

Indice

Link code: <https://github.com/protossw512/AdaptiveWingLoss>

Autores: Xinyao Wang, Liefeng Bo, Li Fuxin

Año: 2020

Pages: 16

"A fixed loss function cannot achieve both properties simultaneously"

Proceso

Se buscara llegar a lo siguiente:

Ground truth $\approx 1 \rightarrow$ comportamiento como Wing (alta sensibilidad)

Ground truth $\approx 0 \rightarrow$ comportamiento como MSE (baja sensibilidad)

Dataset

COFW, 300W and WFLW

Parámetros

En la función de perdida de se usa

$$AWing(y, \hat{y}) = \begin{cases} \omega \ln\left(1 + \left|\frac{y - \hat{y}}{\epsilon}\right|^{\alpha-y}\right) & \text{if } |(y - \hat{y})| < \theta \\ A|y - \hat{y}| - C & \text{otherwise} \end{cases}$$

Con $\alpha = 2.1$, $\omega = 14$, $\epsilon = 1$ y $\theta = 0.5$

From scratch.

The input of the network is 256×256 ,

The output of each stacked HG is 64×64 .

Optimizer: RMSProp|

Learning rate: $1e-4 \rightarrow 1e-5$ (epoch 80) $\rightarrow 1e-6$ (epoch 160)|

Momentum: 0

Weight decay: $1e-5$

Epochs: 240

Data augmentation: Rotación ($\pm 50^\circ$), Translación ($\pm 25px$), Flip (50%), Rescale ($\pm 15\%$), Gaussian blur, Noise, Occlusion

Métricas de Evaluación

- **Normalized Mean Error (NME)**
 - **Failure Rate (FR)**
 - **Cumulative Error Distribution (CED)**
-

Figuras

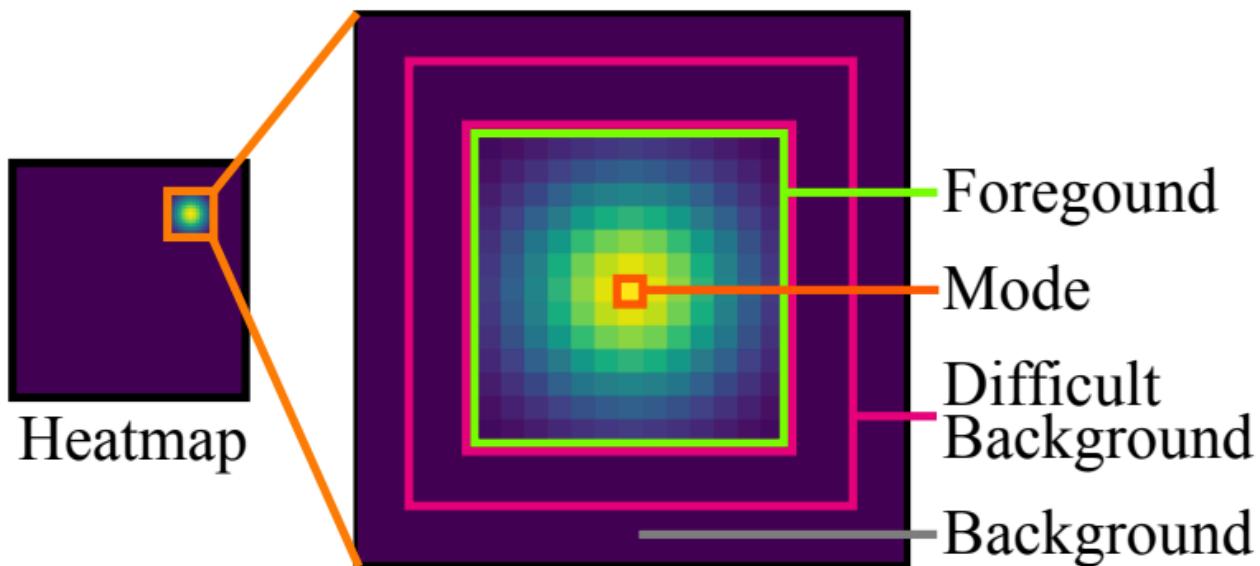
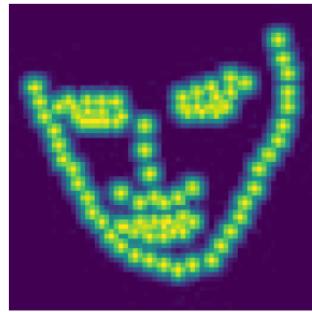


