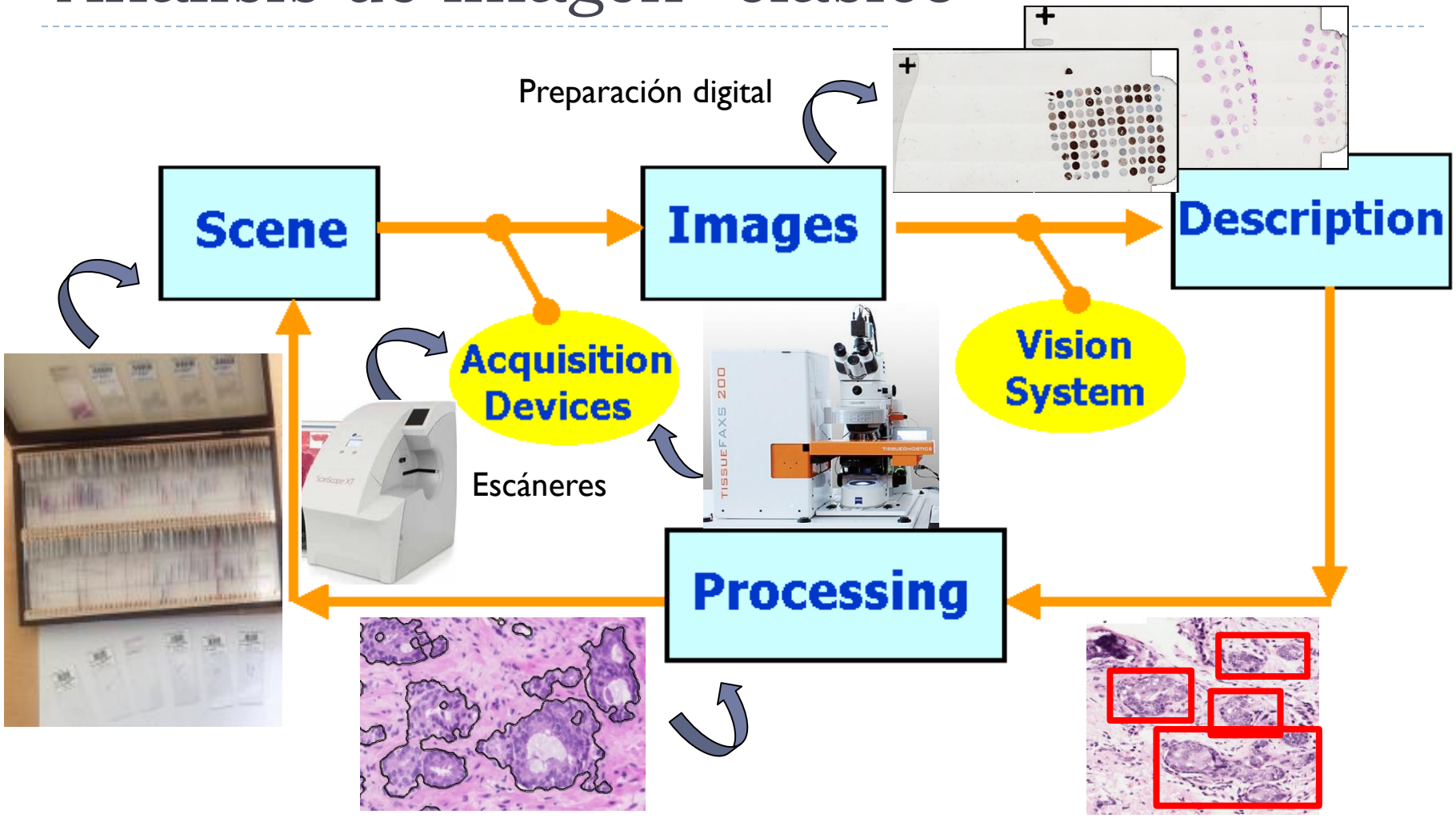
“Deep Learning”. Aplicaciones en Patología

Marcial García Rojo. Patólogo. Cádiz.
Gloria Bueno. VISILAB Group. Universidad de Castilla-La Mancha

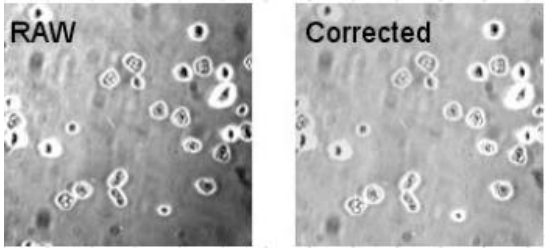
Para qué necesitamos análisis de imagen en Patología?

- Reducir la discrepancia inter/intraobservador (HER2)
- Mejorar la objetividad y claridad de nuestros informes
- Mejoras en la eficiencia ("más y mejor")
- Control de calidad (de técnicas, de áreas revisadas,…)
- Abordar nuevos flujos de trabajo: análisis supervisado o totalmente automatizado.

Análisis de imagen “clásico”



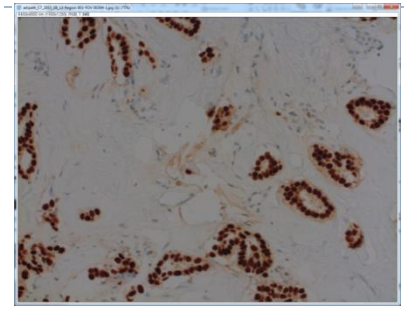
Métodos de análisis de imagen

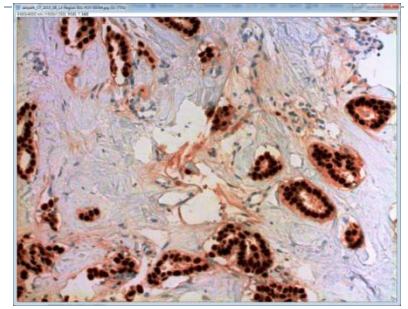


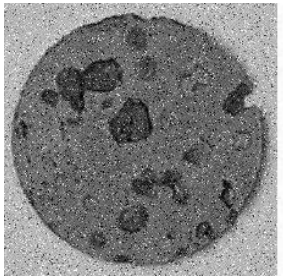
- Preprocesado
  - Filtrado
  - Realzado (ej. eliminar esquinas oscuras)
  - Análisis ➔ Procesado de color
- Segmentación
  - Detección de área de interés (ROI)
  - Análisis morfológico (ej. watershed)
- Registro de imágenes
  - Alineamiento de secciones seriadas (reconstrucción 3D)
- Reconocimiento de patrones – "Machine Learning"
  - Análisis CADx (Computer Aided Detection and Classification)

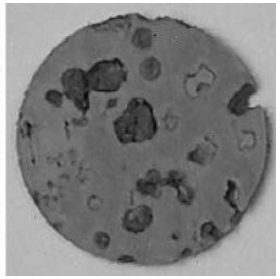
Ejemplos de técnicas: Preprocesado

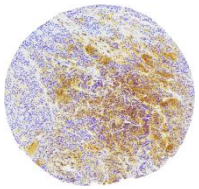
- Dominio espacial:
  - Ecualización de Histograma
  - Filtrado
    - Filtrado espacial no lineal (mediana, mínimo). Útil para eliminar ruido de fondo
  - Aritméticos. Separar canales (FISH: DAPI, …)
  - Procesado de color
    - Normalización









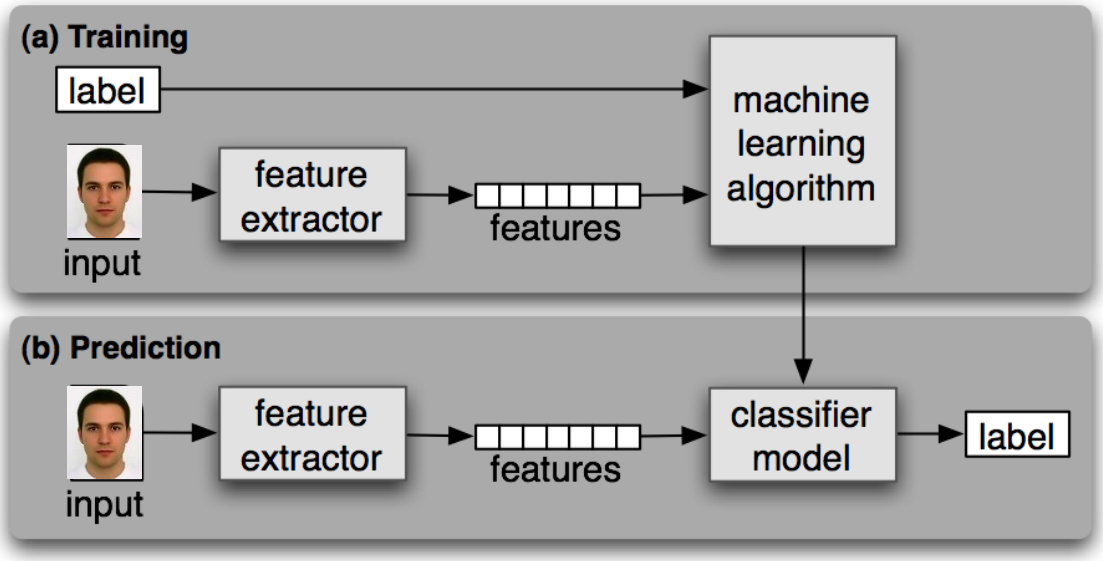


Reconocimiento de patrones

- Métodos tradicionales: parámetros seleccionados manualmente
  - SVM
  - Random Forest
  - Adaboost
  - ……
- Redes neurales convolucionales: "Deep Learning"

