
This is the **published version** of the bachelor thesis:

Muñoz González, Marcos; Parraga, Carlos Alejandro, dir. Métrica de evaluación de simetría global en imágenes naturales basadas en valoraciones de observadores humanos. 2021. (958 Enginyeria Informàtica)

This version is available at <https://ddd.uab.cat/record/248500>

under the terms of the  license

Métrica de evaluación de simetría global en imágenes naturales basadas en valoraciones de observadores humanos

Marcos Muñoz González

Resumen– En este proyecto se estudia la correlación entre la simetría global existente en imágenes naturales con la percepción humana de las mismas mediante la creación de una métrica continua capaz de evaluar la simetría en estas imágenes. Este trabajo se divide en dos partes: una primera parte en la que se realiza un experimento para extraer valoraciones humanas respecto a la cantidad de simetría de una imagen, se extraen características de las imágenes y se crean diversos algoritmos para poder relacionar estas características con las valoraciones humanas; y una segunda parte en la que se realiza un estudio del experimento y se mejoran los algoritmos basándose en este estudio. Adicionalmente, se estudian diversas métricas ya existentes que permiten darle una valoración de cuán simétrica es a una imagen y se utilizan como modelos para poder realizar una comparación de los resultados obtenidos.

Palabras clave– simetría, algoritmo, métrica continua, imágenes naturales, red neuronal, percepción humana, experimento

Abstract– This project studies the correlation between the global symmetry that exists in natural images with the human perception of them by creating a continuous metric capable of evaluating the existing symmetry in these images. This work is divided into two parts: a first part in which an experiment is carried out to extract human evaluations regarding the amount of symmetry of an image, characteristics of the images are extracted and algorithms are created to be able to relate these characteristics with the human evaluations; and a second part in which a study of the experiment is carried out and the algorithms are improved based on this study. Additionally, various existing metrics are studied that allow an assessment of how symmetric an image is and are used as models to be able to make a comparison of the results obtained.

Keywords– symmetry, algorithm, continuous metric, natural images, neural network, human perception, experiment



1 INTRODUCCIÓN

La simetría, tan amplia o estrechamente como se pueda definir su significado, es una idea mediante la cual el hombre a lo largo de los siglos ha tratado de comprender y crear orden, belleza y perfección.

Hermann Weyl

Cuando hablamos de simetría hablamos principalmente de determinadas propiedades que posee un objeto y las cuales lo hacen característico.

La simetría es un rasgo característico de muchas entidades y es utilizada en campos muy diversos como pueden ser la música, el dibujo, la física, las matemáticas, la biología, la pintura o el modelado 3D. Si bien existen muchos tipos de simetría, si las relacionamos con el ámbito geométrico generalmente son tres los tipos que se pueden encontrar en gran cantidad de campos [1]:

- Simetría traslacional: Aquella en la que, aplicando una traslación a un objeto, por ejemplo, un posible movimiento en alguna dirección no cambia el objeto.
- Simetría rotacional: También llamada simetría radial,

● E-mail de contacto: marcos.munozg@e-campus.uab.cat
● Mención realizada: Computación
● Trabajo tutorizado por: Carlos Alejandro Parraga (Computación)
● Curs 2020/2021

es aquella en la que aplicando una rotación a un objeto este se ve igual que en su posición inicial.

- **Simetría bilateral:** La simetría bilateral, también llamada simetría de reflexión, es aquella simetría en la que un objeto puede ser dividido mediante un eje y a cada lado de este eje obtenemos una imagen espejo del otro lado.

Generalmente la simetría en imágenes naturales también puede ser dividida en dos tipos:

- **Simetría global:** Simetría relacionada con el conjunto el cual supone la imagen.
- **Simetría local:** Simetría existente en ciertas regiones de una imagen.



Fig. 1: A la izquierda imagen con gran simetría local, a la derecha imagen con gran simetría global

El hecho de que algo tenga un tipo de simetría o no y su relación con las propiedades del objeto depende también del campo del que estemos hablando y del objeto. Pues por ejemplo en la geometría, la simetría traslacional proviene de realizar una operación en un objeto de tal forma que este no esté rotado y en física puede existir en la invariancia de un sistema de ecuaciones [2].

En otros campos relacionados con el arte como pueden ser la música, el dibujo, la arquitectura o la pintura la simetría es utilizada como herramienta la cual pueden utilizar los artistas para dar impresiones y sensaciones diferentes a su público. Un ejemplo muy característico utilizado en la música existe en las Variaciones de Goldberg de Bach [3] donde el autor utilizó con frecuencia la simetría creando cánones simétricos mediante el ritmo. También existe la música simétrica o palindrómica, la cual está compuesta de canciones que son iguales si se escuchan al revés, como por ejemplo la Sinfonía n47 de Joseph Haydn [4].

La simetría es utilizada también en medicina. Se han utilizado redes neuronales siamesas para detectar asimetrías en el cerebro humano [5]. También, existen virus muy maliciosos con la capacidad de reproducirse muy rápidamente dada su estructura simétrica [6] y por este motivo el hecho de conocer que tan simétrica es una estructura de forma computacional supone una gran ayuda para el estudio de estos elementos.

1.1. Objetivos

El objetivo de este trabajo es crear una métrica de evaluación de simetría global en imágenes naturales basadas en valoraciones de observadores humanos.

Los objetivos de este trabajo son los siguientes:

1. Implementación de un experimento para recoger valoraciones de cuán simétrica es una imagen según los observadores humanos.
2. Creación de un dataset con estas valoraciones.
3. Creación de un software que permita valorar la simetría continua en una imagen que correlacione con los datos obtenidos en el dataset.
4. Creación de un sistema de visualización que permita estimar las partes de la imagen que son relevantes para el cálculo de simetría.

Para poder crear un algoritmo capaz de dar unas valoraciones parecidas a los observadores humanos y que los resultados puedan correlacionar con estas valoraciones, deberemos tener en cuenta sus características y los conocimientos e investigaciones conocidas hasta la fecha para ser capaces de obtener un algoritmo que pueda relacionar estas características con la percepción humana. La tarea será dividida en 2 fases:

- Una primera fase en la que reuniremos información inicial de las características que pueden tener relevancia a la hora de calcular la cantidad de simetría existente en una imagen, la recopilación de valoraciones categóricas de los distintos observadores realizadas en los experimentos y la creación de algoritmos capaces de dar una métrica continua de simetría a una imagen.
- Una segunda fase en la que calibraremos los algoritmos para que sean capaces de dar unas valoraciones similares a los resultados obtenidos mediante la percepción humana en el experimento y crearemos un sistema de visualización para estimar que partes de la imagen son relevantes para el cálculo de simetría.

