

Trabajo Parcial - Computación Gráfica

Food Image Recognition and Modelling through Neural Networks - Reconocimiento y Modelamiento de Comida mediante Redes Neuronales

Authors: Franco Moloche, Daniel Núñez, Sergio Villarruel

Abstract— To help computers recognize the complexity of images, in this , we have used image sorter techniques using Convolutional Neural Networks and 3D modelled the closest prediction of that image. The solution worked generally well with basic images, however, we identified flaws due to automation of the image searching and due to working with similar objects (fruits share certain shapes and colours). These decisions increased AI confusion. In the future, the AI should have more training time the models being analyzed should be changed.

Para ayudar a las computadoras a reconocer la complejidad de las imágenes, en este caso de comida, hemos utilizado técnicas de clasificadores de imágenes utilizando Redes Neuronales Convolucionales y hemos modelado en 3D la predicción más acercada de dicha imagen. La solución funcionó generalmente bien con imágenes básicas. Identificamos fallas debido a la automatización en la búsqueda de imágenes y al trabajar con objetos similares, decisiones que incrementaron la confusión de la IA sea por deformidades o colores distintos. En un futuro, se debería realizar un mayor tiempo de entrenamiento o cambiar los modelos que se analizan.

I. INTRODUCTION - INTRODUCCIÓN

Computers don't understand about physical objects in an image, they can only see and process the values of every pixel contained in the image nor do they understand dimensions in physical objects. Because of this, it's our intent to find a solution to the lack of image interpretation in computers.

Las computadoras no entienden de objetos físicos en una imagen, estas sólo pueden ver y procesar los valores de cada píxel de una imagen. Tampoco tienen el entendimiento sobre las dimensiones de un objeto físico, debido a eso, nuestra intención es lograr encontrar una solución a esta carencia de interpretación por parte de las computadoras.

In this paper, we will detail the implementation process required to link an image recognition application (which uses Artificial Intelligence with Neural Networks) with a graphical application that will allow us to render the results in 3D using the OpenGL library.

El presente informe dará a conocer el proceso que hemos llevado a cabo para poder implementar y conectar una aplicación de reconocimiento de imágenes, que usa inteligencia artificial con redes neuronales, con una aplicación gráfica que nos permitirá mostrar los resultados anteriormente calculados con renderizados en 3D utilizando la librería gráfica OpenGL.

We believe this is an interesting idea to tackle, since it will greatly increase our knowledge on the subject and will serve future generations of students to be able to develop and understand the same problem.

Nos parece una idea bastante interesante de abordar, ya que incrementará ampliamente nuestros conocimientos sobre el tema y servirá a futuras generaciones de estudiantes a poder desarrollar y comprender el mismo problema.

In order to solve the problem, we decided to investigate previous publications that have addressed certain points of our problem. These studies are specified in the State of the Art section of the report.

Para poder llegar a resolver el problema, que nos hemos planteado, decidimos investigar sobre anteriores trabajos que hayan abordado ciertos puntos del problema a resolver. Estos estudios están especificados con mayor precisión en el Estado del Arte del informe.

II. THEORETICAL FRAMEWORK - MARCO TEÓRICO

Para facilitar la comprensión de nuestro documento, se contextualizará sobre los conceptos utilizados en esta sección.

A. Artificial Neural Networks - Redes Neuronales Artificiales

According to Khepri, Artificial Neural Networks (ANN) are born from the inspiration of Neural Networks (NN) that the human brain has; and that is why they are made up of similar elements. These networks learn from experience, generalize previous experiences to apply their knowledge to new experiences, and collect important information from data series.

Según Khepri, las Redes Neuronales Artificiales (RNA) nacen de la inspiración de las Redes Neuronales (RN) que tiene el cerebro humano; y es por ello que están constituidas por elementos similares. Estas redes aprenden de la experiencia, generalizan experiencias previas para aplicar su conocimiento en nuevas experiencias y recopilan información importante de series de datos. [6].

Similarities between Artificial Neural Networks and Neural Networks - Similitud Redes Neuronales y Redes Neuronales Artificiales

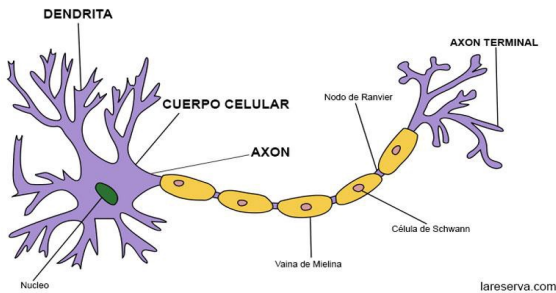


Fig. 1. The Structure of a Human Neural Network - Estructura de una red neuronal humana. [6].