Conceptos básicos

- Algoritmos de aprendizaje: "Clasificadores"
- Métodos para extracción de parámetros



Comparación de resultados

- ImageNet, hasta 2012 usaba técnicas manuales para extraer parámetros. Tasa de error: 25%

- Krizhevsky, Sutskever y Hinton, "ImageNet Classification with Deep Convolutional Neural networks" (2012): Tasa de error de 15%

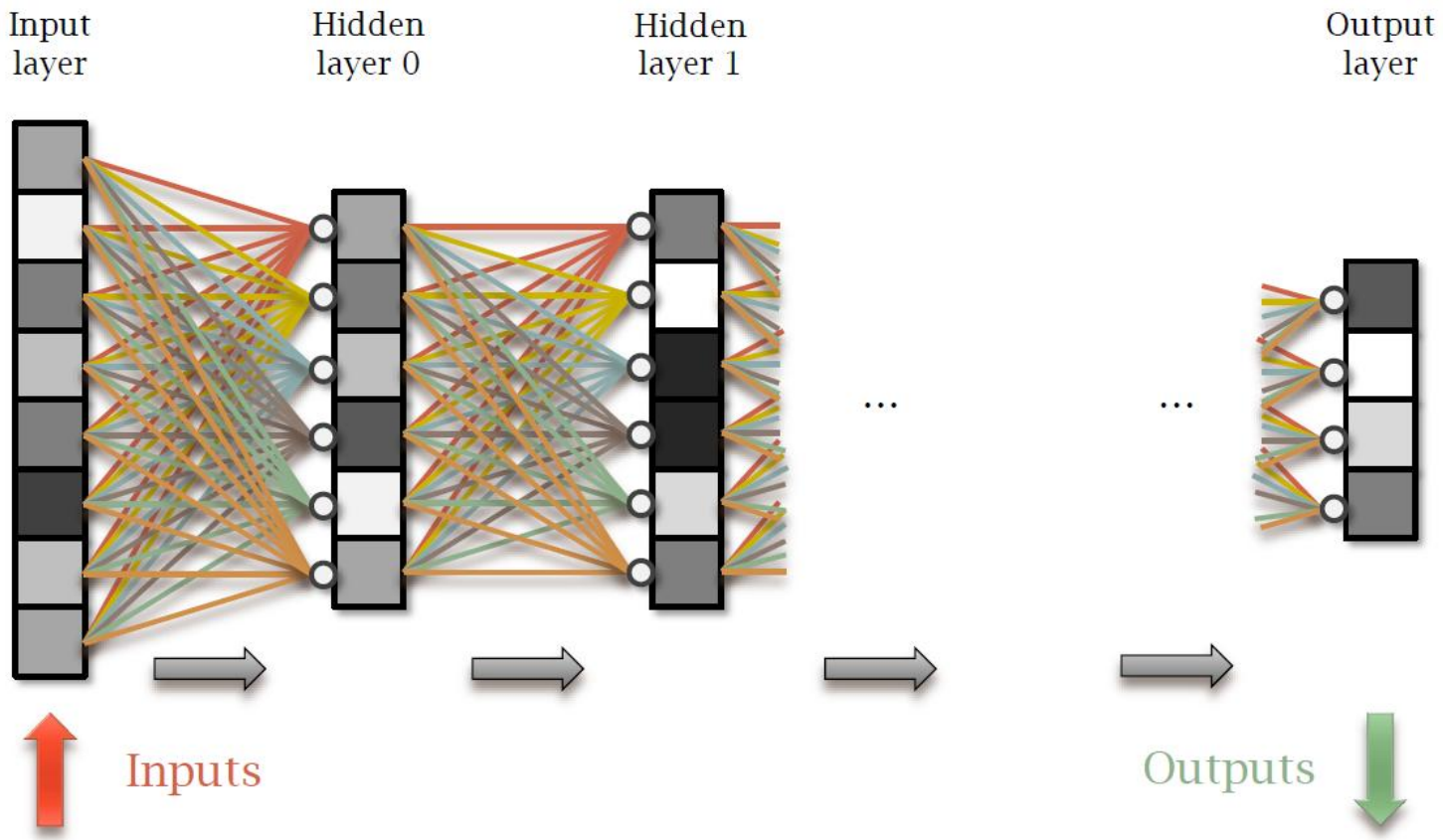
- Otros [Simonyan14, He15]: 5-7% Error

Introducción a Deep Learning

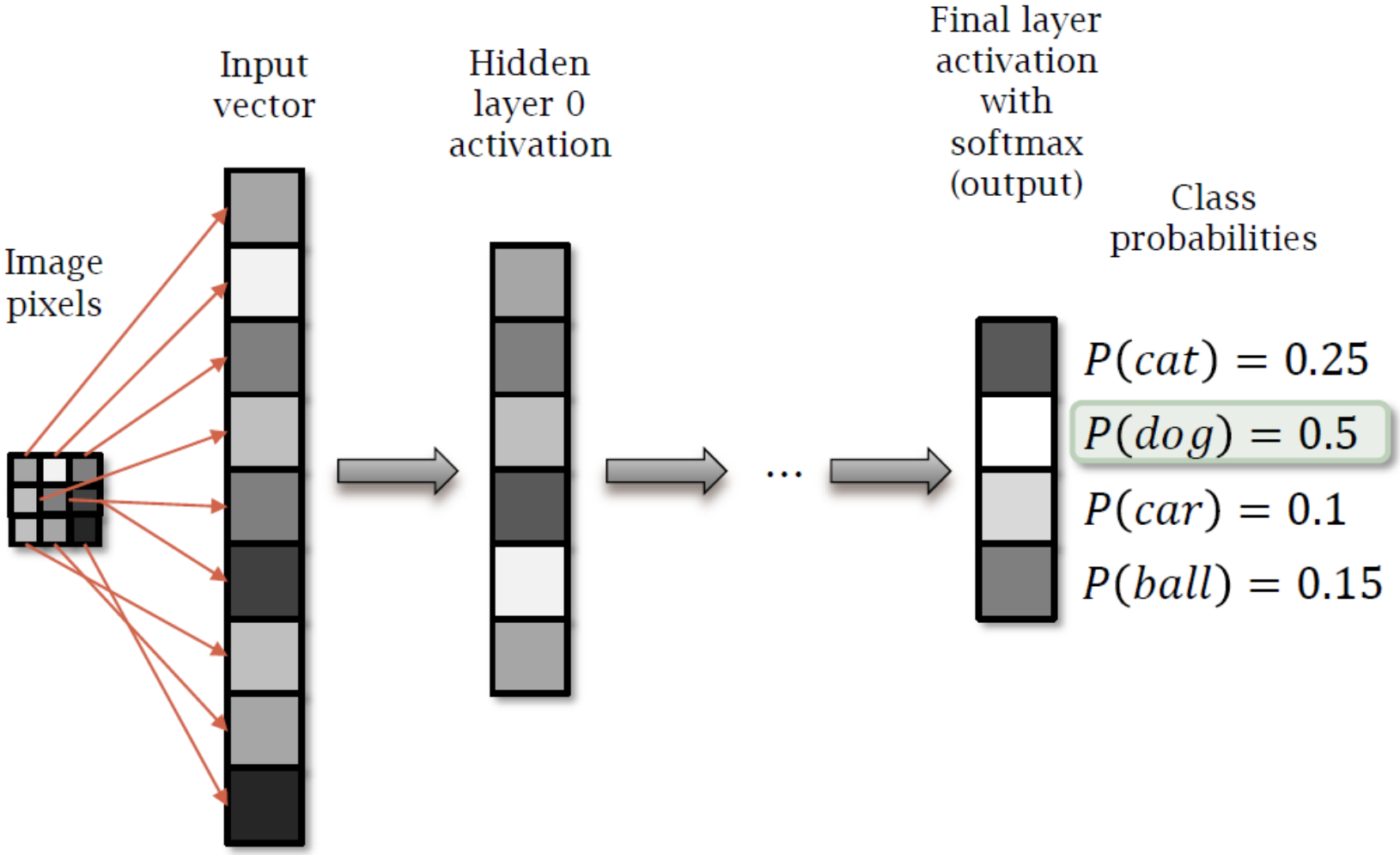
- Hoy día, las principales empresas en Internet y las instituciones de investigación más prestigiosas utilizan *deep learning* para desarrollar sus productos y servicios más innovadores

