13/12/2023

Introducción redes neruonales convolucionales

La imágenes se representan como matrices

Anes de existir la CNN también se podían resolver problemas de clasificación de imáginenes. Una persona especialista hacía programas para extraer características de la imagine y luego entreneaba a un modelo con estas características. Sin embargo, este proceso es bastante costoso a nivel de tiempos y conocimientos.

Además toda la variabilidad que existe en el mundo real (Ilumunación, rotación, etc..) hacía que fuera aún más complejo

Con las CNN (Deep learning) se unifica la fase de extracción características y la clasifición. Se añaden filtros para que la red pueden identificar las características de la imagen. Estas arquitecturas son muy complejas, tienen un gran de parámetros que el modelo debe ajustar

Las CNN aprenden características a distintos niveles

* Características de bajo nivel: bordes, contornos
* Características de nivel intermedio: Ojos, orejas
* Características de alto nivel: caras

Redes neuronales convolucionales (CNN)

Modelo disenasdo para trabajar con imágenes. Pueden extraer características de imágenes durante el peiodo de entrenamiento

Filtro convolucionales: región que pasa a lo largo de la imagen para detectar características

Capas principales en una CNN

* Filtro convolucionales: filtros para generar mapas de características
* Función de activación: usualmente relu
* Pooling: reducir el tamaño de la imagen

Capa convolucionales

Stride: Numero de pixel de salto entre ventanas. Siempre suele ser 1

Padding: agregar valore sa la imagen para que sea mas grandes y los bordes no se puierdan eneguida. Con padding same se consigue controlar que la imagen de salida tenga el mismo tanano que la de entredada.

Usualmente el filtro suelve ser 3x3

20/10/2023

Función de activación

Se introducen no linealidades, que se aplican tras cada filtro. Las mas común es RELU. De esta forma se consigue resalta lo que nos importa e ignorar aquello que no

Usualmente en la propia funcion de activación se puede pasar un parámetro de funcion de activación, también se puede poner como su propia capa.

Pooling o agrupación

Capa para reducir el tamaño de los maps de características, perservando lo mas importante. Se puede hacer. Casi siempre se suele reducir el tamaño a la mitad, para hacer esta reducción se utiliza la técnica de “Max pooling” seleccionanado el valor máximo de la región seleccionada. También existe otra técnica utilizando la media, aunque el principal suele ser el máximo.

Esta capa no agrega ningún parámetro adicional

Otras capas populares en una CNN

Capa de sobremuestreo / Capa de deconvolucion: utilizada para redes que reciben una imagen y devuelven también otra imagen

Capa de normalización de lote: en ocasiones se coloca antes de la funcion de activación.

Flatten y global average pooling: convierte la salida de la ultima capa convolucial para transformarlas a un vector

Totalmente conectadas: se utilizan en problemas de clasificación

Flatten

Convierte para cada filtro la matriz de la imagen en un array con los pixeles y saca el máximo de estos

Global average pooling

Saca un pixel de cada filtro y hace la media entre ellos

10/01/2024

Cargar y preparar los datos

Existen dos formas principales para cargar las imágenes:

* Cargar todas la imágenes en memoria
  + Ventaja: no depende de ningún método en concreto para ciertos formatos de imágenes
  + Inconveniente: requiere mas memoria y con una gran base de datos es inviable
* Por lotes: Cargar los datos por lotes durante el entrenamiento
  + Ventaja: Requiere menos memoria
  + Existen métodos que solo trabajan con ciertos formatos de imágenes que no sean los standard

Para problemas de clasificación se utiliza el método image\_datase\_from\_directory, que permite crear un dataset automáticamente según la estructura de directorios.

El shuffle es importante en el entranamietno, ya que si los datos están ordenado por cualquier tipo de sesgo, se elimina para conseguir que los datos sean lo mas homogéneos posibles.

Definición de arquitectura

Una vez cargados los datos, hay que definir la aquitectutra, según estos aspectos

* Conjunto de capas que forman la arquitectura
* Tasa de aprendizaje
* Funcion de coste
* Metricas

17/01/2024

A la hora de estructura las imágenes, es importante clasificar las imágenes a través de agregarlas en carpetas diferentes, correspondiente a su clase. Una vez organizadas existe un método para clasificar imágenes según el directorio en el que se encuentre

También se pueden dividir los directorios y crear datasets differentes

El tamaño de las imágenes deben ser el mismo que la capa de entrada de la red, cuando se crea el dataset se redimensionaras a dicho tamaño

Las imágenes normalmente se almacenan con valores entre 0 y 255, sin embargo, a las redes neuronales es mejor valores entre 0 y 1

Compilacion

Cuando se compila una red neuronal, hay que definir 3 parametros:

* El optimizador: relacionada con la tasa de aprendizaje
* Loss: las métricas de perdida. Si el problema es de clasificación binaria se utiliza el binary\_crossentropy y si no el categorical\_crossentropy
* Metricas: no afectan a la red como tal, pero sirve para monitorizar el comportamiento de la red

Entrenamiento

Los callbacks son funciones que se ejecutan después de cada época

* Model check\_point: guardar copias del modelo durante el entrenamiento. Útil cuando los entrenamientos son muy largos. Permite guardar una copia a mitad del entrenamiento para luego seguir entrenándolo. El modelo solo se guarda si se ha mojorado en validación

Evaluación