

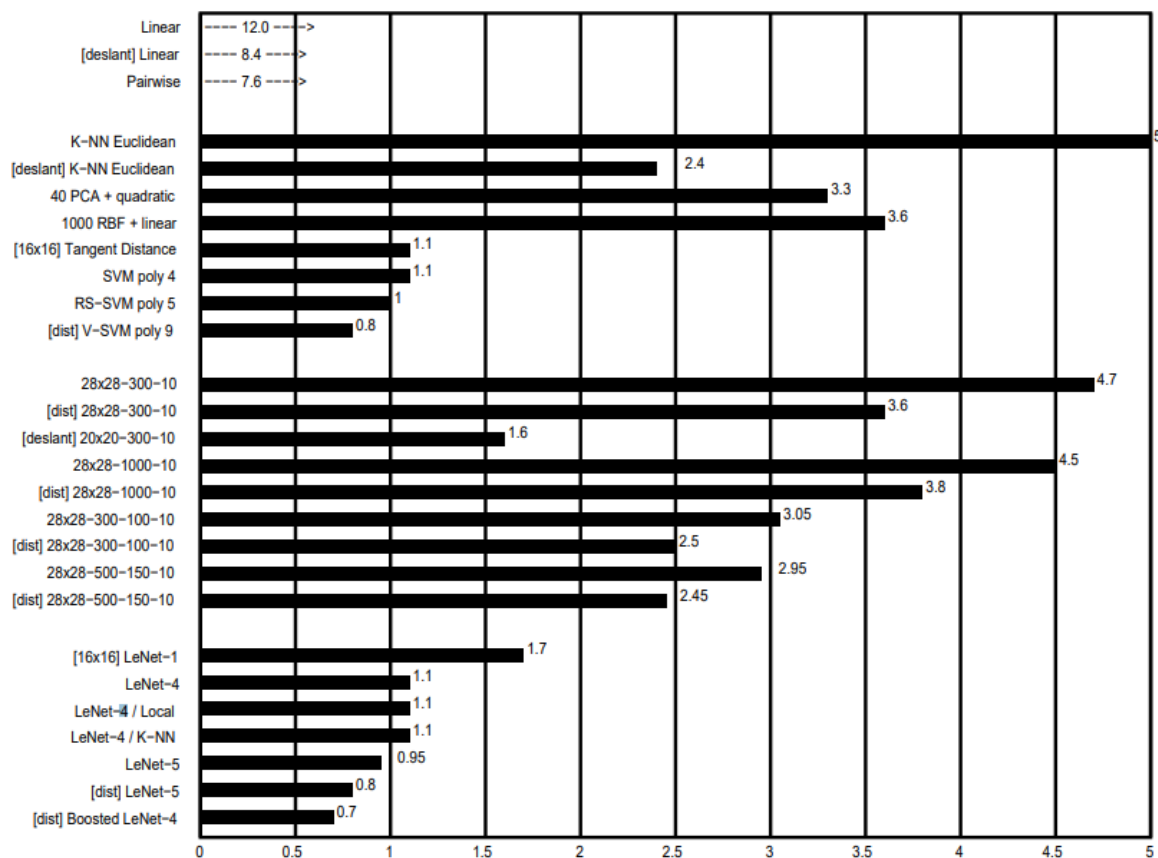
# Análisis de LeNet-5

Esta red neuronal, pionera y en gran medida responsable de la adaptación de las CNN como estándar de facto en materia de reconocimiento de imágenes, apareció en el paper Gradient-Based Learning Applied to Document Recognition, publicado en 1998.

El objetivo de este estudio era el de demostrar que la estrategia de back-propagation junto al uso de técnicas de aprendizaje por gradiente podían dar lugar a nuevas herramientas a considerar de cara a sintetizar superficies de decisión complejas que permitieran clasificar un alto número de patrones.

El estudio se enfoca en probar diversos métodos en tareas de reconocimiento numérico a partir de imágenes. Este problema ha sido históricamente tratado desde diversos enfoques, y el desencadenante de este estudio fue los notables resultados de las redes neuronales en este tipo de problemas en los años previos.

A continuación podemos observar la comparativa de los métodos probados donde el eje de las x muestra el porcentaje de error en el conjunto de test y el eje de las y enumera los diferentes métodos.



Como podemos observar en la imagen anterior los métodos basados en redes neuronales entran directamente a competir con el top de los métodos tradicionales. De ahí la popularidad, a posteriori del artículo, del uso de redes neuronales convolucionales.

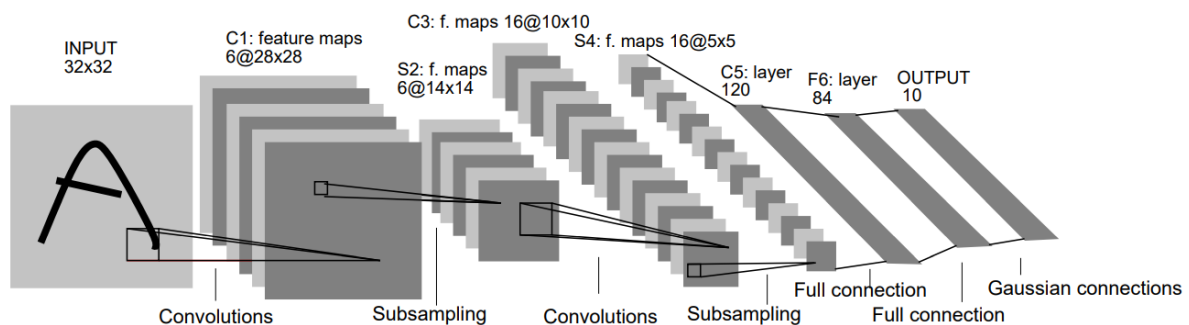
La creciente complejidad de los problemas y la capacidad de usar mejor hardware permitió que los autores del paper dieran al lector un apunte de gran importancia: "Los métodos basado en heurísticas manuales, son una herramienta que caerá en desuso a favor de los métodos basados en aprendizaje automático."

En último lugar, y siendo posiblemente el punto por el que tratamos hoy día de este estudio, hablaremos de los motivos de los autores para utilizar redes convolucionales frente a las redes neuronales y profundizaremos en la arquitectura.

La principal mejora de usar redes convolucionales frente a usar otras técnicas basadas en descenso de gradiente como las redes neuronales, son principalmente:

- Requieren de una menor cantidad de cálculo.
- Mejoran el comportamiento frente a rotaciones y traslaciones.
- Mejoran los resultados de las NN cuando los inputs se someten a ruido.

En cuanto a arquitectura LeNet-5, toma entradas que previamente han recibido un preprocesamiento(se han transformado a un tamaño de 32x32, y se han centrado los elementos)



LeNet-5 como se puede ver en la imagen superior está formada por 7 capas más la entrada. Podemos observar 4 capas de convolución y tres completamente conectadas.

Dos detalles que tenemos que mencionar relativos a su implementación son:

- El valor del fondo( blanco), es de -0.1 mientras que el trazo del carácter tiene un valor de 1.175 lo que genera que la media de los pixeles sea prácticamente 0 y a su vez esto hace que la varianza tienda a 1 lo que acelera el entrenamiento.
- La función de activación en la capa de salida, lejos de ser una función softmax como acostumbramos a ver en la actualidad para problemas multiclase es una RBF. La salida de cada unidad  $y_i$  de RBF puede ser computada mediante:

$$y_i = \sum_j (x_j - w_{ij})^2$$