

Universidad del Valle de Guatemala

Data Science

Sección 20

# **Laboratorio 3**

Brian Carrillo - 21108

Carlos López - 21666

**Guatemala, 11 de agosto del 2024**

# Arquitectura Le-Net

## Capas

### 1. Entrada:

- Tamaño de Entrada: Imágenes de 32x32 píxeles.

### 2. Capa Convolutiva 1 (C1):

- Número de Filtros: 6 filtros.
- Tamaño del Filtro: 5x5.
- Función de Activación: Sigmoid/ReLU
- Salida: 6 mapas de características de 28x28 (cada filtro reduce la dimensión de la imagen en 4 píxeles:  $32 \times 32 \rightarrow 28 \times 28$ ).
- Propósito: Detectar características locales como bordes, texturas y patrones básicos en la imagen.

### 3. Capa de Submuestreo (S2) - Pooling:

- Tipo: Pooling promedio.
- Tamaño del Filtro: 2x2.
- Stride: 2 (esto reduce la dimensión a la mitad).
- Salida: 6 mapas de características de 14x14.
- Propósito: Reducir la dimensionalidad y la sensibilidad a la ubicación de las características.

### 4. Capa Convolutiva 2 (C3):

- Número de Filtros: 16 filtros.
- Tamaño del Filtro: 5x5.
- Función de Activación: Sigmoid/ReLU.
- Salida: 16 mapas de características de 10x10.
- Conectividad: No todos los filtros están conectados a las 6 salidas de S2. Algunos están conectados a combinaciones específicas de los 6 mapas de S2 (esto reduce la complejidad).
- Propósito: Detectar patrones más complejos y específicos.

### 5. Capa de Submuestreo (S4) - Pooling:

- Tipo: Pooling promedio.
- Tamaño del Filtro: 2x2.
- Salida: 16 mapas de características de 5x5.
- Propósito: Continuar la reducción de dimensionalidad y la invariancia a la traslación.

### 6. Capa Convolutiva 3 (C5):

- Número de Filtros: 120 filtros.
- Tamaño del Filtro: 5x5 (completamente conectada a la capa S4).
- Salida: 120 mapas de características de 1x1.
- Propósito: Extraer las características más complejas y obtener una representación final que captura la esencia de la imagen.

### 7. Capa Fully Connected (F6):

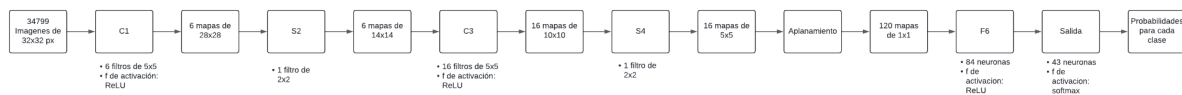
- Neuronal fully connected: 84 neuronas.
- Función de Activación: Sigmoid/ReLU.
- Propósito: Actuar como una red neuronal clásica que combina las características extraídas para tomar decisiones.

### 8. Capa de Salida:

- Neuronal fully connected: 43 neuronas (correspondientes a las 43 clases de señales en el conjunto).
- Función de Activación: Softmax.
- Propósito: Producir probabilidades para cada clase.

## Diagrama

### [Diagrama LucidChart](#)



## Procesos y Modelo

1. Convolución: En las CNNs, la convolución es un proceso esencial que permite a la red identificar patrones locales en las imágenes, como bordes y texturas.

- Filtro/Kernel: Un pequeño conjunto de pesos, como un filtro 3x3 o 5x5, se desplaza sobre la imagen de entrada.
- Operación de Convolución: En cada posición del filtro, los valores de la imagen se multiplican por los correspondientes del filtro, y luego se suman los productos para obtener un único valor.
- Mapa de Características: Al aplicar el filtro a la imagen, se genera un mapa de características que resalta ciertos patrones presentes en la imagen original.

2. Función de Activación: Tras la convolución, se utiliza una función de activación no lineal para introducir no linealidad en el modelo. Esto permite que la red pueda aprender y representar patrones más complejos.

- ReLU (Rectified Linear Unit): Es una de las funciones de activación más utilizadas en las CNNs. ReLU transforma la salida en  $\max(0, x)$ , es decir, convierte a cero los valores negativos y deja inalterados los positivos.
- Sigmoid/Tanh: Estas son otras funciones de activación usadas en redes más antiguas o en ciertos casos específicos. Sigmoid ajusta la salida entre 0 y 1, mientras que Tanh lo hace entre -1 y 1.

3. Pooling: El pooling es una técnica de reducción de dimensionalidad empleada para disminuir el tamaño de los mapas de características, lo que aumenta la eficiencia del modelo y reduce la posibilidad de sobreajuste.

