

Avances 2 del proyecto 2

Algoritmos y/o modelos que se utilizan para resolver problemas similares

- Procesamiento de datos médicos mediante el análisis de de imágenes y la geometría fractal

Los avances topológicos han permitido que los objetos y sus propiedades sean analizados de maneras más complejas. Un ejemplo, es el cambio de perspectiva de análisis de superficies, pasando de Euclidiano a topológico. El análisis topológico permite descomponer el espacio a analizar en partes más pequeñas, para poder representarlo e incluso medirlo con más precisión.

De igual forma, se introdujo el análisis fractal de cualquier sistema complejo o irregular, permitiendo encontrar patrones y similitudes en donde aparentemente solo hay caos. Con estos nuevos conceptos introducidos en el área de la medicina, comenzaron a surgir investigaciones que relacionaban comportamientos bioquímicos luego de traumas, con multifractales. Fue así como el instituto de oncología de Yugoslavia introdujo el análisis multifractal (MFA) para clasificación de imágenes de 256x256 píxeles en escala de grises, para el análisis de traumas epidermicos por quemaduras.

Estas interpretaciones fueron útiles para el Instituto de Medicina Patológica y Forense de la Academia Médica Militar de Belgrade, por lo que profundizaron el estudio y desarrollaron métodos y programas como FALFA Y HISTMF que seguirían mejorando y evolucionando hasta el día de hoy. (Reljin & Reljin, 2012)

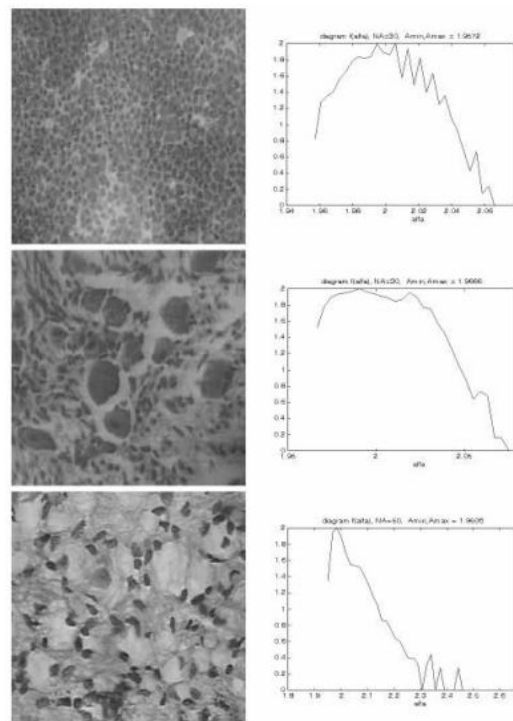


Imagen 1. Análisis de imagen mediante MFA por el Instituto de Oncología de Yugoslavia

- Análisis de imágenes utilizando modelos de redes neuronales convolucionales

Las redes neuronales convolucionales (CNN) es un algoritmo de Deep Learning utilizado principalmente para clasificar imágenes. Este algoritmo recibe como input una imagen, detecta las características principales de la misma y devuelve la categoría de la imagen. La forma en que detecta las características principales es que se le asigna un peso (importancia) a ciertos elementos de la imagen para poder diferenciar unos de otros. Algo importante a recalcar es que la salida se expresa en probabilidades, por lo que se debe tomar el que de mayor probabilidad.

El kernel en las CNN se considera como el filtro que se aplica a una imagen para extraer ciertas características o patrones importantes. Sirve para detectar bordes, enfoque, desenfoque, etc. Esto se logra al realizar una convolución entre el kernel y la imagen. La convolución en las redes neuronales consiste en tomar un grupo de píxeles de la imagen de entrada e ir realizando un producto escalar con un kernel. El kernel recorrerá todas las neuronas de entrada y obtendremos una nueva matriz, la cual se convierte en una de las hidden layers.

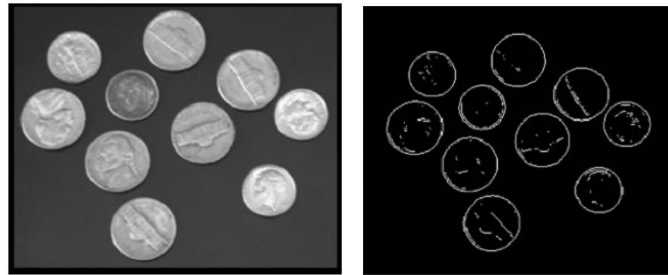


Imagen 2. Ejemplo de convolución entre el kernel y la imagen

Luego, en la capa de pooling, se combinan todas las salidas de los grupos de neuronas en una sola neurona para la siguiente capa. Esta capa es muy importante por el hecho que las imágenes no siempre se encuentran en la misma posición; pueden estar rotadas, posicionadas en la esquina, etc. Pooling reduce el tamaño de la representación de entrada y permite detectar objetos en una imagen sin importar dónde se encuentren. También ayuda a controlar el overfitting. Ya con estas capas creadas, se procede a la parte de flattening que consiste en aplanar estas matrices en un solo vector y ya el perceptrón hace la parte final de clasificación. (Bonner, 2019)

