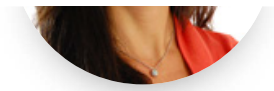


Con inteligencia artificial, logran generar automáticamente informes médicos a partir de radiografías



Utilizando un sistema de redes neuronales, un egresado de la carrera Ciencias de la Computación de la UNC logró implementar un modelo neuronal capaz de describir textualmente imágenes médicas. Se trata de un logro importante para la medicina, ya que la redacción de los informes es un proceso repetitivo y consume gran parte de la jornada laboral del equipo de profesionales en la salud. El desarrollo permitiría reducir notablemente el tiempo dedicado a esa tarea y optimizar la atención de los pacientes. [31.03.2021]





Por **Soledad Huespe**

Redacción **UNCiencia**

Prosecretaría de Comunicación Institucional – UNC

soledad.huespe@unc.edu.ar

Las imágenes médicas, como radiografías y ecografías, se utilizan ampliamente para el diagnóstico de muchas enfermedades. Su lectura e interpretación corresponde a especialistas en radiología o patología, y se estima que escribir los resultados de cada estudio demanda entre cinco y diez minutos aproximadamente, lo que ocupa la mayor parte de su jornada laboral.

Para optimizar esa tarea, Lucas Garay, egresado de la Licenciatura en Ciencias de la Computación de la Facultad de Matemática, Astronomía, Física y Computación (FAMAF) de la Universidad Nacional de Córdoba, diseñó un modelo basado en redes neuronales (inteligencia artificial) que permite obtener una descripción textual de una imagen.

Para ello, utilizó una arquitectura provista para la descripción de imágenes genéricas y la aplicó al dominio médico. “Un texto generado automáticamente puede reducir el trabajo que lleva a cabo el personal médico, que en lugar de redactar el informe completo se enfocará en revisar y modificar uno generado automáticamente”, explica el joven, quien actualmente se dedica a identificar imágenes médicas de los huesos de la pelvis en el Nanjing Technology Research Institute, en China.

La generación automática de descripciones textuales de imágenes –también conocida como etiquetado automático de imagen, indexación textual o image captioning– es el proceso mediante el cual un sistema informático asigna automáticamente metadatos en forma de subtítulos o palabras clave a una imagen digital.

Esta aplicación de técnicas de visión por computadora se utiliza en los sistemas de recuperación de imágenes para organizar y localizar imágenes de interés de una base de datos. También para resolver problemas de preguntas visuales, es

decir, dada una fotografía y una pregunta sobre ella en lenguaje natural, la tarea es proveer una respuesta precisa también en lenguaje natural.

Lucas lo explica de manera sencilla: “A través del lenguaje natural lo que buscamos es que la computadora pueda entender nuestro lenguaje escrito y al mismo tiempo ella también lo pueda generar para que también lo entendamos. Eso es, de manera sencilla, el procesamiento de lenguaje natural, una forma de la inteligencia artificial. Puede aplicarse a imágenes y a textos, pero la idea es la misma. Con las imágenes la idea es que la computadora pueda entender la semántica que hay detrás de ellas”.

“El procesamiento de lenguaje natural es una de las áreas centrales dentro de la inteligencia artificial y la tarea de Lucas, muy bien definida, fue la generación automática de pies de foto. Desde 2015 hay artículos científicos y software asociados con aproximaciones bastante satisfactorias. Lo que hizo Lucas fue aplicarlo al dominio médico y es de mucho interés”, explica Laura Alonso Alemany, docente e investigadora de la FAMAFA y directora de este trabajo de tesis.

Garay utilizó una base de datos abierta compuesta de 3.851 radiografías de la Universidad de Indiana (Estados Unidos) para “entrenar” a la red neuronal.

“Primero le di a la red como entrada una radiografía asociada a un reporte, con tags o etiquetas que, a grandes rasgos, la describen. La clave es la cantidad de datos para que el modelo ‘aprenda’, de manera que pueda aprender los patrones de la imagen y, en base a eso, devolver un reporte”, describe.

