Deep Learning

TIMMD

Objetivos

- Google Colab
- Kaggle
- Imágenes
- Conceptos de ANN
- Keras + Tensorflow

Colab + Kaggle

https://www.kaggle.com/

https://colab.research.google.com/

Computer Vision Tasks

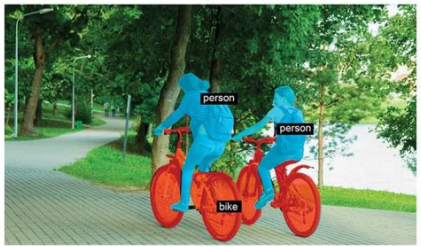
Single-label multi-class classification

- (•) Biking
- ( ) Running
- ( ) Swimming

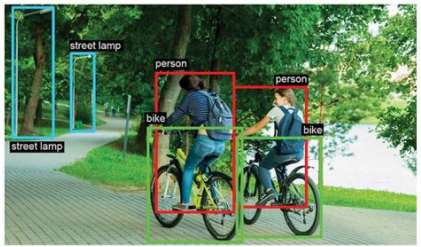
Multi-label classification

- [x] Bike
- [x] Person
- [ ] Boat
- [x] Tree
- [ ] Car
- [ ] House

Image segmentation

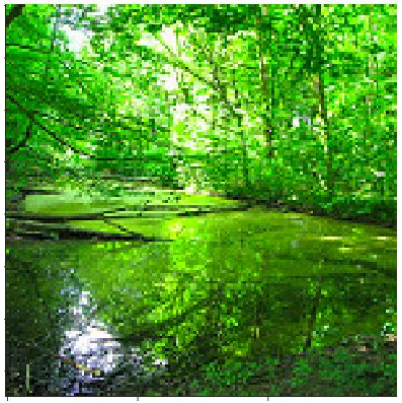


Object detection



Scenes Dataset

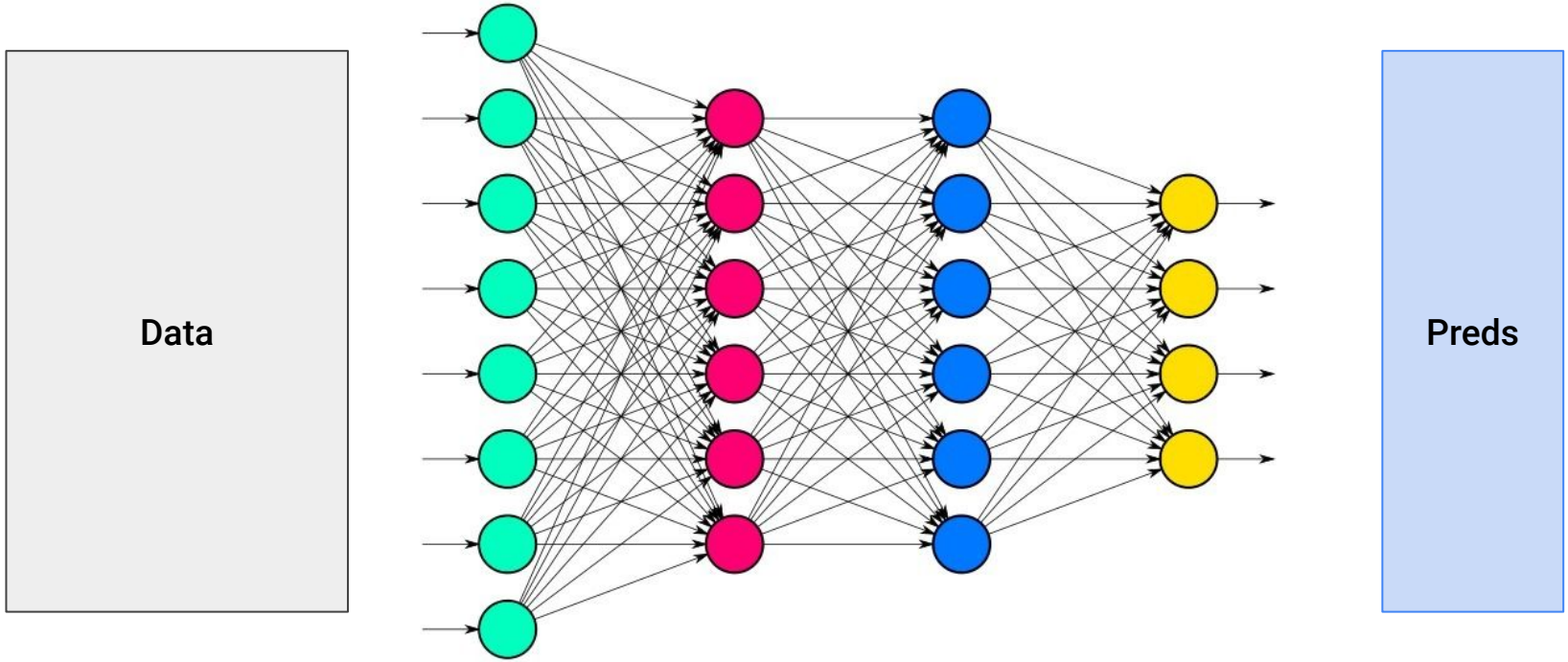
Usamos un subset de datos de 5000 observaciones y predecimos 4 de las 6 clases presentes.