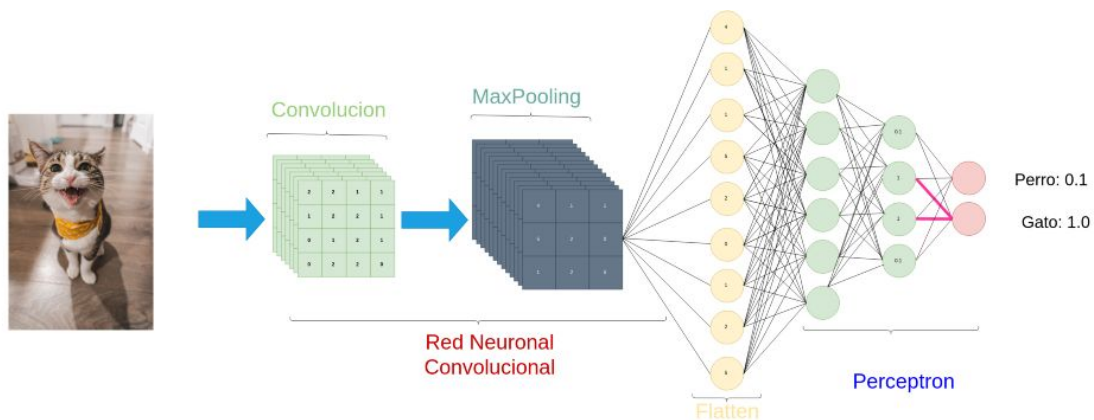


Imagen 3. Pasos de aplicación de una CNN

- Refinamiento de imágenes con el algoritmo Phase Stretch Transform

Existe una extensa cantidad de algoritmos que permiten el procesamiento de imágenes, los cuales permiten que se realicen cambios en ellas con el fin de poder mejorar su calidad, o en otros casos, permitir que la computadora o máquina pueda tener mayor facilidad al momento de reconocer o resaltar detalles. Estos algoritmos realizan acciones sobre las imágenes como por ejemplo cambiar la tonalidad, el contraste, mejorar definición, cambios de colores, rotación de contenido, reconstruir datos, entre otros. Estos algoritmos no proveen un tipo de modelo como tal para clasificación o predicción, pero son útiles al momento de apoyar a un algoritmo que sí haga este tipo de tareas facilitando el procesamiento o mejorando la calidad del modelo.

En este caso, el algoritmo Phase Stretch Transform es un algoritmo de procesamiento y reconocimiento de imágenes que es más o menos reciente, con 3 años de haber sido desarrollado con ingenieros de UCLA con acceso a código abierto. Este tiene como principal característica que mejora las imágenes, reconoce texturas y puede distinguir los bordes y otros detalles que aparecen en las imágenes. Capaz de mostrar detalles no visibles a simple vista, como estructuras de diversos materiales o las estrellas en una imagen del espacio. Algunos de los ejemplos que se han mostrado sobre cómo quedan las imágenes al usar el algoritmo son las siguientes:

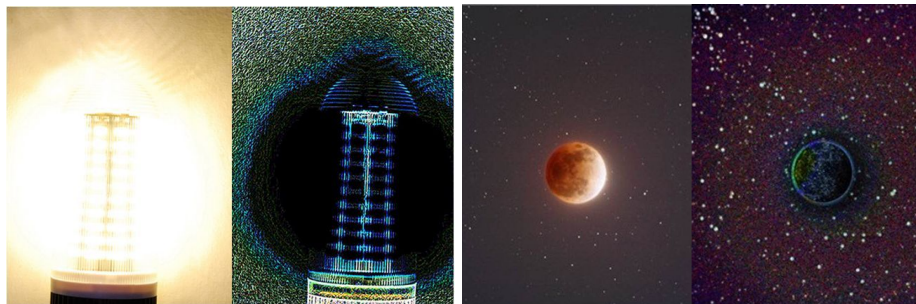


Imagen 4. Aplicaciones del algoritmo Phase Stretch Transform de los ingenieros de la UCLA sobre imágenes

Otro dato interesante del algoritmo es que puede procesar imágenes a una alta velocidad, lo que puede llegar a permitir el análisis e identificación de imágenes de video, o bien el procesamiento de una gran cantidad de imágenes en poco tiempo. Otras aplicaciones del algoritmo que mencionan los investigadores de la UCLA son la detección de células cancerígenas, análisis de huellas dactilares, caras y ojos.

Una pequeña descripción sobre cómo funciona el algoritmo es que la imagen es pasada a través de un filtro de suavizado seguido de la aplicación de una fase dependiente de la frecuencia no lineal descrita por el kernel de la fase PST. La salida de la transformación es la fase en el dominio espacial. El paso principal es la función de fase 2-D (kernel de fase PST) que normalmente se aplica en el dominio de la frecuencia. La cantidad de fase aplicada a la imagen depende de la frecuencia con una mayor cantidad de fase aplicada a las características de frecuencia más alta de la imagen. Dado que las transiciones nítidas, como los bordes y las esquinas, contienen frecuencias más altas, PST enfatiza la información de los bordes (Chin, 2016).