1.2. Trabajos previos

Existen investigaciones que muestran que la percepción de la simetría es un proceso inconsciente y que el cerebro humano no destina muchos recursos para ello, haciéndolo de forma eficaz, fácil y rápida [7]. También ha sido demostrado que la percepción de la simetría humana es robusta a diferentes tipos de estrés visual tales como el difuminado, el aclarado, el oscurecimiento o el ruido en una imagen [8]. Se sabe que la percepción de la simetría aparece en la infancia a la edad de 4 meses, edad en la que los infantes ya son capaces de distinguir el tipo de simetría respecto a un eje [9][10]. También hay estudios que indican que los seres humanos son capaces de detectar la simetría bilateral más rápidamente y de forma más sencilla que otras como la rotacional o la traslacional [11][12]. Las primeras etapas del sistema visual incluyen neuronas que filtran la imagen a diferentes escalas y existe evidencia de que este filtrado espacial precede a la computación de la simetría [9][13][14]. Según ciertos estudios el sistema visual humano trabaja de forma local orientada mediante los contornos fragmentados de una imagen y la combinación de sus componentes de acuerdo con un conjunto de preferencias que describen un “campo de asociación” [15]. Otros muestran que la percepción de la simetría podría ser nada más que un conjunto de operaciones de filtrado ejecutadas en paralelo [14]. Existen

también evidencias que indican la presencia de mecanismos diferentes en la detección de simetría en la visión humana como el procesado de texturas [16]. Se conocen reportes de que el procesado de la simetría es afectado por el número de objetos percibidos y profundidad en una pantalla y de que la orientación de los ejes de la simetría tiene un efecto en cómo se percibe [9]. Estos estudios muestran que la detección humana de la simetría podría ser categorizada como un conjunto de operaciones de bajo coste que permiten el análisis de una imagen mediante su división en diferentes filtros y partes. No obstante, todavía no está claro como el cerebro humano logra percibir la simetría y los procesos neuronales responsables de ello, si bien hay estudios de los procesos involucrados como por ejemplo [16], no hay nada concluyente. En las últimas décadas los investigadores han sido capaces de obtener diferentes características inspiradas en el sistema visual humano y gracias al auge de diferentes técnicas de aprendizaje profundo cada vez se pueden llegar a crear nuevos algoritmos y herramientas para poder tratar con ella tanto en sistemas de 2 dimensiones como de 3 dimensiones [17]. Finalmente, hay que tener en cuenta que la detección de la simetría para una maquina es una tarea difícil de realizar y no pueden obtenerla con la misma eficiencia y rapidez con la que lo haría un ser humano [18].

2 MÉTRICAS EXISTENTES PARA PODER DAR UNA VALORACIÓN DE LA CANTIDAD DE SIMETRÍA EN UNA IMAGEN

Actualmente existen diferentes algoritmos capaces de dar una valoración continua de cuanta simetría tiene una imagen respecto a la percepción humana tanto a color [19]-[22] como en imágenes binarias (imágenes con solamente dos valores posibles para cada pixel) [23]. Los algoritmos que permiten valorar imágenes a color encontrados se muestran seguidamente en orden cronológico. Observamos que los algoritmos más nuevos se han desarrollado en base a algoritmos anteriores y habitualmente han sido comparados entre ellos en sus respectivos artículos.

2.1. Mediante la transformada de coseno discreta (TDC)

Este algoritmo [21] creado por Goksel Gunlu y Hasan S. Bilge es capaz de dar una métrica de como de simétrica es una imagen mediante la transformada de coseno discreta. Una de las características de este algoritmo reside en que solamente compara regiones de la imagen y que por lo tanto será propenso a dar altas valoraciones a aquellas imágenes con regiones totalmente simétricas y muy bajas en casos en los que haya mucha asimetría en esas regiones.

2.2. Mediante la intensidad

Este algoritmo [22] creado por Eelco den Heijer es capaz de dar diferentes métricas de simetría mediante la intensidad de una imagen y la entropía.

Principalmente ofrece cuatro métricas:

- AMsym1: Simetría mediante diferentes intensidades de partes de la imagen.

- AMsym2: Simetría mediante AMsym1 y una función gaussiana.
- AMsym1*: Simetría mediante AMsym1 y la entropía de la imagen.
- AMsym2*: Simetría mediante AMsym2 y la entropía de la imagen.

Hemos implementado este algoritmo y será utilizado como característica para los algoritmos, además de como modelo de comparación. Este algoritmo presenta cuatro métricas, las cuales en algunos casos correlacionan positivamente con las valoraciones humanas. Uno de los problemas de este algoritmo reside en que cada métrica puede dar valoraciones muy diferentes. La métrica ofrece valoraciones del 0 al 10.

En la tabla número 1 se muestran las distintas valoraciones de las métricas de intensidad de la imagen en la figura 2

Tabla 1: Valoraciones de la figura número 2

AMsym1	AMsym2	AMsym1*	AMsym2*
6.34	10	3.83	6.05



Fig. 2: Imagen del dataset

2.3. Mediante filtros de Gabor

Mediante este algoritmo [20] creado por Fariba Shaker y Amirhassan Monadjemi se extraen un conjunto de líneas de la imagen mediante filtros de Gabor y se realizan comparaciones entre diferentes partes de esta para computar una métrica de simetría continua. Hemos implementado este algoritmo y será utilizado como característica en los algoritmos y modelo de comparación. Uno de los problemas de este algoritmo es que depende principalmente de una serie de parámetros modificables tales como sigma, lambda, psi o gamma en los filtros de Gabor o el tamaño de las imágenes. Este algoritmo da valoraciones entre un número del 0 al 1.

2.4. Mediante la activación de filtros existentes en una red neuronal

Este algoritmo [19], [24] desarrollado por Anselm Brachmann y Christoph Redies proporciona una valoración de como de simétrica es una imagen en base a los resultados de

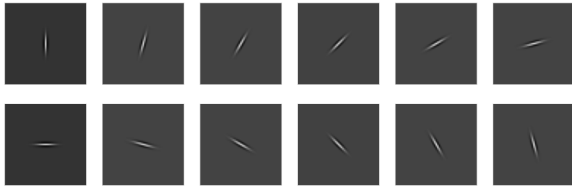


Fig. 3: Banco de filtros de Gabor utilizados para la métrica



Fig. 4: Valoración de la imagen según la métrica implementada: 0.98 (mínimo: 0, máximo: 1)

que filtros es capaz de activar está a la hora de realizar una predicción en una red neuronal convolucional (CNN) preentrenada en datasets muy grandes como puede ser CaffeNet [25]. En el artículo se muestra que es utilizado para dar una métrica de como de simétrica es una imagen en base a la simetría bilateral. Este proporciona tres métricas de simetría basadas en diferentes partes de la imagen. Las métricas tienen un rango de valores del 0 al 1. Este algoritmo lo hemos obtenido y será utilizado como modelo y métrica de comparación.

3 MÉTODOS

3.1. Etiquetado de imágenes por observadores humanos

Primeramente realizamos un etiquetado de las imágenes mediante una encuesta, de tal forma que podamos obtener una evaluación humana de las mismas. Estas imágenes serán valoradas mediante estrellas y en total habrá 5 categorías.

Hemos decidido utilizar estrellas en lugar de textos como “Muy simétrica” o “Poco simétrica” debido a que mediante el uso de iconos el observador es capaz de intuir y asociar una determinada cantidad de estrellas con su valoración.

Para la realización del experimento se ha utilizado una aplicación web creada mediante MATLAB Web Apps y alojada en un servidor MATLAB Web App Server. La encuesta podía ser realizada tanto en inglés como en castellano.

