

Método para la Clasificación de Imágenes de Moda con TensorFlow

Method to classify fashion images with TensorFlow

Paredes-Miranda Fernanda Teresa¹, Muñoz-Zamora Guillermina¹, García-Alva Sigifredo¹, Cruz-Rentería Jesús Raúl¹

¹Tecnológico Nacional de México Campus Nogales, Posgrado e Investigación, Ave. Instituto Tecnológico #911, Nogales, Sonora, México, C.P. 84065.

Autor para la correspondencia: Muñoz Zamora Guillermina, guillermina.mz@nogales.tecnm.mx.

Resumen

En el negocio de la moda existe una gran variedad de ropa, cada una con un nombre propio para identificar de qué tipo de prenda se trata. El presente artículo describe un método para generar un modelo que identifique el tipo de ropa en base a 15 categorías, aplicando la tecnología de identificación de imágenes y aprendizaje automático como son las redes neuronales convolucionales bajo la utilización de la librería de Google TensorFlow, para elaborar un clasificador de ropa de moda. Concluyendo después de la fase de pruebas se alcanzó un 95 % de exactitud. Mientras que con un conjunto de prueba externo se logró un 70%, considerando la calidad de las imágenes como uno de los posibles motivos de falla.

Palabras clave: TensorFlow, clasificación de moda, clasificador de ropa.

Abstract

In the fashion business there is a wide variety of clothing, each with its own name to identify what type of garment it is. This article describes a method to generate a model that identifies the type of clothing based on 15 categories, applying the technology of image identification and machine learning such as convolutional neural networks under the use of the Google TensorFlow library, to develop a fashion clothing classifier. Concluding after the testing phase 95% accuracy was achieved. While with an external test set 70% was achieved, considering the quality of the images as one of the possible reasons for failure.

Key words: TensorFlow, Fashion classification, clothes classification.

DOI 10.46588/invurnus.v18i1.88

Recibido 13/09/2022

Aceptado 11/11/2022

Publicado 23/04/2023

Introducción

La moda es un gran negocio, ya sea que se venda en línea o en un espacio físico, la ropa es uno de los mayores impulsores de las ventas minoristas. Imagine tener un clasificador impulsado por inteligencia artificial que las tiendas podrían usar para identificar qué visten las personas que ingresan a la tienda. Esta información podría usarse de forma pasiva, para recopilar información agregada sobre el tipo de ropa que suelen usar los clientes o incluso podría usarse activamente como, por ejemplo, un detector de moda podría alertar al personal de ventas cada vez que un cliente ingresa a la tienda con un vestido o traje. Dado que esta persona ya lleva ropa elegante y de alto valor, es más probable que gaste más dinero en la tienda hoy. El primer paso para hacer posible lo anterior es utilizar un modelo clasificador de modas eficaz y preciso. En este artículo, mostraremos un sistema de aprendizaje profundo impulsado por inteligencia artificial que nos dé la clasificación más cercana del tipo de moda de la ropa de acuerdo con una serie de categorías establecidas.

El documento está organizado de la siguiente manera: La sección 2 contiene una breve descripción de las soluciones existentes que se relacionan con las técnicas implementadas en este artículo. La sección 3 contiene el desarrollo del proceso que se utilizó para llegar a un resultado del proyecto. Finalmente, la sección 4 está reservada

Trabajos Relacionados

1. La tarea de clasificación de imágenes es útil y ampliamente trabajada. A continuación, se describen trabajos que se vinculan a nuestro trabajo fuertemente.
2. En (Kayed et al., 2020) se establece la utilización de redes neuronales convolucionales (CNN LeNet-5) en el comercio electrónico para resolver problemas, como el reconocimiento de ropa, la búsqueda y recomendaciones de esta. Un paso fundamental para todas estas implementaciones es la clasificación de imágenes. Sin embargo, la clasificación de la ropa es una tarea desafiante ya que la ropa tiene muchas propiedades y la profundidad de la categorización de la ropa es muy complicada. Esta profundidad complicada hace que las diferentes clases tengan características muy similares, por lo que el problema de clasificación se vuelve muy difícil.
3. Otro trabajo relacionado (Shajini & Ramanan, 2020) propone una técnica basada en la atención para abordar el análisis visual de ropa de moda en imágenes, con el objetivo de lograr la clasificación de categorías de ropa y la predicción de atributos mediante la producción de diseños regularizados. Para mejorar la clasificación de ropa, su modelo de moda incorpora dos líneas de atención: atención basada en puntos de referencia y atención de canal espacial. Los cuales permiten que su modelo represente información contextual multiescala de puntos de referencia, mejorando así la eficiencia de la clasificación al identificar las características importantes y ubicar dónde existen en una imagen de entrada. Los resultados experimentales de este artículo muestran que la arquitectura propuesta que involucra una red neuronal profunda supera a otras técnicas de vanguardia informadas recientemente en la clasificación de la ropa de moda.