ANN

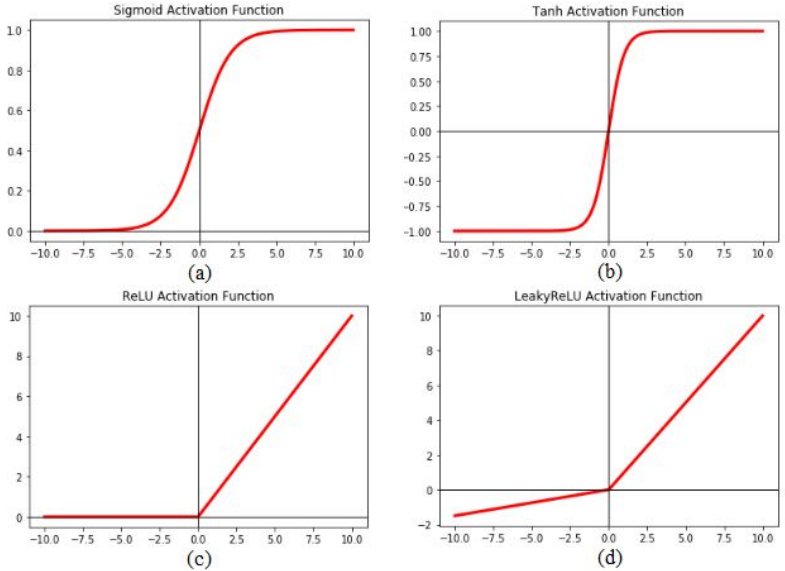
- Aproximadores universales.
- Transformación sucesiva de los datos para generar representaciones que facilitan la resolución del problema.
- Importancia de activaciones No Lineales.

Redes Neuronales Densas



Activaciones

- Le dan mayor flexibilidad a las representaciones creadas por la red en las capas ocultas.

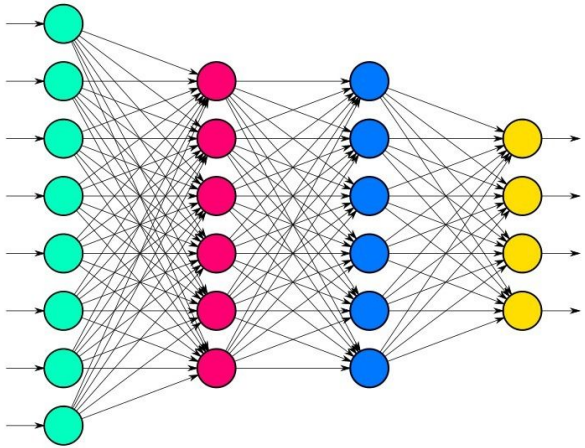


Conceptos ANN

- Datos Normalizados
- Entrenamiento - Backpropagation
- Loss Function
- Learning Rate
- Batch Size
- Epochs
- Activaciones