La principal dificultad que debió resolver Garay fue la escasez de datos, porque si bien se generan grandes volúmenes de imágenes y reportes, estos no siempre se encuentran disponibles libremente para su uso.

“Para alimentar el sistema son necesarias imágenes con y sin patologías. Por ejemplo, radiografías de pulmones sanos y enfermos. A partir de esta necesidad surgió un problema denominado desbalanceo de clases: mi dataset tenía muchos más datos de pulmones sanos. Entonces muchas veces el reporte que devolvía la computadora estaba equivocado, decía que los pulmones estaban bien, pero en realidad la imagen tenía problemas”, manifiesta Lucas Garay.

Para solucionar este inconveniente, se balancearon las clases, es decir, se introdujeron la misma cantidad de datos de imágenes buenas y de imágenes con patologías a través de dos técnicas: suprasampling y subsampling.

La primera permite multiplicar las imágenes introduciendo algún cambio como rotación o filtros. Se trata de la misma imagen, con la misma semántica, aplicada a la clase con menos ejemplos. La segunda se aplicó a la clase mayoritaria y consiste en extraer un subconjunto de los datos con el objetivo de nivelar la cantidad de ejemplos con respecto a la otra clase.

El desafío futuro es realizar transferencia de conocimiento o transfer learning entrenando el modelo con un conjunto de datos en español, ya que por disponibilidad de datos, este modelo fue realizado en inglés.

“La buena noticia es que hace tres meses se publicó un gran conjunto de datos en castellano de radiografías asociadas a informes textuales. Creemos que con eso vamos a poder entrenar sistemas que hagan esta tarea automáticamente en nuestro idioma con mayor confiabilidad”, completa Alonso Alemany.

La importancia de curar los datos

La generación de descripciones a partir de imágenes aplicando inteligencia artificial está en crecimiento. Es un área muy valiosa para personas con problemas de visión y para entornos donde no es posible descargar las imágenes, pero sí acceder a su descripción.

También es utilizada para la recuperación de imágenes, ya que se rastrean a partir del texto asociado a ellas. Esto facilita significativamente la búsqueda y ubicación de fotografías de archivos, por ejemplo.

“La generación de descripciones de imágenes tiene muchas utilidades. Sin embargo, uno de los grandes problemas que estamos detectando es que –como todos estos sistemas se construyen a partir de datos–, los sistemas también

reproducen prejuicios o sesgos que hallamos en esos datos”, advierte Alonso Alemany.

En ese sentido, el sistema termina así alimentando y profundizando estereotipos de género y sector sociocultural, entre otros. Hay muchos ejemplos. “Cuando estos sistemas encuentran una imagen de una persona en una cocina, van a tender a decir que esa persona es una mujer, incluso si muy claramente esa persona es un hombre”, apunta la investigadora.

“Otro ejemplo: si encuentran una imagen de alguien en una biblioteca, el sistema tenderá a decir que es una persona de clase acomodada y no de clase pobre. Si no curamos esos datos, van a amplificar prejuicios y sesgos que se encuentran en nuestra sociedad y van a perjudicar de una forma mucho más fuerte a los colectivos que se discriminan”, analiza la directora del trabajo.

Tesis | “**Procesamiento de imágenes médicas para generación automática de reportes**” es el trabajo final para obtener el título de Licenciado en Ciencias de la Computación de Lucas Gonzalo Garay bajo la dirección de Laura Alonso Alemany.

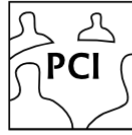


UNCiencia es la agencia de comunicación pública de la ciencia, el arte y la tecnología de la Universidad Nacional de Córdoba. Es una iniciativa de la Prosecretaría de Comunicación Institucional.

✉ unciencia@pci.unc.edu.ar | ☎ (0351) 5353730.



Universidad
Nacional
de Córdoba



Prosecretaría
de Comunicación
Institucional