4. Actualmente, la tecnología de reconocimiento de objetos que utiliza el aprendizaje profundo ya está siendo aplicado en varios campos. En el trabajo (Son,H. Y. et al., 2020) se utilizó el aprendizaje de transferencia de AlexNet para clasificar diferentes tipos de prendas de ropa. Recopilaron datos de ropa en formato imagen que fueran claras y limpias. Las imágenes de la ropa de los datos se clasificaron en nueve grupos: cárdigan, chaquetas, camisas, Camisetas, tejidos, jeans, pantalones de algodón, pantalones cortos y faldas. El resultado de clasificación que utilizan AlexNet mostraron una precisión de clasificación de aproximadamente 69,28% y cuando se probaron los datos de imagen limpia, la clasificación precisión aumentada al 76,67%. Por lo tanto, se confirmó la posibilidad de una mayor segmentación de ropa y aplicaciones para otros objetos en el futuro, donde También se puede utilizar librerías con un mejor rendimiento.
5. En este estudio (Shajini & Ramanan, 2021), se detalla un enfoque de aprendizaje multitarea semi-supervisado con la intención de lograr la clasificación de categorías de ropa y la predicción de atributos. Para intensificar el análisis de ropa de moda semi-supervisado, se adoptó un modelo de pares maestro-alumno que utiliza la minimización ponderada de la pérdida mientras se comparte el conocimiento entre el maestro y el alumno. Los autores evaluaron el enfoque propuesto de etiquetado en el conjunto de datos DeepFashion-C a gran escala y el conjunto de datos combinado sin etiquetar obtenido de seis conjuntos de datos disponibles públicamente. Los resultados experimentales mostraron que la arquitectura emparejada propuesta que involucra redes neuronales profundas es comparable a las técnicas más modernas en el análisis de ropa de moda.

Materiales y métodos

El proyecto se llevó a cabo con la finalidad de validar si un enfoque de aprendizaje profundo (Deep Learning), podría clasificar en alguna de las 15 categorías predefinidas una imagen de prendas de moda escogida al azar de una colección descargada de internet.

3.1.Desarrollo

así como una versión reducida del Conjunto de datos (Dataset) de DeepFashion¹, para la generación del clasificador de ropa. El proyecto se realizó en el entorno Jupyter Notebook, el cual es una herramienta de la suite anaconda para generar el entorno de programación. Antes de codificar se debe verificar que se tengan instaladas las librerías que necesarias, de no ser así se procede a instalarlas desde la consola del software de Anaconda. Una vez instaladas las librerías correspondientes abrimos el apartado de JupyterLab para comenzar la codificación, importando las librerías al proyecto y especificando la ubicación de nuestro grupo de imágenes, que en este caso fueron 9,425 obtenidas de DeepFashion para entrenamiento, como se muestra en la Figura 1..

¹ <https://mmlab.ie.cuhk.edu.hk/projects/DeepFashion.html>



```
import os
os.chdir(r'C:\Users\Luna\Desktop\Code and Data\DeepFashion\Train')

# import the needed packages
import matplotlib.pyplot as plt
import matplotlib.image as img
import tensorflow.keras as keras
import numpy as np
import tensorflow as tf
config = tf.compat.v1.ConfigProto()
config.gpu_options.allow_growth = True
tf.compat.v1.keras.backend.set_session(tf.compat.v1.Session(config=config))
```

Figura 1. Importación de librerías y especificación de la ruta del set de entrenamiento.

