

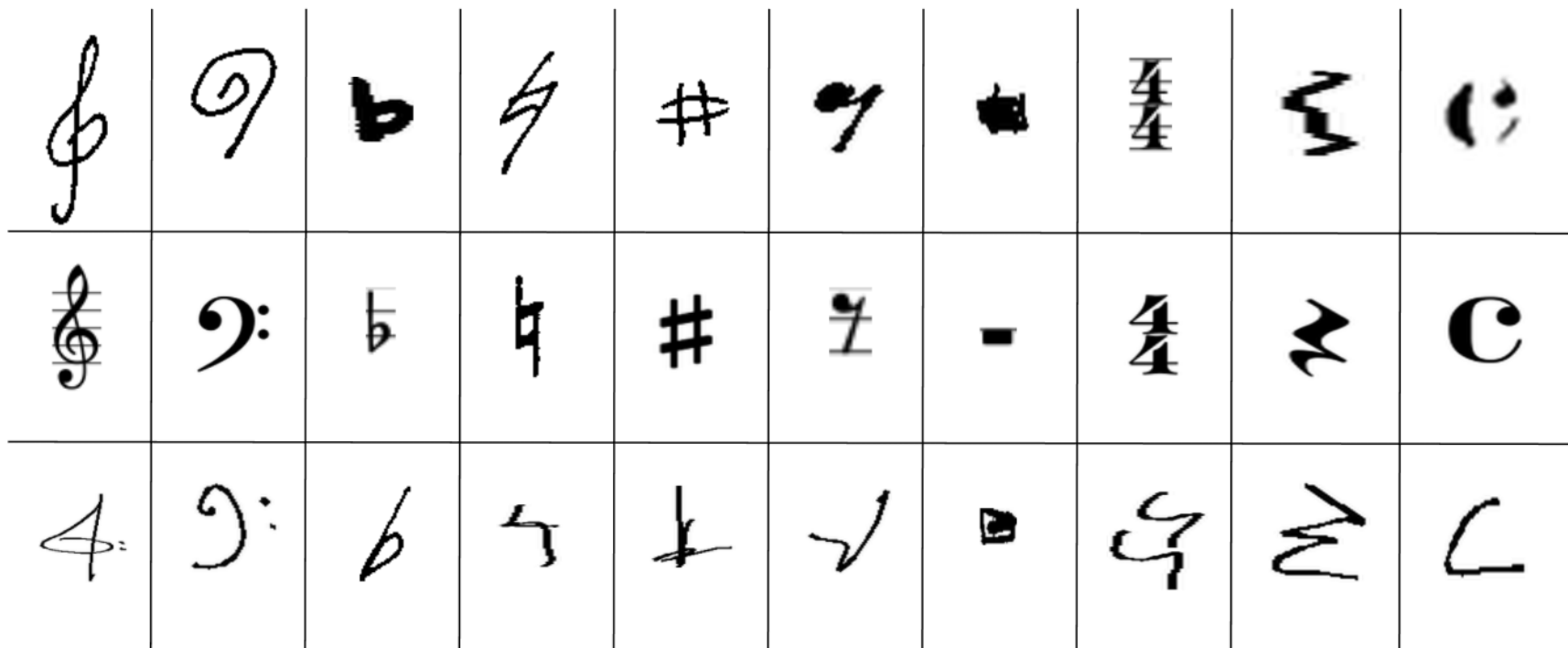
Music Notation CNN

Red Neuronal para clasificación de símbolos musicales



¿En qué consiste el proyecto?

- El proyecto consiste en analizar distintas imágenes con distintos símbolos de la notación musical y clasificarlas satisfactoriamente.



Aplicaciones en la vida cotidiana

- Su principal aplicación sería incorporarla a un software de transcripción digital de melodías.
- Puede usarse para composición musical.



