

**CENTRO DE ENSEÑANZA TÉCNICA INDUSTRIAL**



**Investigación 8**

**Profesor: Mauricio Alejandro Cabrera Arellano**

**Alumno: Ángel de Jesús Mejía Suárez**

**21310182**

**6°G**

**Materia: Visión Artificial**

**12 de mayo de 2025**

*Explicame acerca de las Redes Neuronales: Introducción a Redes Neuronales. Funciones de activación. Función de costo. Optimizador. Perceptrón. Perceptrón Multicapa. Redes Neuronales Convolucionales.*

Aquí tienes una explicación formal, clara y detallada sobre Redes Neuronales, incluyendo cada uno de los temas solicitados:

## ✦ 1. Introducción a Redes Neuronales

Las redes neuronales artificiales (**ANN**, Artificial Neural Networks) son modelos computacionales inspirados en el funcionamiento del cerebro humano. Están compuestas por capas de **neuronas interconectadas** que aprenden patrones mediante entrenamiento supervisado o no supervisado.

### Componentes principales:

- **Neuronas:** unidades básicas que reciben entradas, aplican una transformación y producen una salida.
- **Capas:** organizan las neuronas en estructuras secuenciales o jerárquicas.
  - Entrada (Input layer)
  - Capas ocultas (Hidden layers)
  - Salida (Output layer)
- **Pesos (weights):** parámetros ajustables que determinan la influencia de cada entrada.

## ✦ 2. Funciones de Activación

La función de activación introduce no linealidad en las neuronas, permitiendo que la red aprenda relaciones complejas.

### Funciones comunes:

- **Sigmoide:**

$$f(x) = \frac{1}{1 + e^{-x}}$$

- Salida entre 0 y 1; útil en clasificación binaria.

- **Tangente Hiperbólica (tanh):**

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

- Salida entre -1 y 1; mantiene valores centrados.

- **ReLU (Rectified Linear Unit):**

$$f(x) = \max(0, x)$$

- Muy popular por su simplicidad y eficiencia en entrenamiento.

- **Softmax:**

$$f(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}}$$

- Utilizada en clasificación multiclase; produce probabilidades.

### ✦ 3. Función de Costo (Loss Function)

Mide el error entre las predicciones de la red neuronal y los valores reales esperados. El entrenamiento busca minimizar esta función.

#### Funciones comunes:

- **Error cuadrático medio (MSE):**

$$J = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

- **Entropía cruzada (Cross-Entropy) (clasificación):**

$$J = - \sum_{i=1}^n (y_i \log(\hat{y}_i))$$

#### ✦ 4. Optimizador

Es el algoritmo que ajusta los pesos de la red para minimizar la función de costo durante el entrenamiento.

**Optimizadores comunes:**

- **Descenso del gradiente (Gradient Descent):** Método básico, actualiza los pesos según el gradiente del error.

$$w = w - \eta \frac{\partial J}{\partial w}$$

- **SGD (Stochastic Gradient Descent):** Usa muestras individuales o pequeños lotes; es eficiente y rápido.
- **Adam:** Combinación adaptativa de tasas de aprendizaje, ampliamente usado por su rapidez y precisión.

#### ✦ 5. Perceptrón

Es la red neuronal más simple y consiste en una sola neurona artificial que realiza una clasificación lineal binaria.

**Funcionamiento básico:**

- Calcula una suma ponderada de entradas:

$$z = \sum_i w_i x_i + b$$

- Aplica una función escalón o sigmoide para clasificar la entrada en una de dos clases.

**Limitación:**

Sólo puede resolver problemas linealmente separables.

## ✦ 6. Perceptrón Multicapa (MLP)

Extensión del perceptrón simple, con múltiples capas ocultas. Capaz de aprender relaciones no lineales complejas.

### **Características principales:**

- Tiene al menos una capa oculta.
- Emplea funciones de activación no lineales (ReLU, tanh, sigmoide).
- Entrenado mediante el algoritmo de **propagación hacia atrás (backpropagation)**.

### **Aplicaciones típicas:**

- Clasificación de datos complejos.
- Regresión de series temporales.
- Reconocimiento de patrones.

## ✦ 7. Redes Neuronales Convolucionales (CNN)

Son redes especializadas en el procesamiento de datos estructurados en forma de matrices, como imágenes. La estructura convolucional permite extraer patrones espaciales eficientemente.