Lo siguiente es realizar la codificación para la generación del modelo primario de imágenes a partir del Dataset, así como también predefinir la lista con las categorías en las que se pretenden clasificar los elementos del Dataset. Este nos regresa un resumen como el que se muestra en las figuras 2 y 3.

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, None, None, 3)]	0
block1_conv1 (Conv2D)	(None, None, None, 64)	1792
block1_conv2 (Conv2D)	(None, None, None, 64)	36928
block1_pool (MaxPooling2D)	(None, None, None, 64)	0

Fig. 2. Fragmento del modelo primario generado.

block5_pool (MaxPooling2D)	(None, None, None, 512)	0
=====		
Total params: 14,714,688		
Trainable params: 14,714,688		
Non-trainable params: 0		

Figura 3. Resumen del modelo primario generado.

Parámetros de configuración

Se etiquetaron las imágenes con nuestras 15 categorías predefinidas (Blazer, Blouse, Cardigan, Dress, Jacket, Jeans, Jumpsuit, Romper, Shorts, Skirt, Sweater, Sweatpants, Tank, Tee, Top), y definimos en 3 el número de capas que se utilizaran para entrenar el modelo. Compilamos nuevamente ahora con las etiquetas incluidas y finalmente vemos el modelo completo con el que se va a entrenar el Sistema, el cual debe regresar el modelo y un resumen del modelo como se muestra en la figura 4.



```

=====
Total params: 17,245,303
Trainable params: 17,245,303
Non-trainable params: 0
=====

```

Figura 4. Resumen del modelo completo.

El siguiente paso para realizar es entrenar el sistema con el modelo completo y esperar a que termine la compilación como se muestra en la figura 5. El proceso puede tardar de pocos minutos hasta varias horas dependiendo del equipo donde se esté corriendo el sistema y el número de capas predefinidos para el entrenamiento.

```

  | . . . |
Train for 2552 steps, validate for 281 steps
Epoch 1/3
2552/2552 [=====] - 125s 49ms/step - loss: 0.2182 - acc: 0.9345 - val_loss: 0.1864 - val_acc: 0.9374
Epoch 2/3
2552/2552 [=====] - 110s 43ms/step - loss: 0.1550 - acc: 0.9462 - val_loss: 0.1934 - val_acc: 0.9369
Epoch 3/3
2552/2552 [=====] - 110s 43ms/step - loss: 0.1300 - acc: 0.9537 - val_loss: 0.1794 - val_acc: 0.9409

```

Figura 5. Proceso de entrenamiento del modelo completo.

Una vez terminado el entrenamiento se procede a validar el resultado, se realizaron corridas de pruebas con 1,009 imágenes del Dataset escogidas para probar el modelo y se revisa para saber si se encontraron imágenes mal clasificadas durante nuestras corridas de pruebas con los elementos de pruebas. El valor de exactitud o “accuracy” alcanzado fue de 0.9537, lo que es un buen resultado.

A continuación, se muestran la gráfica de la exactitud conforme avanzan los números de época. Podemos observar que el conjunto de validación comienza por encima del conjunto de entrenamiento y al final lleva una tendencia constante positiva. Mientras el conjunto de entrenamiento se acerca a 1

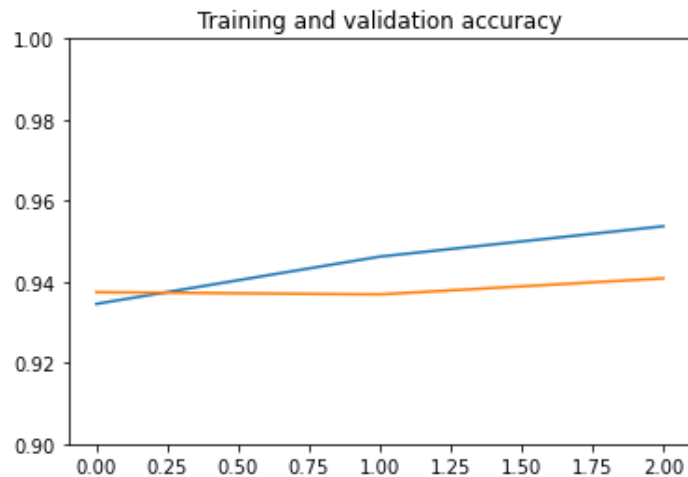


Figura 6. Exactitud para los grupos de entrenamiento y validación.