Deep Learning se basa en redes "neuronales"

- Son algoritmos con múltiples capas
- Propagan los datos a través de las diferentes capas
- Los datos se transforman en cada capa



Red neuronal como clasificador

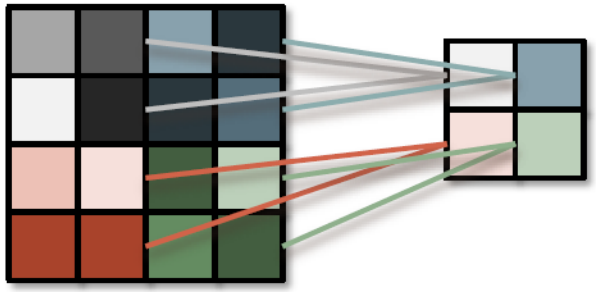


Entrenamiento de red neuronal

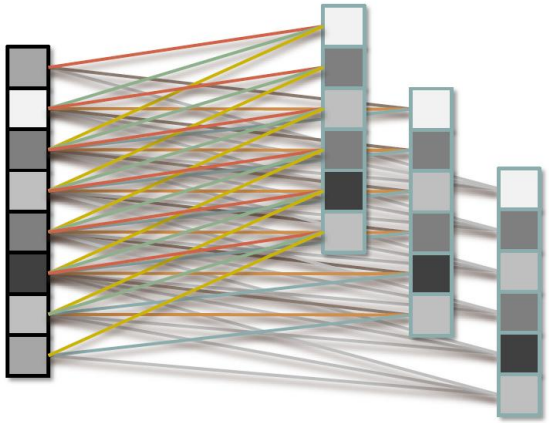
- Aprende valores por los parámetros (para cada capa)
- Se utiliza un método llamado "propagación hacia atrás" (método repetitivo)
- Una función de coste (o pérdida) mi dela diferencia entre la predicción y el resultado corregido
- Para imágenes complejas se requiere entrenar con un gran número de imágenes (numerosos parámetros en numerosas posiciones diferentes). Para representar esto → REDES NEURONALES CONVOLUCIONALES

Qué son las redes neuronales convolucionales o Deep Learning

- Son redes con 3 tipos de capas:
  - Capa convolucional

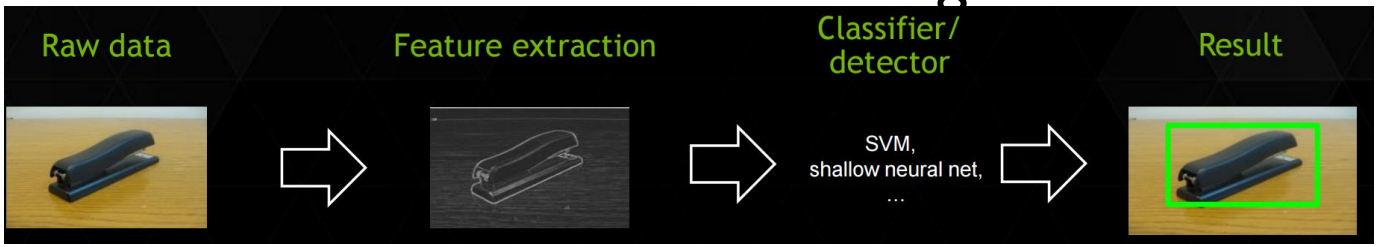


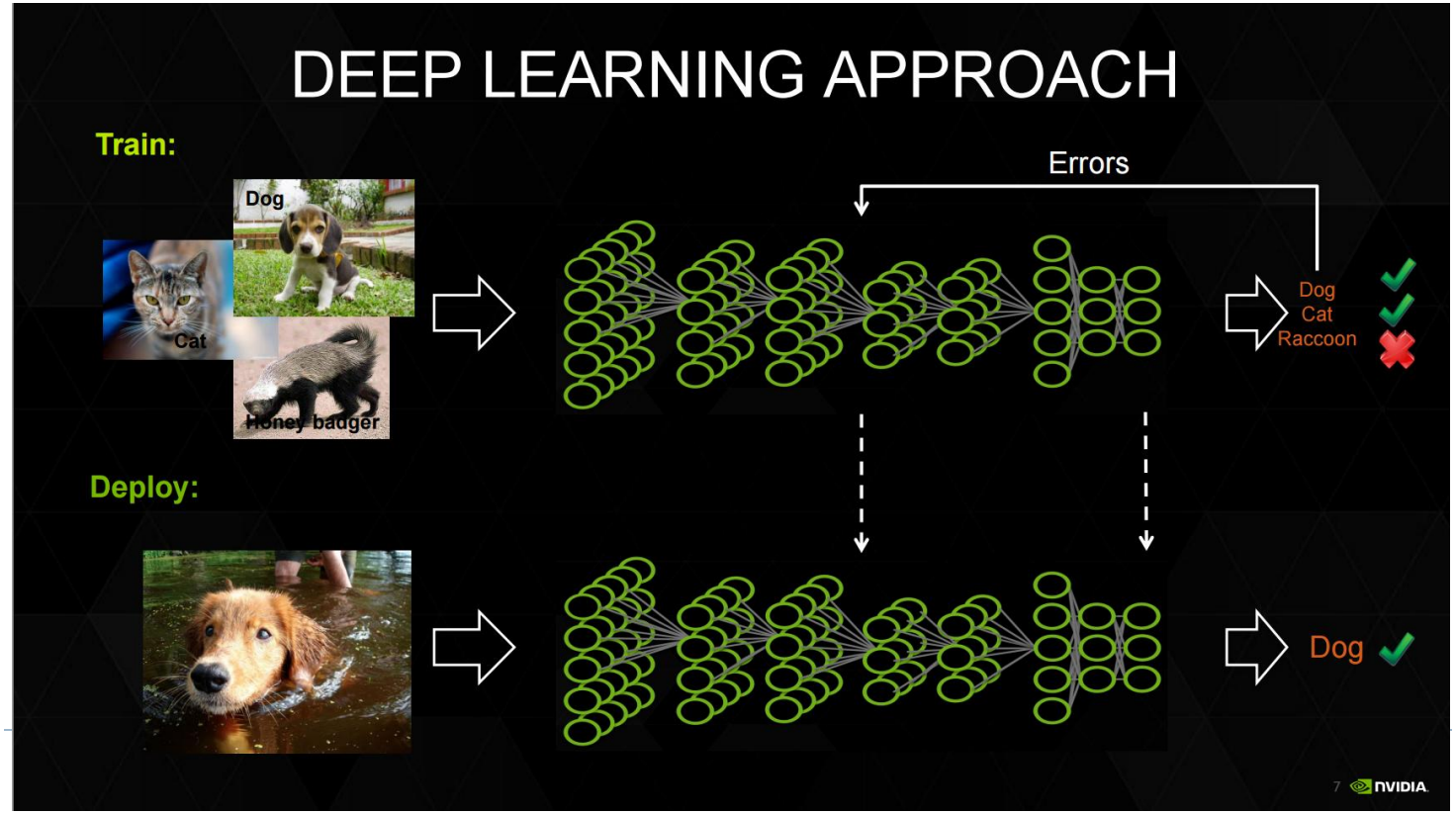
  - Capa de reducción o pooling (reducir nº parámetros a los más comunes)
  - Capa clasificadora totalmente conectada que ofrece el resultado



En la práctica, supone una diferencia

- Método tradicional de machine learning



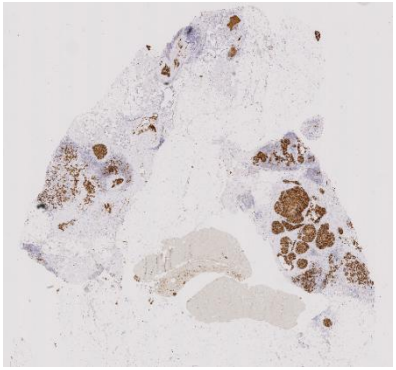


Software disponible

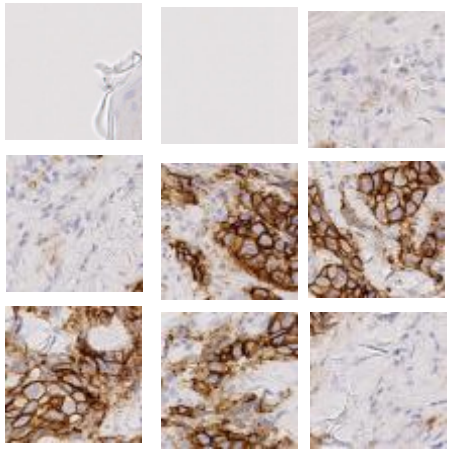
- El más popular es CAFFE (de Berkeley)
- http://caffe.berkeleyvision.org/
- Puede usarse : desde línea de comandos, Matlab o Python

Adorable! Let's classify it!

