

Instituto Tecnológico y de Estudios Superiores de Monterrey

Investigación Parcial 1: Deep Learning

Métodos Numéricos

Mtro. Adolfo Centeno T

Integrantes:

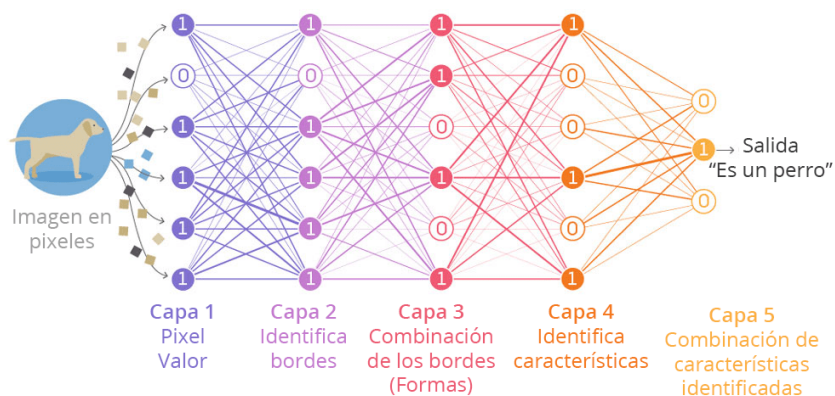
Maximiliano Macotela Nava, Luz Karen Hernández Hernández, Andres Serrato
Barrera, Julio Rodríguez Salcedo, César David Rosales Álvares, Jim Kevin Holguín
Rodríguez

A01366347, A01275802, A01701263, A01635920, A01368386, A01411784

Introducción

Lo que se conoce como *Deep Learning* (Aprendizaje Profundo, en español) es un tipo de *Machine Learning* (Aprendizaje Automático) que nos ayuda a entrenar a un dispositivo para que sea capaz de ejecutar tareas de una manera similar a como lo hacemos los seres humanos, tales como poder identificar voces, sonidos, imágenes, movimientos, e incluso poder predecir situaciones en un caso determinado. Es una de las bases importantes en lo que ha sido el desarrollo de lo que hoy conocemos como *Inteligencia Artificial*.

El concepto de Deep Learning se basa en configurar parámetros básicos acerca de los datos con los que se van a trabajar para poder entrenar a la computadora para que sea capaz de reconocer patrones de comportamiento mediante el uso de muchas capas de procesamiento, en lugar de organizar datos para que se ejecuten a través de ecuaciones predefinidas por una persona.



Fuente: <https://www.quantamagazine.org/>

La definición de Deep Learning apareció por primera vez en 1974, en la tesis doctoral de Paul Werbos, pionero en describir el proceso de entrenamiento de una red neuronal artificial a través de la propagación hacia atrás de errores o retropropagación. En 1980 el término fue re-descubierto por Geoffrey E. Hinton, un licenciado de la Universidad de Psicología Experimental de Cambridge y doctorado en ciencias de la computación, quien se dedicó a intentar aplicar a una máquina el modelo de aprendizaje que tenemos los humanos. A éste último se le atribuye el mérito de ser el padrino del Deep Learning, incluso recibiendo el Premio Fundación BBVA Fronteras del Conocimiento en 2017.

En nuestro día a día podemos encontrarnos con diferentes ejemplos de la inclusión del Deep Learning en las aplicaciones y los sitios web que usamos, tales como Netflix y Spotify, donde el algoritmo compara el material que hemos consumido con el que otras personas también han hecho, a fin de poder recomendarnos títulos, canciones y artistas similares a lo que ya hemos escuchado y lo que el sistema ha aprendido a predecir en base a nuestro gusto musical.

Los teléfonos inteligentes, asistentes virtuales, consolas de videojuegos y más son capaces de reconocer nuestras voces y acatar instrucciones dentro de los comandos que sepa recibir. Los automóviles de conducción autónoma son capaces de reconocer imágenes en sus cámaras frontales y traseras para poder advertir a los pasajeros de algún objeto en el camino y poder conducir acorde a la situación. Quizá nos hemos adaptado ya a cómo funcionan algunos de los dispositivos y programas mencionados anteriormente, pero no teníamos el conocimiento de qué era lo que se encontraba detrás de ellos para hacerles funcionar.

