

Clasificación de imágenes para fotos de ambientes de propiedades en venta con algoritmos de aprendizaje supervisado

Jorge Collinet
collinetjorge@gmail.com

Resumen. La compra-venta de inmuebles en un medio online requiere de una administración constante sobre la información de cada aviso publicado en el sitio. Aumentar la cantidad de información que se tiene sobre cada aviso resulta de vital importancia para el sitio de venta porque le permite brindar esa información a los usuarios que buscan. Las fotos de los avisos contienen información valiosa que puede ser utilizada para este fin. En este trabajo se presenta un algoritmo de clasificación de imágenes para intentar determinar en qué categoría pertenece cada una de las fotos presentes en un aviso. A partir del procesamiento de un dataset específico y aplicación de algunas herramientas para enriquecerlo, se presenta un algoritmo basado en redes neuronales. Esta nueva herramienta intenta agregar información adicional a las imágenes, información que luego puede ser usada para accionar sobre estas y al aviso al cual pertenecen las fotos.

1 Introducción

La utilización de internet para la compra venta online está creciendo de forma acelerada [5], dando nuevas formas de interactuar a la hora de realizar las operaciones, esto incluye a las operaciones de compra venta de inmuebles online.

Inmobiliarias y usuarios particulares publican avisos en los sitios web, en ellos especifican la propiedad con sus características y fotos, potenciales compradores entran en esos sitios web y buscan su propiedad deseada. Dependiendo del modo de operación del sitio, se puede realizar la transacción en el portal o se realiza el contacto entre el vendedor y el comprador para que posteriormente ellos realicen la transacción.

Los avisos provienen de inmobiliarias, vendedores particulares y/o integradores (empresas encargadas de buscar propiedades en venta desde distintos medios).

Cuando un aviso se publica en un sitio, este empieza a aparecer en los listados del mismo. Este puede ser encontrado utilizando filtros de búsqueda específicos, como por ejemplo: precio, metros cuadrados, cantidad de ambientes, antigüedad, ubicación geográfica, amenities, etc. Cuando un potencial comprador entra al portal, aplica estos

filtros y visualiza la propiedad, en la ficha de la propiedad se observan todas las características de la propiedad y las fotos de la misma.

Vale remarcar que la información que se muestre en la ficha del aviso es crucial para que el potencial comprador decida si esa propiedad es de su interés o no, las fotos juegan un rol fundamental en esa decisión. Debido a lo anteriormente mencionado, se considera sumamente interesante la posibilidad de encarar el desarrollo de un algoritmo de clasificación de imágenes que detecte que categoría corresponde a la foto y utilizar esta nueva información.

Las posibles aplicaciones del algoritmo pensadas hasta el momento pueden ser:

- Chequeo de la información del aviso con respecto a las categorías detectadas en las fotos para accionar sobre los vendedores (Ej: si el aviso es una casa, y en las imágenes no se detecta ninguna categoría del tipo “living”, entonces pedirle al vendedor que agregue una foto que sea del living).
- Chequeo de políticas de condiciones y uso de los portales, muchos de los portales de venta de propiedades no permiten la publicidad contacto en las imágenes, si se detecta una foto que tiene la categoría del tipo “publicidad”, se podría bloquear esa imagen o pedirle al vendedor que la quite).
- Ordenamiento y/o filtrado en los listados por categoría de las fotos, el usuario que está navegando por los listados podría elegir qué categoría quiere ver primero en los avisos.
- Estadísticas de qué tipo de categoría performa mejor en los listados y campañas de marketing según el tipo de propiedad que se está publicitando.

En este trabajo se elaboró un algoritmo de clasificación de imágenes basado en la fotos presentes en los sitios de compra-venta de inmuebles. Para ello, se experimentó con diferentes parámetros y técnicas de aprendizaje automático. El algoritmo base es Xception[1], este algoritmo fue elegido para validar validar la idea y puede haber otro que se adecue mejor para la resolución del problema, utilizando transfer learning[6] se obtuvieron muy buenos resultados.

En la sección siguiente se presenta la metodología utilizada. Luego, se muestra la experimentación efectuada y los resultados computacionales obtenidos. Por último, se presentan las conclusiones del trabajo y futuras líneas de avance.

2 Metodología

En este trabajo se elaboró un algoritmo de clasificación de imágenes basado en fotos obtenidas de los sitios de compra-venta de inmuebles. Para ellos, se utilizaron técnicas de aprendizaje automático para implementar un algoritmo de clasificación.

Esas imágenes obtenidas de los portales de compra-venta fueron posteriormente manualmente catalogadas en las siguientes categorías:

- Exterior
- Dormitorio
- Baño
- Cocina
- Living
- Publicidad
- Plano

La idea del presente trabajo es utilizar esta información y analizar las imágenes de los avisos que todavía no fueron catalogadas. Y de esa manera permitir a los sitios de compra venta obtener información adicional sobre las fotos que tienen cada aviso.