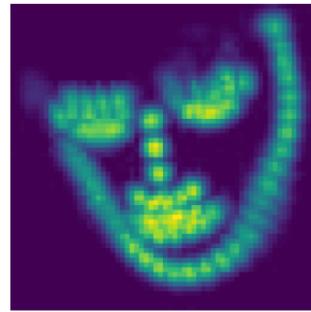
Figure 1: Definición de tipos de pixel



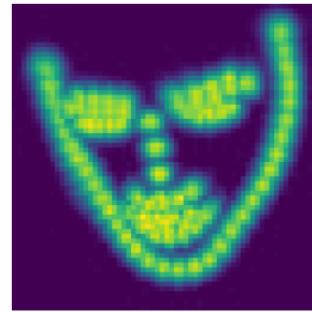
(a) Blurry
Large Pose



(b) GT
Heatmap



(c) MSE
NME: 6.27%



(d) AWing
NME: 4.23%

Figura 2: Comparación de la calidad del mapa de calor predicho. El modelo entrenado con MSE no logró predecir con precisión el mapa de calor alrededor de la mejilla izquierda, la parte inferior de la mejilla derecha y las cejas. Con la pérdida Adaptive Wing propuesta (Fig. 2d), el mapa de calor se vuelve mucho más nítido en los puntos de referencia.

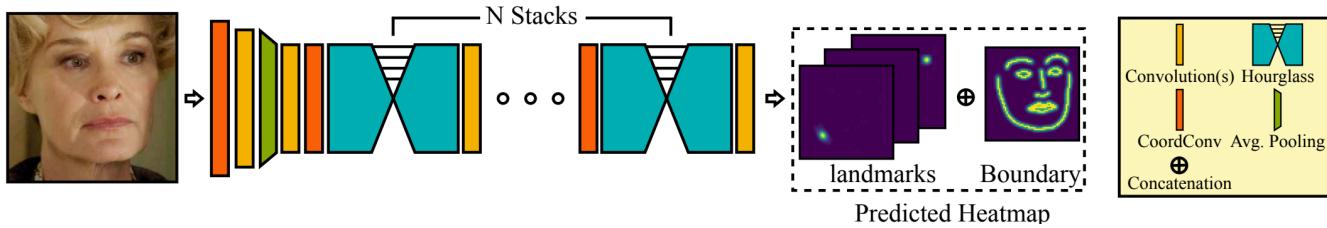
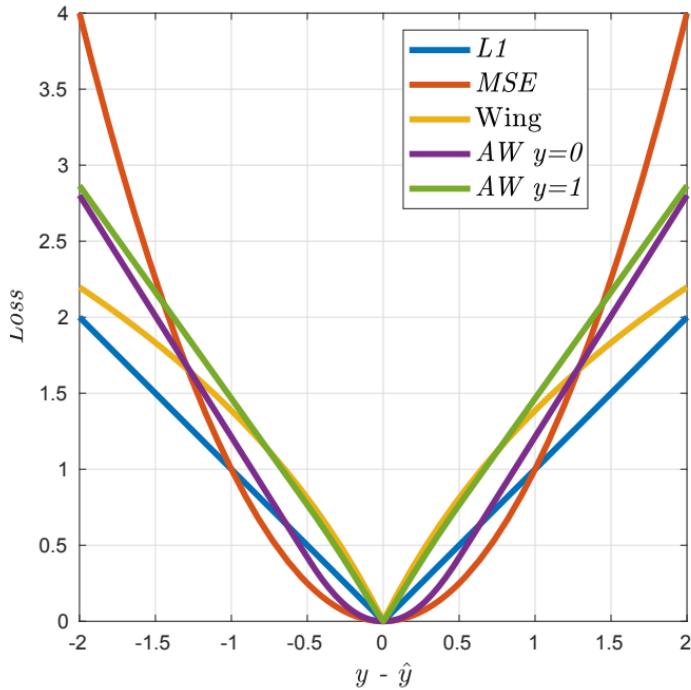
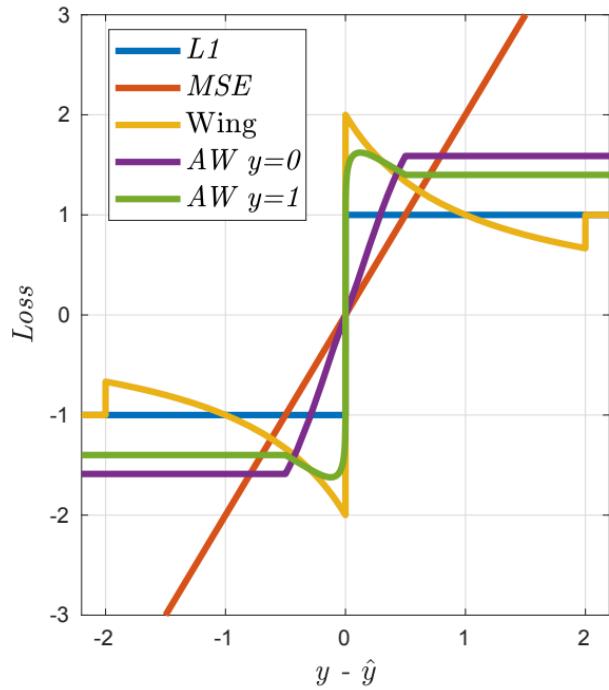