Deep Learning en Inmunohistoquímica



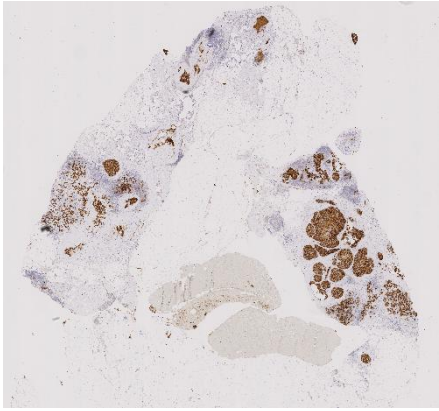
Whole Slide Images (**WSI**)



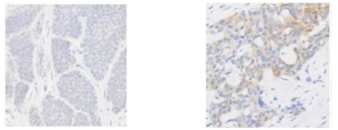
68x68 patches

Score the entire **WSI**

Class 3

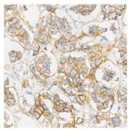


Determine the % of patches per class ➔

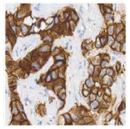


45 %

35 %



5%

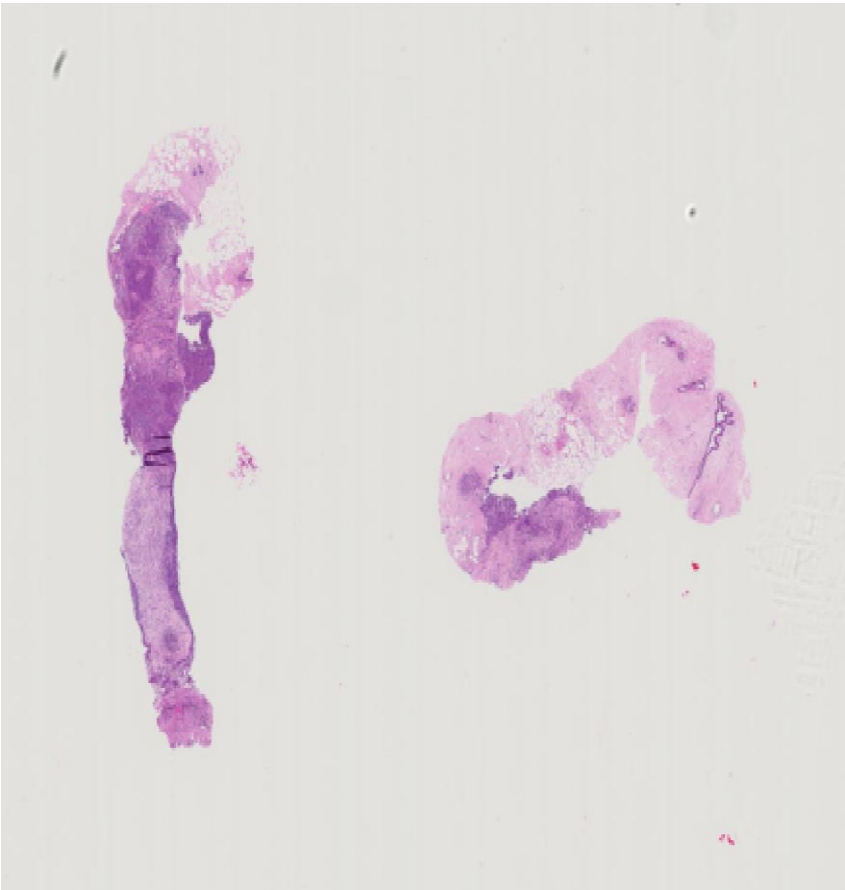


15 %

- <u>Heterogenous Tumour</u>
- From Class 3 to Class 0
  - Max % and % > 10

Estudio de correlación

Case Number	Ground Truth	Path_1	Path_2	Path_3	Algort.1	Algort.2	Algort. DL
20	3	3	3	3	3	3	3

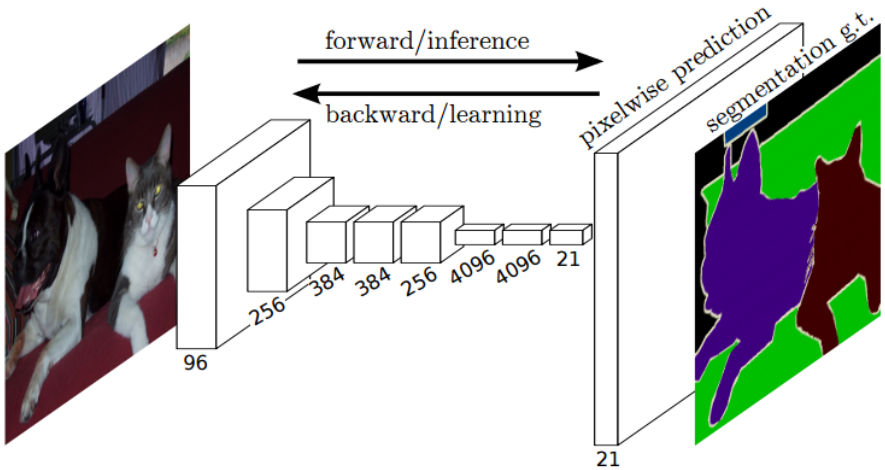


Estudio de correlación

Case Number	Ground Truth	Path_1	Path_2	Path_3	Algort.1	Algort.2	Algort. DL
56	1	1	1	2	0	1	1

Podemos usar DL para segmentar imágenes

- Se entrena en píxeles:



- # de salidas = # de píxeles
- La función de pérdida se define para la segmentación real (groundtruth).

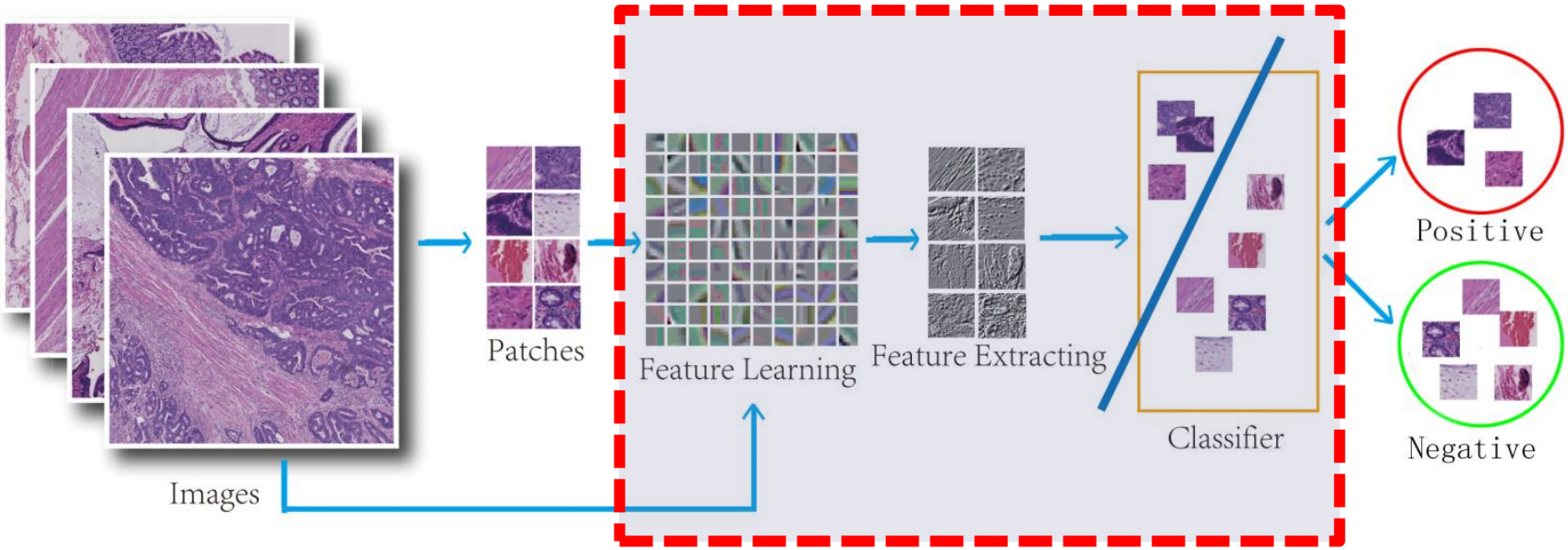
DEEP LEARNING OF FEATURE REPRESENTATION WITH MULTIPLE INSTANCE LEARNING FOR MEDICAL IMAGE ANALYSIS