Las imágenes mostradas en el experimento han sido extraídas de Pixabay en <https://pixabay.com> y Pexels en <https://www.pexels.com/>. Todas las imágenes tienen una licencia de uso totalmente gratuito, incluso para el ámbito comercial siempre que se cumplan ciertas restricciones que no afectan a este proyecto. Se han recopilado un conjunto de 200 imágenes con un ancho de 960 píxeles, un alto mínimo de 499 píxeles y un alto máximo de 720 píxeles. Todas las imágenes tenían formato landscape. Además, se ha creado un balance compuesto de imágenes muy simétricas, poco simétricas y nada simétricas. Principalmente hemos realizado este balance mediante el uso de las categorías de las páginas mencionadas anteriormente utilizando etiquetas

como ‘Symmetry’ o ‘Asymmetry’. Las últimas separaciones en grupos han sido realizadas de forma subjetiva por nosotros.

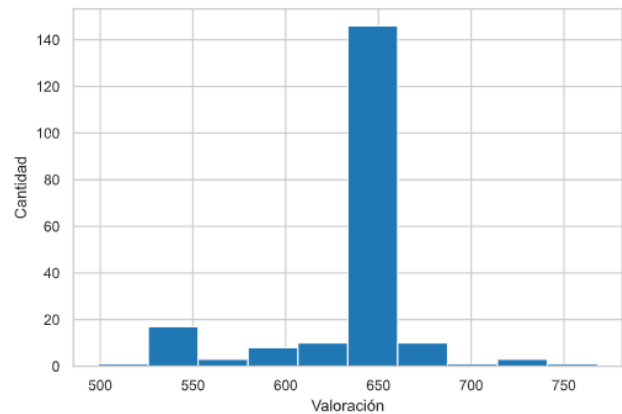


Fig. 5: Distribución de altura de las imágenes

El experimento transcurre de la manera siguiente. Primero el observador debe rellenar una serie de datos tales como la edad, el género o si tiene alguna deficiencia visual. Seguidamente se le muestran unas instrucciones sobre el experimento que va a realizar. Las imágenes son mostradas durante 0.5 segundos debido a que queremos evitar que entren en juego mecanismos cognitivos de alto nivel, ya que nuestros modelos son relativamente sencillos. También existen estudios que muestran que los humanos son capaces de detectar la simetría en imágenes en menos de 0.05 segundos [7]. Inicialmente se inicia una fase de calentamiento en la que hay 5 imágenes de prueba para que el observador se adecue al experimento, estas imágenes también serán mostradas más adelante y servirán para saber si las valoraciones del usuario coinciden con sus observaciones iniciales. Esta fase de calentamiento no se indica al observador si no que transcurre normalmente sin que lo sepa. El observador puede pausar el experimento si lo ve conveniente y continuar más adelante o indicar en caso de no haber podido ver la imagen que se distrajo durante el experimento. Las observaciones son guardadas conforme el observador realiza el experimento.

Seguidamente se filtran los datos provenientes del experimento para utilizarlos en los algoritmos. Se han eliminado aquellas encuestas con un número de votos menor a 40, las valoraciones en las que el usuario se distrajo o las valoraciones en la que en aquel momento se pausó el experimento. También se han eliminado las valoraciones dadas por los observadores cuya implicación en el experimento sea dudosa, como por ejemplo las valoraciones que difieran mucho de las de la fase de calentamiento. Las valoraciones de la fase del calentamiento también son descartadas. No todas las personas terminaron el cuestionario. Finalmente, hemos conseguido aproximadamente 4519 valoraciones las cuales provienen de 25 personas encuestadas. De estos observadores 17 eran hombres y 8 mujeres. La edad mínima ha sido de 18 años, la máxima de 61 y el promedio de edad ha sido de 37 años. Las encuestas fueron realizadas en inglés por 4 observadores y en castellano por 21. Doce de los usuarios encuestados indicaron que padecían de algún tipo de deficiencia visual.

Para comprobar la relación entre las valoraciones dadas por los observadores utilizamos el coeficiente alfa de Krippendorff que además puede ser utilizado con observadores que no han valorado algunas imágenes. Este coeficiente mide cual es el porcentaje de acuerdo entre los diferentes observadores respecto a las valoraciones dadas en el experimento y nos sirve para comprobar cuanta es la fiabilidad del experimento en una escala del 0 al 1. Este coeficiente nos da un valor menor a 0.6, por ese motivo decidimos utilizar la media truncada eliminando un 10 % de los datos de los extremos. En la figura 6 mostramos las valoraciones dadas por los observadores observando que parecen dar una distribución más o menos uniforme de las valoraciones.

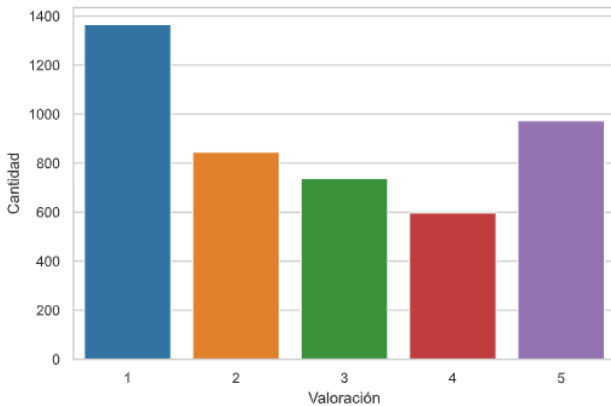


Fig. 6: Cantidad de votos y valoraciones dadas por los observadores

3.2. Algoritmos de extracción de características

Seguidamente mostramos los diferentes algoritmos utilizados para extraer características de las imágenes. Estas características serán utilizadas para estimar cuál es su correlación respecto a las valoraciones de simetría dadas en el experimento y como datos para algunos algoritmos.

3.2.1. Características extraídas mediante la comparación de lados de la imagen

Basándonos en los algoritmos que permiten dar valoraciones de simetría en imágenes naturales mediante las comparaciones entre diferentes partes de éstas [22] se han extraído características mediante la realización de ciertas comparaciones entre diferentes partes de las imágenes y componentes como pueden ser el color, las texturas o la intensidad. De esta manera se pueden obtener más características para los algoritmos.

Principalmente las partes de la imagen que vamos a comparar son 8, 4 áreas y 4 lados. Los lados serán el derecho, el izquierdo, el superior y el inferior de la imagen. La división de las áreas puede ser visualizada en la figura número 11.

Para hacer estas comparaciones se han utilizado diferentes métricas de distancia de una imagen respecto a otra tales como la distancia euclídea [26] con los histogramas de diferentes partes de la imagen.

3.2.2. Colores

Una de las características extraídas consiste en la comparación de colores de las áreas de la imagen. Para ello realizamos las comparaciones en base a la composición RGB. Primeramente, obtenemos la media del color RGB a extraer, lo cual resultara en un vector de 3 valores (R, G y B) y lo comparamos con los resultados obtenidos en la otra parte de la imagen mediante una substracción. Finalmente sumamos los valores del vector y dividimos entre 3 para obtener la media. De esa manera tenemos la diferencia de color existente entre una parte y otra.

3.2.3. Texturas mediante Local Binary Patterns

Una de las formas de conseguir extraer las texturas en una imagen es mediante el operador Local Binary Patterns (LBP) [27]. Este operador puede ser utilizado para obtener vectores de características y clasificar texturas. En este caso, para comparar las diferentes texturas existentes en la imagen hemos convertido la imagen a blanco y negro y hemos creado el histograma resultante de aplicar LBP en las diferentes partes de la imagen. Seguidamente hemos utilizado la distancia euclidiana para comparar los histogramas.