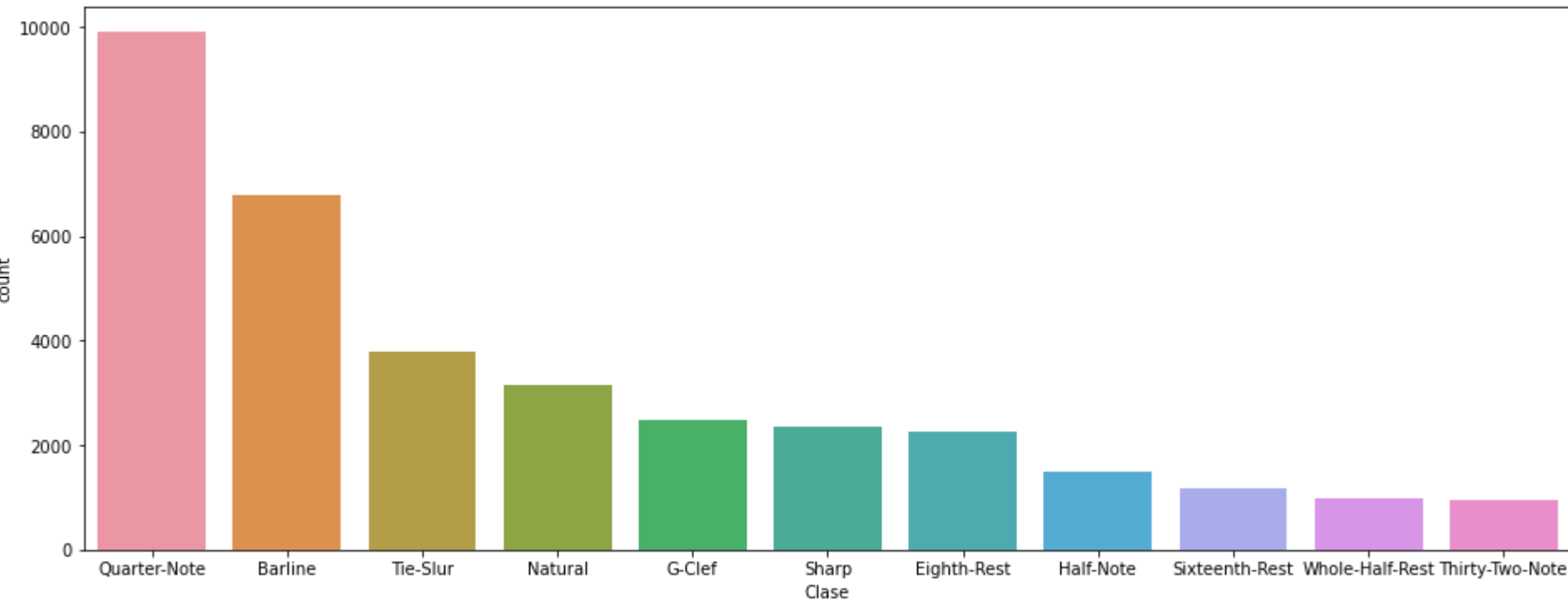
Obtención de los datos

- Los datos fueron obtenidos de OMR.
- Originalmente son imágenes de distintos tamaños que fueron procesadas con CV2 con Python y luego reescaladas al tamaño 96×96.
- Tras ser codificadas fueron guardadas en CSV con tamaños 24×24 debido a que con el tamaño 96×96 mi máquina quedaba trabada.
- El dataset quedó con 39518 observaciones tras retirar las marcas de compás (ya que estas en una partitura pueden deducirse por el compás) y 577 atributos.



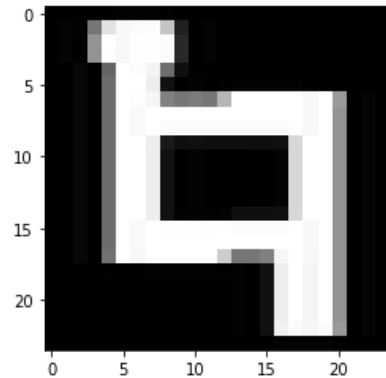
Entrenando el modelo

- Al notar que el dataset originalmente tiene 79 categorías podría ser demasiado pesado para el entrenamiento, por lo que se tomaron las primeras 11 clases que tenían elementos más abundantes.