Entrenamiento

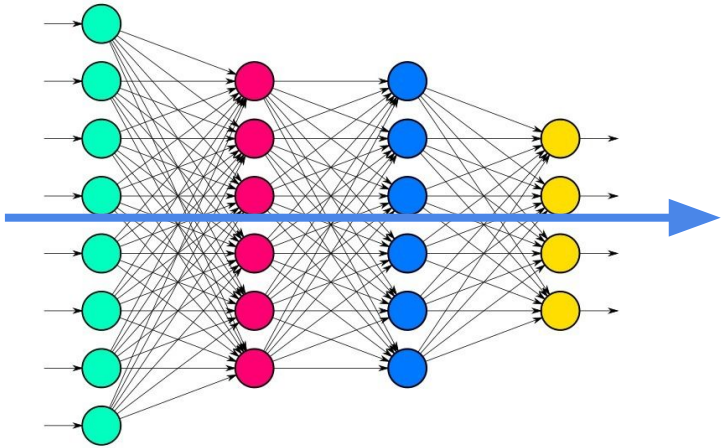
Train Data



Train Targets

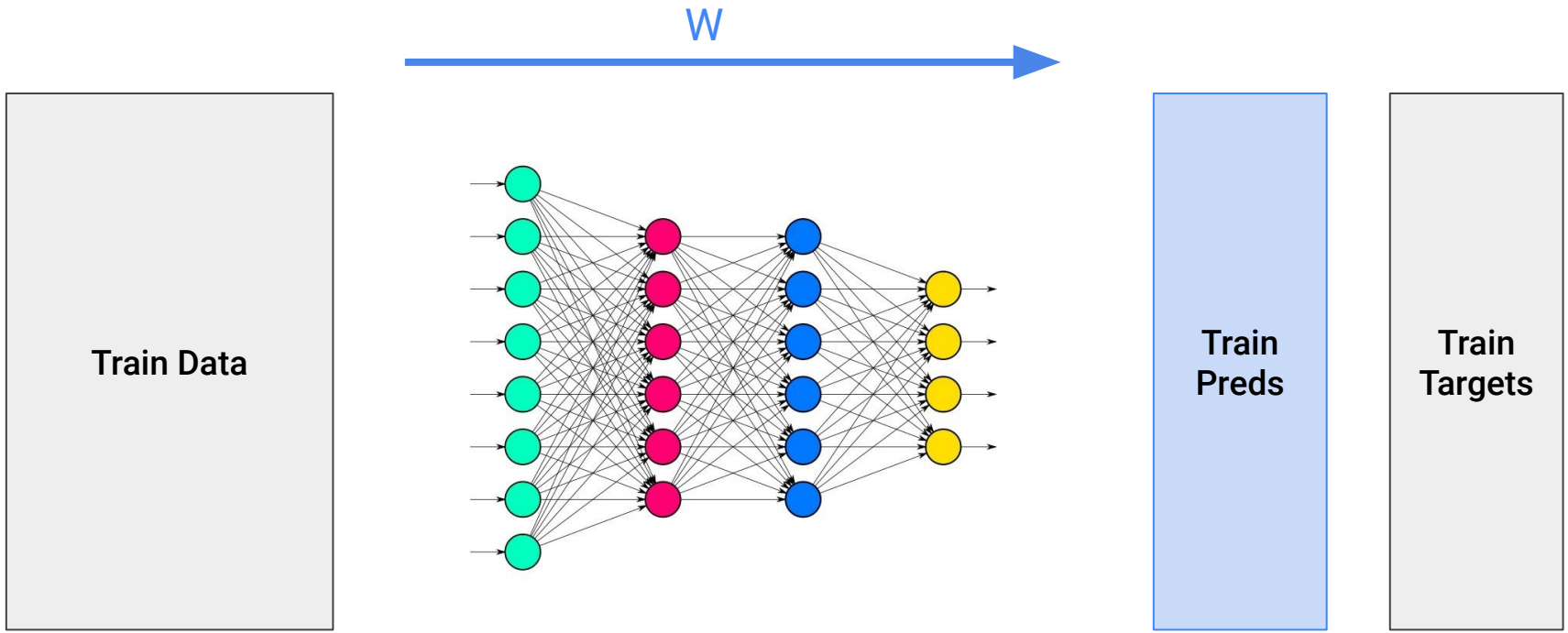
Entrenamiento

Train Data



Train Targets

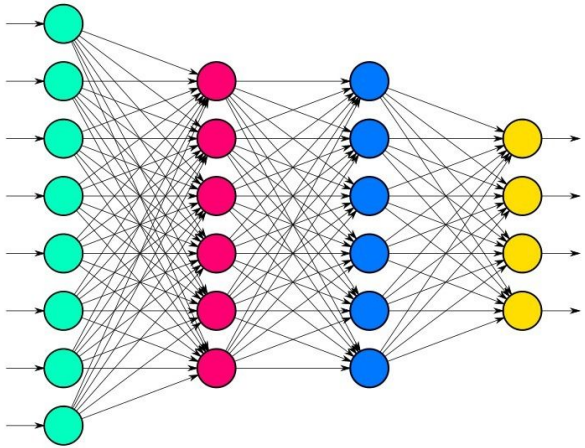
Entrenamiento



Entrenamiento

Loss

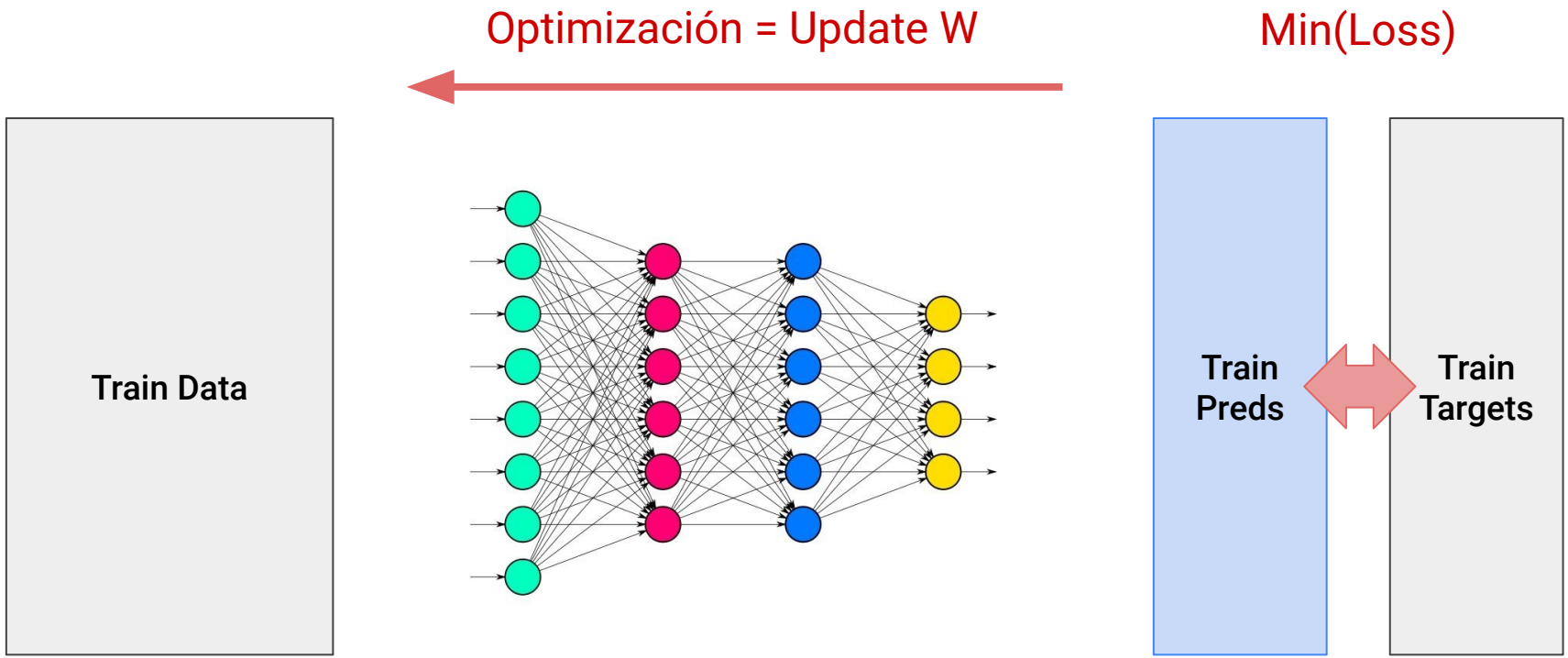
Train Data



Train Preds

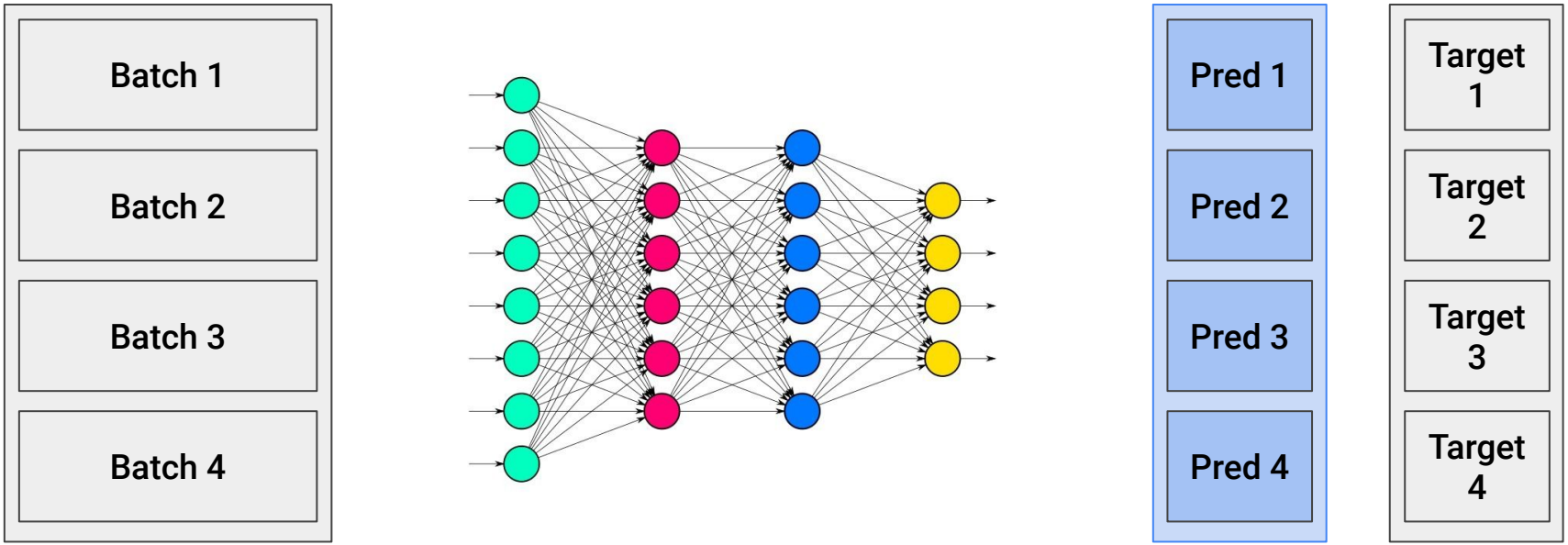
Train Targets

Entrenamiento