Figura 3: Descripción general de nuestro modelo. El HG apilado toma una imagen facial recortada con el cuadro delimitador de referencia y genera un mapa de calor predictivo para cada punto de referencia. Se utiliza un canal adicional para predecir los límites faciales. Debido a limitaciones de espacio, omitimos la estructura detallada de la arquitectura del HG apilado.

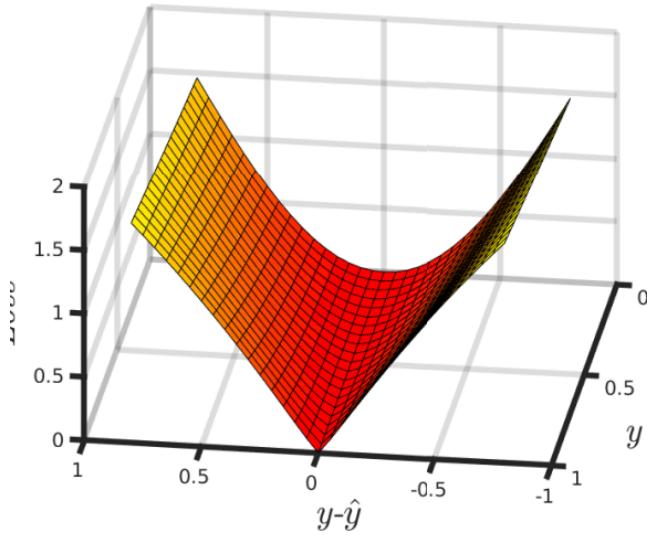


(a) Loss functions

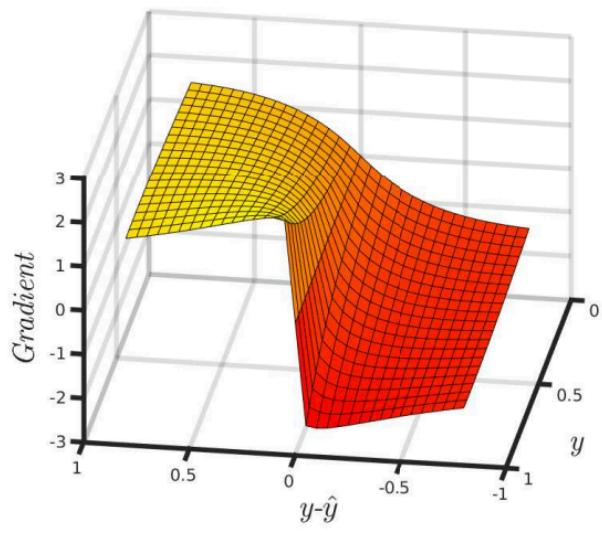


(b) Gradient of losses

Figura 4: Diferentes funciones de pérdida. Cuando $y = 0$, la pérdida de Wing adaptativa (morado) se comporta de forma similar a la pérdida MSE (rojo). Cuando $y = 1$, la pérdida de Wing adaptativa (verde) se comporta de forma similar a la pérdida de Wing (amarillo), pero el gradiente de la pérdida de Wing adaptativa es suave en el punto $y = \hat{y}$, como se muestra en la Figura 4b (se aprecia mejor en color).