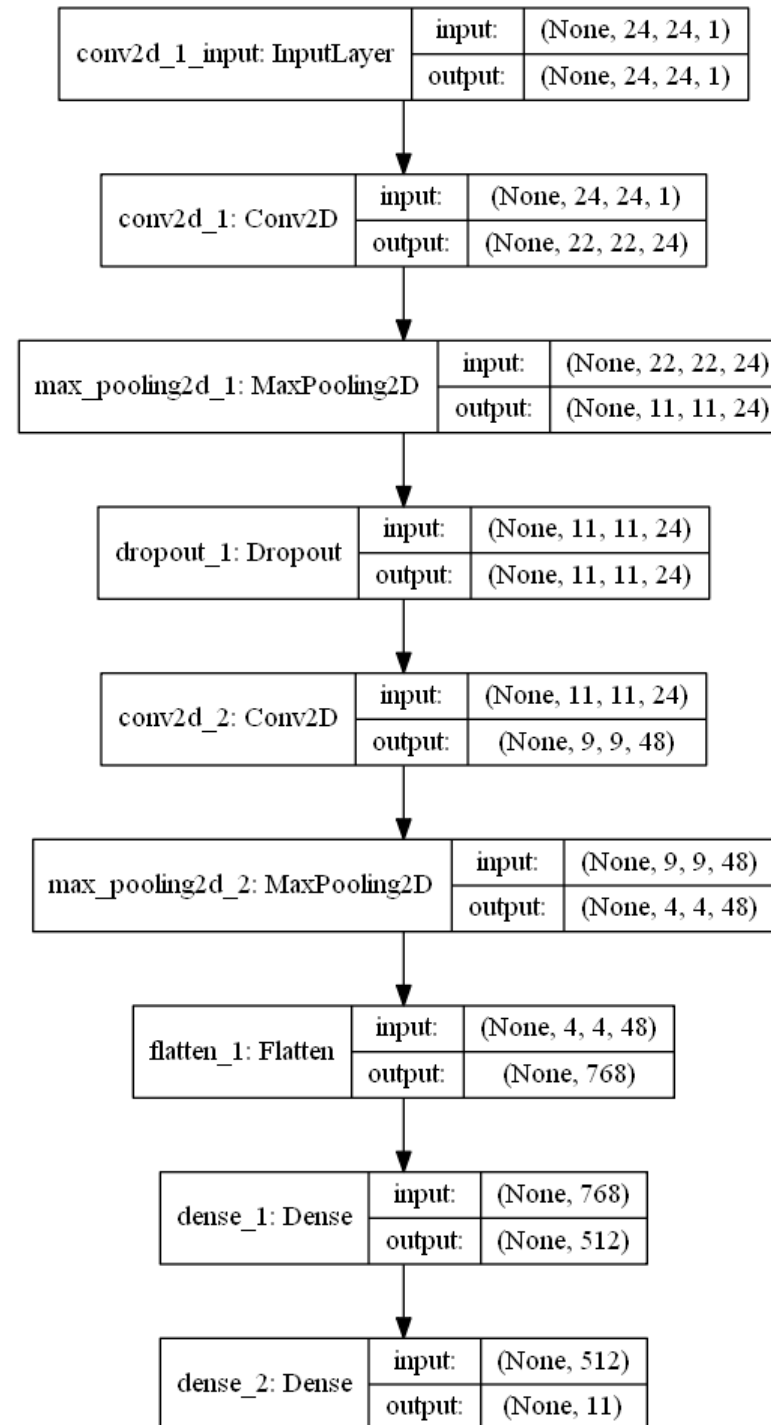
Estructura del modelo

Ingreso de la CNN

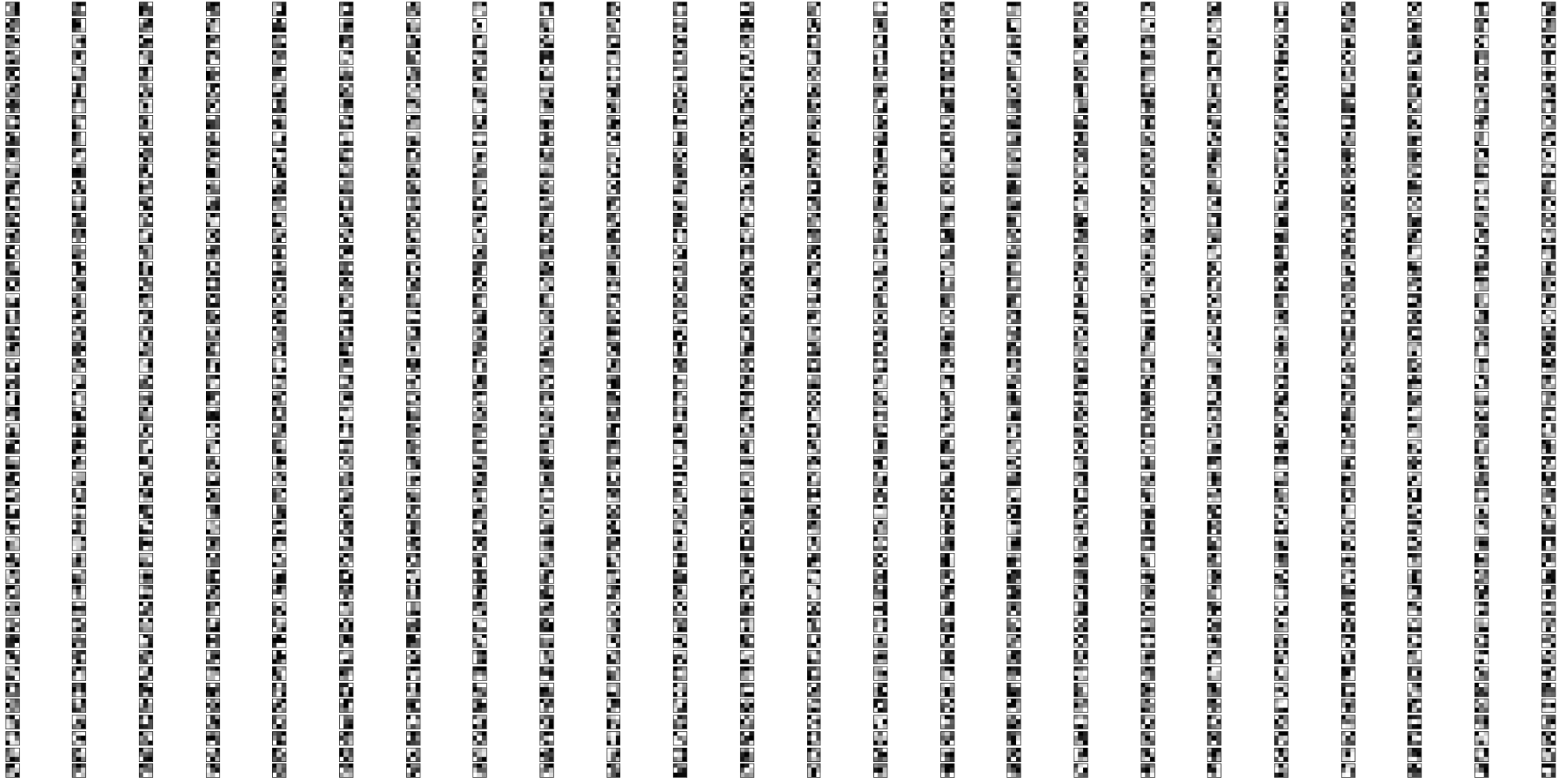
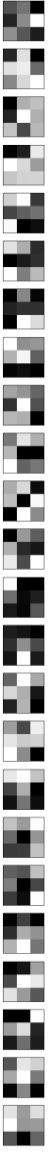


Salida de la CNN

Natural (Becadro)