El Deep Learning promete poder generar sistemas predictivos que sean capaces de adaptarse a los nuevos datos que vayan recibiendo, sin la necesidad de que un humano tenga que editar una fórmula o algoritmo por sí mismo. Siempre se están recibiendo nuevos datos, a diferencia de los sistemas de analítica antiguos, en los que solamente se podía trabajar con el grupo de datos que tuviesen disponibles para estudiar en el momento.

El Deep Learning es más dinámico y eficaz, con el costo de necesitar de un hardware lo suficientemente capaz de poder trabajarlo; hardware que en la actualidad se está volviendo cada vez más potente y accesible incluso para personas que no son empleados de alguna compañía grande.

Alex Net

El ejemplo a trabajar para esta investigación es de la red neuronal llamada *Alex Net*, la cual, en palabras pequeñas, es capaz de utilizar un algoritmo de apenas unas cuantas líneas anexo a una red neuronal para poder identificar imágenes en momento real con la ayuda de una cámara capaz de procesar video.

En el año 2010 se lanzó el *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC). El ILSVRC es una competencia anual de visión por computadora desarrollada sobre un subconjunto de un conjunto de datos de visión por

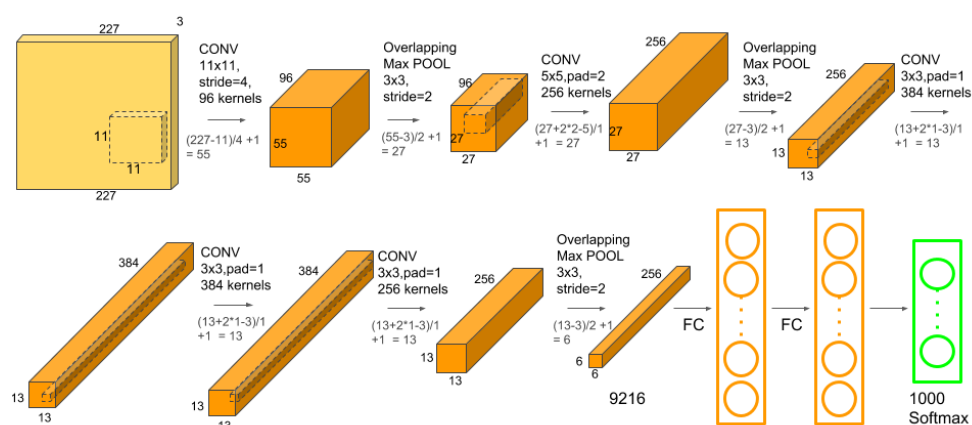
computadora disponible al público llamado ImageNet, que es un conjunto de imágenes anotadas para que una computadora sea capaz de trabajar con ellas.

AlexNet, desarrollada por Alex Krizhevsky en colaboración con Ilya Sutskever y el antes mencionado Geoffrey E. Hinton, fue la entrada ganadora en el ILSVRC del 2012. La red neuronal es capaz de resolver el problema de la clasificación de imágenes donde la entrada es una imagen de una de 1000 clases diferentes y la salida es un vector de 1000 números. El i -ésimo elemento del vector de salida se interpreta como la probabilidad de que la imagen de entrada pertenezca a la i -ésima clase. Por lo tanto, la suma de todos los elementos del vector de salida es 1.

La entrada a AlexNet es una imagen RGB de tamaño 256×256 . De lo contrario, debe convertirse a 256×256 antes de usarla para entrenar la red. Para lograr esto, la dimensión más pequeña se redimensiona a 256 y luego la imagen resultante se recorta para obtener una imagen de 256×256 .



AlexNet tiene 60 millones de parámetros y 650.000 neuronas, y entrenar la red tomó de cinco a seis días usando dos GPU GTX 580 de 3GB. Para su época, la red era bastante más grande que las que solían usarse para la época, aunque ahora ya no son números muy grandes en comparación a una red actual. La arquitectura de AlexNet consta de 5 capas convolucionales y 3 capas completamente conectadas.