### **Capas principales de una CNN:**

- **Capa convolucional:**  
Aplica filtros para detectar patrones locales.
  - Extracción de características visuales (bordes, texturas).
- **Capa de activación (ReLU):**  
Añade no linealidad tras la convolución.

- **Capa de Pooling (Agrupación):**  
Reduce dimensionalidad espacial y mantiene características esenciales.
  - Ejemplo: Max Pooling, Average Pooling.
- **Capas completamente conectadas (Fully Connected):**  
Clasifican características extraídas por capas convolucionales.

#### Aplicaciones populares:

- Reconocimiento de imágenes.
- Detección de objetos.
- Clasificación facial y biométrica.

#### ✦ Resumen comparativo:

Tipo de red	Complejidad	Capacidad	Aplicaciones típicas
<b>Perceptrón</b>	Baja	Limitada (lineal)	Clasificación simple
<b>Perceptrón Multicapa</b>	Media-Alta	Alta (no lineal)	Problemas generales de clasificación y regresión
<b>CNN</b>	Alta	Muy alta (patrones espaciales)	Procesamiento visual, imágenes, video

*Ahora explicame acerca de la clasificación y localización de objetos en imágenes, así como de la clasificación binaria y la clasificación multiclase.*

Aquí tienes una explicación clara, detallada y formal sobre la **clasificación y localización de objetos en imágenes**, así como la **clasificación binaria y multiclase** en el contexto del procesamiento de imágenes y aprendizaje automático.

#### ✦ Clasificación y Localización de Objetos en Imágenes

##### Conceptos básicos:

- **Clasificación:** Identificar o etiquetar el objeto principal presente en una imagen.
- **Localización:** Además de identificar qué objeto aparece, se determina su posición exacta mediante un rectángulo delimitador (**bounding box**).

#### Proceso general:

1. **Extracción** **de** **características:**

Las redes neuronales extraen patrones visuales de la imagen.

2. **Clasificación:**

Una red neuronal asigna una clase específica (persona, vehículo, animal, etc.) al objeto.

3. **Localización:**

Se predicen coordenadas que describen la ubicación exacta del objeto (normalmente usando coordenadas (x, y, ancho, alto)).

#### Modelos populares:

- **YOLO (You Only Look Once)**
- **SSD (Single Shot Detector)**
- **Faster R-CNN**

Estos métodos combinan ambas tareas en un único modelo eficiente.

#### ✚ **Clasificación Binaria vs Clasificación Multiclase**

En procesamiento de imágenes, ambos tipos de clasificación son fundamentales:

##### **Clasificación Binaria:**

- Implica decidir entre **dos categorías** o clases.
- Ejemplos prácticos:
  - Detección facial (**cara/no cara**).
  - Diagnóstico médico (**enfermedad/no enfermedad**).

### Modelos típicos:

- Perceptrón simple
- Regresión logística
- Redes neuronales sencillas

### Clasificación Multiclase:

- Consiste en identificar un objeto entre **tres o más clases posibles**.
- Ejemplos:
  - Clasificación de animales (**perro, gato, caballo**).
  - Reconocimiento de caracteres (**A-Z**).

### Modelos típicos:

- Redes neuronales convolucionales (CNN)
- Máquinas de soporte vectorial (SVM multiclase)
- Clasificadores basados en árboles (Random Forest)

### 📌 Métricas de evaluación importantes:

Para ambos casos, se emplean métricas específicas que miden la calidad del modelo:

- **Exactitud (Accuracy):** Proporción de predicciones correctas.

$$\text{Accuracy} = \frac{\text{Predicciones correctas}}{\text{Total de predicciones}}$$

- **Precisión (Precision):** Proporción de verdaderos positivos sobre todos los positivos predichos.

$$\text{Precision} = \frac{TP}{TP + FP}$$



- **Recall (Sensibilidad):** Proporción de verdaderos positivos detectados respecto al total real de positivos.

$$\text{Recall} = \frac{TP}{TP + FN}$$

- **Matriz de confusión:** Herramienta para evaluar desempeño y tipos de errores cometidos.

