

Informe de avance 3

Damián Quiroz
201321056-2

1. Propuesta de Modelo de Machine Learning para reconocimiento de letras

Como propuesta principal se considera una opción óptima, el hecho de usar redes neuronales convolucionales (CNN) para el reconocimiento de letras. Se considera éste algoritmo ya que a diferencia de SVM (Support Vector Machine, la propuesta inicial), éste tipo de algoritmo no es susceptible a diferencias en el brillo de la imagen, lo que es un problema bastante grande por parte de SVM, ya que para su entrenamiento se usan imágenes ya sea en formato GRAY o en formato RGB y las diferencias de iluminación de las imágenes de entrenamiento generan una detección con alta probabilidad de fallo.

El funcionamiento general de una CNN se muestra en la Figura 1.

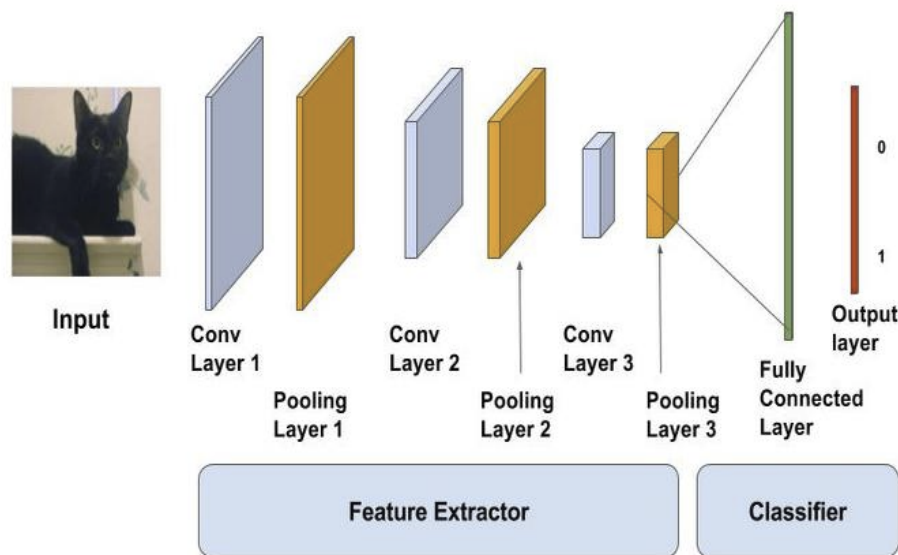


Figura 1: Funcionamiento general de una CNN.

Todas las capas convolucionales sirven para encontrar las características de la imagen que se ingresa, mientras que la capa de clasificación sirve para discernir si la imagen ingresada es o no el elemento deseado. Una de las grandes ventajas que tiene éste método, es que puede aprender características que otros métodos no detectan, como la relación entre la sombra generada y el objeto detectado, o características en la posición de los elementos.

Para la generación de éste tipo de redes neuronales en C++, se usan las bibliotecas de Keras y TensorFlow. Por lo que debido al poco tiempo para la implementación y adaptación de éste algoritmo (se debe adaptar Keras para usarse en C++ y se debe aprender a usar TensorFlow), se descarta éste algoritmo. Otra razón por la que se descarta es por el excesivo tiempo que se debe dedicar a entrenar el algoritmo, además del excesivo tiempo de procesamiento que se requiere para procesar una imagen, por lo que no se podría usar de forma continua y fluida en una comunicación por Video a menos que se tenga un equipo con el hardware que soporte tales velocidades de procesamiento.