

TP Final - Skintone

Schiavinato Mauro

Procesamiento digital de imágenes · Universidad de Buenos Aires

1 Introducción

Vamos a analizar los tipos de piel asignados automáticamente por el algoritmo Individual topology angle para distintas iluminaciones de una persona. Luego para la clasificación de piel vamos a utilizar el Fitzpatrick rating.

2 Algoritmo

1. Obtención de bounding box de cara mediante clasificador de cascada entrenado (opencv).
2. Pasamos la fotos al espacio de color YCbCr y obtenemos los píxeles de la cara mediante la siguiente función:

$$pixel = \begin{cases} piel & 136 \leq Cr \leq 173 \text{ \& } 77 \leq Cb \leq 127 \\ no\ piel & cc \end{cases}$$

3. Promediamos los píxeles de la piel, lo convertimos al espacio de color CIELab y por último calculamos el ITA como:

$$ITA = \frac{\arctan(\frac{L-50}{b}) \cdot 180^\circ}{\pi}$$

4. Mapeamos según el ITA a la categoría de piel según el siguiente cuadro:

Condición	Categoría
$ITA \geq 50$	I
$25 \leq ITA < 50$	II
$0 \leq ITA < 25$	III
$-25 \leq ITA < 0$	IV
$-50 \leq ITA < -25$	V
$ITA < -50$	VI

3 Análisis

Utilizamos el dataset **CasualConversations**. En este dataset se tomaron varias fotos a cada persona. Siempre vamos a analizar las fotos agrupando los resultados de las fotos de cada persona. El skin type de una persona es el promedio del skin type en todas las fotos que aparece. Definimos como **oscuras** a las imágenes donde se tomaron con una iluminación más oscura que el resto en un grupo de fotos de una persona. En la figura 3 vemos la distribución de skintones del dataset.

Para pasar de RGB a CIELAB utilizamos la iluminancia D50 obteniendo un resultado mejor que la iluminación D65 que es la más general.

Obtuvimos un accuracy del 0,26 y un accuracy del 0.57 si tomamos que la predicción puede tener hasta una diferencia de uno. En la figura 2 vemos un gráfico de como cambia el accuracy dejando más diferencia entre el label y la predicción.

En la figura 4 observamos la distribución de la predicción de skintones de nuestro algoritmo. Vemos que obtuvimos muchos más casos en la categoría 6 que lo que realmente son. Esto puede deberse a las imágenes oscuras: en la figura 5 vemos que la predicción de las imágenes oscuras tuvo una mayor diferencia que las imágenes no oscuras.

En la figura 6 podemos ver a donde fueron a parar las predicciones para cada categoría. Observamos que muchas imágenes con label 1, 2 y 3 se predijeron con label 4. También vemos una tendencia de que las predicciones caen a la derecha de la diagonal, es decir que generalmente se predijo un skintone mayor al verdadero. Esto puede deberse a las imágenes oscuras. Si ahora vemos la figura 8, casi todas las imágenes se predijeron al skintone 6.

Por último en la figura 7 vemos la matriz de confusión sin imágenes oscuras y vemos que en este caso también encontramos que las predicciones terminan en una categoría mayor que las reales, aunque en menor medida que utilizando todas las fotos. Esto puede ser debido a la segmentación de la cara, la cual muchas veces no era perfecta y se agregan partes de la cara más oscuras, como las cejas, la barba, anteojos.

Categoría	Accuracy
I	0.13
II	0.12
III	0.22
IV	0.14
V	0.36
VI	0.46

Fig. 1: Accuracy del método para las distintas categorías de piel.

