

Informe Proyecto 1

Analisis de ondas gravitacionales

Por: Javier Brito, *PUCV, Escuela de Ingenieria Electrica,*

Abstract—En el presente informe se mostrara detalladamente el proceso realizado para predecir la clase de una imagen, esto utilizando 3 imagenes de referencia para predecir si la imagen sera caja, circulo o cruz, para ello se utilizara como principal herramienta la convolución.

Index Terms—Clase, imágenes, predicción, convolución, etc.

I. INTRODUCCIÓN

La visión por computadora es un campo que investiga algoritmos o métodos para reconocer objetos y patrones en las imágenes. En este mismo campo se encuentra presente la idea de correlación entre imágenes, la cual puede ser utilizada para identificar patrones y relaciones entre los datos que, con un correcto uso, sirven para etiquetar objetos y clases.

La correlación entre las imágenes se centra en la idea de que clases parecidas tendrán características en común, mientras que en clases significativamente diferentes estas características diferirán rotundamente. Siguiendo esta idea, nacieron las Redes Neuronales Convolucionales (CNN), las cuales son ampliamente utilizadas para resolver problemáticas asociadas con la visión artificial. Cabe destacar que en la presente experiencia no se pueden utilizar este tipo de arquitecturas neuronales para la resolución de la problemática en cuestión.

II. MARCO TEÓRICO (FORMULAS Y TEORIA)

A. Media de la Imagen

La media de una imagen se puede calcular mediante la siguiente expresion:

$$\mu_I = \frac{1}{N} \sum_{i=1}^N I(i)$$

Donde N es el número total de píxeles en la imagen respectivamente. Con esta expresion se obtiene como resultado una medida de tendencia central la cual indica el valor promedio de los píxeles.

B. Desviación estandar de la imagen

La desviación estándar es la medida con la cual se puede representar la cantidad de variación o dispersión en un conjunto de valores. Para calcular este factor utilizaremos la siguiente expresión:

$$\sigma_I = \sqrt{\frac{1}{N} \sum_{i=1}^N (I(i) - \mu_I)^2}$$

Este documento fue producido por la Escuela de Ingenieria Electrica. PUCV, Junio de 2024

Donde :

- $I(i)$ es el valor del píxel en i.
- μ_I es la media de los valores de los pixeles con la ventana seleccionada.
- N es el número de pixeles de dicha ventana.

C. Normalización en Imagenes

La normalización de una imagen implica ajustar los valores de los pixeles para obtener una media de 0 y una desviación estándar de 1. Esto puede lograrse utilizando la siguiente expresión:

$$I_{\text{norm}}(i) = \frac{I(i) - \mu_I}{\sigma_I + 1 \times 10^{-5}}$$

Donde :

- $I(i)$ es el valor del píxel en i.
- μ_I corresponde a la media de los valores en la imagen.
- σ_I corresponde a la desviación estandar.
- 1×10^{-5} tiene como proposito evitar la división por cero.

D. Correlación Normalizada

La correlación normalizada consiste en comparar dos señales o imagenes y evaluar su similitud, en este caso utilizaremos la siguiente expresión:

$$\text{NCC}(x, y) = \frac{1}{N} \sum_{i=1}^N I_{\text{norm}}(x + i, y + i) \cdot T_{\text{norm}}(i)$$

Donde :

- N es el número total de pixeles en la ventana utilizada.
- I_{norm} corresponde a la imagen normalizada previamente.
- T_{norm} corresponde a la ventana o template utilizado.

III. METODOLOGÍA APLICADA

Pasos en orden del procesamiento de la imagen:

1. Primero: Se transforma la imagen a escala de grises manteniendo su resolución inicial, esto para que el sistema no pierda precisión.
2. Segundo: Se transformaron las imágenes a tensores.
3. Tercero: A las imágenes a etiquetar se les realizó un cálculo de la media y desviación estándar.
4. Cuarto: Se realizó la normalización de la imagen.
5. Quinto: Se realizan las mismas operaciones a las plantillas o "templates".
6. Sexto: Cálculo de la convolución normalizada entre la plantilla y la imagen de interés a clasificar.
7. Séptimo: Por último, se clasifica mediante el criterio de la imagen con mayor puntaje de correlación.

IV. RESULTADOS

1) *Carga de datos:* Se presentaran los datos del modelo con mayor precisión obtenida, el cual se logro utilizando las imagenes con una resolución de 512x512 pixeles.

Se presenta la tabla con los valores obtenidos para cada resolución respectivamente:

Resoluciones	(512x512)	(256x256)	(128x128)
Tiempo(Segundos)	247.24	148.78	86.40
Presición(Porcentaje)	0.58	0.54	0.51

TABLE I

TABLA COMPARATIVA ENTRE VALORES OBTENIDOS EN LAS DISTINTAS RESOLUCIONES

El codigo encargado de hacer la clasificacion de las clases guarda los valores obtenidos en una variable llamada "predictions", dicha variable guarda los datos de la siguiente forma:

```
('box_01_freq_reflect_09a2f2c2-ad7f-11ee-bb2a-047c16a08772_0-0_80-90.png', 'box', 'box_template', 867.48083)
('box_01_freq_reflect_2a1a6b92-ea35-11ee-b922-047c16a08772_0-513_80-90.png', 'box', 'box_template', 472.2976)
('circ_01_freq_reflect_536ce6b7-e07b-11ee-9fac-047c16a08772_0-10_80-90.png', 'circ', 'circ_template', 255.8855)
('circ_01_freq_reflect_56dee422-b317-11ee-bb58-047c16a08772_0-129_80-90.png', 'circ', 'box_template', 398.48383)
('circ_01_freq_reflect_56dee422-b317-11ee-bb58-047c16a08772_0-1344_80-90.png', 'circ', 'box_template', 247.13239)
```

Fig. 1. Impresión en jupyter notebook de la variable predictions.