(a) AWing loss



(b) Gradient of AWing

Figura 5: La parte no lineal de la pérdida de ala adaptativa puede adaptar su forma según los diferentes valores de y . A medida que y aumenta, la forma se asemeja más a la pérdida de ala, y la influencia de los pequeños errores (cerca del eje y) sigue siendo fuerte. A medida que y

disminuye, la influencia de estos errores disminuye y la función de pérdida se comporta más como el error cuadrático medio (MSE).

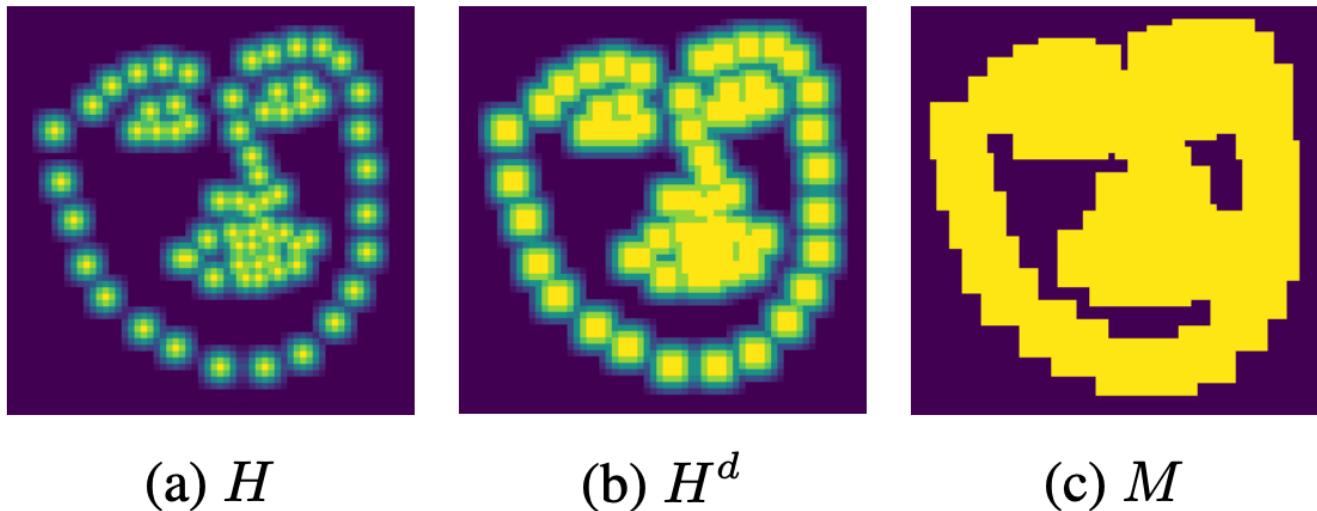


Figura 6: Los píxeles importantes se generan dilatando H de la Figura 6a con una dilatación de 3x3 y luego binarizando a la Figura 6c con un umbral de 0,2. Para fines de visualización, todos los canales se combinan mediante max-pooling en un solo canal.