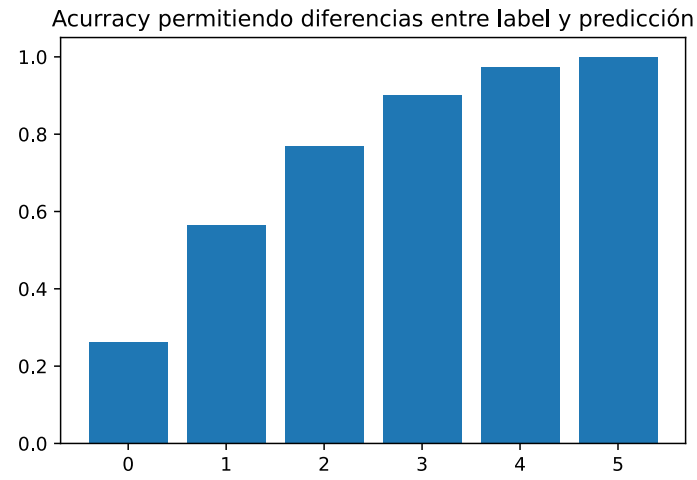


Fig. 2

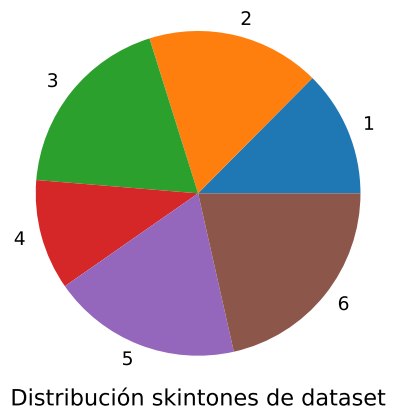
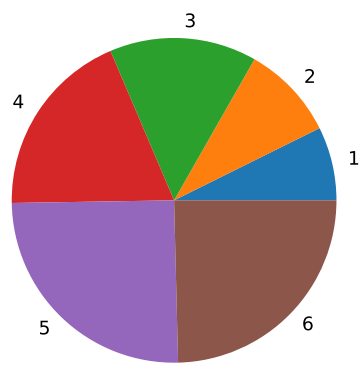
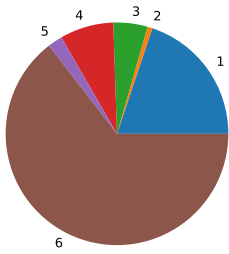


Fig. 3

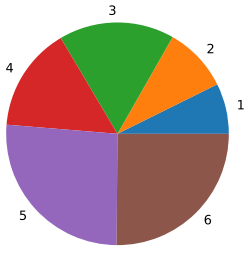


Distribución de predicción

Fig. 4

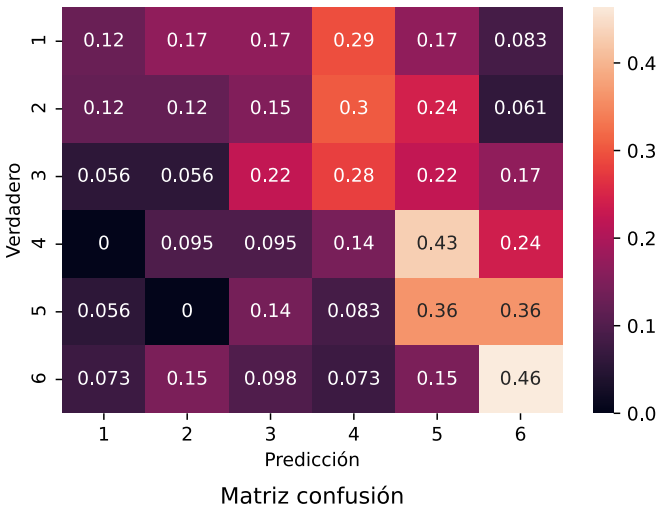


Distribución de predicción solamente imágenes oscuras



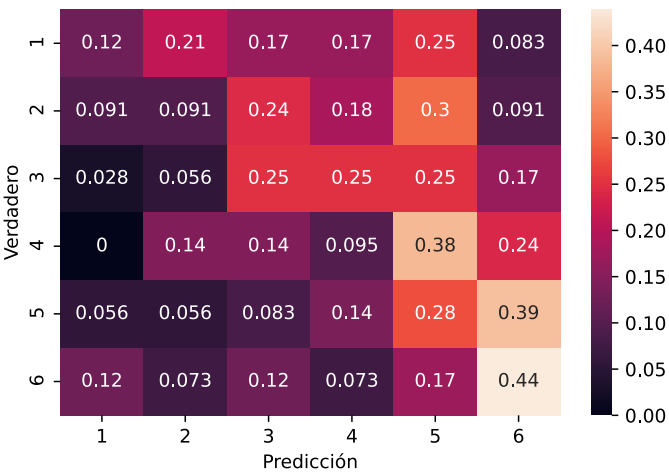
Distribución de predicción sin imágenes oscuras