B. Perceptrons - Perceptrones

Neural networks are worked with perceptrons. This represents an artificial neuron that processes the input data and generates an output based on a function. *Las redes neuronales son trabajadas con perceptrones. Este representa una neurona*

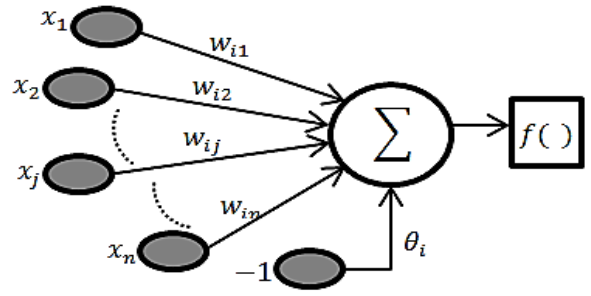


Fig. 2. The Structure of an Artificial Neural Network - Estructura de una red neuronal artificial. [6].

artificial que procesa los datos de entrada y genera una salida en base a una función.

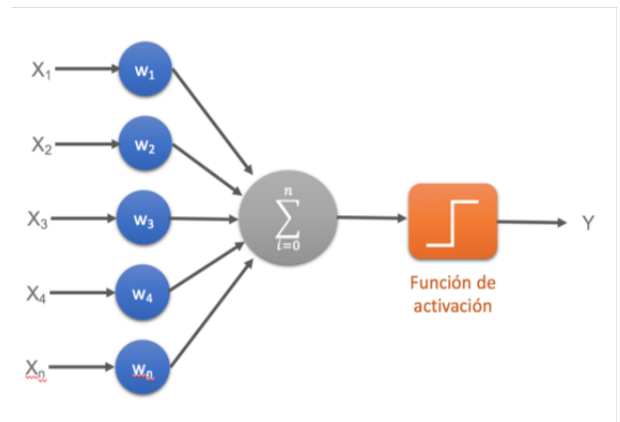


Fig. 3. The Structure of a simple perceptron - Estructura de un perceptrón simple. [3].

Perceptrons can become more complex (allowing them to solve more advanced problems) when working in conjunction with layers. These perceptrons are known as multilayers. They have three sections: the input layer, hidden layer(s) and the output layer. Each output of a perceptron represents the input of the next layer, and that allows complex problems to be solved. To solve the image recognition problem, a multilayer perceptron variance with two-dimensional matrices called a Convolutional Neural Network (CNN) is being used. It specializes in artificial vision problems and image management, which is important for our work.

Los perceptrones pueden llegar a ser más complejos (lo que les permite solucionar problemas más avanzados) al trabajar en conjunto con capas.

Estos perceptrones son conocidos como multicapas. Cuentan con tres secciones: la capa de entrada, capa(s) oculta(s) y la capa de salida. Cada salida de un perceptrón representa la entrada de la siguiente capa, y eso permite que se resuelvan problemas complejos. Para la resolución del problema de reconocimiento de imágenes, se está utilizando una varianza del perceptrón multicapa con matrices bidimensionales llamado una Red Neuronal Convolutiva (RNC). Esta se especializa en problemas de visión artificial y manejo de imágenes, lo cual es importante para nuestro trabajo.

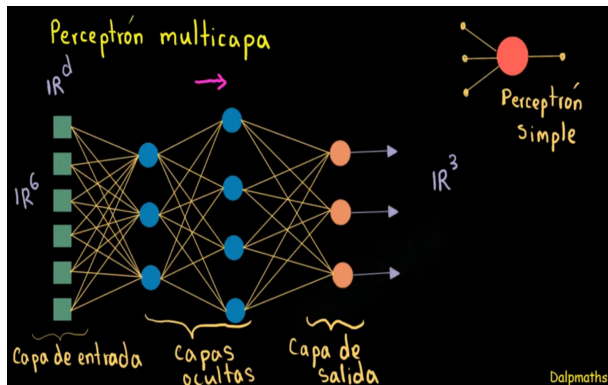


Fig. 4. The Structure of a multi-layered perceptron - Estructura de un perceptrón multicapa.

[8].

III. STATE OF THE ART - ESTADO DEL ARTE

In the search to solve our topic of interest (and trying to find a definition of the problem that fits our reality), we have given ourselves the task of investigating the existing publications on the topic, through projects and similar articles that will allow us have better bases when developing our project, so that our results can stay relevant, allowing for further knowledge and added value to our research for those interested.