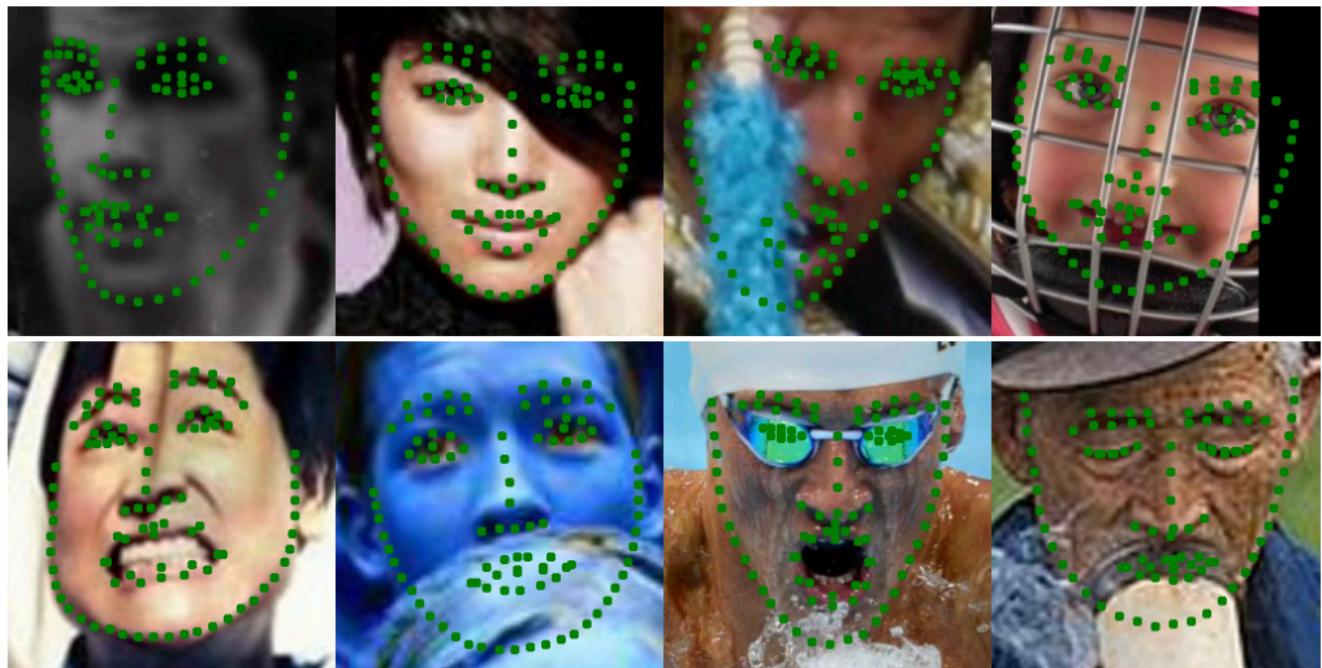


Figura 7: Visualizaciones en el conjunto de datos de prueba WFLW.

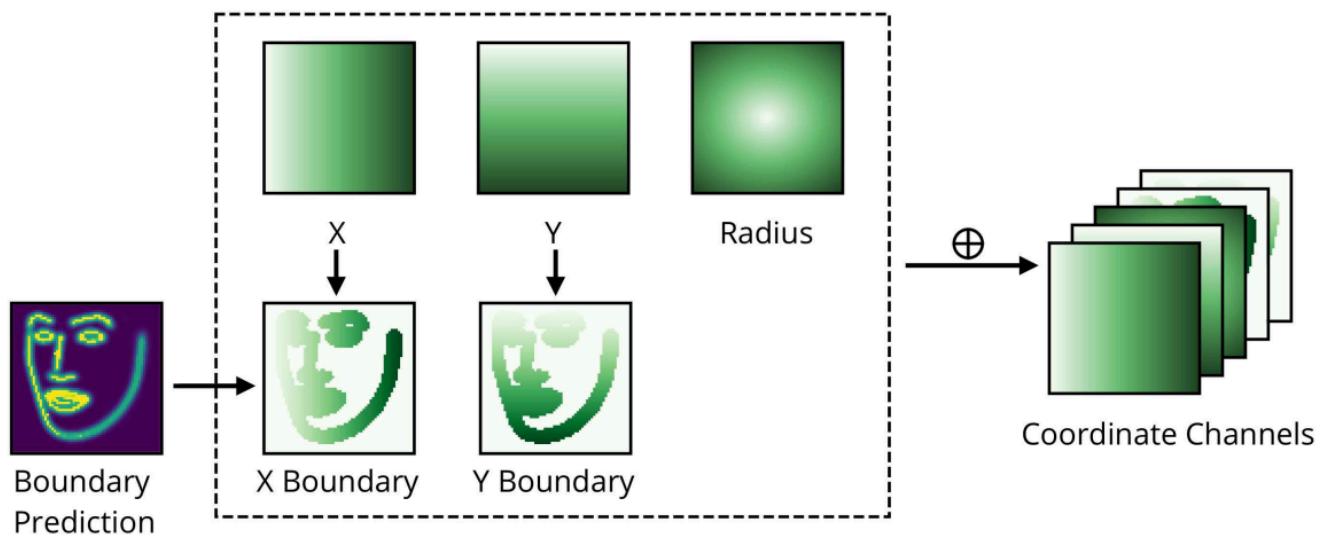


Figura 8: CoodConv con información de contorno. Los contornos X e Y se generan a partir de los canales de coordenadas X e Y, respectivamente, mediante una máscara binaria creada a partir de la predicción de contorno del módulo Hourglass anterior. La máscara se genera aplicando un umbral de 0,05 a la predicción de contorno. (Se visualiza mejor en color).