- Generar descripciones de imágenes con redes neuronales recurrentes

Se le conoce redes neuronales recurrentes a los algoritmos que implementan neuronas retroalimentadas, es decir neuronas que tienen capacidades de memoria en el tiempo. Para los humanos, es fácil describir una inmensa cantidad de detalles sobre una imagen. Sin embargo, esto es muy difícil realizarlo con modelos clásicos de reconocimiento visual. Normalmente el reconocimiento visual se ha centrado en etiquetar imágenes con un conjunto fijo de elementos visuales categorías. Sin embargo, mientras que los vocabularios cerrados de conceptos visuales nos otorgan modelos más sencillos de implementar, son muy restrictivos en comparación con los enormes cantidad de descripciones que un humano puede dar.

Para dar una descripción completa de una imagen utilizando redes neuronales recurrentes y convolucionales se analiza la imagen por regiones, identificando poco a poco el contexto de los posibles objetos que forman toda la imagen.

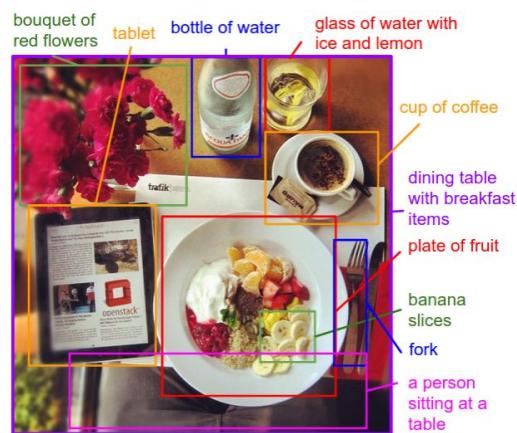


Imagen 5. Identificación de contexto por sectores

Luego basado en estos resultados se pueden unir las diferentes inferencias de la imagen en una descripción coherente y más completa de todo lo que se encuentra en la imagen original. El proceso de inferir a qué sectores corresponden diferentes descriptores se puede observar en la siguiente imagen. (Karpathy, et. al 2015)

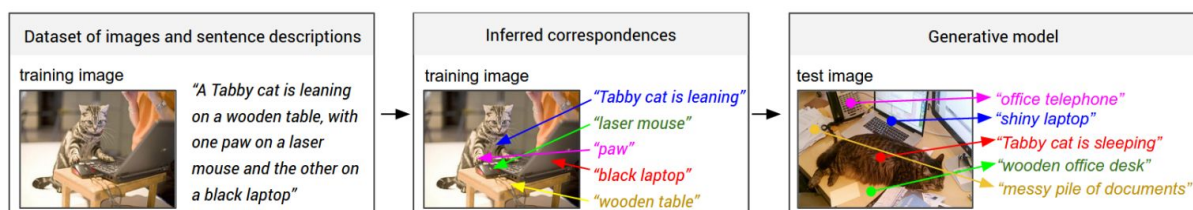


Imagen 6. Inferencia de descriptores

- Data Augmentation

Un problema en Deep Learning es que para entrenar correctamente un modelo se necesita una gran cantidad de datos. En esto nos ayuda Data Augmentation, nos permite aumentar el dataset de imágenes de 2 formas: Introduciendo perturbaciones en las imágenes, así como replicando la imagen descentrada o invirtiendo los ejes y también utilizando distintas distribuciones, por ejemplo, si se quiere entrenar un modelo de clasificación de imágenes de

alta resolución, se le añadirá imágenes con ruido, baja resolución, etc, pero siempre manteniendo la mayor proporción en las imágenes de alta resolución.

Referencias

Chin, M. (10 de febrero de 2016). *UCLA researchers release open source code for powerful image detection algorithm*. Science and Technology. Extraído de: <https://newsroom.ucla.edu/releases/ucla-researchers-release-open-source-code-for-powerful-image-detection-algorithm>

Gandhi, A. (2018). *Data Augmentation | How to use Deep Learning when you have Limited Data — Part 2*. Extraído de: <https://nanonets.com/blog/data-augmentation-how-to-use-deep-learning-when-you-have-limited-data-part-2/>

Reljin, I., & Reljin, B. (4 de octubre de 2012). *Fractal Geometry and Multifractal in Analyzing and Processing Medical Data and Images*. Belgrade, Yugoslavia: Facultad de Ingeniería.

Bonner, A. (2 de febrero del 2019). *The Complete Beginner's Guide to Deep Learning: Convolutional Neural Networks and Image Classification*. Extraído de: <https://towardsdatascience.com/wtf-is-image-classification-8e78a8235acb>

Karpathy, A. & Fei-Fei, L. (2015). *Deep Visual-Semantic Alignments for Generating Image Descriptions*. Extraído de: https://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Karpathy_Deep_Visual-Semantic_Alignments_2015_CVPR_paper.pdf