En la búsqueda por resolver nuestro tema de interés, y tratando de encontrar una definición del problema que se ajuste a nuestra realidad, nos hemos dado el trabajo de investigar sobre la literatura existente respecto al tema, a través de proyectos y artículos similares que nos permitirán tener mejores bases al momento de desarrollar nuestro proyecto, y que de esta manera nuestros resultados puedan ser pertinentes, permitiendo la

generación de conocimiento y un valor agregado para los interesados en nuestra investigación.

One of the documents found, 'Food Image Recognition Using CNNs', by Yuzhen Lu, tells us that using good data pre-processing and advanced image enhancement techniques can improve the accuracy of a classifier. This study also points out that overfitting can be avoided by making advanced use of image enhancement.

Uno de los documentos encontrados Food Image Recognition Using CNNs, de Yuzhen Lu, nos dice que utilizando un buen pre-procesamiento de los datos y técnicas avanzadas de mejoramiento de imagen se puede mejorar la precisión de un clasificador. Este estudio también apunta a que se puede evitar el sobre-ajuste haciendo un uso avanzadas de mejoramiento de imagen.

IV. METHODOLOGY - METODOLOGÍA

A classification process is to determine the class to which a certain entity belongs. In our case, the class to which an entity (an image) could belong in is fruit types.

Un proceso de clasificación es determinar la clase a la que pertenece una determinada entidad. En nuestro caso, la clase a la que podría pertenecer una entidad (una imagen) es a sus tipo de fruta.

Initially, a support vector machine (SVM) can do the job of sorting with 50 % accuracy the raw images.

En un principio, una máquina de vectores de soporte (MVS), puede hacer el trabajo de clasificación con un 50% de precisión las imágenes en bruto.



Fig. 5. Image Recognition for an Apple - Reconocimiento de imágenes para una Manzana.

[8].

Convolutional Neural Networks (CNNs) seek to highlight characteristics (in our case, from an image) to facilitate the training process. This helped improve accuracy to 75 %. All this avoiding overfitting.

Las redes neuronales convolucionales (CNNs, Convolutional Neural Networks) buscan resaltar características (de una imagen en nuestro caso) para facilitar el proceso de entrenamiento. Esto ayudó a mejorar la precisión a un 75%. Todo ello evitando el sobre-ajuste (overfitting).

Finally, data augmentation techniques were applied to the training images, which increased their size. This helped to significantly improve the accuracy of the classifier to 90 % and thanks to the increase in data, the over-adjustment generated by CNNs was avoided.

Por último, se le aplicó a las imágenes de entrenamiento técnicas de aumento de datos, lo que aumentó su tamaño. Esto ayudo a mejorar significativamente la precisión del clasificador a un 90% y gracias al aumento de datos se evitó el sobre-ajuste generado por las CNNs.

Cereza	Durazno
Fresa	Lechuga
Limón	Mango
Manzana	Maracuya
Naranja	Palta
Pera	Piña
Sandía	

TABLE I

CLASSES THAT CAN BE EVALUATED

Once the image recognition result has been obtained, our graphic application will automatically start with its modeling, so that the final result is presented in a visually attractive way. This is achieved by executing shaders and mathematical calculations that allow us to shape the different objects in our collection.

Una vez obtenido el resultado del reconocimiento de la imagen automáticamente nuestra aplicación gráfica empezará con el modelado del mismo, para que el resultado final se presente de una manera visualmente atractiva. Esto se logra hacer mediante la ejecución de shaders y cálculos matemáticos que nos permiten dar forma a los diferentes objetos de nuestra

colección.

V. RESULTS / RESULTADOS

In general terms, our solution is effective when classifying most images. A continuación se presentarán imágenes y resultados de nuestras selecciones. *En términos generales, nuestra solución es efectiva al clasificar la mayoría de las imágenes. A continuación se presentarán imágenes y resultados de nuestras selecciones.*

