

Dress up!

Jorge Luis Tapia Peñaloza

Septiembre 2019

1 Introducción

Dress up! es una aplicación creada a partir de dos arquitecturas de red, combinando el poder de una unet, y un modelo pix2pix.

El objetivo de este proyecto es plantear una solución a la incertidumbre por la cual pasan muchos hombres y mujeres actualmente al pedir ropa por internet. A través de modelos de segmentación y redes adversarias condicionadas, se crea un algoritmo capaz de mostrar cómo luciría una persona con una prenda de interés.

El flujo del algoritmo es como sigue; Se recibe la imagen de la persona que quiere probarse una prenda, la cual debe escoger de entre las opciones mostradas, la unet detecta la misma, se aplican una serie de funciones que permiten transformar dicha detección en un color dependiendo de la prenda que se desee probar, Seguido a esto entra en acción el modelo pix2pix cuyo output es sustituido en la imagen original de acuerdo a los píxeles de la unet.

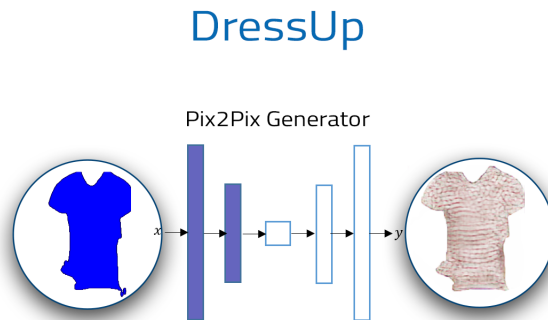


Figure 1: Generador del modelo pix2pix

2 Modelo de Segmentación

El primer paso para llevar a cabo este proceso fue la segmentación de las prendas con ayuda de una UNET, en este paso se utilizó el dataset *clothing-co-parsing-master*.

Para alcanzar el objetivo de esta arquitectura no fue requerido un etiquetado, sólo se llevó a cabo un parseo de las prendas superiores e inferiores de las cuales se serviría el modelo. Se hizo uso de el framework Keras con backend de Tensorflow para el entrenamiento de esta red convolucional en la que se busca dividir la imagen de entrada en diversos objetos de interés, en este caso el modelo es capaz de distinguir el fondo y otros objetos de las prendas superiores e inferiores que fueron especificadas con anterioridad.

El modelo fue entrenado hasta alcanzar el resultado siguiente.

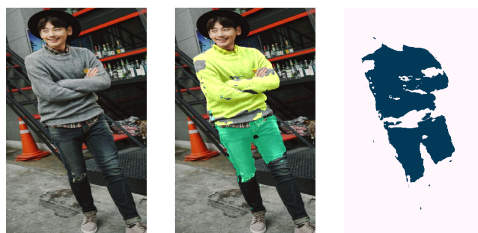


Figure 2: Resultado

Una vez teniendo el mejor resultado en el set de entrenamiento, el modelo de segmentación creado con la unet fue probado en imágenes no vistas mostrando un buen desempeño como se ilustra a continuación.

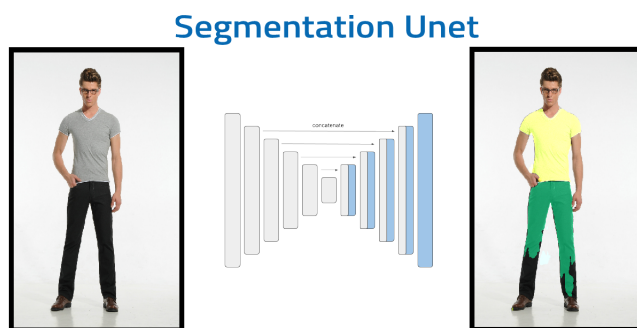


Figure 3: Segmentación con unet

Al conseguir un modelo que sea capaz de segmentar las prendas, se prosiguió a entrenar el modelo pix2pix.

3 Modelo PIX2PIX

Haciendo uso de las redes generativas adversarias se continuó con el proceso para alcanzar el objetivo de tener un probador de ropa inteligente. Para el entrenamiento de este modelo se hizo uso del dataset *lookbook*, de las cuales fueron ocupadas aquellas prendas que tenían una mayor presencia con el objetivo de que el modelo pudiera aprender mejor los patrones en los datos. Las imágenes disponibles para elegir en el probador Dress up! con las que fue entrenado el generador fueron las siguientes:



Figure 4: Prendas disponibles en el probador

Como también puede apreciarse en la Figura 3 cada una de las prendas del dataset de entrenamiento pix2pix fue etiquetada con un color diferente.



Figure 5: Ejemplo de etiqueta

En el flujo completo del modelo, es necesario que se tenga como input la imagen de la persona que quiere probarse una de las prendas disponibles, seguido a esto la unet entrenada anteriormente segmentará la imagen, dicha segmentación será coloreada dependiendo la prenda por la cual se quiera sustituir. El generador

de la GAN ahora se encargará de producir playeras de este estilo y el discriminador de la red la comparará con la original hasta que se consiga un nivel de precisión alto, como paso final el output de modelo se ajusta en la imagen original para dar como resultado la vista de la persona con la ropa deseada. Durante el entrenamiento de este modelo fue necesario llevar a cabo algunos ajustes para obtener un desempeño óptimo, por ejemplo, en un principio el entrenamiento de todas las prendas al mismo tiempo causaba ruido al modelo y no se obtenían resultados favorables, por tal razón se entrenó modelo por cada prenda de ropa, más aún esto fue mejorado con más entrenamiento. Después de varias épocas se vio un avance significativo del modelo como se ilustra en la figura siguiente.

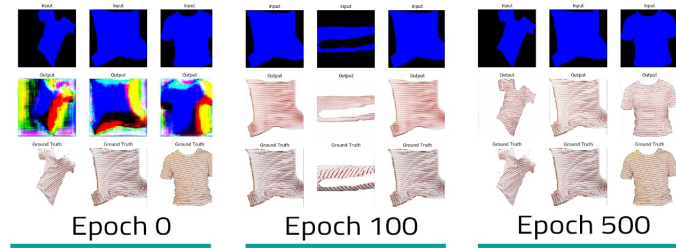


Figure 6: Avance del entrenamiento del modelo PIX2PIX

Además es de gran importancia mencionar que para mejores resultados en el modelo es necesario que se cumplan con ciertas características, pues los mejores resultados se obtienen al mostrarse la persona de frente, en algunas otras posturas y posiciones afectarán en el performance del modelo.

A continuación se presentan algunos ejemplos de los resultados conseguidos.

4 Conclusión

Ante los avances en las arquitecturas de redes neuronales y modelos de inteligencia artificial que permiten trabajar cada vez más con lo que hace años sólo parecía un sueño, como imágenes que den como resultado otras imágenes, texto a imágenes, voz a texto, etc, es importante tomar este conocimiento y encontrar una aplicación práctica que nos permita crecer, ayudar y seguir desarrollando, con miras en un futuro más brillante.

En este caso, se hace uso de dos arquitecturas complejas y con un gran sustento matemático para ayudar en actividades que se llevan a cabo usualmente, como elegir las prendas que se acomoden mejor a una persona en una sociedad en la que cada vez más y más gente hace uso de aplicaciones y de herramientas tecnológicas para vivir el día a día. Este algoritmo podría ser empleado en muchas aplicaciones y páginas web de tiendas de ropa y comercio electrónico, dispersando



Figure 7: Output del modelo

la incertidumbre y minimizando el riesgo de clientes insatisfechos debido a que una prenda no les quedó como ellos esperaban.