2) *Obtención de la matriz de confusión:* Para esta actividad se probaron 3 resoluciones distintas para verificar cuánto influiría este parámetro en la precisión final del sistema. Para ello, se obtuvieron las matrices de confusión para cada una de esas resoluciones:

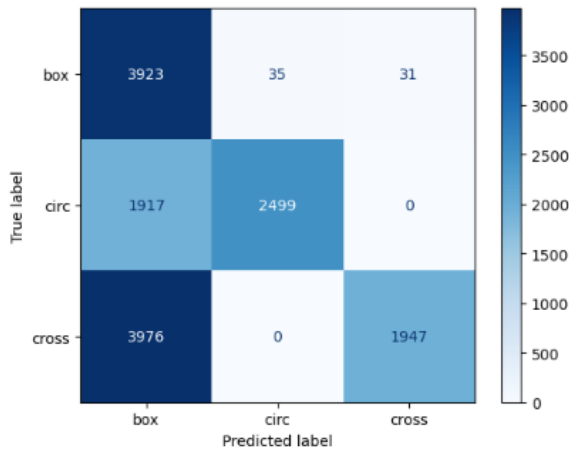


Fig. 2. Matriz de confusión para imágenes de 512x512 píxeles.

El propósito de la matriz de confusión es ilustrar cuánto se confunde el algoritmo con las clases dadas. En las tres matrices obtenidas, se puede comprobar que la clase donde más se confunde el algoritmo es con las cruces. Con esta representación, se puede apreciar que dicha clase se confunde mucho con la clase de las cajas.

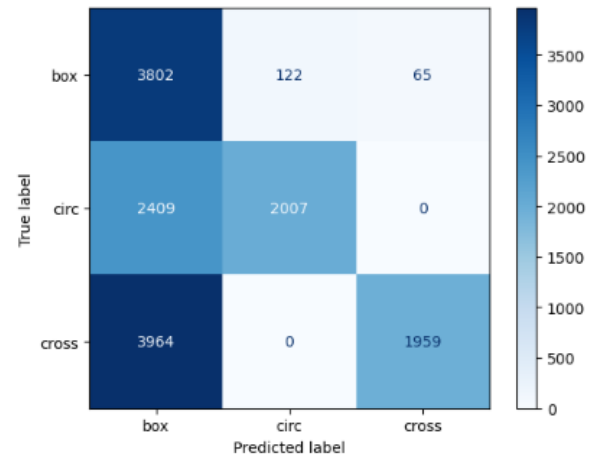


Fig. 3. Matriz de confusión para imágenes de 256x256 píxeles.

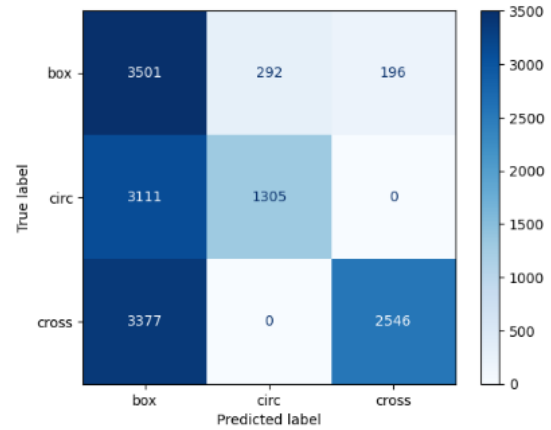


Fig. 4. Matriz de confusión para imágenes de 128x128 píxeles.

V. CONCLUSIONES

Se pudo comprobar que, mediante la variación de parámetros, la resolución de las imágenes afecta directamente a la precisión del sistema. Observando los datos obtenidos, se puede inferir que al disminuir la resolución de las imágenes, se obtenía una disminución en la precisión y el tiempo de procesamiento.

Verificando las matrices de confusión obtenidas para cada resolución, se presentó el modelo con mayor precisión, que fue el que utilizaba las dimensiones originales de las imágenes a clasificar.

Disminuyendo la resolución de la imagen, el algoritmo propuesto empieza a confundirse más con la clase de los círculos, prediciéndola como caja y con ello disminuyendo la precisión final del algoritmo.

Bibliografía:

- [1]. "Correlación - Wikipedia, la enciclopedia libre". By:Wikipedia, the free encyclopedia.
- [2]. "torch - PyTorch 2.3 documentation" By: pytorch.org.

[3]. "Transforming and augmenting images". By: pytorch.org.

[4]. "Lecture Notes: Basic Image Processing". By: Computer Science — Rice University.

Links:

<https://es.wikipedia.org/wiki/Correlaci%C3%B3n>

<https://pytorch.org/docs/stable/torch.html>

<https://pytorch.org/vision/stable/transforms.html>

[https://www.cs.rice.edu/~vo9/recognition/notebooks/
image_processing_lab.html](https://www.cs.rice.edu/~vo9/recognition/notebooks/image_processing_lab.html)