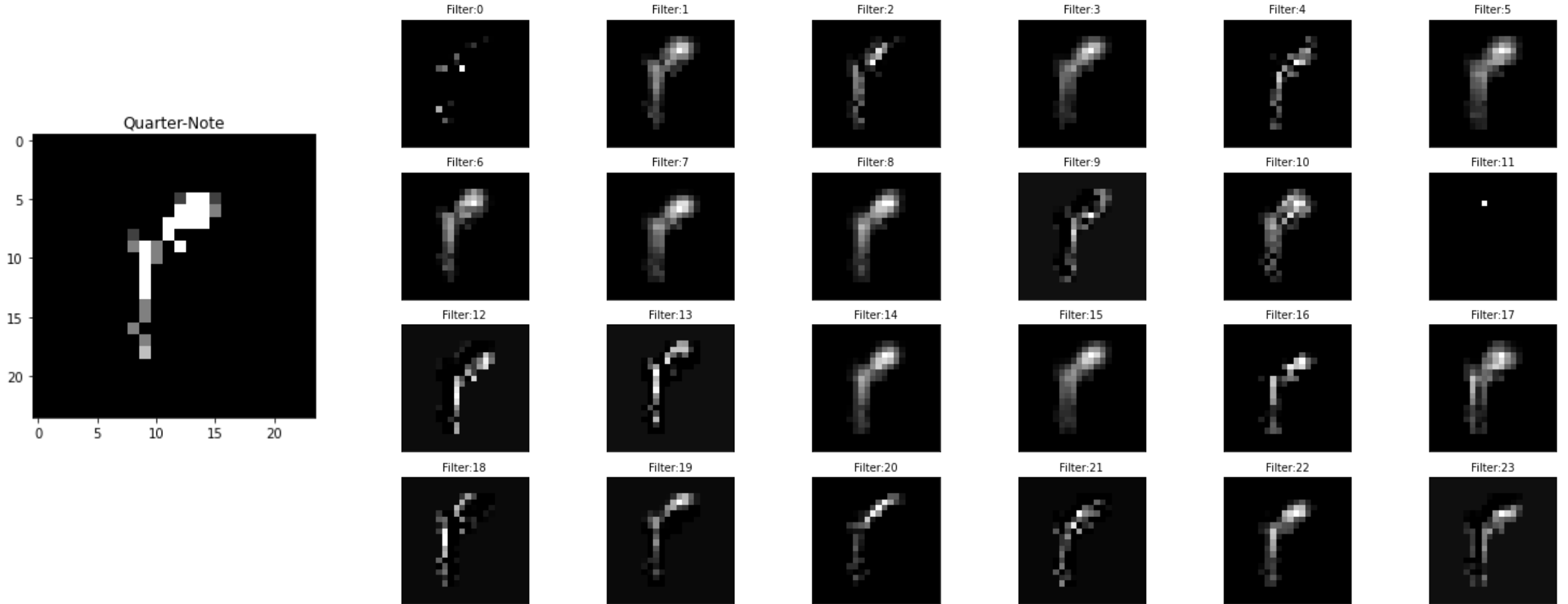


Filtros de la CNN



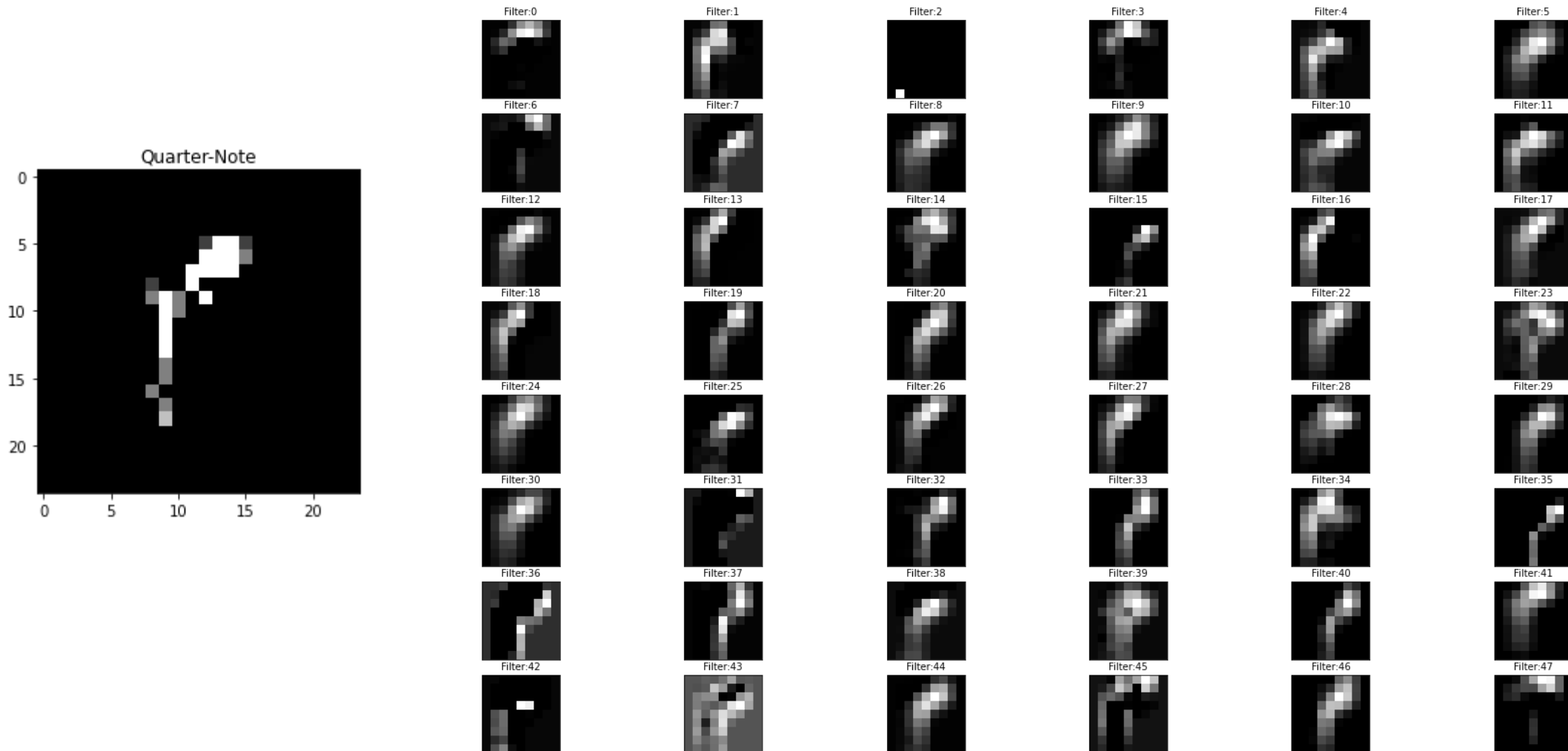
Avances de las convoluciones

Capa 0



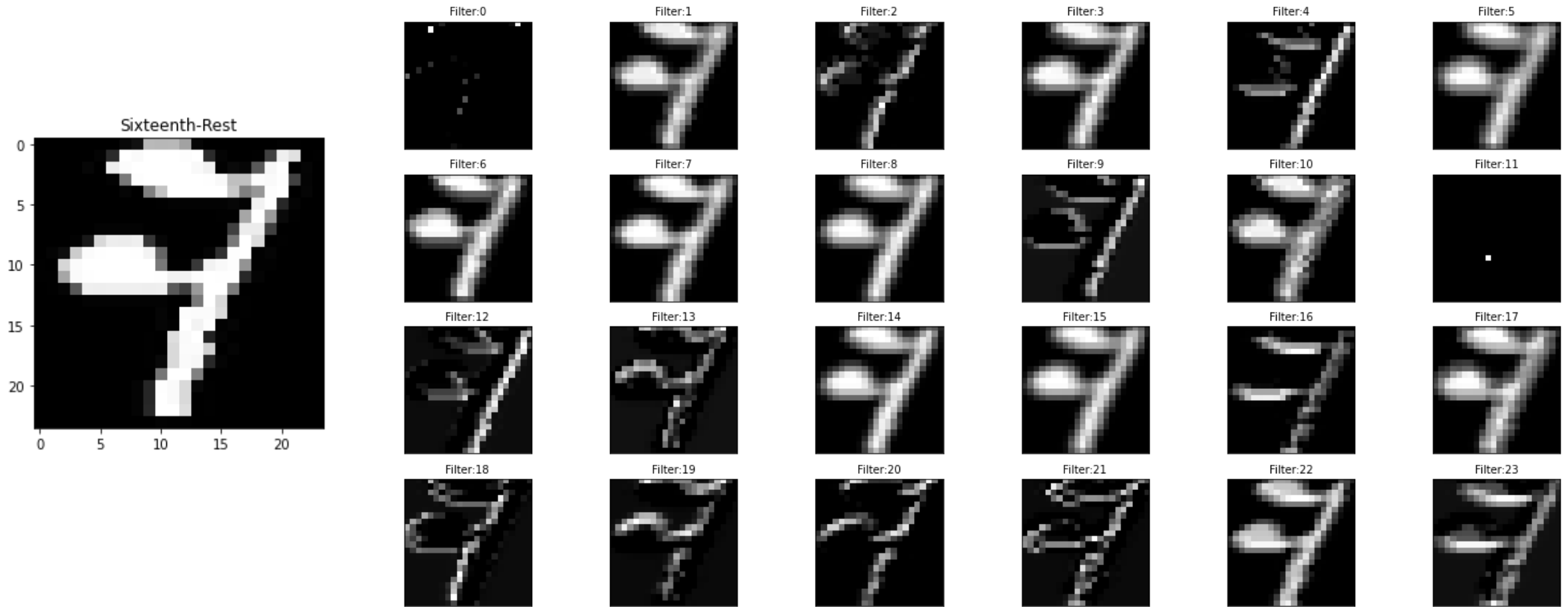
Avances de las convoluciones

Capa 1



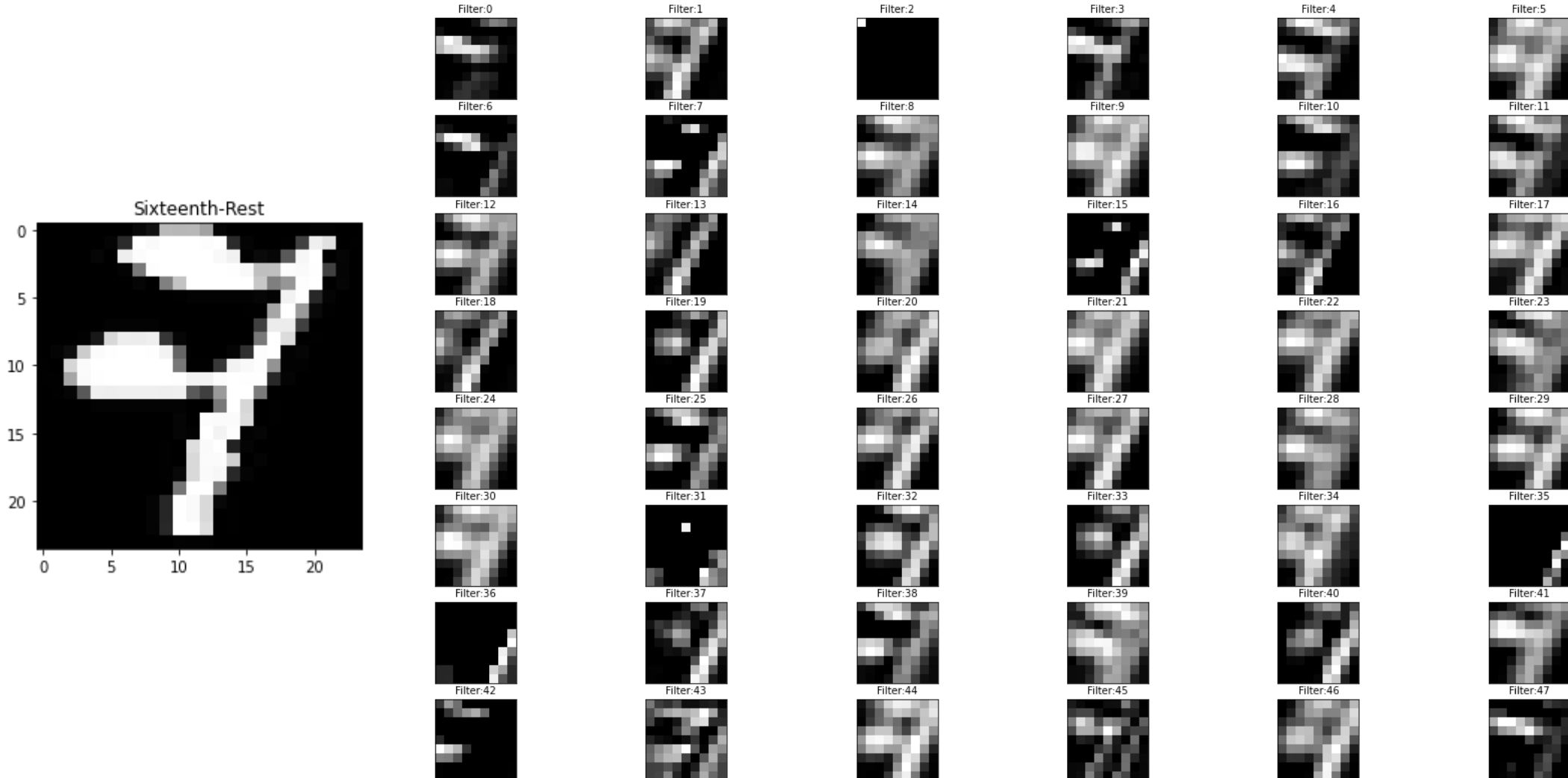
Avances de las convoluciones