A continuación, se muestra una prueba con una sola imagen seleccionada del Dataset de Prueba. Al terminar el proceso se imprime la etiqueta predicha, que sería la etiqueta en la que el sistema clasificó la prenda y la etiqueta real, que es la etiqueta que, conocida previamente, como se muestra en la Figura 7.

Found 1009 images belonging to 15 classes.



True label: Top
 Predicted label: Tank

Figura 7. Resultado de corrida prueba.

Resultados

Se realizó una prueba al modelo con un conjunto de prendas seleccionadas previamente de la nube. A continuación, se muestra una imagen al azar encontrada en internet diferente a cualquiera existente en el Dataset, (Figura 8). El resultado que el método arroja se considera aceptable ya que muestra la clase más genérica en la que se puede clasificar la prenda (Figura 9).

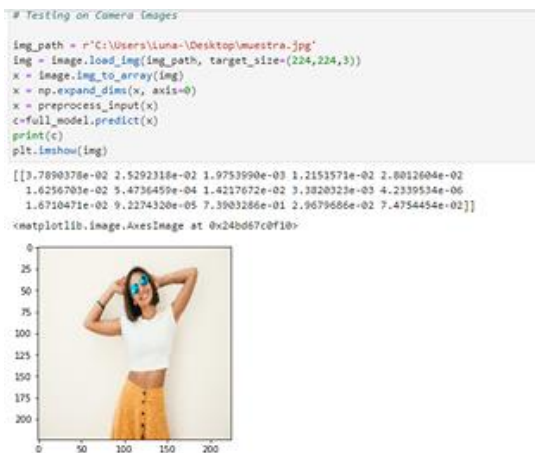


Figura 8. Imagen aleatoria clasificada.



```

Predicted_Class=np.argmax(c, axis = 1)
print('Predicted_Class is:', Predicted_Class)    #Get the rounded value of the predicted class

Predicted_Class is: [12]

classes[8]

print(classes[12])

Tank

```

Figura 9 Etiqueta resultante.

Conclusiones

El modelo alcanza un 95% de exactitud, siendo un resultado comparable con los trabajos relacionados. Después de la evaluación del modelo con un conjunto de validación diferente al conjunto de prueba, se obtuvo un 70% de exactitud es decir prendas clasificadas correctamente, ya que el sistema clasifica solo la prenda con mejor visibilidad y ofrece una de las etiquetas correspondientes a la clasificación de la prenda, la cual sería la clasificación general de la prenda, pero no la más adecuada. El modelo puede llegar a mejorarse y quizás tardar menos tiempo en la compilación si se corre en un dispositivo con mejor rendimiento y con un Dataset de mayor volumen. Podemos concluir que el enfoque de aprendizaje profundo puede ser útil para clasificación de grandes volúmenes de imágenes.

Referencias

1. Kayed, M., Anter, A., & Mohamed, H. (2020). Classification of Garments from Fashion MNIST Dataset Using CNN LeNet-5 Architecture. In 2020 International Conference on Innovative Trends in Communication and Computer Engineering (ITCE) (pp. 238-243).
2. Shajini, M., & Ramanan, A. (2020). An improved landmark-driven and spatial-channel attentive convolutional neural network for fashion clothes classification. *The Visual Computer*, 37(6), 1517-1526. <https://doi.org/10.1007/s00371-020-01885-7>
3. Shajini, M., & Ramanan, A. (2021). A knowledge-sharing semi-supervised approach for fashion clothes classification and attribute prediction. *The Visual Computer*. <https://doi.org/10.1007/s00371-021-02178-3>
4. Son, H. Y., Lee, M. S., & Noh, S. K. (2020). A study on the clothes classification using alexnet deep learning. In Proceedings of the 9th International Conference on Smart Media and Applications (SMA 2020), from https://manuscriptlink-society-file.s3.amazonaws.com/kism/conference/sma2020/presentation/SMA-2020_paper_85.pdf.