La distancia euclidiana [28] viene dada por

$$\|\mathbf{u} - \mathbf{v}\|_2$$

Dónde

u : Colección de entrada

v : Colección de salida

3.2.4. Estructura mediante Structural Similarity Index Measure

El Structural Similarity Index Measure (SSIM) [29] es una métrica que puede ser utilizada para medir la similitud entre dos imágenes. Esta tiene su fortaleza en dar más peso a aquellos píxeles que son espacialmente más cercanos.

La fórmula utilizada viene dada por

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

Dónde

μ : Media

σ : Varianza

c_1, c_2 : Variables de estabilización

3.2.5. Probabilidades mediante la distancia Jensen-Shannon

La distancia Jensen-Shannon [30] es un método utilizado para medir la similitud de dos distribuciones de probabilidad. Además, con esta distancia obtendremos un valor finito del 0 al 1. Se trata de una extensión de la divergencia Kullback-Leibler [31] que mediante la puntuación de divergencia de una distribución de probabilidad a otra obtiene una puntuación simétrica y su distancia.

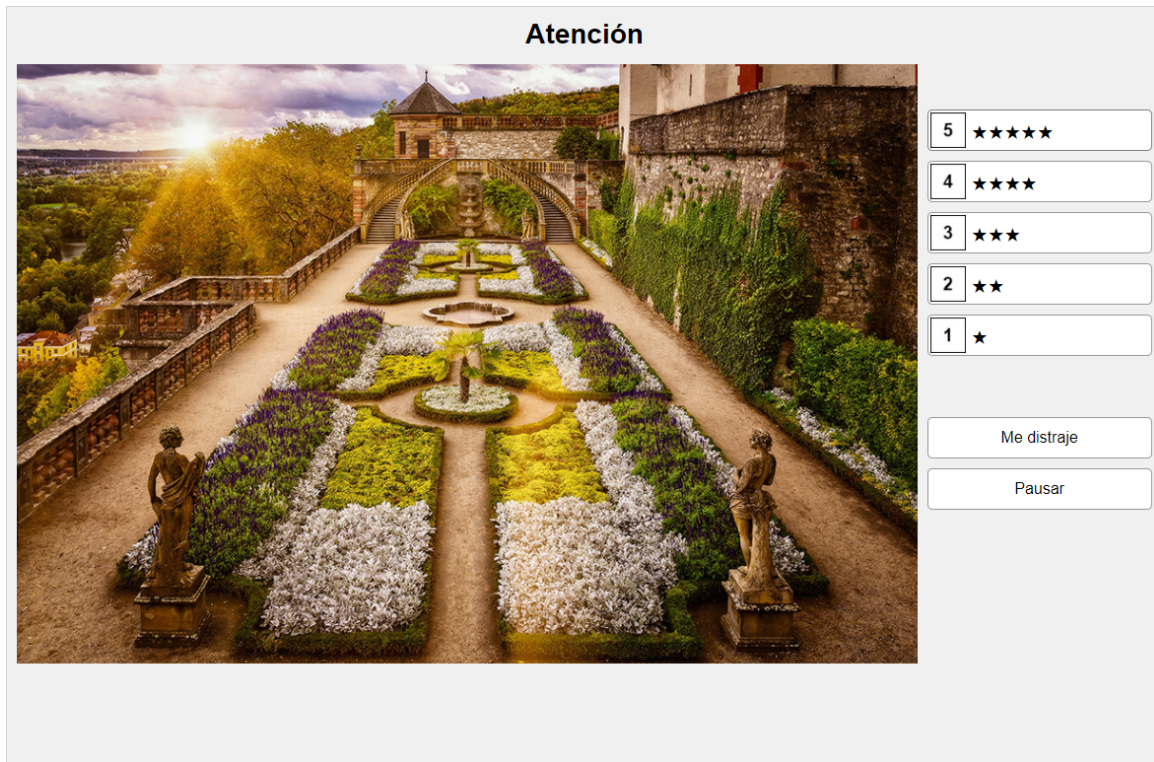


Fig. 7: El experimento en el escritorio

La fórmula utilizada viene dada por

$$\sqrt{\frac{D(p||m) + D(q||m)}{2}}$$

Dónde

p : Vector de probabilidad

q : Vector de probabilidad

m : Media del producto de p y q

D : Divergencia Kullback-Leibler

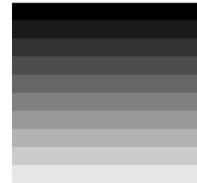


Fig. 8: Delentropy: 0.19 - Entropía skimage: 3.45

3.3. Características relacionadas con la simetría

Otras características que hemos extraído de las imágenes son aquellas relacionadas con el orden como pueden ser la entropía o la cuantificación del orden.

3.3.1. Delentropy

Actualmente las implementaciones de MATLAB y algunas librerías de Python como scikit-image no ofrecen un valor intuitivo de cómo es la entropía en una imagen. Esta extensión de la entropía de Shannon llamada Delentropy [32] ofrece una métrica más intuitiva en imágenes y es llamada así por el operador Del utilizado para la computación de esta.

En la figura 8 y 9 mostramos respectivamente las diferencias entre las computaciones de la entropía en una imagen con cierto orden y una imagen creada mediante ruido.

3.3.2. Cuantificación del orden

Este algoritmo [33] permite dar una valoración de cuanto orden existe en una imagen con una serie de puntos utilizando entre otras cosas la triangulación de Delaunay. Este

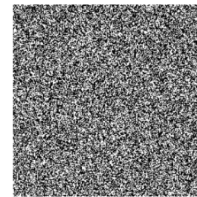


Fig. 9: Delentropy: 5.49 - Entropía skimage: 4.82

algoritmo tiene una implementación creada por los autores en MATLAB disponible. Para utilizarlo ha sido necesario transformar la imagen en una nube de puntos como en [34]. Principalmente hemos obtenido dos métricas con diferentes nubes de puntos mediante los contornos de la imagen con Sobel y mediante las características principales con KAZE [35]. En la figura 10 mostramos las transformaciones de una imagen para su cuantificación del orden. La imagen con Sobel da un valor de 7.53, mientras que la imagen con KAZE da un valor de 0.14. La gran diferencia respecto a la valoración reside en la forma de obtener la nube de puntos. Con Sobel hemos extraído los bordes y eliminado aquellos puntos cuya intensidad no fuera mayor a un threshold determinado. Con KAZE hemos obtenido los puntos clave y hemos posicionado las coordenadas mediante la sustitución en blanco de la coordenada del punto clave en la imagen.