Entrenamiento

- Batch = Lote ~ N Observaciones para actualizar W.
- Epoch = Iteración ~ N Veces que se repite el proceso.
- Learning Rate = Magnitud de actualización de W.



Keras

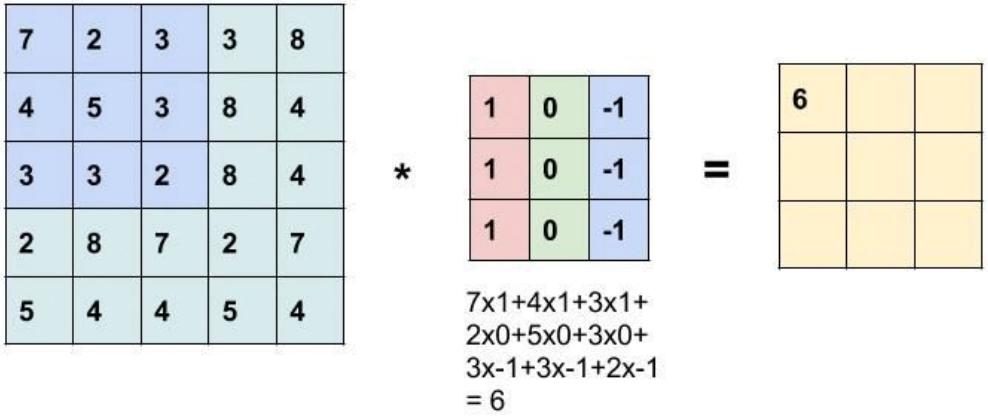
- Layers - https://keras.io/api/layers/
- Loss Function - https://keras.io/api/losses/
- Metrics - https://keras.io/api/metrics/
- Optimizers - https://keras.io/api/optimizers/