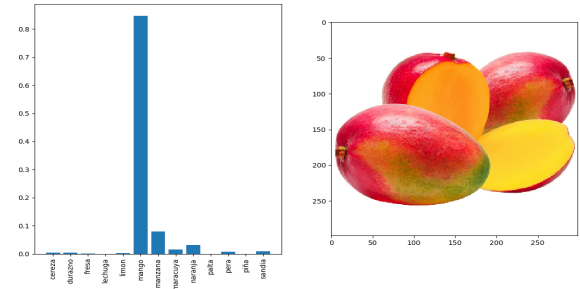


Fig. 6. Mango Image. Mango Highest

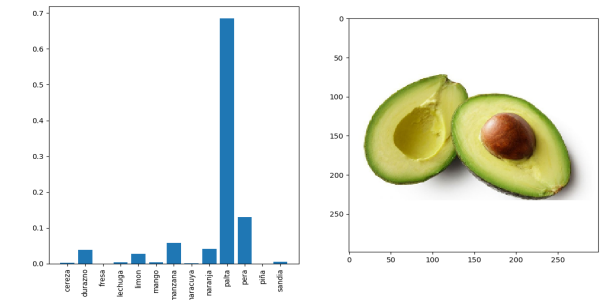


Fig. 7. Avocado Image. Avocado Highest

Contrary to the previous solution, the objects (fruits) we're using are very similar in shape and colour, which has resulted in a higher confusion rate for some of our models.

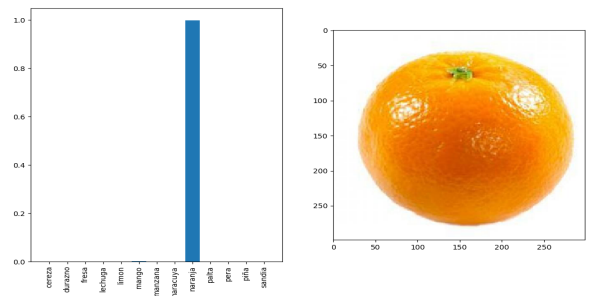
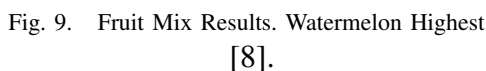


Fig. 8. Orange Image. Orange Highest

As a result of our testing with mixed fruits, we can observe that the most recognized fruit is the watermelon. This could be because apples and pineapples don't have those markings on them and since the orange is sliced. *Como resultado de nuestros experimentos en mezclas de frutas, se observa que la fruta más reconocida es la sandía. Esto puede deberse a que las manzanas y piñas no cuentan con esos relieves en líneas y porque la naranja es una rodaja.*



Finalmente, hemos incluido un modelo renderizado para la fruta reconocida en la imagen. Esto se ha hecho con el fin de ofrecerle al usuario una representación visual del resultado de la clasificación.

In this report we have managed to add to what our ancestors have done by not only recognizing the image, but also outputting a fully rendered model of the recognized object. For experimentation reasons, we decided to focus mainly on fruits or fruit-derived food. To this end, we obtained a large amounts of fruit training data for the AI. As a result of the similarities in the objects, we have determined that image AIs require or a lot of training data for each object to be precise and that



En este informe hemos logrado agregar a lo que han hecho nuestros antecesores al no solo reconocer la imagen, sino mostrar como salida el modelo reconocido totalmente renderizado. Por motivos de experimentación, decidimos concentrarnos principalmente en frutas o comida derivada de frutas. Con este fin, conseguimos muchos datos e información de frutas para entrenar a la IA. Por consecuencia del tipo de datos similares, hemos determinado que las inteligencias de reconocimiento de imágenes requieren o muchos datos de entrenamiento para cada objeto a reconocer o no depender de colores y formas en exceso entre estos objetos. En un futuro, se podría utilizar este mismo programa pero para modelos más distintos.

BIBLIOGRAFÍA

- [1] Cárdenes R. (s.f). "Inteligencia Artificial". Publicado por la Universidad de la Palmas de Gran Canaria.
https://www2.ulpgc.es/hege/almacen/download/38/38584/practica_ia_2.pdf
- [2] Numerentur (s.f) "CNN-RNA Convolutacional". Publicado por Numerentur
<http://numerentur.org/convolucionales/>
- [3] Calvo D. (2018). "Perceptrón - Red Neuronal". <http://www.diegocalvo.es/perceptron/>
- [4] Calvo D. (2018) "Función de Activación - Redes Neuronales". <http://www.diegocalvo.es/funcion-de-activacion-redes-neuronales/>
- [5] Centro Informático Científico de Andalucía "Introducción a la Computación Neuronal".
<https://thales.cica.es/rd/Recursos/rd98/TecInfo/07/capitulo1.html>
- [6] Khepri W. "Redes Neuronales, Qué son?". Publicado por Medium
<https://medium.com/@williamkhepri/redes-neuronales-que-son-a64d022298e0>
- [7] Food Recognition via Convolutud Neural Networks
<https://arxiv.org/pdf/1612.00983v2.pdf>
- [8] DalpMaths, Video Perceptrones Multicapa https://www.youtube.com/watch?v=Fz4Uo8_uUfg