- Pooling Promedio (Average Pooling): Calcula el promedio de los valores dentro de cada región del mapa de características.
- Pooling Máximo (Max Pooling): Selecciona el valor más alto en cada región del mapa de características. Max pooling es más común que el pooling promedio porque suele mantener las características más relevantes.

#### Importancia de la función de pérdida

La función de pérdida sirve como una referencia esencial en el proceso de aprendizaje. Un valor elevado de pérdida sugiere que el modelo no está realizando predicciones acertadas y requiere ajustes. A lo largo del entrenamiento, se busca reducir al máximo esta función, lo cual indica que el modelo está afinando su habilidad para hacer predicciones más precisas.

#### Importancia del optimizador

El optimizador juega un papel fundamental al definir la forma y la velocidad con que la red neuronal adquiere conocimiento. Un optimizador eficiente puede acelerar la convergencia del modelo y mejorar su rendimiento final. Asimismo, contribuye a prevenir dificultades como el estancamiento en mínimos locales o el sobreajuste del modelo.

### Entrenamiento

Para el proceso de entrenamiento, se utilizaron los siguientes hiper parámetros:

- Tamaño de batch: 256. Es decir, el modelo procesa 256 imágenes por iteración, previo a actualizar los parámetros de la CNN.
- Número de épocas: 20. Es decir, el modelo entrenará con el training set completo un total de 20 veces.
- Optimizador: Adam. Es un optimizador adaptativo de tasa de aprendizaje, diseñado para entrenamiento de redes neuronales.
- Función de pérdida: categorical cross-entropy. Generalmente, es la función utilizada en problemas de clasificación multiclase, como en este caso.
- Métricas: Precisión. Se especificó el seguimiento de la métrica de precisión durante el entrenamiento del modelo.

### Pruebas

Métrica	Valor
Accuracy	0.91
Precision	0.91
Recall	0.91
F1-score	0.91

**Accuracy:** La exactitud mide la proporción de predicciones correctas realizadas por el modelo sobre el total de predicciones. En este caso, el modelo acertó en aproximadamente el 90.84% de las señales de tráfico. Esto significa que de todas las señales que el modelo clasificó, casi el 91% fueron correctamente identificadas.

**Precision:** La precisión es la proporción de verdaderos positivos entre todos los casos en los que el modelo predijo la clase positiva. En otras palabras, mide cuántas de las señales que fueron etiquetadas como una determinada clase por el modelo eran realmente de esa clase. Con una precisión de 91.19%, esto indica que cuando el modelo predijo una señal de tráfico en particular, fue correcto aproximadamente el 91.19% de las veces.

**Recall:** El recall mide la proporción de verdaderos positivos entre todos los casos que realmente pertenecen a la clase positiva. Es decir, indica qué tan bien el modelo puede identificar todos los casos de una clase en particular. Un recall de 90.84% significa que el modelo fue capaz de identificar correctamente el 90.84% de todas las señales de tráfico de la clase relevante.

**F1:** El F1 Score es la media armónica de la precisión y el recall. Proporciona un balance entre estos dos, especialmente útil cuando tienes un desequilibrio entre clases o cuando tanto la precisión como el recall son importantes. Un F1 score de 90.75% indica un buen equilibrio entre precisión y recall. En otras palabras, el modelo tiene un buen rendimiento general tanto en minimizar los falsos positivos como en maximizar los verdaderos positivos.

## Mejoras

### 1. Aumentar la Complejidad del Modelo:

- Agregar más capas convolucionales puede permitir que el modelo capture características más complejas y detalladas. Se agregó una tercera capa convolucional con 128 filtros.
- Incrementar el número de filtros en cada capa convolucional permitirá al modelo capturar más patrones. Para esto, se aumentó el número de filtros de 32 y 64 en las dos primeras capas, respectivamente.

### 2. Agregar Capas de Regularización:

- El uso de Dropout previene el sobreajuste al desactivar aleatoriamente neuronas durante el entrenamiento. Se agregaron capas dropout luego de las capas densas.
- La regularización L2 penaliza grandes pesos en el modelo, lo que también ayuda a reducir el sobreajuste. Se añadió esta característica a algunas capas densas.

### 3. Capas de Normalización:

- Se agregaron capas de normalización por lotes, luego de las capas convolucionales y antes de las funciones de activación para estabilizar y acelerar el entrenamiento.

## Resultados

Métrica	Valor
Accuracy	0.95
Precision	0.95
Recall	0.95
F1-score	0.95

Existe una clara mejora en todas las métricas. Son resultados muy buenos que reflejan la capacidad del modelo para identificar las clases de las señales de tránsito provistas. Este modelo, con mayores mejoras, podría ser utilizado en automóviles inteligentes que sean capaces de una conducción autónoma, y puedan interpretar las señales que ven en su entorno.

## Referencias

TensorFlow. (2022). Red Neuronal Convolucional. Obtenido de <https://www.tensorflow.org/tutorials/images/cnn?hl=es-419>