

Predicción de dispositivo médico utilizando Convolutional Neural Network (CNN)

Agustín Bignú

Marzo 2019

Resumen

Este es el primer *paper* de un proyecto más ambicioso de Stening® que pretende revolucionar el mercado y la producción en el sector de la medicina y las prótesis a medida. En este escrito introductorio del proyecto entrenamos una red neuronal para que sea capaz de distinguir seis clases diferentes de dispositivos. Los resultados obtenidos son muy esperanzadores y demuestran el potencial que tiene este tipo de tecnología en el futuro de la medicina.

Abstract

This is the first paper of a more ambitious project by Stening® which goal is to revolutionize the market and production in the field of medicine and custom prostheses. In this very first work we train a neural network to be able to distinguish six different kinds of devices. The results obtained are very encouraging and demonstrate the potential that this type of technology has in the future of medicine.

1. Introducción

En Stening® siempre nos interesó la investigación y la innovación. Es por eso que nos renovamos continuamente. La *inteligencia artificial* es una disciplina de la computación que está creciendo a pasos agigantados y cada vez tiene más impacto en la medicina.

Los objetivos principales del estudio son los siguientes:

- potenciar el nivel innovador e investigador de la empresa
- marcar una diferencia en el sector del mercado en el cuál compite Stening®, siendo pioneros en este tipo de investigación
- dar a conocer la inteligencia artificial y el *machine learning*

Atendiendo a los objetivos mencionados, en este escrito se introducirá de forma breve lo que es la inteligencia artificial y el machine learning así como los tipos de aprendizaje que subyacen dentro de la disciplina. Seguiremos por el sector del aprendizaje profundo (*deep learning*) introduciendo las redes neuronales, para luego hablar de la que se utilizará en este estudio: la *Convolutional Neural Network*. Luego pasaremos a explicar cómo se realizó el dataset y por qué se hizo de esa manera. A su vez, discutiremos los resultados obtenidos mostrando algunos ejemplos y gráficos. Por último, se darán las conclusiones obtenidas y las vistas a futuro de la investigación.

2. Fundamento teórico

El objetivo de esta sección es introducir el tipo de tecnología utilizada. De esta manera, se procederá a presentar lo que es la *Inteligencia Artificial (IA)*, las *redes neuronales* y el tipo de red utilizada para este caso práctico, la *Convolutional Neural Network (CNN)*.

2.1. Inteligencia artificial

La IA es una rama de la computación que se centra en la realización de programas de computadora que tienen como objetivo realizar diversas operaciones y tareas que se consideran propias de la inteligencia humana, como el autoaprendizaje.

En concreto nos centraremos en la disciplina del aprendizaje de máquina o *machine learning*. Esta disciplina, dentro de la IA, es la encargada de crear todo este tipo de programas. Dentro del machine learning encontramos tres ramas bien diferenciadas de aprendizaje:

- *supervisado*
- *no-supervisado*
- *por refuerzo*

El primero consiste en algoritmos que reciben una serie de datos de los que luego aprenden, ya que saben la respuesta. De esta forma, utilizan lo aprendido para luego hacer predicciones. El segundo tipo de aprendizaje se diferencia del primero ya que no sabemos las respuestas, por lo que el enfoque del problema es diferente. Los algoritmos pertenecientes a este tipo de aprendizaje se centran en analizar los *datasets*¹. El tercer tipo no utiliza grandes cantidades de datos, como los otros dos, sino que aprende a base de prueba y error en un entorno. Este tipo de aprendizaje es el que se utiliza en conducción autónoma, robótica o IA aplicada a juegos.

En este estudio se utilizó el aprendizaje supervisado. Se recopiló un conjunto de datos, en este caso imágenes de dispositivos médicos de seis diferentes clases. Luego, se introdujeron los datos a una red neuronal que analizó las imágenes, aprendiendo de ellas y pudiendo predecir nuevos dispositivos de estas clases. En la siguiente sección se explicará lo que es una red neuronal.

2.2. Redes Neuronales [1]

Las redes neuronales artificiales se inspiran en el funcionamiento del cerebro. Los seres humanos somos capaces de ver una imagen y describir lo que aparece en ella. También somos capaces de escuchar una pieza musical y saber de qué género se trata. Todo esto lo sabemos porque lo hemos escuchado en otra ocasión y aprendimos de ello. Estamos aplicando el aprendizaje supervisado.

El objetivo principal de la IA es recrear este tipo de comportamientos y aprendizajes, como se dijo anteriormente. Para ello se crean algoritmos, como las redes neuronales. Este tipo de algoritmos conectan nodos denominados *neuronas*. El motivo de esta denominación es porque intentan emular a las neuronas del cerebro que actúan como unidades de procesamiento de la información que nos entra por los sentidos: vista, oídos, tacto, etc.