Fig. 5



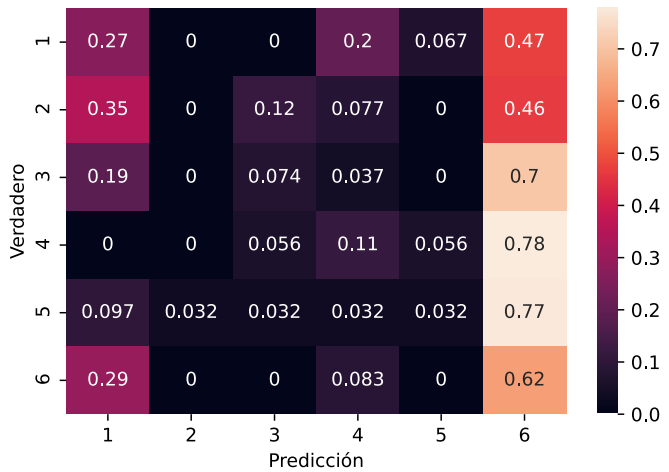
Matriz confusión

Fig. 6



Matriz confusión sin imágenes oscuras

Fig. 7



Matriz confusión solamente imágenes oscuras

Fig. 8

4 Corrección de color

A partir del color de fondo del grupo de fotos de una persona, retocamos el color de las imágenes oscuras. Cómo las imágenes del dataset no fueron tomadas en el mismo ambiente, el proceso de obtener el color del fondo es manual y solo fue realizado a un conjunto pequeño de muestras.

A partir del color del fondo, los pasos para corregir el color de la imagen fueron:

1. Obtenemos el color del fondo objetivo (el que vamos a usar como referencia) mediante un promedio de muestras tomadas del fondo, lo llamamos $V_{referencia}$.
2. Obtenemos el color promedio del fondo de la imagen oscura, lo llamamos $V_{promedio}$.
3. Calculamos el factor de corrección como:

$$V_{factor} = \frac{V_{referencia}}{V_{promedio}}$$

4. Corregimos cada píxel como:

$$V_{corregido} = V_{factor} \cdot V_{original}$$

Observamos en las figuras 9 y 10 ejemplos de corrección de color (obteniendo el color de fondo de referencia manualmente).

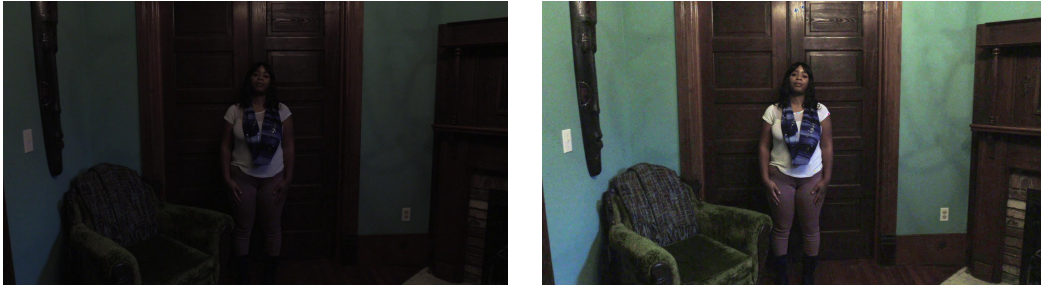


Fig. 9: Corrección de color. Izquierda antes de corregir y derecha luego de corregir.

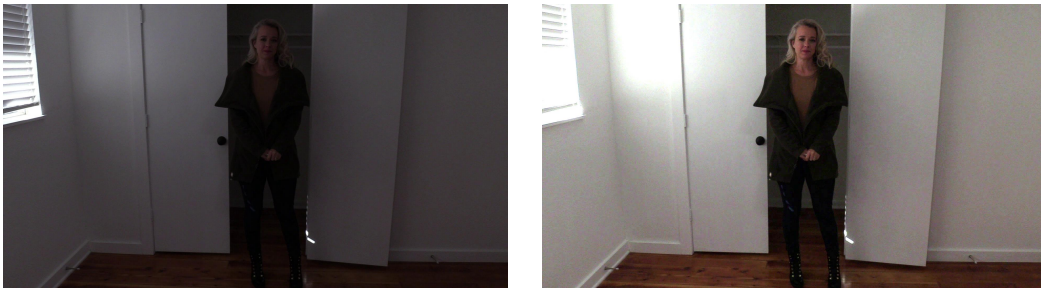


Fig. 10: Corrección de color. Izquierda antes de corregir y derecha luego de corregir.

5 Conclusión

Concluimos que la categorización de tipo de piel es un problema difícil de resolver automáticamente. No es solo difícil para un algoritmo sino que también es difícil para un ser humano clasificar de manera formal y constante como se analiza en el paper[1]. Vimos que la luz es un factor determinante para obtener un resultado preciso. Obtendríamos mejores resultados en un ambiente controlado, donde utilicemos la misma luz. Para finalizar propusimos un algoritmo para la corrección de color.

References

1. Analysis of Manual and Automated Skin Tone Assignments for Face Recognition Applications (2021) (<https://arxiv.org/abs/2104.14685>).