Fig. 10: Transformación de una imagen para la cuantificación del orden

Tabla 2: Correlación respecto a las valoraciones del dataset de las características extraídas

Algoritmo	Basado en	Correlación
GB NN LR	Activaciones CNN (2.4)	0.73
GB NN MEAN	Activaciones CNN (2.4)	0.52
GB NN TB	Activaciones CNN (2.4)	0.45
JensShann LR	Probabilidad (3.2.5)	0.45
Delentropy LR	Entropía (3.3.1)	0.38
JensShann MEAN	Probabilidad (3.2.5)	0.35
Gabor Symmetry	Filtros de Gabor (2.3)	0.34
TextDiff LR	Texturas (3.2.3)	0.32
TextDiff MEAN	Texturas (3.2.3)	0.32
JensShann D	Probabilidad (3.2.5)	0.32
TextDiff D	Texturas (3.2.3)	0.29
Color LR	Colores (3.2.2)	0.29
StructSym D	Similitud Estructural (3.2.4)	0.26
Color MEAN	Colores (3.2.2)	0.26
TextDiff TB	Texturas (3.2.3)	0.26
Delentropy	Entropía (3.3.1)	0.25
JensShann TB	Probabilidad (3.2.5)	0.20

3.3.3. Otras características

Otras características que hemos extraído de las imágenes son la media, la implementación de la entropía de Shannon de OpenCV, el número de Euler, la desviación estándar, la varianza y la mediana. Hemos podido extraer estas características fácilmente gracias a las librerías de Python. Las características que no han podido ser extraídas de la imagen ya sea por su dificultad, porque no hemos tenido acceso al código de los artículos pertinentes o porque no había recursos suficientes para poder añadirlas han sido las siguientes:

- Las formas geométricas existentes en la imagen: Esta característica es muy difícil de obtener en imágenes de objetos naturales, puesto que hay mucha variación de elementos.
- Cantidad de esfuerzo necesario para crear una forma a su forma más simétrica: Como las formas geométricas existentes en la imagen, esta característica no es fácilmente extraíble en imágenes naturales y por ello hemos decidido no implementarla.

3.4. Correlación entre valoraciones filtradas y características

Mostramos en la tabla 2 la correlación absoluta existente entre las valoraciones filtradas y las características obtenidas en esta fase. También mostramos en que están basadas las características y su respectiva sección entre paréntesis. No mostramos aquellas características con una correlación menor a 0.2.

Algunas de las características mostradas en este mapa de calor como pueden ser Gabor Symmetry, Strict Symmetry,

Strict Symmetry Liveliness, Relaxed Symmetry Liveliness o Relaxed Symmetry son características extraídas de algoritmos ya creados para poder dar una métrica de como de simétrica es una imagen [19], [22] y las hemos utilizado debido a la posibilidad de que la combinación de las mismas pueda dar una correlación aun mayor con las valoraciones dadas por los observadores.

En la tabla 2 observamos que existe una gran correlación entre las características de Gabor Symmetry, la distancia entre dos distribuciones de probabilidad, el color y Delentropy con la valoración de simetría dada por los seres humanos.

Algunas extensiones de las características vienen dadas por las partes de la imagen utilizadas para extraerlas. Por ejemplo, Color LR es la diferencia de color existente entre la parte izquierda y la parte derecha de la imagen. Las extensiones de algunas características vienen dadas por:

- Left-Right (LR): Comparaciones entre el lado izquierdo y el lado derecho de la imagen.
- Top-Bottom (TB): Comparaciones entre el lado de arriba y el lado de abajo de la imagen.
- Diagonal (D): Substracción de las comparaciones entre área 1 y área 4 con área 2 y área 3 mostradas en la tabla 2.
- Mean (MEAN): Media entre la característica LR de la imagen, la característica TB y la característica D.



Fig. 11: Diversas áreas de la imagen

De cara al algoritmo, las características han sido escaladas. Por ejemplo aquellas métricas que daban un valor del 0 al 1, ahora dan un valor del 1 al 10 para poder dar un valor más relacionado con el de los observadores.

4 ALGORITMOS

Con el objetivo de encontrar el algoritmo que más correlacione con las valoraciones dadas por los observadores hemos creado diversas redes neuronales. Estas redes neuronales son utilizadas como modelos de base cuya precisión puede ser mejorada. Estos algoritmos serán comparados con los resultados de los algoritmos desarrollados con algunas de las métricas actuales para darle una valoración a una imagen.

Existen otros tipos de redes neuronales que podrían ser probadas para dar valoraciones respecto a la simetría en una imagen más parecidas al sistema visual humano como por ejemplo LGN-CNN [36]. También, ya que los filtros iniciales de las redes neuronales entrenados en datasets muy grandes suelen estar compuestos de filtros de Gabor, [19]

es posible que las GaborNET [37], redes neuronales cuyos filtros iniciales son filtros de Gabor puedan dar buenos resultados en este aspecto.

Para poder valorar los algoritmos hemos utilizado validación cruzada con 5 iteraciones y hemos dividido el conjunto de 200 imágenes valoradas por los observadores en dos conjuntos:

- Conjunto de entrenamiento: Compuesto por 160 muestras.
- Conjunto de validación: Compuesto por 40 muestras.

Como métrica para valorar la precisión del algoritmo utilizamos la siguiente formula:

$$f(h, m) = \begin{cases} \text{correcta} & : m \geq h - 1, m \leq h + 1 \\ \text{incorrecta} & : f(h, m) \notin \{\text{correcta}\} \end{cases}$$

Dónde

h : Valoración humana

m : Valoración del algoritmo

El objetivo de utilizar esta fórmula como métrica de precisión en lugar de la utilizada para la función de pérdida Mean Absolute Percentage Error (MAPE) es comprobar si los algoritmos implementados serían capaces de dar una valoración muy similar a los observadores de forma cuantitativa. Para todos los algoritmos relacionados con redes neuronales el tamaño del lote (batch size) ha sido 32.

Mostramos en la tabla 3 la precisión de los algoritmos utilizando la formula mostrada anteriormente, los algoritmos han sido renombrados para facilitar su visualización en la tabla de la siguiente manera:

- InceptionResNetV2: Red neuronal InceptionResNetV2 entrenada en ImageNet [38] congelada, pero con las capas de normalización (batch normalization) entrenables. Hemos añadido la capa 'Global Average Pooling 2D' y activación 'SELU' como cabezas. Además, se ha realizado aumento de datos mediante la modificación aleatoria del brillo.
- Gabor Symmetry NN [LR, TB o MEAN]: Métricas obtenidas mediante las activaciones de la red neuronal [19].
- Gabor Symmetry: Métrica obtenida utilizando filtros de Gabor del algoritmo [20] implementada por nosotros.
- Multilayer perceptron (MLP): Red neuronal artificial (ANN) que utiliza las características extraídas mostradas anteriormente.
- Red Neuronal Convolutiva (CNN): Red Neuronal Convolutiva simple.

Finalmente se entrenó ResNet-50 preentrenada en ImageNet [38] congelando todas las capas excepto las de normalización (batch normalization) y se añadieron la capa 'Global Average Pooling 2D' y activación 'ReLU' como cabezas. Esta red fue entrenada haciendo validación cruzada con 200 epochs por cada k, de esta manera en cada

iteración la red neuronal tenía 160 imágenes de las cuales algunas no había visto durante la iteración anterior. La red neuronal en ningún momento entrenaba desde el principio si no que entrenaba con los pesos de la iteración anterior. El objetivo era ver donde se estaba centrando la red hasta la parte convolutiva para hacer las predicciones mediante la creación de mapas de calor obtenidos mediante la suma de las salidas de la última capa convolutiva de ResNet-50 al predecir las imágenes del entrenamiento. Los resultados pueden ser observados en el apéndice A.1. Estos resultados son prometedores pues muestran que la red neuronal sí que está utilizando las simetrías de la imagen para hacer predicciones, por lo menos hasta la parte convolutiva. Curiosamente observamos también que la red neuronal está centrándose en los ojos de los animales para hacer las predicciones, pensamos que esto se debe a que es la característica más simétrica de estos animales. Por último, intentamos ver si la red neuronal InceptionResNetV2 también estaba centrándose en simetrías de la imagen para hacer las predicciones pero esta vez en las imágenes de validación no vistas por el algoritmo ya que es el algoritmo con mayor precisión y si bien se observa que en algunos casos esta centrándose en simetrías de la imagen como ResNet-50 en el entrenamiento, no acaba de encontrar estas simetrías. Atribuimos este hecho a la poca cantidad de datos que no permiten a la red neuronal generalizar. Los resultados de InceptionResNetV2 pueden ser observados en el apéndice A.1.6.