Convoluciones

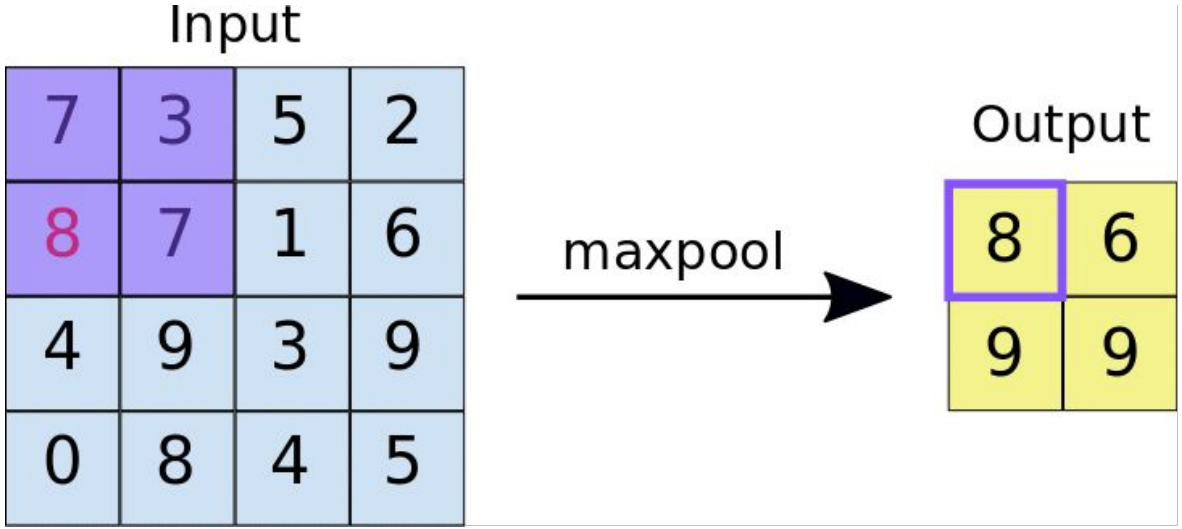
- En Imágenes tenemos una **alta** dimensionalidad (gran cantidad de pixeles) que se traducen a grandes cantidad de parámetros en DNN.
- En Imágenes hay patrones que se repiten que son invariantes a traslación.

- Buscamos una mejor manera de detectar variables en imágenes, donde es importante la espacialidad pero no la posición.