Yan Xu[1,2], *Tao Mo*[2,3], *Qiwei Feng*[2,4], *Peilin Zhong*[2,4], *Maode Lai*[5], *Eric I-Chao Chang*[2*]

- **Ejemplo: Buscar imágenes con cáncer**

Segmentación de imágenes con cáncer

- 132 pacientes. Imágenes 10.000 x 10.000 píxeles → parches de 200 x 200 píxeles con superposición (9.801 parches en cada imagen). 4.500 imágenes para entrenamiento y 4.500 para test. Tras 4,6 horas….



CNN

97.30% correcto

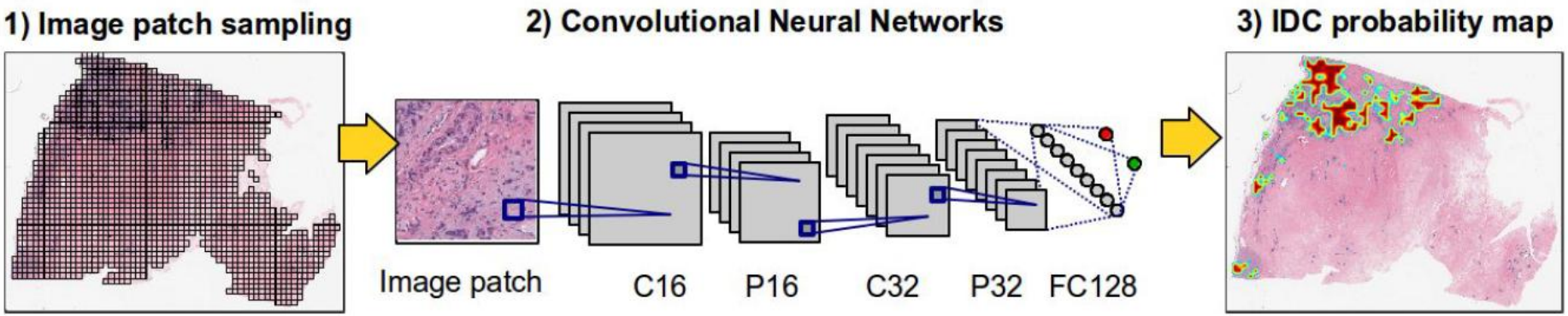
Revisión de la literatura

- Sensibilidad 79.60%, Especificidad: 88.86%

Proc. of SPIE Vol. 9041, 2014

Automatic detection of invasive ductal carcinoma in whole slide images with Convolutional Neural Networks

Angel Cruz-Roa[a], Ajay Basavanhally[b], Fabio González[a], Hannah Gilmore[c], Michael Feldman[d], Shridar Ganesan[e], Natalie Shih[d], John Tomaszewski[f] and Anant Madabhushi[g]



a Universidad Nacional de Colombia, Bogotá, Colombia

2017 PATHOLOGY VISIONS

DIGITAL PATHOLOGY:
THE POWER OF PIXELS

OCTOBER 1 - 3 | SAN DIEGO, CA

#PATHVISIONS17

Cloud-Based GPU Computing in Support of Real-time Differential Diagnosis Generation from WSI Data Using Deep Learning

Ulysses J. Balis, MD, FCAP, FASCP, FAIMBE
Professor of Pathology and Bioinformatics
Division of Pathology Informatics
Department of Pathology
University of Michigan

ulysses@med.umich.edu

Exemplar One: Finding Glomeruli

- Goal: Use the smallest possible number of cases, along with the fewest essential structural classifiers to assemble a generalizable computational pipeline that can find all glomeruli, across disparate cases
- Preliminary Data: twenty-seven feature vectors are adequate to complete this goal

Contribution of the Most Important Variables

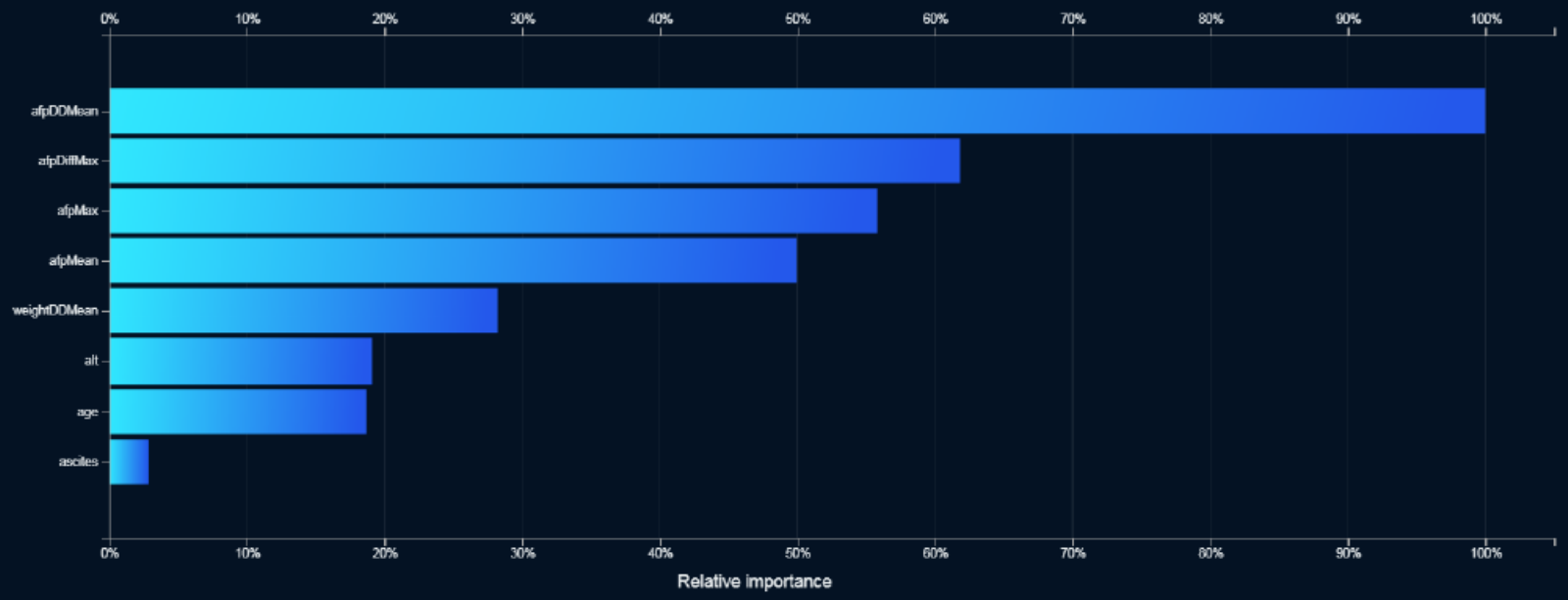
DataRobot Data Models Insights Jupyter Repository HCC Prediction