5 RESULTADOS

La red neuronal InceptionResNetV2 es la que da mejores aproximaciones en cuanto a precisión. Puede ser un algoritmo con gran margen de crecimiento dado que hay sobreentrenamiento (overfitting) y creemos que esto se debe a la poca cantidad de datos. Se han probado técnicas para reducir el overfitting tales como Spatial Dropout [39] o aumento de datos, pero en los experimentos ninguna ha dado resultados. También según lo visto en la suma de las salidas de la última capa convolutiva las redes neuronales con transferencia de aprendizaje parecen estar centrándose en elementos simétricos de las imágenes para dar sus valoraciones. Esperamos que en caso de obtener más datos este algoritmo pueda ser validado más profundamente. La red neuronal MLP da una valoración con más precisión que las características extraídas, por lo tanto, podemos observar que la combinación del conjunto de características con un modelo de red neuronal ha dado buenos resultados. En cuanto a la red neuronal convolutiva y los diferentes algoritmos, no han dado por sí solos una buena precisión. Pero, el hecho de utilizarlos con nuevas técnicas y estructuras (como en el caso de la transferencia de aprendizaje con InceptionResNetV2) o fusionando las características ha permitido que los resultados sean mejorados.

6 CONCLUSIÓN

En este proyecto hemos realizado un experimento con el cual hemos creado un dataset con valoraciones de como de simétrica es una imagen según la percepción humana. Hemos recopilado diversas métricas que permiten dar una va-

Tabla 3: Rendimiento de los algoritmos implementados

Algoritmo	Epochs	Entrenamiento	Validación	Entrenamiento (MAPE)	Validación (MAPE)
InceptionResNetV2	80	90.87 %	64.00 %	07.53 %	19.74 %
MLP	2000	67.37 %	53.00 %	16.29 %	22.29 %
CNN	80	52.75 %	49.00 %	28.71 %	29.93 %
GB NN LR	-	-	47.50 %	-	24.51 %
GB NN MEAN	-	-	42.50 %	-	29.42 %
GB NN TB	-	-	41.50 %	-	31.54 %
Gabor Symmetry	-	-	12.00 %	-	77.66 %

loración de como de simétrica es una imagen y hemos estudiado su correlación con las valoraciones obtenidas en el experimento. Finalmente hemos creado diversos algoritmos los cuales utilizan características extraídas de las imágenes para dar valoraciones de como de simétrica es una imagen y hemos comparado la precisión de los diferentes algoritmos en el dataset con validación cruzada. La insuficiencia en la cantidad de datos puede haber sido un factor determinante a la hora de medir el rendimiento de estos algoritmos, no obstante, se espera que como con todo modelo de aprendizaje de máquina a mayor cantidad de datos más precisión puedan tener estos algoritmos. Las aportaciones de este proyecto son las siguientes:

- Un software creado con MATLAB Web Apps que permite realizar un experimento para recoger valoraciones de como de simétrica es una imagen según los observadores humanos.
- Un dataset con valoraciones de como de simétrica es una imagen según los observadores humanos.
- Un software que permite dar valoraciones de simetría continuas en una imagen parecidas a las obtenidas en el dataset y que permite visualizar donde se centra en la parte convolucional la red neuronal InceptionResNetV2 para dar sus predicciones.

AGRADECIMIENTOS

Quiero agradecer a mi tutor, Alejandro, por su gran labor realizada durante este proyecto. Ha sido un tutor excepcional y me ha ayudado y brindado muchas oportunidades. También, quiero agradecer a mi familia y a mis amigos por estos 4 años en los que me han motivado y ayudado siempre que lo he necesitado, en especial a mi madre, por todos estos años en los que nunca me ha dejado de apoyar.

REFERENCIAS

- [1] *Rotación, traslación y simetría (o reflexión) - Escuelas Pías Valencia. Matemáticas*, <https://sites.google.com/site/escuelaspiascarniceros/home/geometria/rotacion-traslacion-y-simetria>, (Accedido en 04/27/2021).
- [2] *Simetrías en física. Parte I: Qué son, cómo se describen y qué tipos hay – Física Tabú*, <https://fisicatabu.com/simetrías-en-física-parte-i-que-son-como-se-describen-y-que-tipos-hay/>, (Accedido en 05/30/2021).
- [3] O. Jander, «Rhythmic Symmetry in the "Goldberg Variations"», *The Musical Quarterly*, vol. 52, n.º 2, págs. 204-208, 1966, ISSN: 00274631, 17418399. dirección: <http://www.jstor.org/stable/741037>.
- [4] Wikipedia, *Sinfonía n.º 47 (Haydn) — Wikipedia, La enciclopedia libre*, [Internet; descargado 25-abril-2021], 2020. dirección: [https://es.wikipedia.org/w/index.php?title=Sinfon%C3%ADa_n.%C2%BA_47_\(Haydn\)&oldid=127349939](https://es.wikipedia.org/w/index.php?title=Sinfon%C3%ADa_n.%C2%BA_47_(Haydn)&oldid=127349939).
- [5] C.-F. Liu, S. Padhy, S. Ramachandran, V. X. Wang, A. Efimov, A. Bernal, L. Shi, M. Vaillant, J. T. Ratnathar, A. V. Faria, B. Caffo, M. Albert y M. I. Miller, «Using deep Siamese neural networks for detection of brain asymmetries associated with Alzheimer's Disease and Mild Cognitive Impairment», *Magnetic Resonance Imaging*, vol. 64, págs. 190-199, 2019, Artificial Intelligence in MRI, ISSN: 0730-725X. DOI: <https://doi.org/10.1016/j.mri.2019.07.003>.
- [6] S. Baron, *Medical Microbiology. 4th edition*. Galveston (TX): University of Texas Medical Branch at Galveston, 1996.
- [7] *Perception, symmetry of art discussed at brain lecture (03/22/02)*, <https://reporter.newsarchive.vumc.org/index.html?ID=2007>, (Accedido en 06/20/2021).
- [8] I. Atadjanov y S. Lee, «Robustness of Reflection Symmetry Detection Methods on Visual Stresses in Human Perception Perspective», *IEEE Access*, vol. 6, págs. 63 712-63 725, oct. de 2018. DOI: 10.1109/ACCESS.2018.2876882.
- [9] M. S. Treder, «Behind the Looking-Glass: A Review on Human Symmetry Perception», *Symmetry*, vol. 2, n.º 3, págs. 1510-1543, jul. de 2010. DOI: 10.3390/sym2031510.
- [10] V. Giannouli, *VISUAL SYMMETRY PERCEPTION. ENCEPHALOS* 50, 31-42, 201.
- [11] «Symmetry detection by categorization of spatial phase, a model», *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 263, n.º 1366, págs. 105-110, ene. de 1996. DOI: 10.1098/rspb.1996.0017.
- [12] A. D. Makin, M. M. Wilton, A. Pecchinenda y M. Bertamini, «Symmetry perception and affective responses: A combined EEG/EMG study», *Neuropsychologia*, vol. 50, n.º 14, págs. 3250-3261, 2012, ISSN: 0028-3932. DOI: <https://doi.org/10.1016/j.neuropsychologia.2012.10.003>.