¹Conjunto de datos que se le introducen al algoritmo.

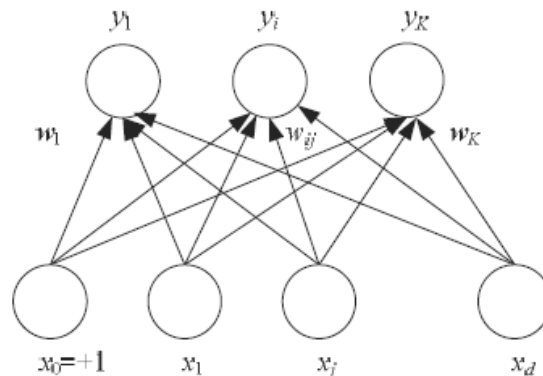


Figura 1: red neuronal [2].

En la figura 1 se puede ver una estructura bastante simple de una red neuronal. Los círculos son los nodos o neuronas. Como se puede apreciar, la red consta de dos capas, una de entrada (la inferior) y otra de salida (la superior). En la imagen, las "x" son los datos que introducimos a la red y las "y" son los resultados que obtenemos de la red. Las "w" son unos números que se van actualizando llamados *pesos*. Estos marcan el aprendizaje de la red. Se irán actualizando a medida que la red neuronal vaya entrenando. La red dejará de entrenar cuando haya alcanzado el porcentaje de certeza adecuado para nuestro objetivo.

Cabe destacar que una red neuronal puede tener más capas de neuronas, no solamente dos. Las redes que tienen capas internas, llamadas *capas ocultas*, son las más utilizadas. En este caso de estudio se utiliza un tipo de red neuronal ideal para tareas de reconocimiento de imágenes llamada *Convolutional Neural Network* o red neuronal de convolución. Se explicará en la siguiente sección.

2.3. Convolutional Neural Network [3]

Una CNN es un tipo de red neuronal que se basa en la estructura y trabajo de las primeras capas de visión del cerebro. En 1998, se introduce este tipo de red (LeCun et Al.) [4]. Una CNN se compone de dos partes principales: una que se encarga de sacarle información a la imagen y otra que se encarga de la clasificación de ésta.

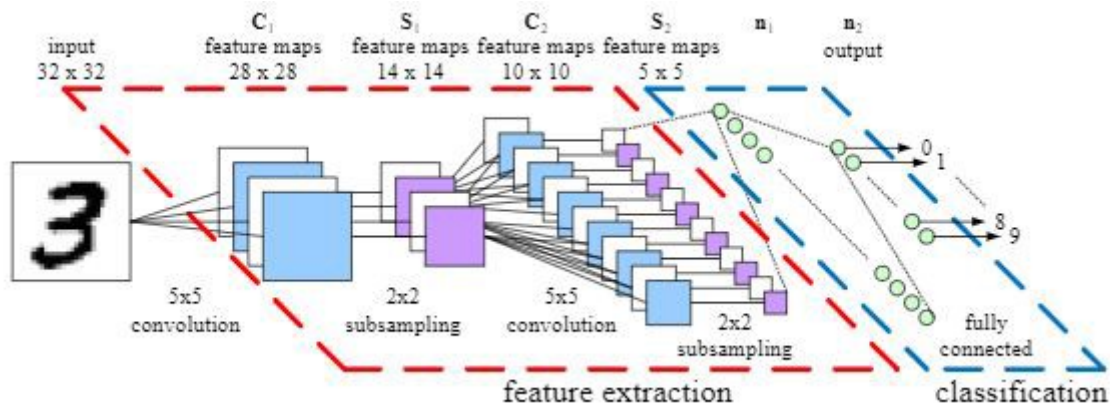


Figura 2: componentes básicos de una CNN [5].

En la primera parte de esta red, *feature extraction*, extraemos la información de la imagen. Le realizamos un método llamado *convolution*. Intentaremos entender, de forma somera, esta primera parte. Aquí extraemos las cualidades principales de la imagen, en nuestro caso: forma del dispositivo, posición, largo, etc.

El proceso de convolution, a grandes rasgos, consiste en ir reduciendo las dimensiones de la imagen de forma tal que nos quedamos con lo más relevante, lo que más información nos aporta para poder clasificarla de manera correcta. Esta primera fase contiene el mayor número de capas de la red.

Una vez superada la primera fase, pasamos a la segunda: clasificación. Esta última fase contiene tan sólo dos capas de neuronas: la primera recibe la información que nos llega de la primera fase y la pasa a la capa de salida que nos indica de que clase se trata la imagen que ingresó en un principio.