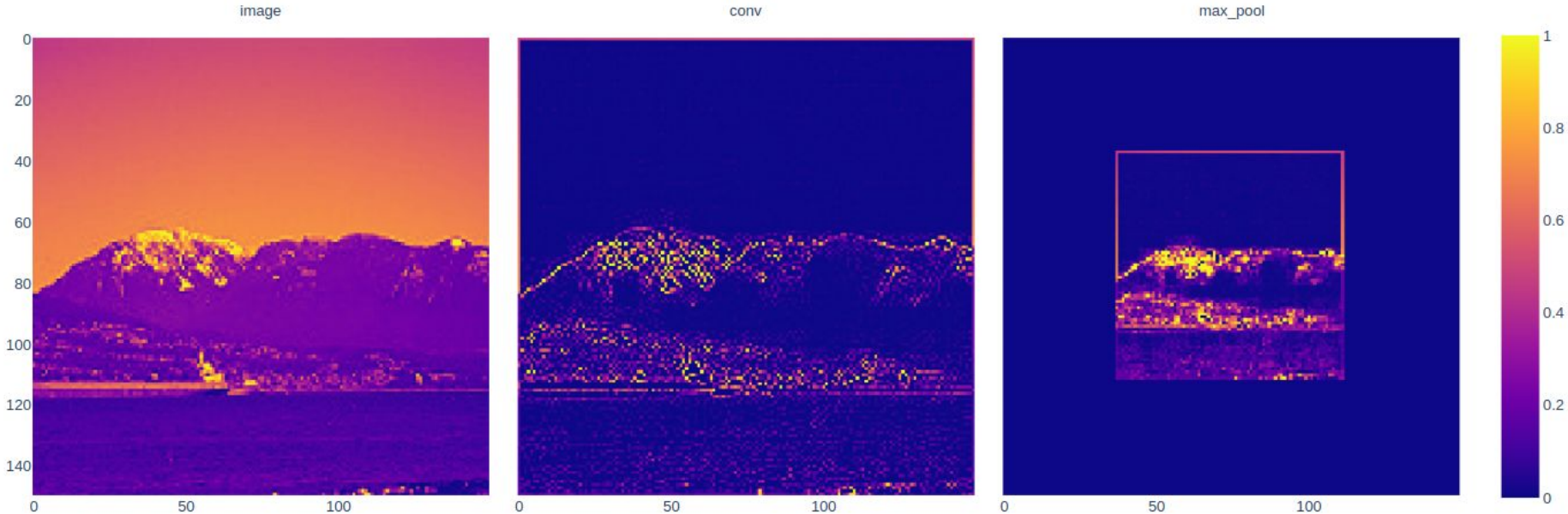
Convolución



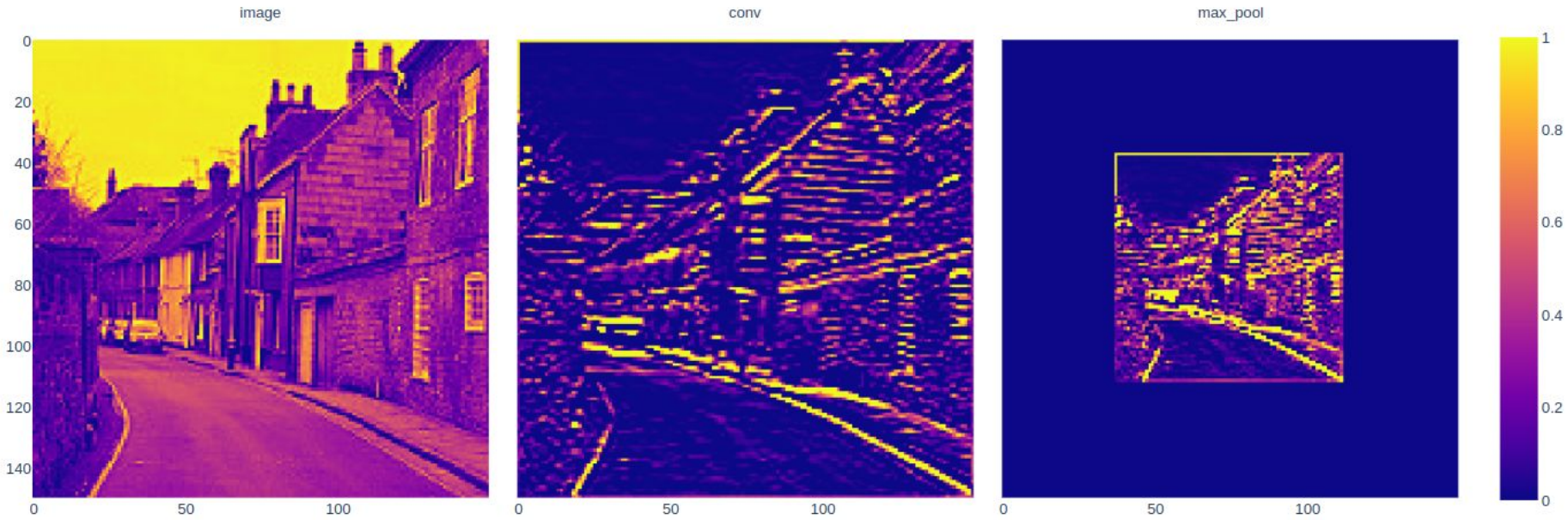
Max Pooling



Convolución + Max Pooling



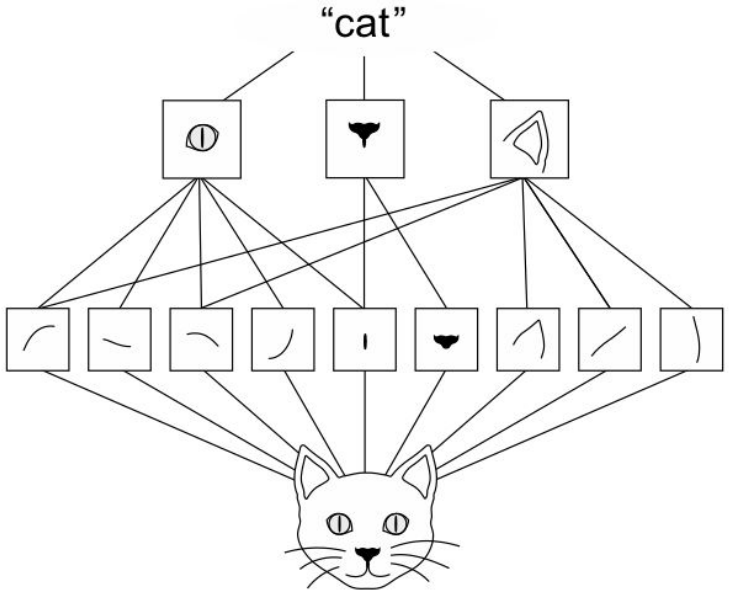
Convolución + Max Pooling

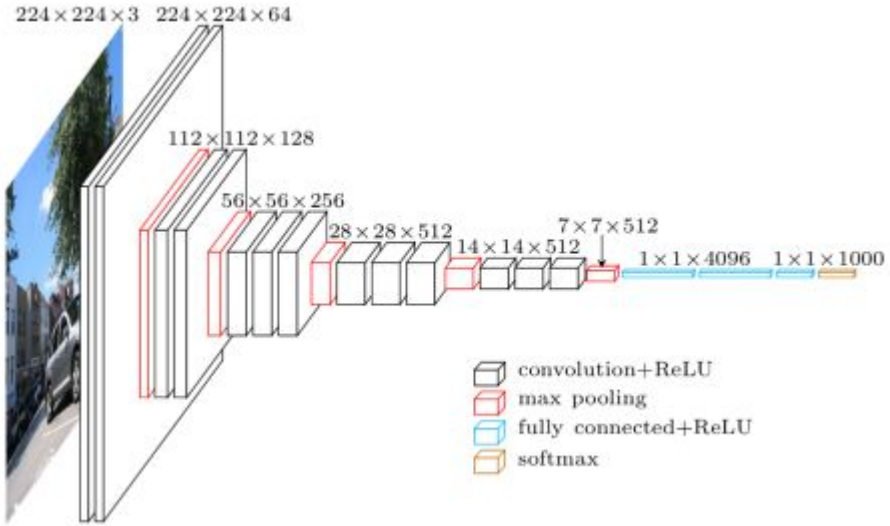


CNN

- Aplicando Convoluciones:
  - Detectamos patrones locales (las DNN aprenden patrones globales)
  - Detectamos patrones invariantes a traslación: Si el mismo patrón aparece en otro lugar de la imagen será detectado.
  - Aprendemos patrones jerárquicos.
  - Requiere menor cantidad de parámetros.

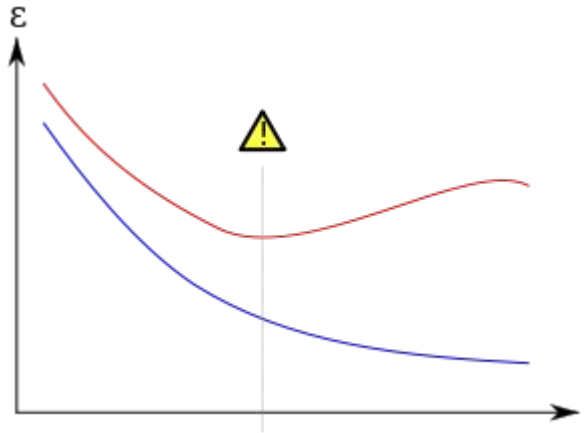
CNN


“cat”


224 × 224 × 3
224 × 224 × 64