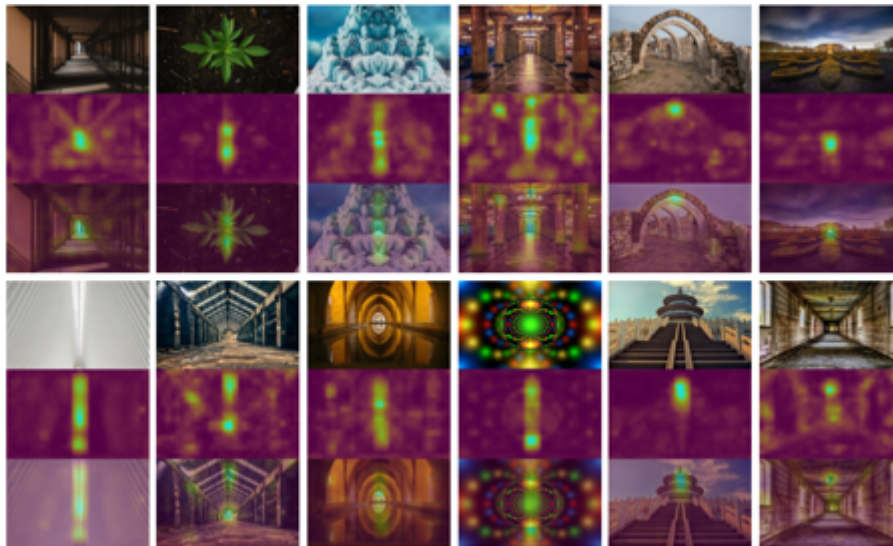
- [13] S. J. M. Rainville y F. A. A. Kingdom, «Spatial-scale contribution to the detection of mirror symmetry in fractal noise,» *Journal of the Optical Society of America A*, vol. 16, n.º 9, pág. 2112, sep. de 1999. DOI: 10.1364/josaa.16.002112.
- [14] J. Wagemans, «Characteristics and models of human symmetry detection,» *Trends in Cognitive Sciences*, vol. 1, n.º 9, págs. 346-352, dic. de 1997. DOI: 10.1016/s1364-6613(97)01105-4.
- [15] R. Hess, A. Hayes y D. Field, «The roles of polarity and symmetry in the perceptual grouping of contour fragments,» *Spatial Vision*, vol. 13, n.º 1, págs. 51-66, 2000. DOI: 10.1163/156856800741018.
- [16] H. R. Wilson y F. Wilkinson, «Symmetry perception: a novel approach for biological shapes,» *Vision Research*, vol. 42, n.º 5, págs. 589-597, mar. de 2002. DOI: 10.1016/s0042-6989(01)00299-1.
- [17] T. Suk y J. Flusser, «Recognition of Symmetric 3D Bodies,» *Symmetry*, vol. 6, n.º 3, págs. 722-757, 2014, ISSN: 2073-8994. DOI: 10.3390/sym6030722.
- [18] C. Funk e Y. Liu, «Symmetry reCAPTCHA,» jun. de 2016. DOI: 10.1109/CVPR.2016.558.
- [19] A. Brachmann y C. Redies, «Using Convolutional Neural Network Filters to Measure Left-Right Mirror Symmetry in Images,» *Symmetry*, vol. 8, n.º 12, 2016, ISSN: 2073-8994. DOI: 10.3390/sym8120144.
- [20] F. Shaker y A. Monadjemi, «A new symmetry measure based on Gabor filters,» en *2015 23rd Iranian Conference on Electrical Engineering*, 2015, págs. 705-710. DOI: 10.1109/IranianCEE.2015.7146305.
- [21] G. Gunlu y H. S. Bilge, «Symmetry analysis for 2D images by using DCT coefficients,» en *2009 Fifth International Conference on Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control*, 2009, págs. 1-4. DOI: 10.1109/ICSCCW.2009.5379480.
- [22] E. den Heijer, «Evolving Symmetric and Balanced Art,» en *Computational Intelligence: International Joint Conference, IJCCI 2012 Barcelona, Spain, October 5-7, 2012 Revised Selected Papers*, K. Madani, A. D. Correia, A. Rosa y J. Filipe, eds. Cham: Springer International Publishing, 2015, págs. 33-47, ISBN: 978-3-319-11271-8. DOI: 10.1007/978-3-319-11271-8_3.
- [23] A. I. Grenov, A. V. Tuzikov y O. V. Krivonos, «Symmetry measure computation for binary images,» en *Nonlinear Image Processing VI*, E. R. Dougherty, J. T. Astola, H. G. Longbotham, N. M. Nasrabadi y A. K. Katsaggelos, eds., International Society for Optics y Photonics, vol. 2424, SPIE, 1995, págs. 394-405.
- [24] C. Redies, *Python code to calculate variances, symmetry and self-similarity based on low-level CNN filter responses*, ago. de 2019. dirección: osf.io/xb983.
- [25] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama y T. Darrell, *Caffe: Convolutional Architecture for Fast Feature Embedding*, 2014. arXiv: 1408.5093 [cs.CV].
- [26] Wikipedia, *Distancia euclidiana* — *Wikipedia, The Free Encyclopedia*, <http://es.wikipedia.org/w/index.php?title=Distancia%20euclidiana&oldid=129997920>, [Accedido en 04/25/2021], 2021.
- [27] K.-C. Song, Y.-H. YAN, W.-H. CHEN y X. Zhang, «Research and Perspective on Local Binary Pattern,» *Acta Automatica Sinica*, vol. 39, págs. 730-744, jun. de 2013. DOI: 10.1016/S1874-1029(13)60051-8.
- [28] *scipy.spatial.distance.euclidean* — *SciPy v1.6.2 Reference Guide*, <https://docs.scipy.org/doc/scipy/reference/generated/scipy.spatial.distance.euclidean.html>, (Accedido en 04/25/2021).
- [29] Wikipedia, *Structural similarity* — *Wikipedia, The Free Encyclopedia*, <http://en.wikipedia.org/w/index.php?title=Structural%20similarity&oldid=1017571671>, (Accedido en 04/25/2021), 2021.
- [30] —, *Jensen–Shannon divergence* — *Wikipedia, The Free Encyclopedia*, <http://en.wikipedia.org/w/index.php?title=Jensen%20divergence&oldid=1018225207>, (Accedido en 04/25/2021), 2021.
- [31] *Kullback-Leibler (KL) Divergence and Jensen-Shannon Divergence*, <https://yongchaohuang.github.io/2020-07-08-kl-divergence/>, (Accedido en 04/25/2021).
- [32] K. G. Larkin, *Reflections on Shannon Information: In search of a natural information-entropy for images*, 2016. arXiv: 1609.01117 [cs.IT].
- [33] E. D. Protonotarios, B. Baum, A. Johnston, G. L. Hunter y L. D. Griffin, «An absolute interval scale of order for point patterns,» *Journal of The Royal Society Interface*, vol. 11, n.º 99, pág. 20140342, oct. de 2014. DOI: 10.1098/rsif.2014.0342.
- [34] M. Burch, J. van Lith, N. van de Waterlaet y J. van Winden, «Voronoi,» en *Proceedings of the 13th International Symposium on Visual Information Communication and Interaction*, ACM, dic. de 2020. DOI: 10.1145/3430036.3430043.
- [35] P. F. Alcantarilla, A. Bartoli y A. J. Davison, «KAZE Features,» en *Computer Vision – ECCV 2012*, Springer Berlin Heidelberg, 2012, págs. 214-227. DOI: 10.1007/978-3-642-33783-3_16.
- [36] F. Bertoni, G. Citti y A. Sarti, *LGN-CNN: a biologically inspired CNN architecture*, 2020. arXiv: 1911.06276 [cs.NE].
- [37] A. Alekseev y A. Bobe, *GaborNet: Gabor filters with learnable parameters in deep convolutional neural networks*, 2019. arXiv: 1904.13204 [cs.CV].
- [38] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li y L. Fei-Fei, «Imagenet: A large-scale hierarchical image database,» en *2009 IEEE conference on computer vision and pattern recognition*, Ieee, 2009, págs. 248-255.
- [39] J. Tompson, R. Goroshin, A. Jain, Y. LeCun y C. Bregler, *Efficient Object Localization Using Convolutional Networks*, 2015. arXiv: 1411.4280 [cs.CV].

