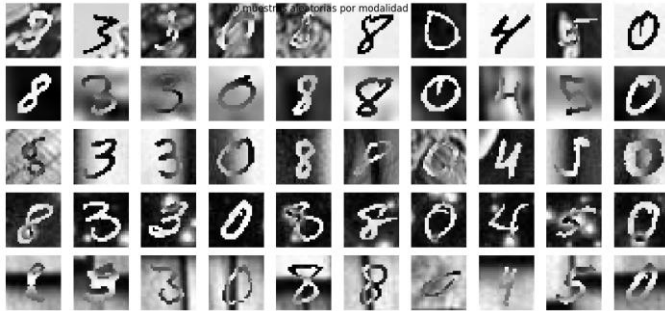


## Laboratorio 3. Deep Learning

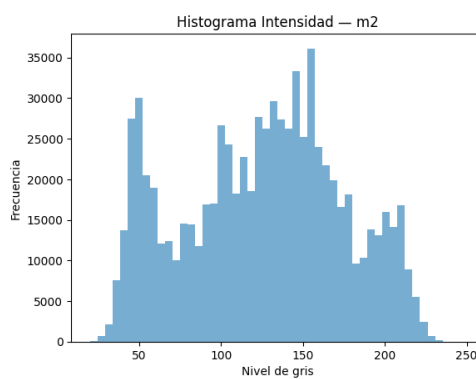
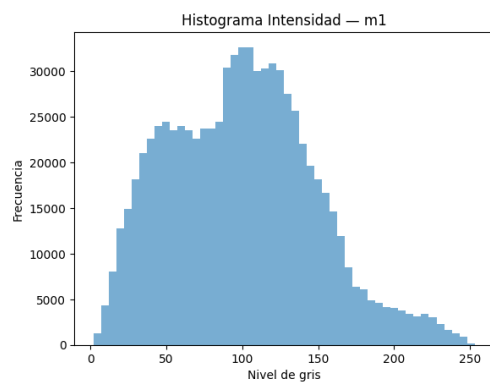
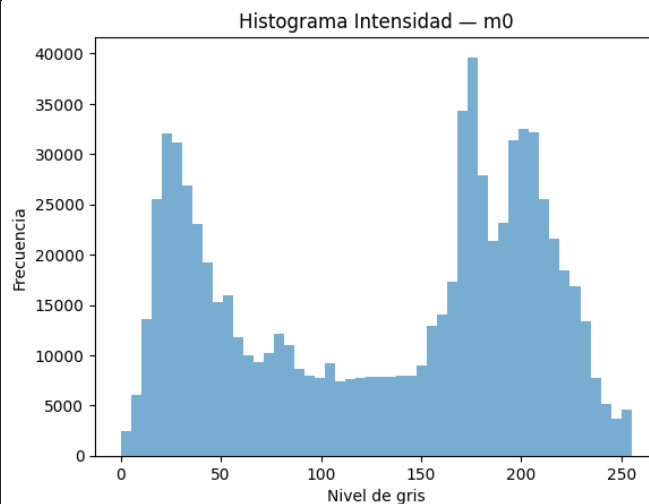
### Análisis exploratorio de los datos

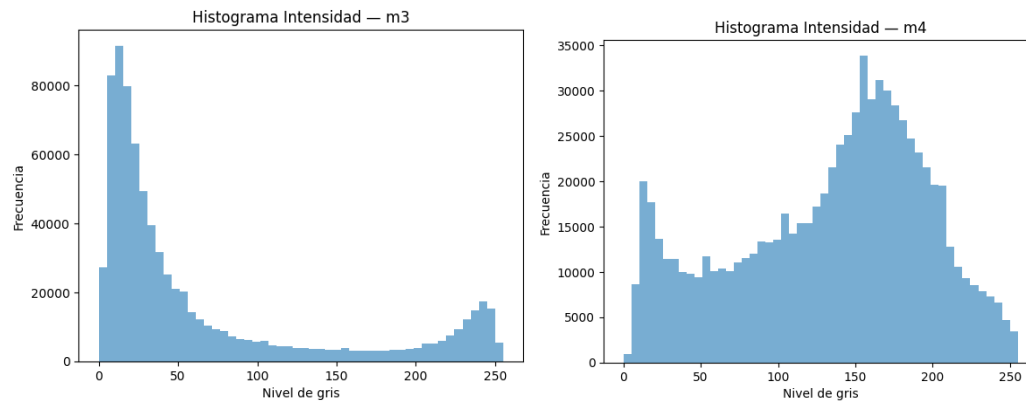


```
-- Primeras filas del DataFrame --
split  modality      filepath  label
0  train      m0  PolyMNIST/MNIST/train/m0\0.0.png    0
1  train      m0  PolyMNIST/MNIST/train/m0\0.1.png    0
2  train      m0  PolyMNIST/MNIST/train/m0\0.2.png    0
3  train      m0  PolyMNIST/MNIST/train/m0\0.3.png    0
4  train      m0  PolyMNIST/MNIST/train/m0\0.4.png    0

-- Conteo imágenes por split/modality --
split  modality
test   m0      10000
      m1      10000
      m2      10000
      m3      10000
      m4      10000
train  m0     60000
      m1     60000
      m2     60000
      m3     60000
      m4     60000
dtype: int64

-- Distribución de clases en TRAIN --
label
0      50
...
6740    5
6741    5
Length: 6742, dtype: int64
```





## Descripción de los modelos

**PolyMNIST** presenta cinco modalidades distintas. El fondo de cada modalidad se compone de parches aleatorios recortados de una imagen de mayor tamaño, con el dígito colocado aleatoriamente dentro de dichos parches. Esta configuración aporta a cada modalidad información única procedente de su respectiva imagen de fondo, mientras que el dígito actúa como información compartida entre todas ellas. Además, a diferencia del PolyMNIST original, aquí los dígitos también se desplazan de forma aleatoria dentro del parche, lo que introduce un reto extra para los modelos de visión por computador.

## Extensión y detalles adicionales sobre PolyMNIST

- **Modalidades y fondos:** Cada modalidad puede extraer sus parches de fuentes muy diversas (fotografías urbanas, texturas naturales, escenas interiores, etc.), de manera que un mismo dígito negro sobre fondo blanco pase a incrustarse en contextos muy distintos.
- **Objetivo de la variabilidad:** Al cambiar tanto el fondo como la posición del dígito, PolyMNIST sirve para evaluar la capacidad de los modelos de separar la señal relevante (el dígito) del “ruido” contextual. Se emplea frecuentemente para estudiar **adaptación de dominio** y **robustez** frente a variaciones del entorno visual.
- **Formato y particiones:** Suele distribuirse en modos de entrenamiento y prueba equilibrados, manteniendo cada uno ejemplos de las cinco modalidades. Con ello se mide no solo la **exactitud** global de clasificación, sino también la habilidad del modelo para generalizar a modalidades vistas y no vistas.

Karen Pineda  
Dolan Cuellar

- **Aplicaciones prácticas:** Más allá de dígitos, el enfoque de PolyMNIST inspira benchmarks en reconocimiento de objetos en entornos variables (p. ej., vehículos en distintas condiciones de iluminación o cámaras de seguridad).