# Datasets used

[**chest\_xray\_pneumonia**](https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia)

test [normal : 234, pneumonia : 390]

train [normal : 1341, pneumonia : 3875]

val [normal : 8, pneumonia : 8]

[**Open-i service of the National Library of Medicine**](https://openi.nlm.nih.gov/faq#collection)

# Big Dataset

## Skin Lesion Classification

<https://challenge2018.isic-archive.com>

[The ISIC Archive](http://ww.isic-archive.com/) contains the largest publicly available collection of quality controlled dermoscopic images of skin lesions.

## EYEPACS

<http://www.eyepacs.com/data-analysis>

Our retinal image database is comprised of over 3 million retinal images of diverse populations with various degrees of diabetic retinopathy. Because of the diversity, it helps algorithms recognize the different retinas that exist in real world settings.

# Small Dataset

## STARE Dataset

<http://cecas.clemson.edu/~ahoover/stare/>

The STARE (STructured Analysis of the Retina) Project

This research concerns a system to automatically diagnose diseases of the human eye. The system takes as input information observable in a retinal image. This information is formulated to mimic the findings that an ophthalmologist would note during a clinical examination. The main output of the system is a diagnosis formulated to mimic the conclusion that an ophthalmologist would reach about the health of the subject.

Our approach breaks the problem into two components. The first component concerns automatically processing a retinal image to denote the important findings. The second component concerns automatically reasoning about the findings to determine a diagnosis. Additional outputs include detailed measurements of the anatomical structures and lesions visible in the retinal image. These measurements are useful for tracking disease severity and the evaluation of treatment progress over time. By collecting a database of measurements for a large number of people, the STARE project could support clinical population studies and intern training.

## DRIVE Dataset

<https://www.isi.uu.nl/Research/Databases/DRIVE/>

The DRIVE database has been established to enable comparative studies on segmentation of blood vessels in retinal images. The photographs for the DRIVE database were obtained from a diabetic retinopathy screening program in The Netherlands. The screening population consisted of 400 diabetic subjects between 25-90 years of age. Forty photographs have been randomly selected, 33 do not show any sign of diabetic retinopathy and 7 show signs of mild early diabetic retinopathy. Each image has been JPEG compressed.

## Messidor

<http://www.adcis.net/en/third-party/messidor/>

MESSIDOR stands for Methods to Evaluate Segmentation and Indexing Techniques in the field of Retinal Ophthalmology (in French).

Within the scope of Diabetic Retinopathy, the primary purposes of the Messidor project is to compare and evaluate:

* Various segmentation algorithms developed for the detection of lesions present in color retinal images;
* Tools to index and manage image databases.

The 1200 eye fundus color numerical images of the posterior pole of the Messidor database were acquired by 3 ophthalmologic departments using a color video 3CCD camera mounted on a Topcon TRC NW6 non-mydriatic retinograph with a 45 degree field of view. Images were captured using 8 bits per color plane at 1440\*960, 2240\*1488 or 2304\*1536 pixels.

## VICAVR database

<http://www.varpa.es/research/ophtalmology.html>

The VICAVR database is a set of retinal images used for the computation of the A/V Ratio. The database currently includes 58 images. The images have been acquired with a TopCon non-mydriatic camera NW-100 model and are optic disc centered with a resolution of 768x584. The database includes the caliber of the vessels measured at different radii from the optic disc as well as the vessel type (artery/vein) labelled by three experts.

## OCTAGON dataset

<http://www.varpa.es/research/ophtalmology.html>

The OCTAGON dataset is a set of Angiography by Octical Coherence Tomography images (OCT-A) used to the segmentation of the Foveal Avascular Zone (FAZ). The dataset includes 144 healthy OCT-A images and 69 diabetic OCT-A images, divided into four groups, each one with 36 and about 17 OCT-A images, respectively.