APÉNDICE

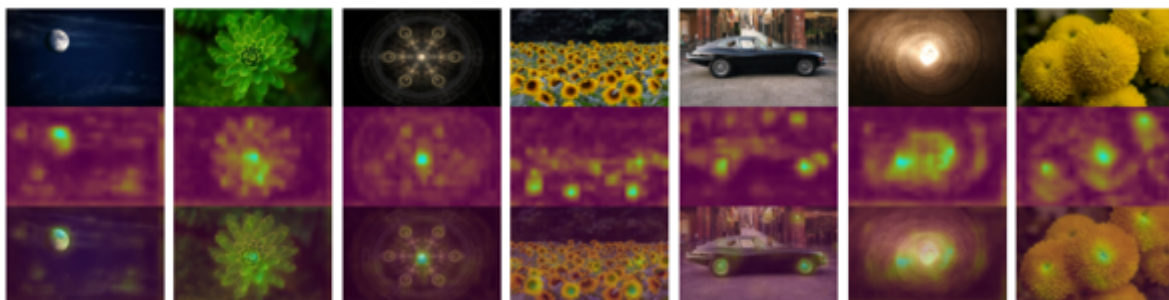
A.1. Mapas de calor

Se muestran seguidamente los mapas de calor obtenidos mediante la suma de las salidas de la última capa convolucional al hacer las predicciones de los datos de entrenamiento en la red neuronal ResNet-50 con transferencia de aprendizaje y entrenamiento de las capas de normalización de lotes. La penúltima subsección muestra más mapas de calor que no han sido etiquetados por nosotros y la última muestra las predicciones dadas por la red neuronal InceptionResNetV2 entrenada por nosotros al realizar la predicción de los datos de validación, mostrando también como a veces es capaz de centrarse en características simétricas en datos no vistos anteriormente. Cuanto más fluorescente es el color más se está centrando la red neuronal en la última capa convolucional en esa parte de la imagen al hacer la predicción, pues existen en esa parte más mapas de características.

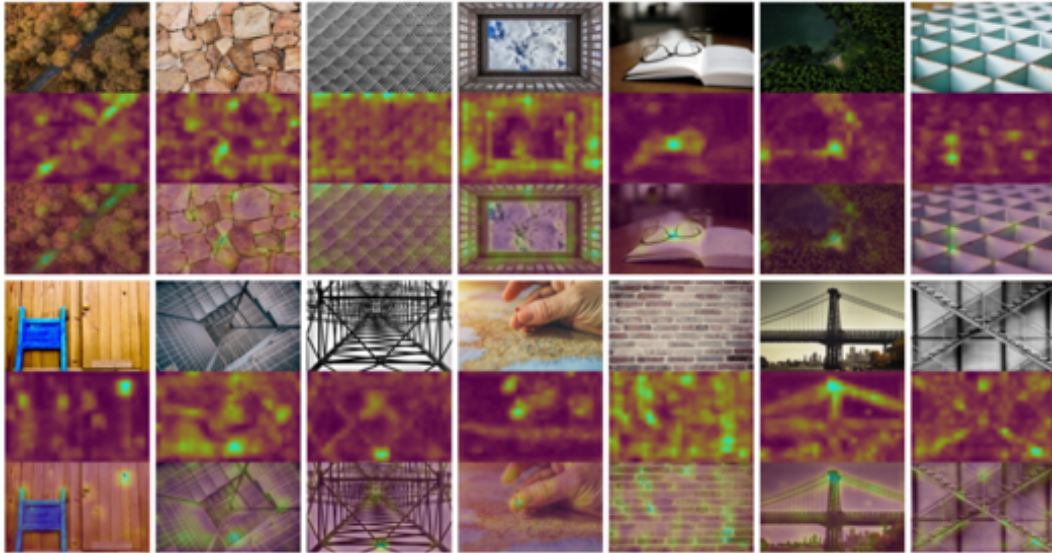
A.1.1. Simetría bilateral



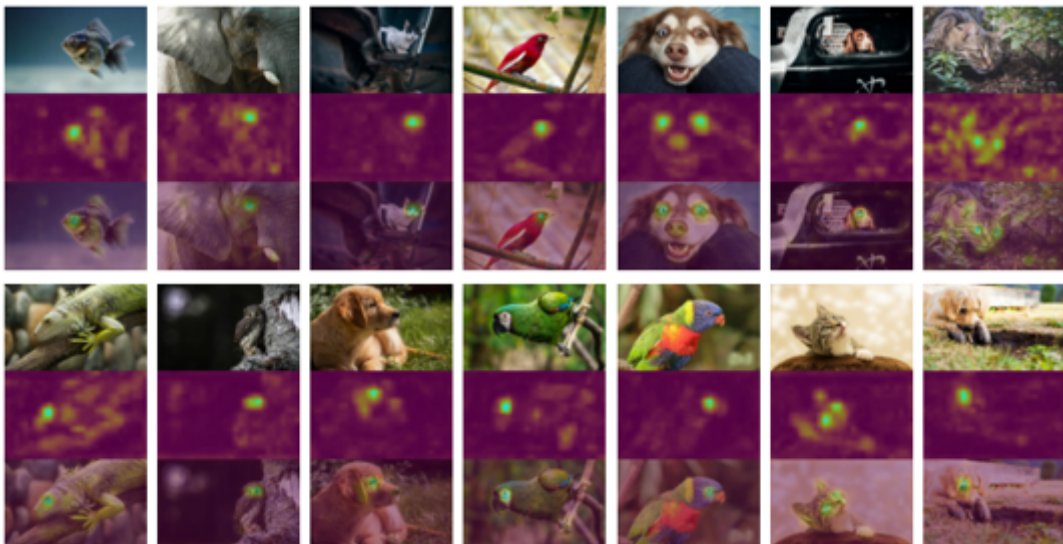
A.1.2. Simetría radial



A.1.3. Simetría compleja



A.1.4. Simetría mediante los ojos de los animales



A.1.5. Más mapas de calor del entrenamiento



A.1.6. Mapas de calor de las predicciones en validación de InceptionResNetV2