Para implementar el algoritmos de clasificación se experimentó con diversos parámetros sobre la red neuronal y sobre los parámetros sobre las imágenes antes de que sean procesadas por la red. Para implementar el algoritmo de clasificación se utilizó el software Tensorflow[2] y Keras[3], la red neuronal Xception está presente en keras.

3 Experimentación y Resultados Computacionales

Se utilizó cómo dataset imágenes obtenidas a partir de los sitios de compra-venta de inmuebles, se variaron la cantidad y tipo de categorías a utilizar, finalmente se eligieron las anteriormente mencionadas en base a una exploración manual sobre el dataset obtenido. Luego de obtener las fotos, estas fueron clasificadas manualmente. Se consiguió una cantidad de 2.369 imágenes.

Luego de dividir en dos partes el dataset, una de entrenamiento y otra de validación, dando un total de 1.955 ejemplos para el set de entrenamiento y 414 para el de validación. Quedaron las siguientes distribuciones:

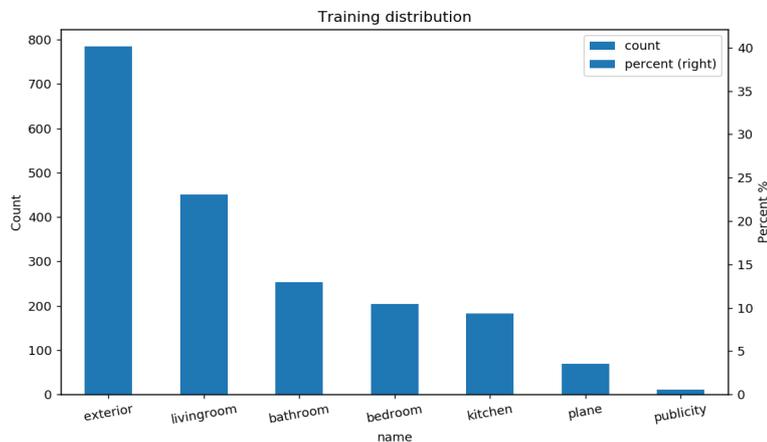


Fig. 1. Análisis de la distribución de clases en el set de entrenamiento.

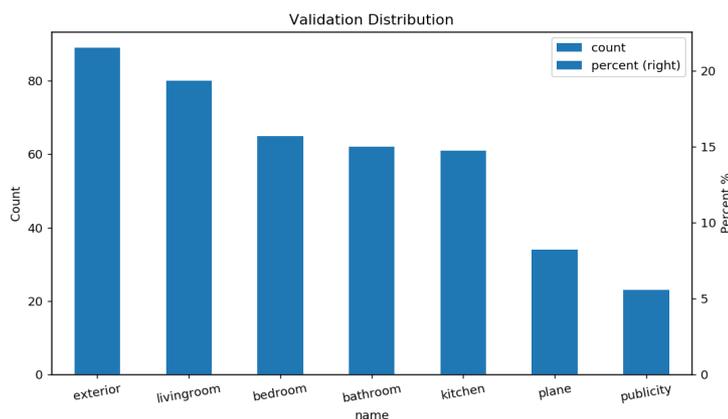


Fig. 2. Análisis de la distribución de clases en el set de validación.

Para el entrenamiento de la red neuronal se aplicó el método de transfer-learning, en el cual se usó una red previamente entrenada y se modificó solamente las últimas capas, en este trabajo se utilizó una red previamente entrenada con imágenes de ImageNet([7], [8]), en nuestro caso se re-entrenaron las últimas capas cambiando sus pesos pero sin cambiar la estructura de la red (a excepción de la última capa que originalmente era una softmax de tamaño 1000 y se cambió a una de tamaño 7), se fue variando la cantidad de capas a modificar con 2, 5, 15 y 20 últimas capas a re-entrenar. A continuación, podemos ver de los gráficos que muestran las métricas de accuracy, precision y recall del set de entrenamiento (acc / precision / recall) y del de validación (val_acc / val_precision / val_recall) a medida que se realizan los epochs:

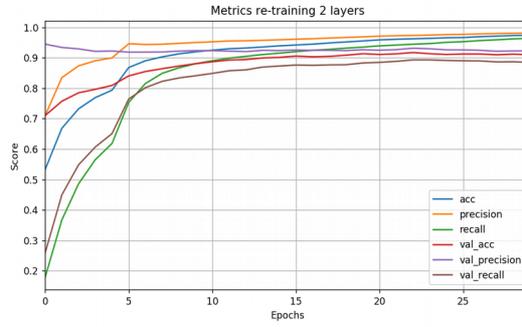


Fig. 3. Métricas re-entrenando 2 últimas capas

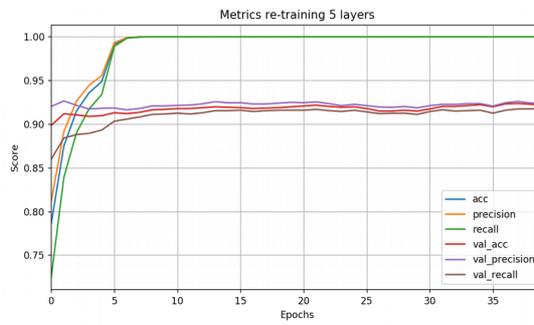


Fig. 4. Métricas re-entrenando 5 últimas capas

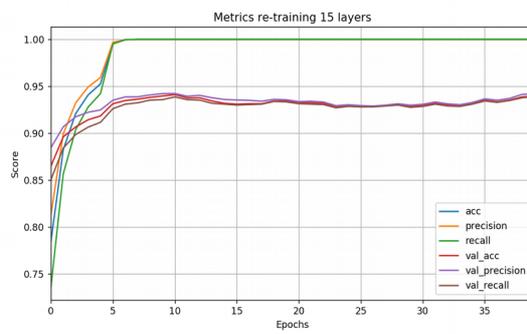


Fig. 5. Métricas re-entrenando 15 últimas capas

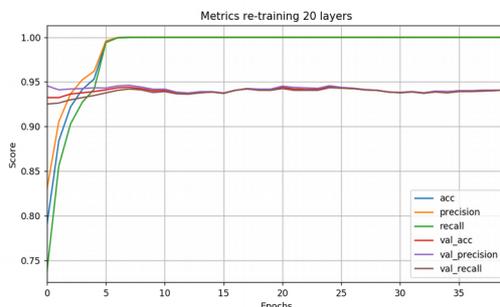


Fig. 6. Métricas re-entrenando 20 últimas capas

Cómo podemos ver en los gráficos, a medida que se aumenta la cantidad de capas a re-entrenar, el clasificador llega en menos epochs a lograr un error cercano a 0 en el set de entrenamiento.

Tabla 1. Comparativa de métricas (accuracy / precision / recall) del set de validación para los distintos valores de re-entrenamiento de las últimas capas.

Nro Capas	Máximo valor accuracy	Máximo valor precision	Máximo valor recall
2	0.932 - epoch 19	0.945 - epoch 1	0.908 - epoch 19
5	0.937 - epoch 36	0.941 - epoch 36	0.929 - epoch 36
15	0.946 - epoch 8	0.948 - epoch 38	0.944 - epoch 8
20	0.951 - epoch 17	0.953 - epoch 20	0.949 - epoch 17

A partir de la anterior tabla podemos concluir que a medida que permitimos a la red re-entrenar más capas, está obtiene mejor performance, en nuestro caso, el mejor score lo consiguió la red neuronal con las últimas 20 capas re-entrenadas siendo el accuracy de 95%, precision del 95% y recall del 95% para el set de validación.

4 Conclusiones

Cómo puede observarse en los resultados obtenidos, el clasificador implementado logra un porcentaje de certeza cercano al 95%. En el presente trabajo se analizan distintas cantidad de capas de la red a entrenar consiguiendo el mejor resultado cuando se re-entrenan las últimas 20. Se inició el estudio con un dataset elemental

compuesto de solamente fotos de propiedades en venta y se clasificó manualmente las mismas asignando categorías en base a un análisis manual.

3 Trabajo Futuro

Actualmente se ha realizado un trabajo sobre un dataset de imagenes de aproximadamente 3.000 ejemplos, se proyecta aplicar este algoritmo a una cantidad de aproximadamente 70 millones de imagenes, para ellos se tendrá que realizar modificaciones sobre el algoritmo y sobre la arquitectura sobre la que actualmente se ejecuta para poder procesar esa cantidad en un tiempo que se considere razonable. También se buscará aumentar la certeza del clasificador aplicando distintas técnicas y probando otros modelos, a su vez se buscará aumentar el set de entrenamiento. Para resolver estos desafíos se planea implementar un sistema de mensajes utilizando Apache Kafka [4] que avise a nuestro sistema cuando hay una nueva imagen a clasificar, luego asincrónicamente se clasificarán las imágenes nuevas, no se tiene un número exacto de la cantidad de imágenes que se tendrían que procesar en un determinado momento pero podemos estimarlo en 3.000 imágenes por minuto.

Otro futuro trabajo es la incorporación de categorías que sean más específicas y que sean un desglose de las anteriormente mencionadas, este posiblemente sea el desafío más grande que tengamos ya que se va a tener que re-procesar todas las imágenes que el modelo anterior fue clasificando, que rondarán en 70 millones de imágenes.

Referencias

1. Xception: Deep Learning with Depthwise Separable Convolutions, <https://arxiv.org/abs/1610.02357>
2. Tensorflow, <https://www.tensorflow.org>
3. keras, <https://keras.io>
4. Apache Kafka, <http://kafka.apache.org>
5. The Commercialization of the Internet, A Progress Report
6. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4890616>
7. Imagenet: a large-scale hierarchical image database, <https://arxiv.org/pdf/1409.0575.pdf>
8. Imagenet large scale visual recognition challenge, <https://arxiv.org/pdf/1409.0575.pdf>