

Desafío 9

**Clasificación de radiografías COVID
usando una CNN**

Iyán Álvarez

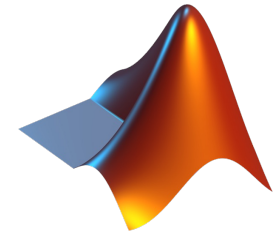


Trabajo realizado

Un script en Matlab.

El script “ClasificacionRadiografias_main” carga el conjunto de imágenes de las radiografías, las muestrea, las separa en el conjunto de train y validación y se entrena una CNN. Finalmente, se reporta el accuracy obtenido.

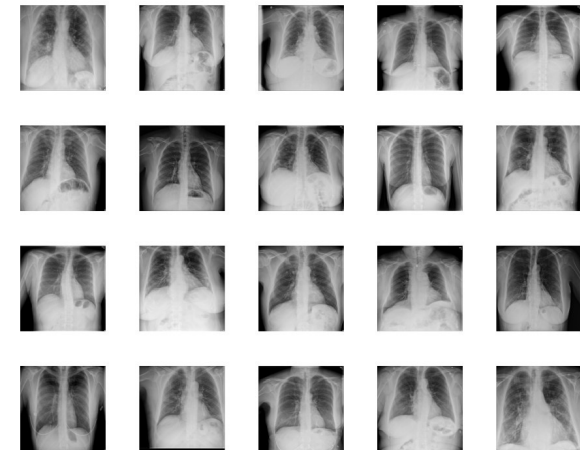
 ClasifiacionRadiografias_main.m



Datos

COVIDGR es un conjunto de imágenes de rayos X para asistir en el diagnóstico de la enfermedad COVID-19, construido con la estrecha colaboración de expertos radiólogos en España.

La primera versión del dataset COVIDGR contiene 426 radiografías positivas y 426 radiografías negativas (852 en total), realizadas en hospitales de Granada (España).



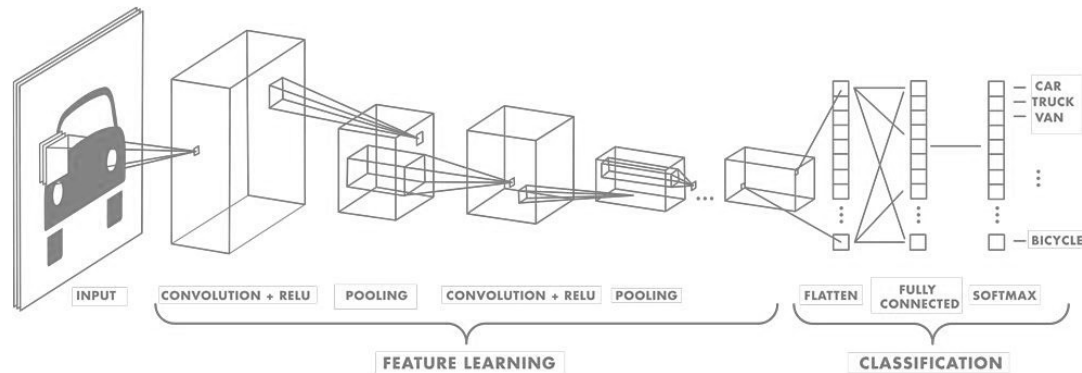
Definición de la CNN

Una CNN consta de una capa de entrada, una capa de salida y varias capas ocultas entre ambas. Estas capas realizan operaciones que modifican los datos, con el propósito de comprender sus características particulares.

Las 3 capas más comunes son:

- Convolución
- Activación o ReLU
- Agrupación o Pooling

Además utilizamos capas de normalización por lotes para acelerar el entrenamiento y reducir la sensibilidad a la iniciación de la red.



```
layers = [  
    imageInputLayer([128 128 1])  
  
    convolution2dLayer(3,8,'Padding','same')  
    batchNormalizationLayer  
    reluLayer  
    maxPooling2dLayer(2,'Stride',2)  
    convolution2dLayer(3,16,'Padding','same')  
    batchNormalizationLayer  
    reluLayer  
    maxPooling2dLayer(2,'Stride',2)  
    convolution2dLayer(3,32,'Padding','same')  
    batchNormalizationLayer  
    reluLayer  
    maxPooling2dLayer(2,'Stride',2)  
    convolution2dLayer(3,64,'Padding','same')  
    batchNormalizationLayer  
    reluLayer  
    maxPooling2dLayer(2,'Stride',2)  
    convolution2dLayer(3,128,'Padding','same')  
    batchNormalizationLayer  
    reluLayer  
    maxPooling2dLayer(2,'Stride',2)  
    convolution2dLayer(3,256,'Padding','same')  
    batchNormalizationLayer  
    reluLayer  
    maxPooling2dLayer(2,'Stride',2)  
  
    fullyConnectedLayer(2)  
    softmaxLayer  
    classificationLayer  
];
```