3. Resultados

En esta sección analizaremos la elaboración del dataset, describiremos los programas con los que obtuvimos los resultados y estudiaremos los resultados obtenidos.

3.1. Dataset

La elaboración del dataset fue propia. Dada la demanda de datos para conseguir una precisión alta en la clasificación, se tuvo que proceder de una manera innovadora. Es decir, para obtener un porcentaje de clasificación alto (>90%) es necesario un alto número de imágenes por clase (>500). De esta forma, al no disponer de tal cantidad de fotos, procedimos a realizar los tubos con un programa de diseño 3D llamado

Blender versión 2.9 obtenido de forma gratuita de su página web (Blender: www.blender.org). Todos ellos están basados en los *stents* fabricados por Stening®. Los resultados de los diseños fueron los siguientes:



A)

B)

C)



D)

E)



F)

Figura 3: dispositivos 3D generados a partir del programa Blender.

En la figura 3, A) HE (Stent Pared Fina), B) SET (Stent Estenosis Traqueal), C) SY13 (Stening® Y), D) SY16 (Stening® Y), E) TF12 (Tubo Faríngeo) y F) TM (Tubo Montgomery).

De cada clase se diseñaron 1400 imágenes, de las cuales 1000 se utilizaron para entrenamiento y las otras 400 para realizar predicciones (testing). Para darle versatilidad a las imágenes se fueron rotando los dispositivos para tenerlos en diferentes posiciones y motivar a la red para que encuentre distintas características en las que apoyarse.

3.2. Programas y resultados

Se realizaron dos programas escritos en Python, versión 3.5. Uno de los programas contiene la red neuronal y realiza el entrenamiento, el otro recibe una imagen y realiza la predicción (sobre imágenes que la red nunca vio). Nos apoyamos en la biblioteca *Keras* para machine learning así como *OpenCV* para visualizar las imágenes.

En el primer programa se sitúa la red neuronal. Ésta se construyó sobre una CNN llamada VGG16 [6]. Es una red neuronal de 16 capas capaz de reconocer diferentes clases de imágenes. Se le retiraron las últimas capas de la red correspondiente a clasificación y se le introdujeron las nuestras propias para que pueda reconocer las imágenes que deseamos.

Los resultados del entrenamiento y de las predicciones fueron los siguientes:

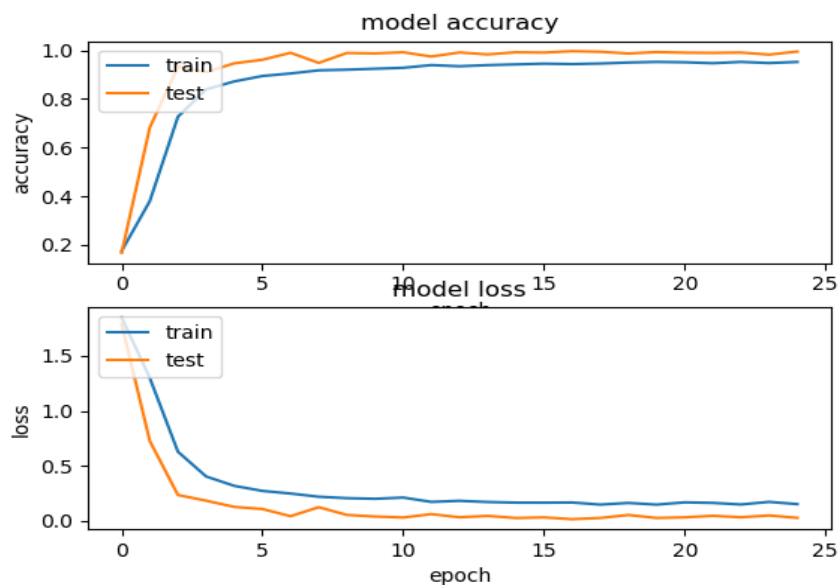


Figura 4: resultados del entrenamiento y de las predicciones.

En la figura 4 vemos dos gráficas. En la primera se muestra la progresión de la predicción, tanto durante el entrenamiento como durante el testeo. Se realizaron un total de 25 *epochs*². Vemos en la gráfica que al llegar al último epoch la predicción, tanto en el entrenamiento como testing, superó el 90% de certeza. En la segunda gráfica tenemos el entrenamiento y el testing. Aquí medimos el error cometido entre la predicción de la red neuronal y el resultado correcto de la predicción. También se puede apreciar como en el primer epoch el error cometido es muy alto y a medida que pasan los epochs éste va disminuyendo hasta casi alcanzar el cero.