Figura 9: Visualización de resultados 1. Filas 1-2: Conjunto de datos AFLW, filas 3-4: Conjunto de datos COFW, filas 5-6: Conjunto de datos 300W.

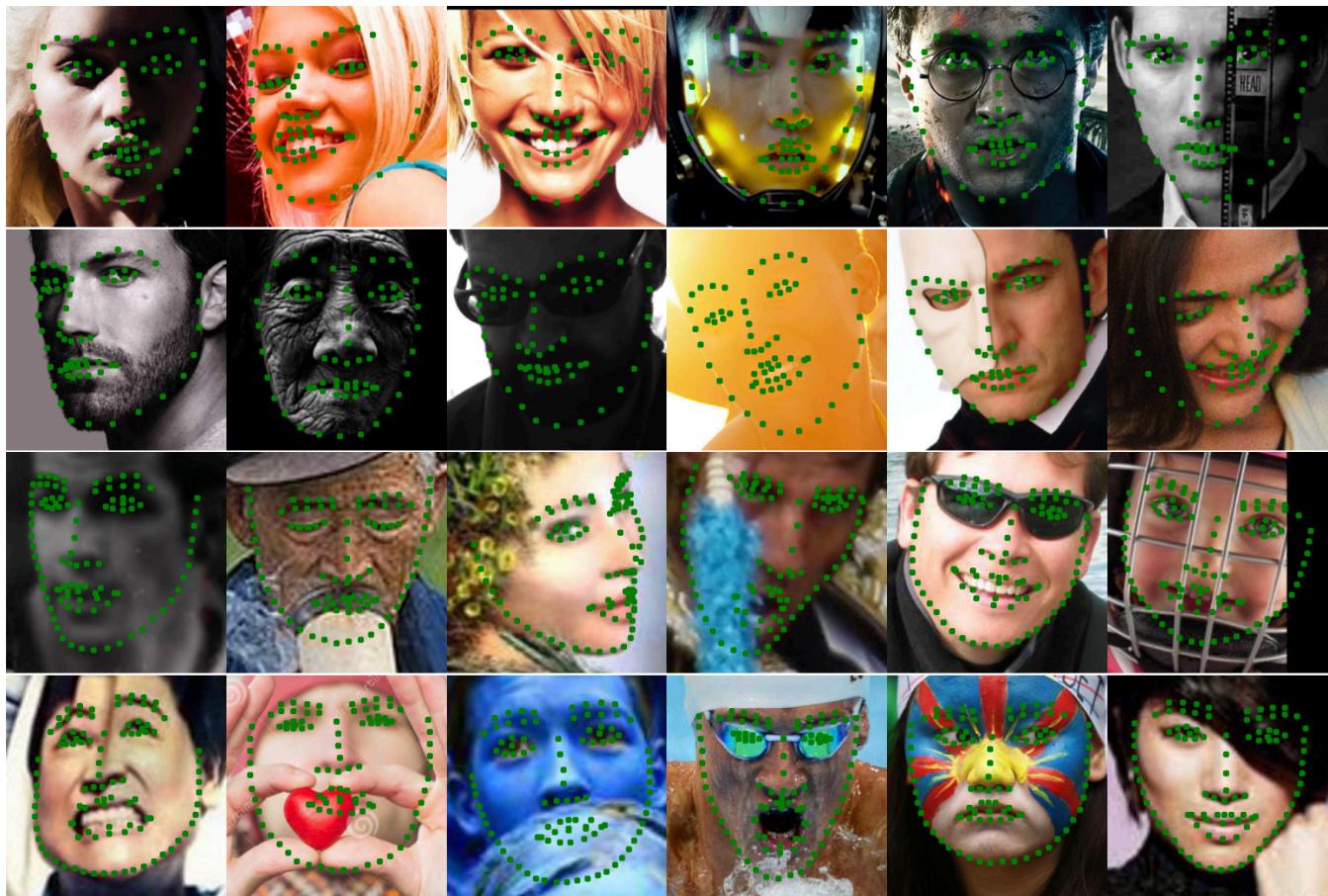


Figura 10: Visualización de resultados 2. Filas 1-2: conjunto de datos privado 300W, filas 3-4: conjunto de datos WFLW.