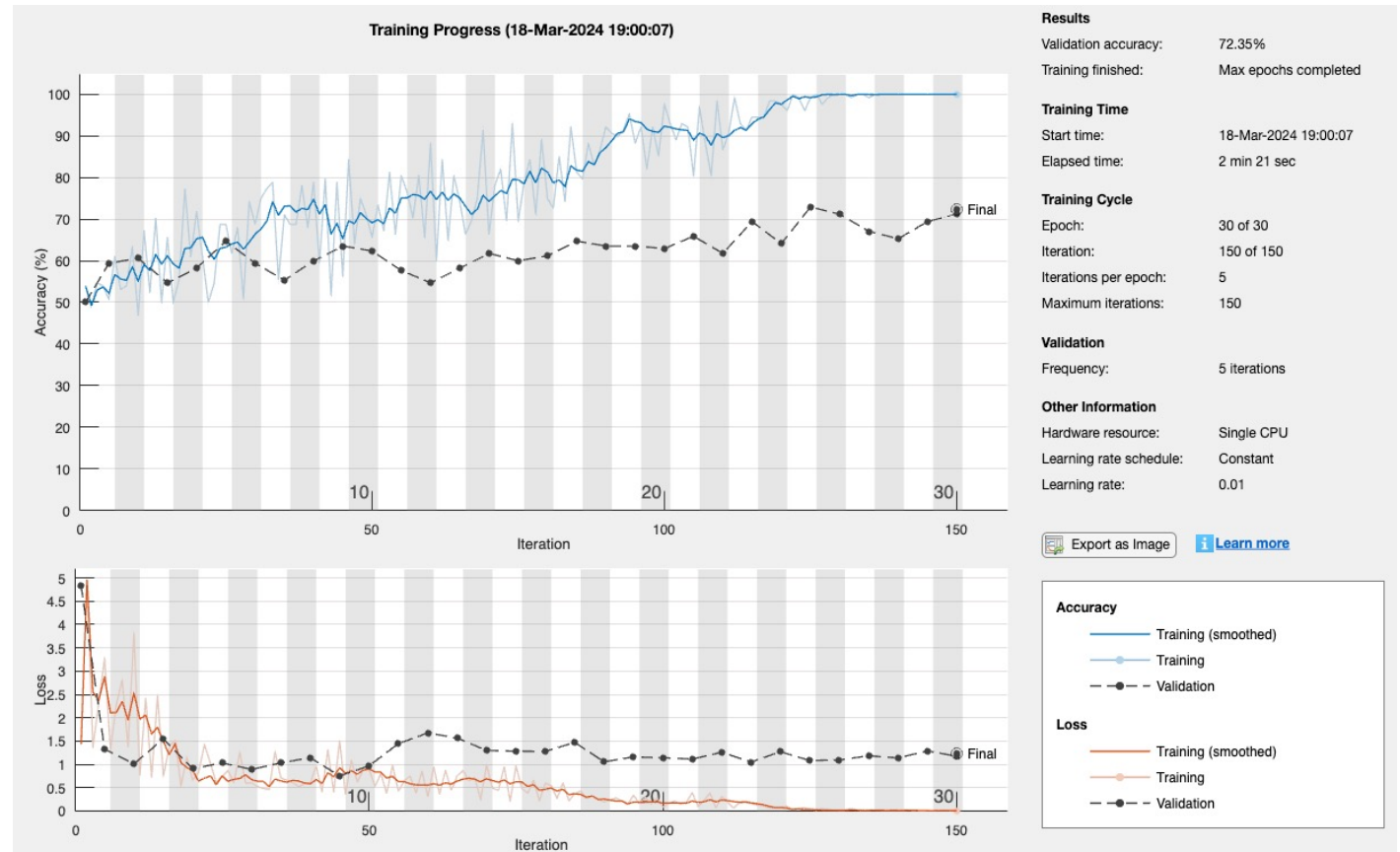
Entrenamiento de la CNN

El entrenamiento de la CNN varía mucho en función de la inicialización de la red.

Se ha utilizado el 80% de las datos para train y el 20% restante para validación.

La accuracy promedio obtenida tras varias ejecuciones es de 71.5% y está en el rango [0.63, 0.76].

```
options = trainingOptions('sgdm', ...  
    'InitialLearnRate', 0.01, ...  
    'MaxEpochs', 30, ...  
    'Shuffle','every-epoch', ...  
    'ValidationData', imds_validacion, ...  
    'ValidationFrequency', 5, ...  
    'Verbose', false, ...  
    'Plots', 'training-progress');
```



Conclusión personal

La toma de decisión cruciales, por ejemplo en la medicina, no deben ser sustituidas por sistemas que no cumplan unos estándares de fiabilidad y hayan sido investigados y testeados en profundidad.

Se deben establecer regulaciones y estándares claros y estrictos para el despliegue de estos sistemas.

En el caso de ser implementados estos sistemas, es importante que las personas conozcan que están siendo utilizados en su diagnóstico y las implicaciones que puede tener.