Capa 0



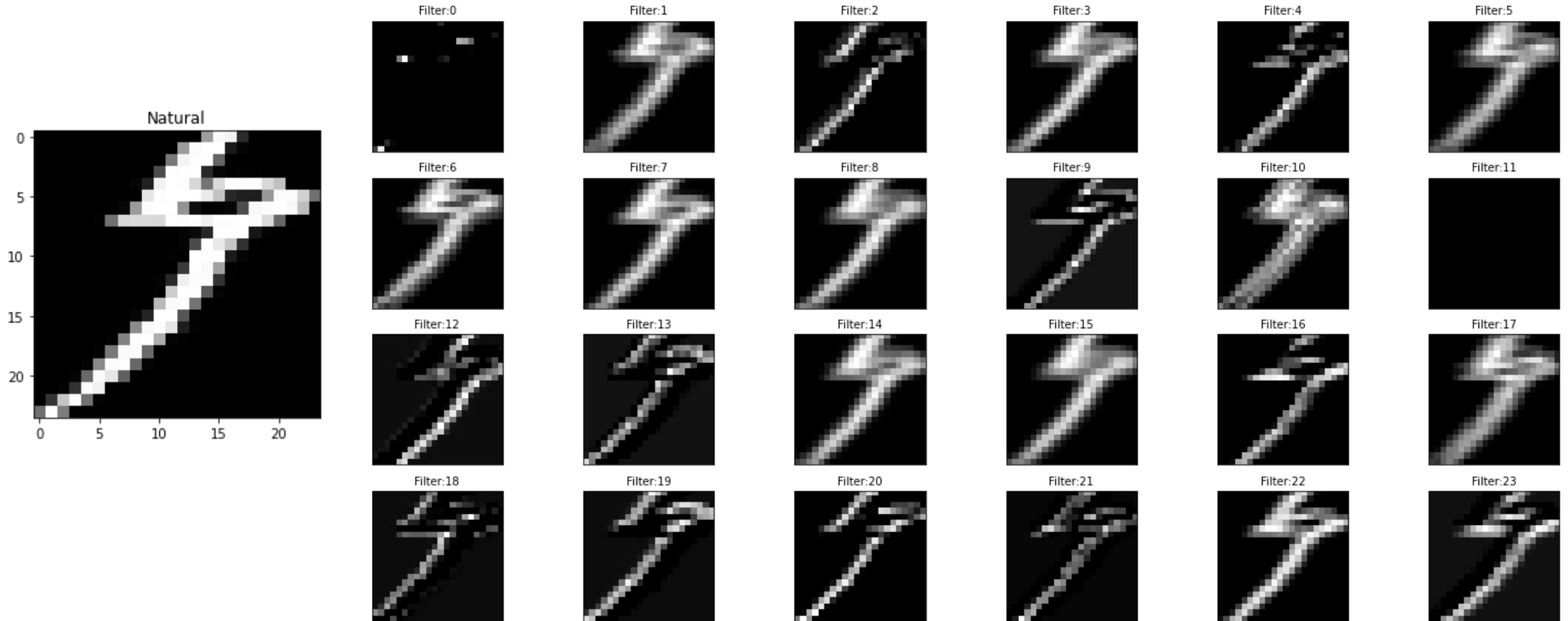
Avances de las convoluciones

Capa 1



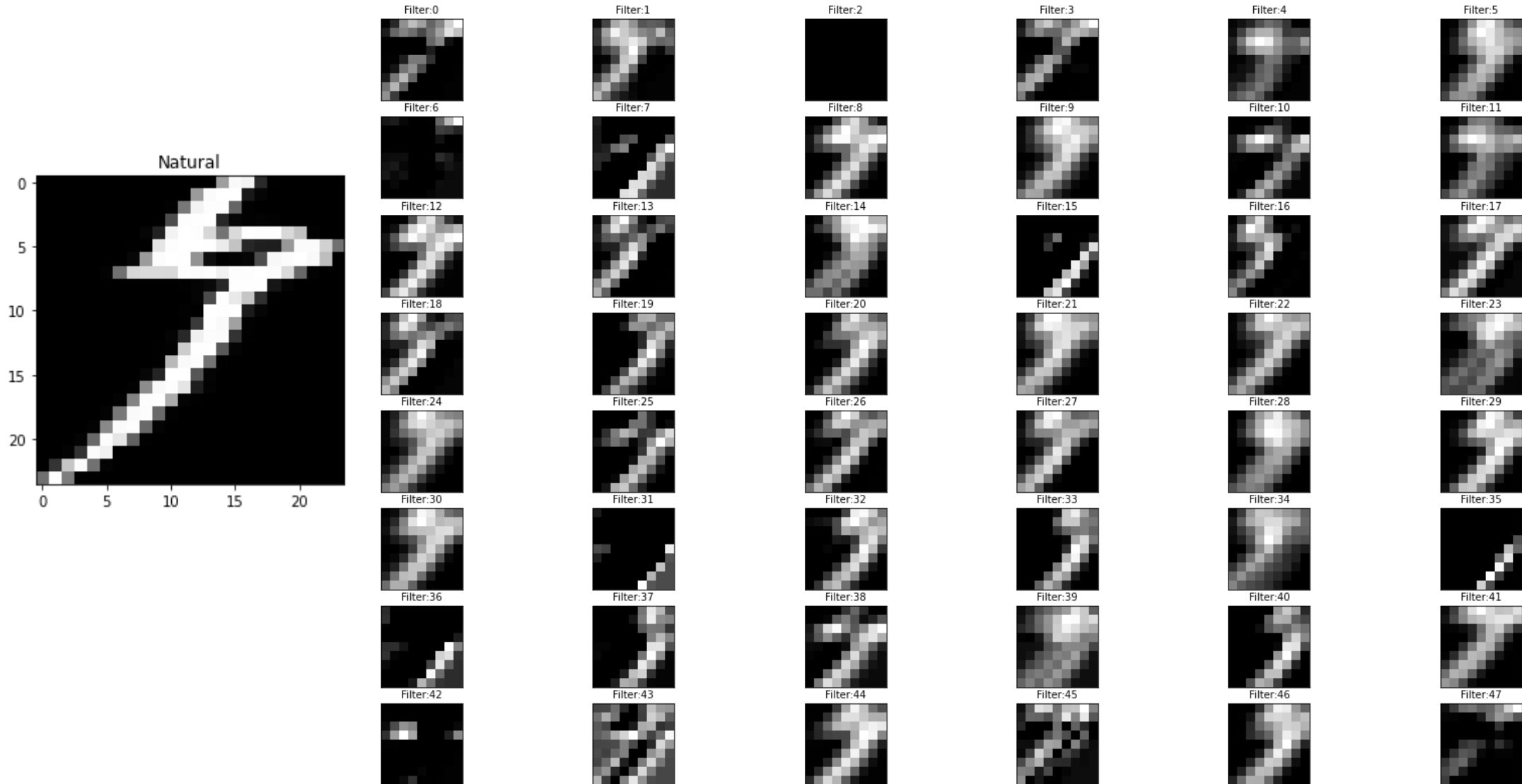
Avances de las convoluciones

Capa 0



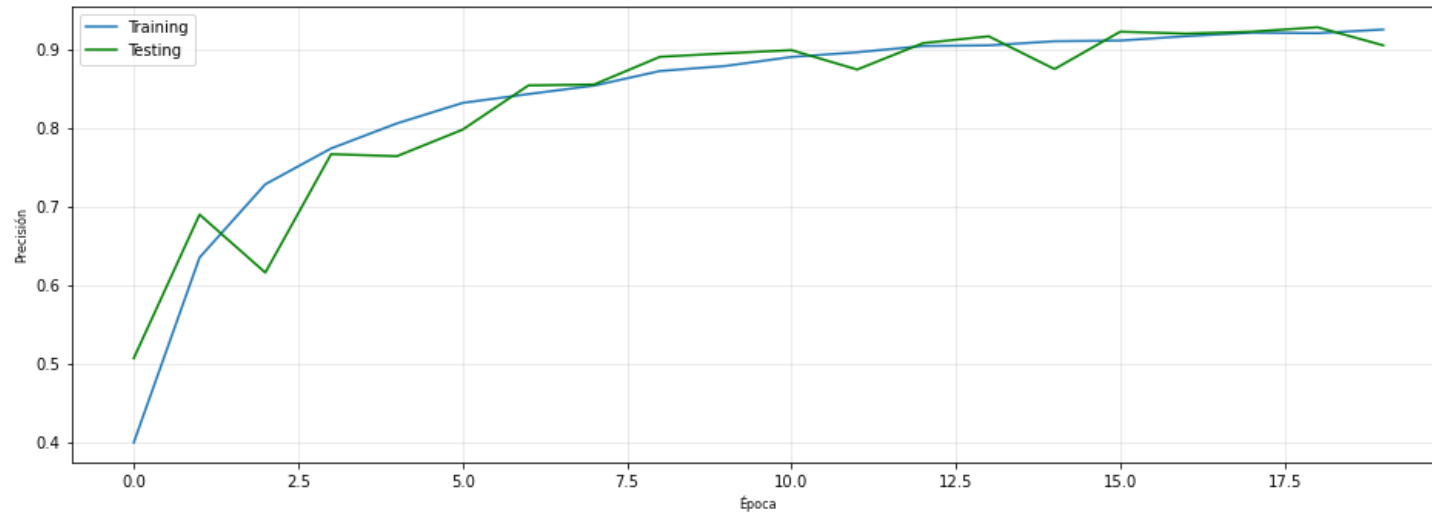
Avances de las convoluciones

Capa 1



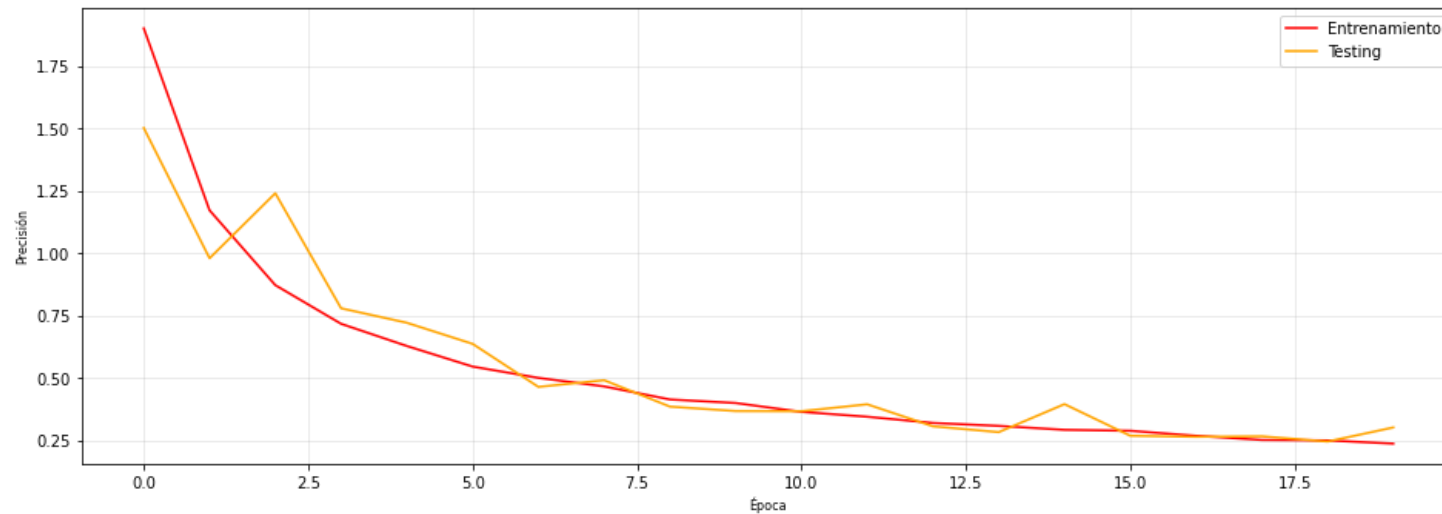
Resultados del modelo

Precisión del modelo



91%

Pérdida de error



28%



Conclusiones y recomendaciones

- Probar a entrenar la red con más símbolos resultantes, parece que los datos son bastante buenos para entrenar una diversidad de objetos.
- Usar otras estrategias de DeepLearning para ver su diferencia.
- Conseguir datasets con imágenes a color para ver como se sensibiliza el modelo a los colores, en el caso algunas notaciones el color de la nota identifica su duración y expresión.



Poniendo en práctica el primer punto

Se entrenó una red neuronal con las 39518 observaciones con 30 clases predictivas

24*24

Capa
Convolutacional

24*48

Capa
Convolutacional

48*64

Capa
Convolutacional

1024

Intermedia

512

Intermedia

20

Salida



Resultados obtenidos

Nueva CNN para clasificación de 30 símbolos

Precisión: 95%

Pérdida: 13%