Práctica

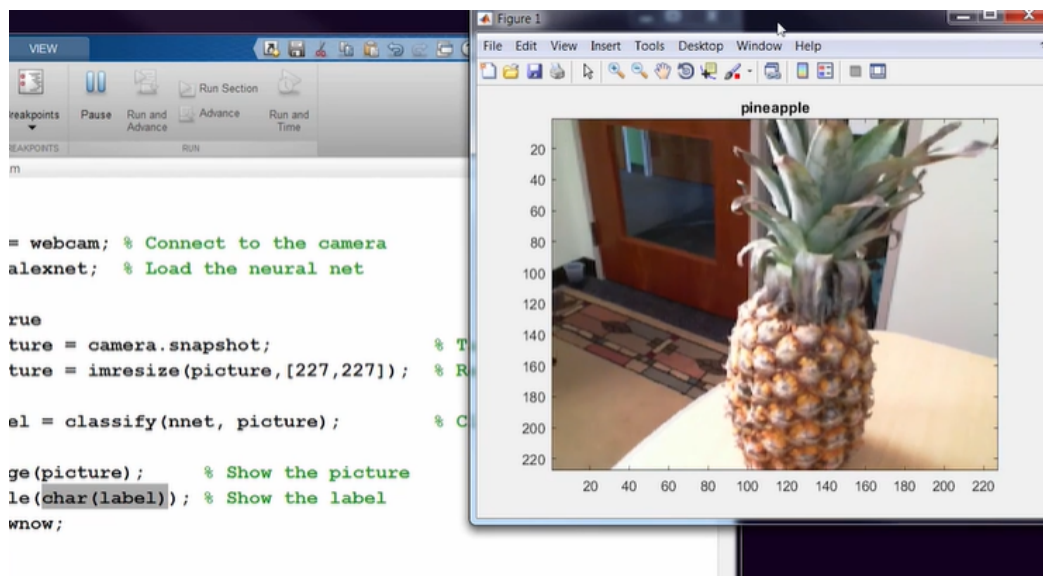
El código a trabajar en la práctica es el siguiente.

```
clear, clc

camera = webcam; % Conecta a la cámara
nnet = alexnet; % Carga la red

while true
    picture = camera.snapshot; % Toma una foto
    picture = imresize(picture, [227,227]); % Ajusta el tamaño de la imagen
    label = classify(nnet, picture); % Clasifica la foto
    image(picture); % Enseña la foto
    title(char(label)); % Enseña la etiqueta
    drawnow; % Hace que la cámara enseñe video en lugar de fotos
end
```

A pesar de nuestros intentos, nos fue imposible poder conseguir que AlexNet funcionara en nuestras computadoras. Afortunadamente, la práctica incluía un video anexo de cómo es el funcionamiento del programa.



Con el código que se nos provee es suficiente para poder llevar a cabo la actividad, ya que el proceso de identificar la imagen captada por la cámara es llevado a cabo por AlexNet.

Conclusiones

Creemos que gracias al aprendizaje automático y profundo es que podemos gozar de muchas de las comodidades del mundo moderno, aunque a veces las pasemos por alto y las tratemos como algo simple e incluso cotidiano. Solemos olvidar el cómo era nuestra tecnología hace apenas menos de una década, e ignoramos el qué tan diferente sería nuestra vida estudiantil y diaria sin lo que ahora nos es disponible al alcance de nuestros dedos, en nuestros celulares y computadoras.

Ya contamos con los sistemas de reconocimiento de imágenes, voz, sonido. todos estos para el uso recreativo, laboral, profesional. Es interesante pensar en cómo será el futuro con estas nuevas tecnologías, y sobre todo el uso que puede tener en áreas como la medicina y la biotecnología en general. De por sí la automatización de procesos en la actualidad parece bastante eficaz y avanzada, nos es posible creer que conectados a una red neuronal algunas tareas puedan ser resueltas con mayor eficacia y eficiencia.

Fuentes

<https://www.mathworks.com/videos/deep-learning-in-11-lines-of-matlab-code-1481229977318.html>
<https://www.mathworks.com/discovery/convolutional-neural-network-matlab.html>
<https://learnopencv.com/understanding-alexnet/>
https://www.sas.com/es_ar/insights/analytics/machine-learning.html
https://www.sas.com/es_ar/insights/analytics/deep-learning.html
<https://www.smartpanel.com/que-es-deep-learning/>
<https://www.quantmagazine.org/>
<https://machinelearningmastery.com/introduction-to-the-imagenet-large-scale-visual-recognition-challenge-ilsvrc/>