RandomForest Classifier (Gini)
BP26 M71
63.81% Sample Size, Top Eight

Sort by Feature Importance Export

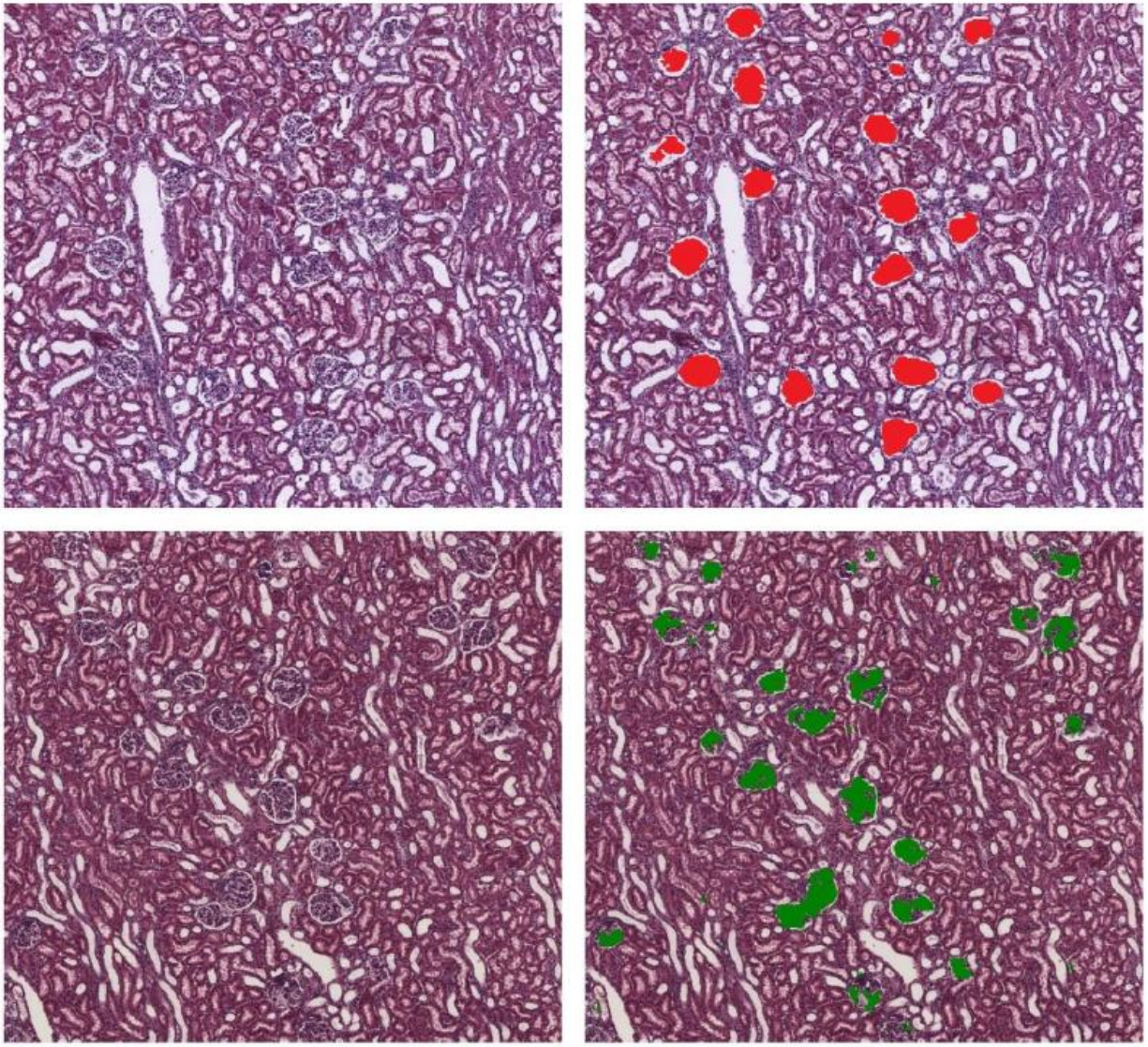
Workers: 004

DataRobot is Thinking



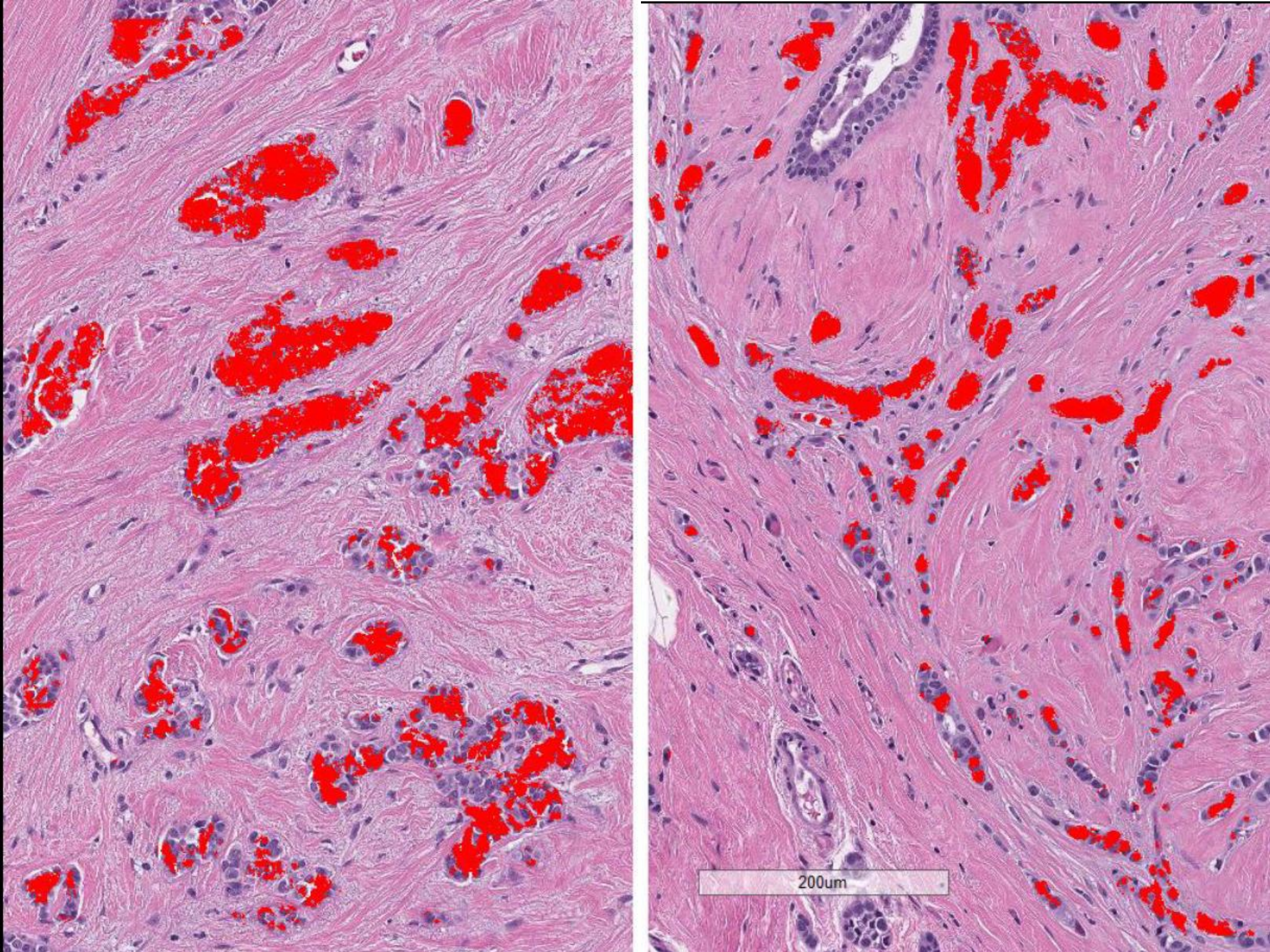
HCClongitudinal_combined_single row.csv
Total features: 47, Datapoints: 420
Target: HCC, Metric: LogLoss

Cross- Case Validation



Exemplar Two: Classifying Breast Cancer

- Goal: Use the smallest possible number of cases, along with the fewest essential structural classifiers to assemble a generalizable computational pipeline that generates a match to the most likely equivalent diagnostic classes from a library of pre-assembled vectors
- Preliminary Data: forty-five feature vectors are adequate to complete this goal


200um

A Web-based, Open-source Data Management Platform Supporting Visualization and Analysis Plugins for Deep Learning in Diagnostic Pathology

C. Charles Law, Ph.D.
Beverly E. Faulkner-Jones, M.D., Ph.D.
Seymour Rosen, M.D.
Preston Law B.S.

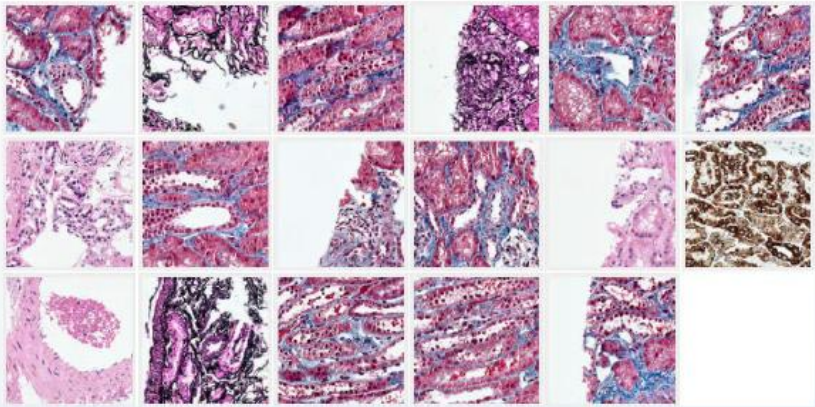
Overview

- Three applications of deep learning on pathology whole slide images.
  - Serial section alignment: Stain invariant feature detection in renal biopsies.
  - Triage: Melanoma cell detection in lymph node biopsies.
  - Diagnosis grading: Endometrial gland segmentation for EIN lesion analysis
- Open source data management application
  - Girder / Resonant ecosystem
  - Making it easier to train deep neural networks on large format images.