### 📌 Ejemplo comparativo de ambos enfoques en visión artificial:

Tipo de Clasificación	Número de clases	Aplicación típica	Modelos comunes
<b>Binaria</b>	Dos (0 y 1)	Rostro / No rostro	Regresión logística, Perceptrón
<b>Multiclase</b>	Múltiples (>2)	Reconocimiento especies, OCR	CNN, SVM, Random Forest

### 📌 Clasificación vs Detección vs Segmentación

Es importante no confundir estos términos:

- **Clasificación:** ¿Qué objeto está presente?
- **Detección:** ¿Qué objeto está presente y dónde está ubicado (bounding box)?
- **Segmentación:** ¿Cuál es la forma exacta del objeto (nivel de píxeles)?

Cada tarea tiene un nivel de complejidad y uso específico en visión artificial.

*Explicame acerca de la Detección de objetos. Así como de: Detección de un objeto. Detección de múltiples objetos.*

Aquí tienes una explicación clara, formal y detallada sobre la **detección de objetos en imágenes**, enfocada en las diferencias entre detección de **un solo objeto** y detección de **múltiples objetos**.

## 📌 Detección de Objetos (Object Detection)

La detección de objetos es un área dentro de la visión artificial y aprendizaje automático que busca identificar y localizar objetos específicos en imágenes o secuencias de video. Esta tarea combina dos elementos fundamentales:

- **Clasificación:** Determinar qué clase de objeto aparece en la imagen.
- **Localización:** Precisar la ubicación del objeto mediante rectángulos delimitadores (**bounding boxes**).

## 📌 Detección de un objeto (Single Object Detection)

La detección de un único objeto consiste en identificar y localizar un solo objeto de interés en una imagen, suponiendo que únicamente aparece un objeto relevante.

### Proceso general:

1. **Selección de características:**  
Identificación de patrones visuales específicos del objeto buscado.
2. **Clasificación:**  
Se determina si la imagen contiene o no el objeto objetivo.
3. **Localización:**  
Se indica la posición exacta del objeto mediante un rectángulo.

### Ejemplo típico:

- Detección facial en fotografías individuales.
- Detección de un automóvil específico en una escena.

### Técnicas usadas:

- Template matching (comparación de plantillas).
- Redes neuronales convolucionales (CNN).
- HOG + SVM (Histogram of Oriented Gradients + Support Vector Machine).

## 📌 **Detección de múltiples objetos (Multiple Object Detection)**

La detección de múltiples objetos es una tarea más compleja, en la cual se busca localizar y clasificar varios objetos simultáneamente en una misma imagen.

### **Proceso general:**

1. **Extracción global de características:**  
El modelo analiza la imagen completa identificando múltiples regiones potenciales.
2. **Localización simultánea:**  
El modelo genera múltiples rectángulos (**bounding boxes**) indicando posiciones y tamaños de objetos.
3. **Clasificación múltiple:**  
Cada bounding box es clasificado en una categoría específica.

### **Ejemplo típico:**

- Detección de múltiples peatones y vehículos en escenas urbanas.
- Detección de diferentes productos en anaqueles comerciales.

### **Técnicas usadas (modelos populares):**

- **YOLO (You Only Look Once):**
  - Procesa toda la imagen una sola vez.
  - Es muy rápido y eficiente para videos en tiempo real.
- **SSD (Single Shot MultiBox Detector):**
  - Similar a YOLO, eficiente y balanceado entre velocidad y precisión.
- **R-CNN, Fast R-CNN, Faster R-CNN:**
  - Técnicas basadas en propuesta de regiones que alcanzan alta precisión a costo de mayor procesamiento.

### ✚ Comparación entre Detección Única y Múltiple:

Aspecto	Detección de un objeto	Detección de múltiples objetos
Complejidad	Baja/Media	Alta
Velocidad	Alta	Media/alta (depende del modelo)
Aplicaciones comunes	Facial, reconocimiento simple	Seguridad, conducción autónoma, vigilancia, comercio
Técnicas comunes	HOG/SVM, simples	CNN YOLO, SSD, Faster R-CNN

### ✚ Consideraciones Importantes en Detección de Objetos:

- **Velocidad vs precisión:**  
Se escoge la técnica según las necesidades del proyecto. Técnicas rápidas (YOLO, SSD) frente a precisas pero más lentas (Faster R-CNN).
- **Tamaño del dataset de entrenamiento:**  
Se necesitan muchas imágenes etiquetadas para entrenar un modelo efectivo.
- **Problemas típicos:**
  - Objetos solapados
  - Cambios de iluminación
  - Diversidad de perspectivas
  - Escala variable