112 × 112 × 128
56 × 56 × 256
28 × 28 × 512
14 × 14 × 512
7 × 7 × 512
1 × 1 × 4096
1 × 1 × 1000
convolution+ReLU
max pooling
fully connected+ReLU
softmax

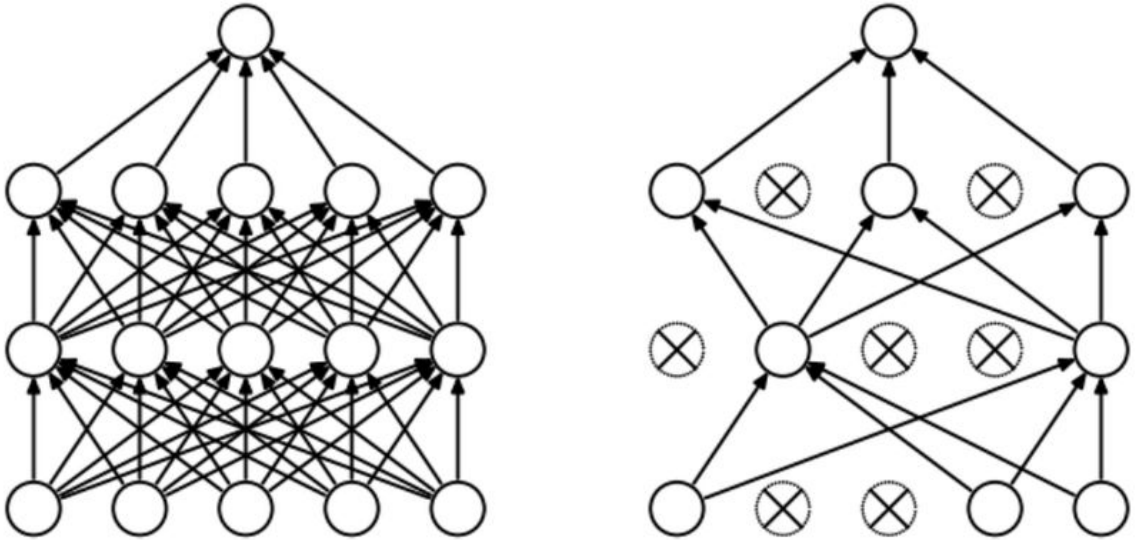
Overfitting

- Cuando la redes neuronales tienen grandes cantidades de parámetros y datos de entrenamiento limitados pueden ajustarse demasiado a esos datos, memorizando el entrenamiento.
- Tendrán menor performance en datos nunca vistos.
- Soluciones:
  - Regularización (Dropout)
  - Data Augmentation
  - Early Stopping



Regularization - Dropout

- Se apagan neuronas de una capa de manera aleatoria durante el **entrenamiento**.