En cuanto al segundo programa, realiza predicciones de una sola imagen, indicando de qué clase se trata. Le introdujimos diseños en 3D que la red neuronal no había visto, ni en el entrenamiento ni en el testing. Los resultados son los siguientes:

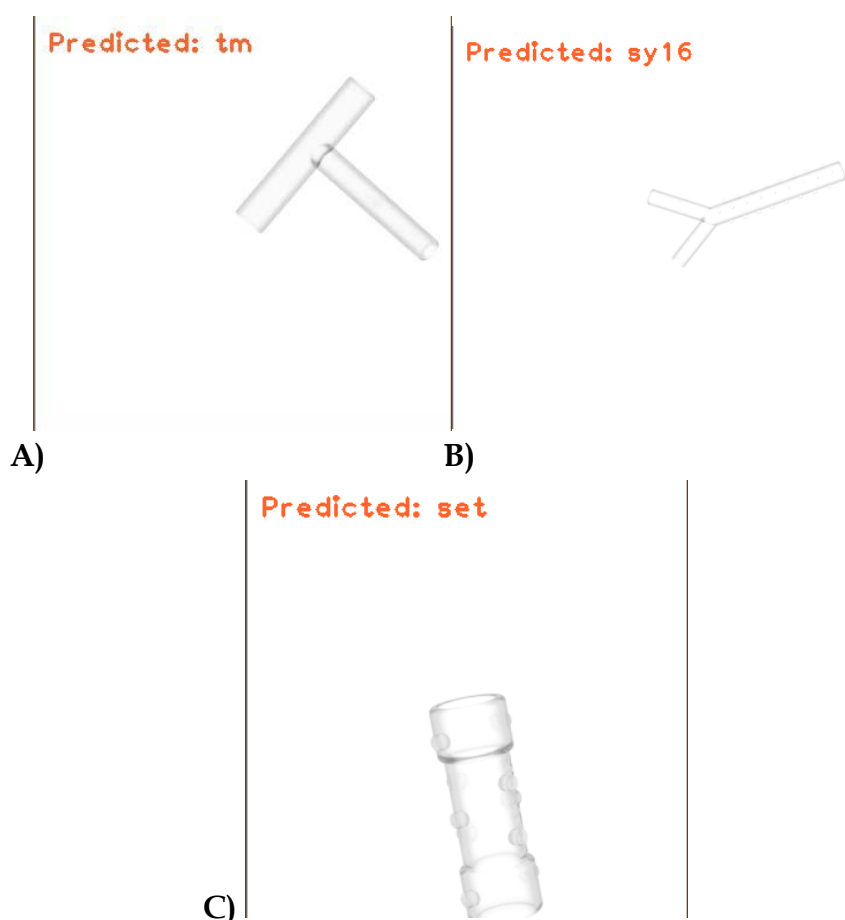


Figura 5: ejemplos de predicción.

²Cantidad de iteraciones de entrenamiento.

4. Conclusiones

Como vimos, los resultados obtenidos fueron muy buenos, llegando a obtener una certeza de más de 90%. En vistas a futuro se podrían hacer mejoras en cuanto al dataset. Es decir, pasar a entrenar con fotos reales de los dispositivos y ver resultados.

Este tipo de tecnología está entrando en auge y Stening® no se quiere quedar atrás, es por eso que apuesta fuertemente por la investigación y la innovación en esta área. Próximamente se seguirán realizando estudios y casos prácticos orientados al sector de la IA y la medicina.

Agradecimientos

Un especial agradecimiento a **Nicolás Bignú** por su colaboración en la elaboración del dataset utilizando el software *Blender*.

Referencias

- [1] E. Alpaydin,
Introduction to Machine Learning,
Second Edition,
The MIT Press, 2010.
- [2] E. Alpaydin,
Introduction to Machine Learning,
Second Edition,
The MIT Press, 2010.
Fig. 11.2.
- [3] Convolutional Neural Networks,
<https://medium.com/x8-the-ai-community/cnn-9c5e63703c3f>
- [4] Y. LeCun, L. Bottou, Y. Bengio and
P. Haffner,
*Gradient-Based Learning Applied to
Document Recognition*,
Proc. Of The IEEE, 1998.
- [5] ResearchGate,
[https://www.researchgate.net/figure/
An-Example-CNN-architecture-for-a-
handwritten-digit-recognition-
task_fig1_220785200](https://www.researchgate.net/figure/An-Example-CNN-architecture-for-a-handwritten-digit-recognition-task_fig1_220785200)
Fig. 1.
- [6] K. Simonyan and A. Zisserman,
*Very Deep Convolutional Networks for Large-
Scale Image Recognition*,
arXiv:1409.1556.