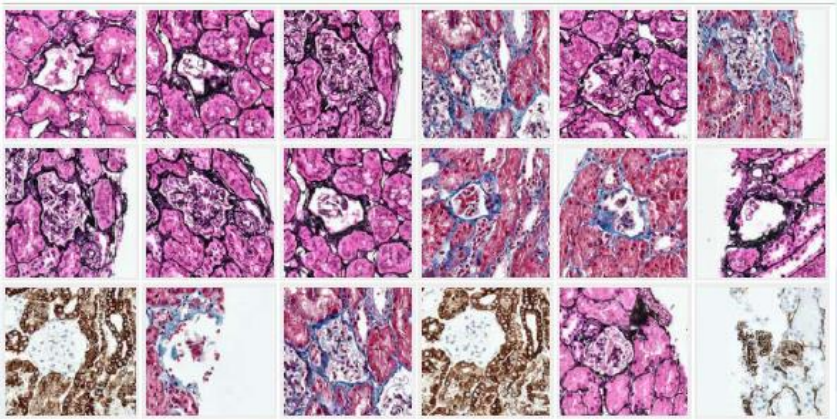
Kitware

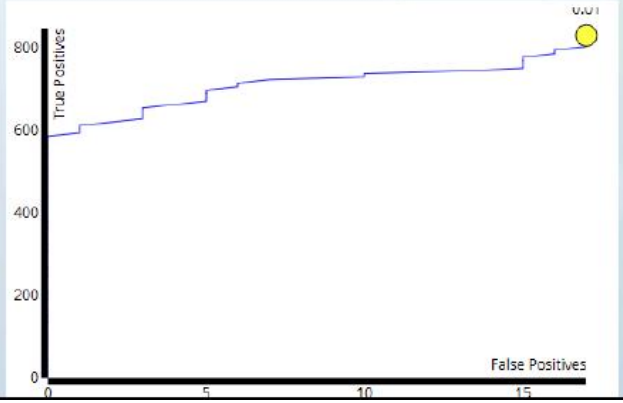
890 Test Glomeruli Detected With 17 False Positives