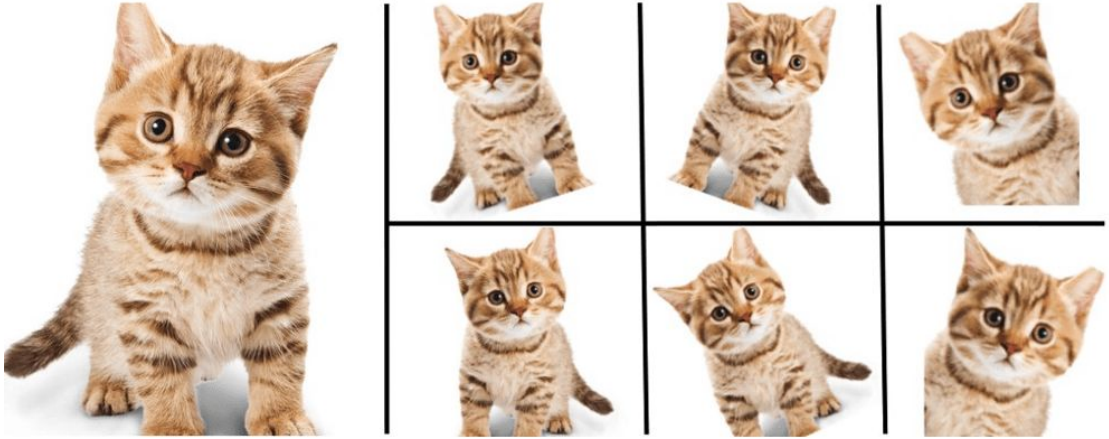


(a) Standard Neural Net

(b) After applying dropout.

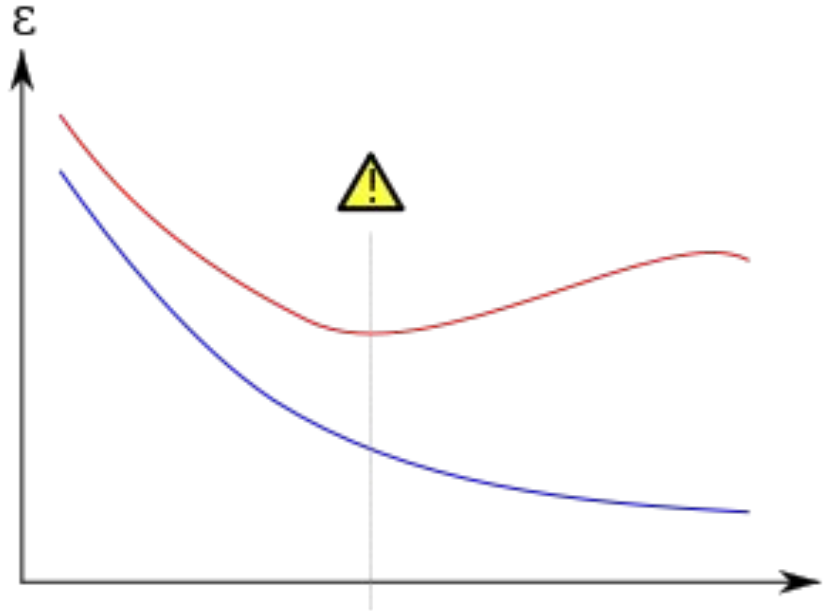
Data Augmentation

- En imágenes, hacer cambios aleatorios no altera su sentido. Podemos introducir cambios **aleatorios** en el input del modelo durante el **entrenamiento** para que en cada iteración el modelo nunca vea 2 veces la misma imagen.



Early Stopping

- Frenamos el entrenamiento en la iteración de mejor resultado en **validación**.

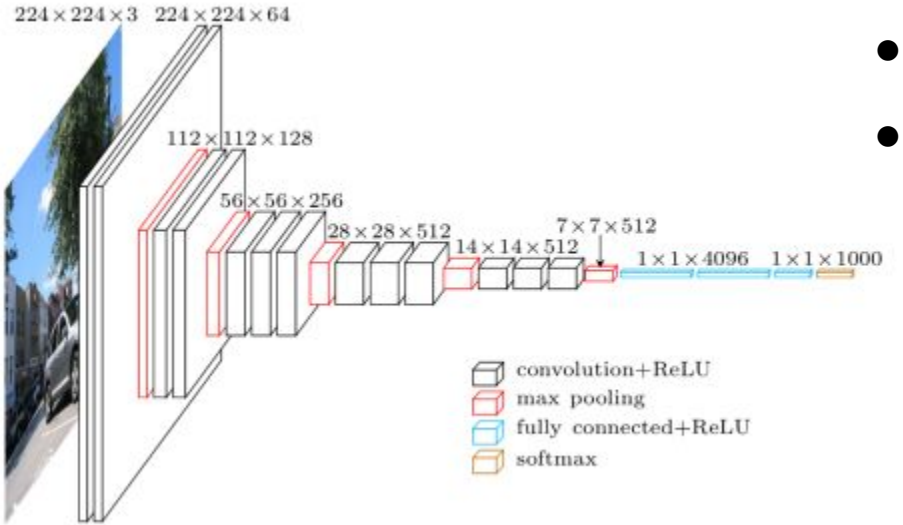


Medimos el error de entrenamiento (azul) y de validación (rojo) durante el entrenamiento del modelo.

Frenamos el entrenamiento cuando el error de validación deja de mejorar.

Transfer Learning

- Es el método de aprovechar los parámetros entrenados de una red y reutilizarlos en un problema diferente, cambiando las últimas capas.



- VGG16
- ImageNet

Transfer Learning

- Podemos aprovechar los parámetros de las capas convolucionales de redes ya entrenadas.
- Estas capas ya deberían captar características y patrones básicos en las imágenes.