17 False Positives



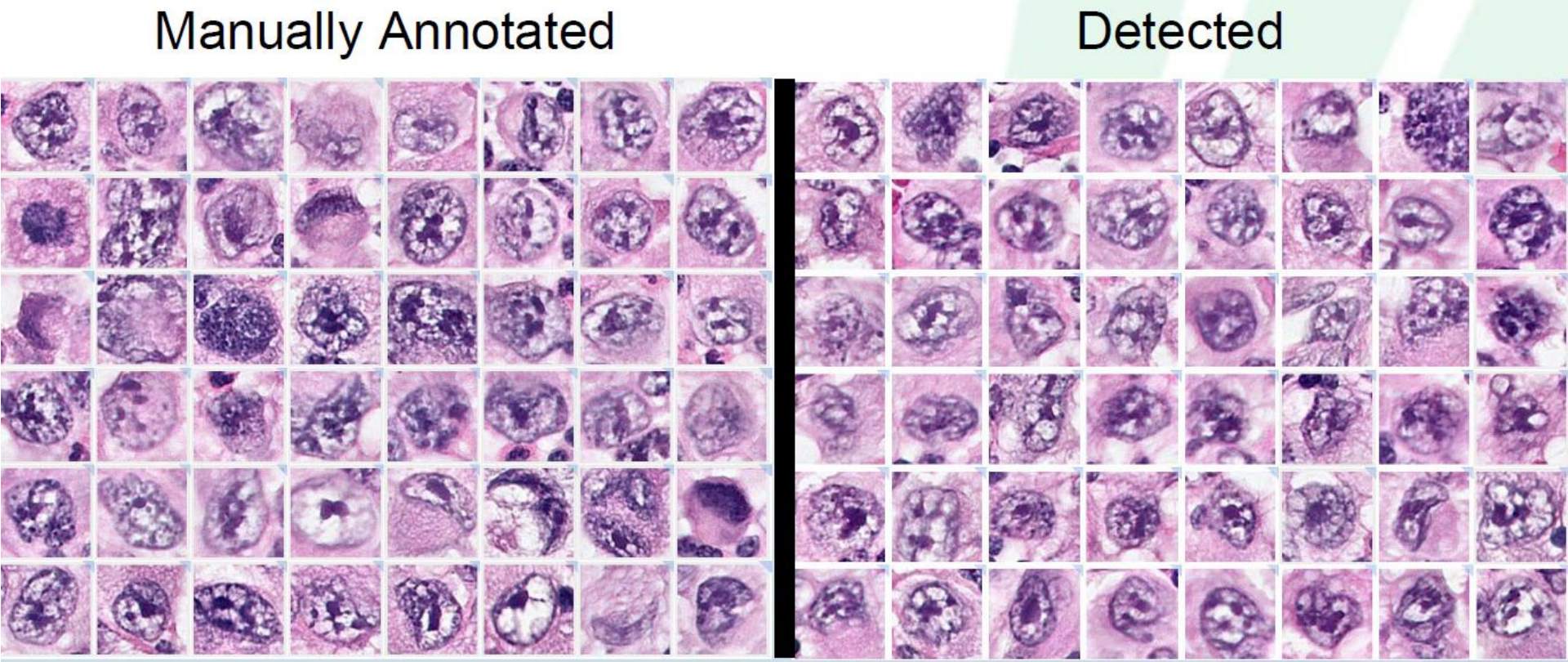
18 False Negatives





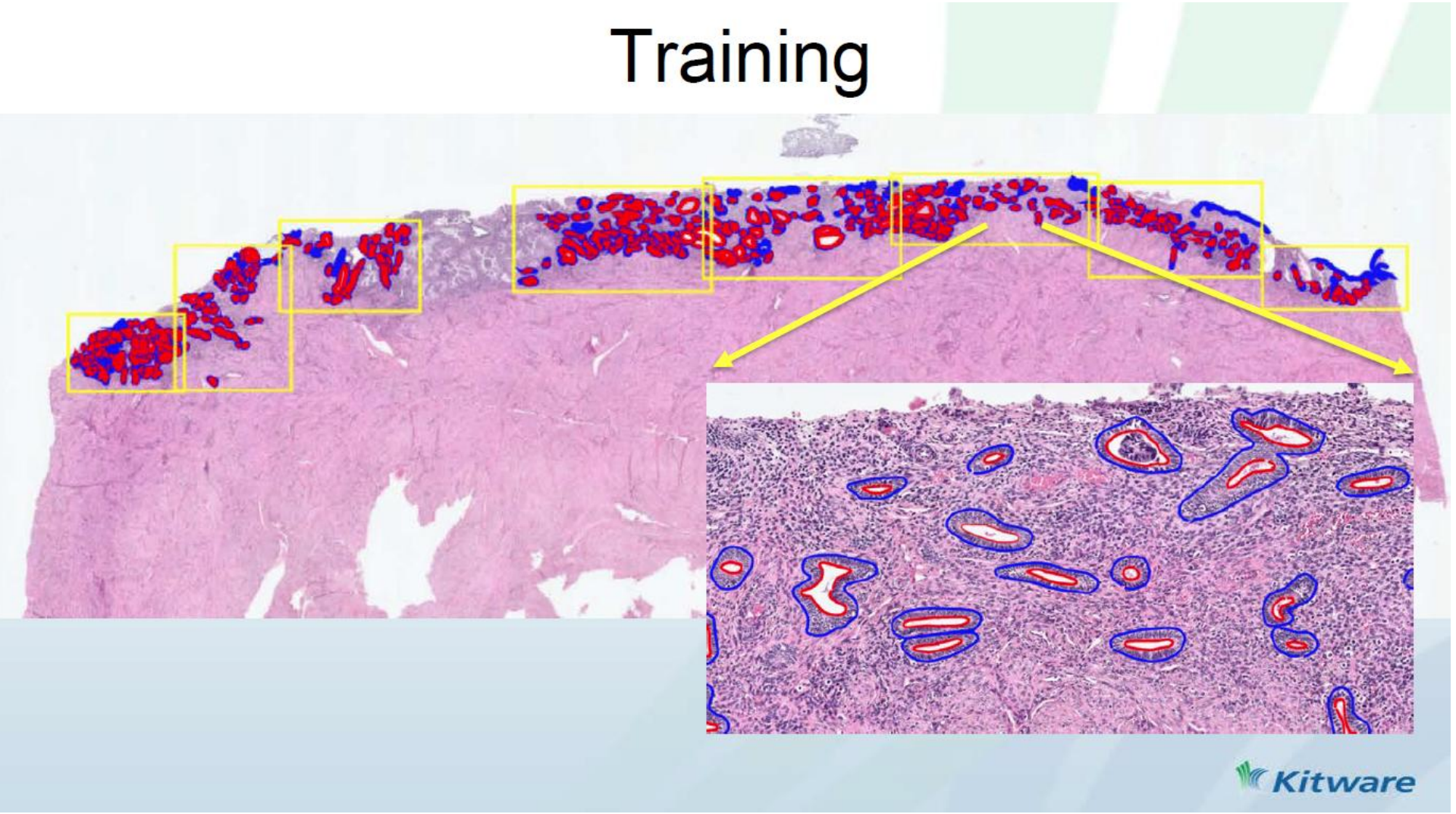
Kitware

Melanoma

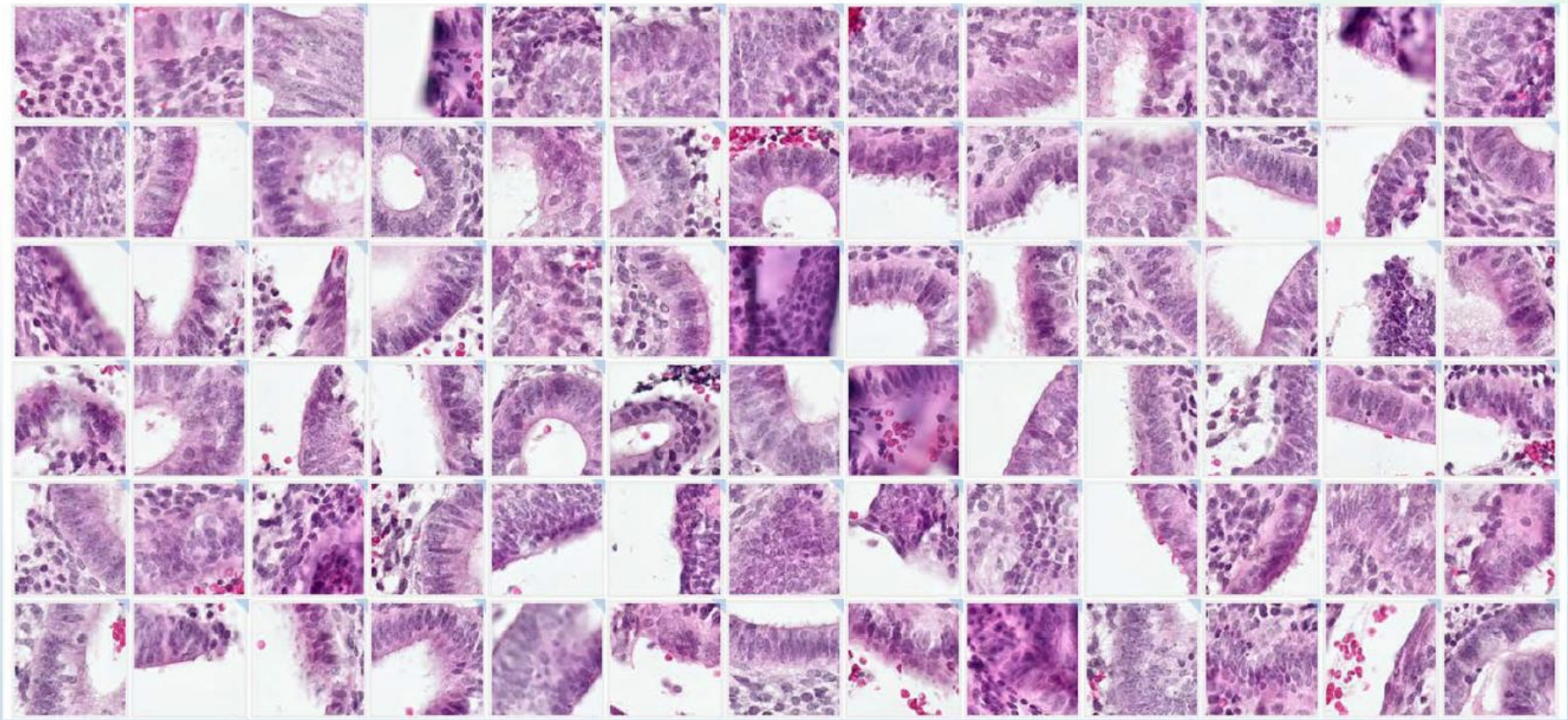


Kitware

Endometrio



Positive Detections



Kitware

www.mirror.co.uk/tech/googles-artificial-intelligence-can-diagnose-9975987

Aplicaciones | Google Calendar | Visilab Viewer | Serendipia | SEAP | Junta de Andalucía | Correo Corporativo :: | AITT-Telemedicina | Biomarker Point, el pu | 3DH CaseCenter | Ot

Mirror

NEWS ▾ POLITICS SPORT ▾ FOOTBALL CELEBS ▾ TV & FILM WEIRD NEWS MORE ▾

Technology ▸ Cancer

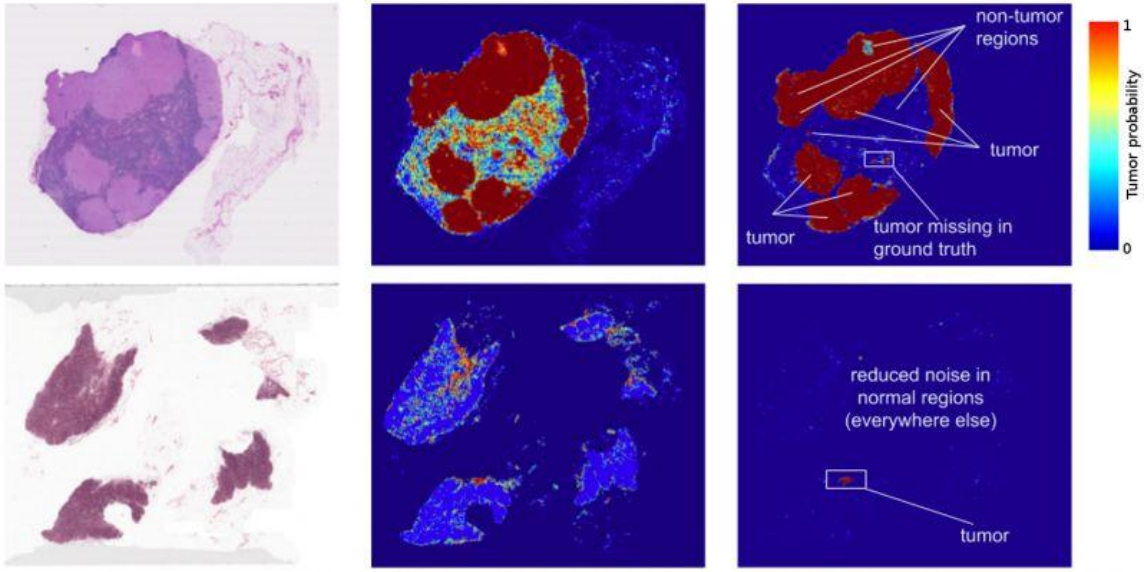
Google's artificial intelligence can diagnose cancer faster than human doctors

The system is able to scan samples to determine whether or not tissues are cancerous

SHARE COMMENTS

BY JEFF PARSONS
16:00, 6 MAR 2017 | UPDATED 10:57, 7 MAR 2017

TECHNOLOGY



https://drive.google.com/file/d/0B1T58bZ5vYa-QlR0QlJTa2dPWVk/view

The clinical use claims describe for the Aperio digital pathology products in the information supplied have not been cleared or approved by the U.S. FDA or are not available in the United States.

Conclusiones

- Es posible realizar de forma fiable análisis de imagen automatizado en imágenes histopatológicas → mayor objetividad

- Existen herramientas muy útiles. El reto es explorar estas nuevas tecnologías en patología digital de forma eficiente para procesar y modelar la enorme heterogeneidad que se extrae de los tejidos.

- Es necesaria la labor de patólogos realizando miles de anotaciones para validar estas herramientas

¿Preguntas?

marcial@cim.es