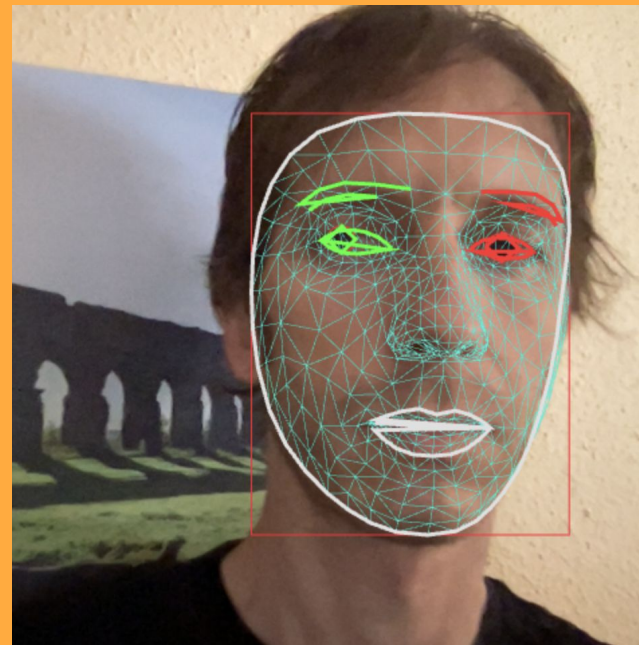


Inteligencia Ambiental

Máster Universitario en Inteligencia Artificial Aplicada

Práctica de reconocimiento de gestos



Contenido

1. Detección de puntos clave de manos con hand-pose-detection
2. Segunda práctica

1. Detección de puntos clave de manos con hand-pose-detection

hand-pose-detection

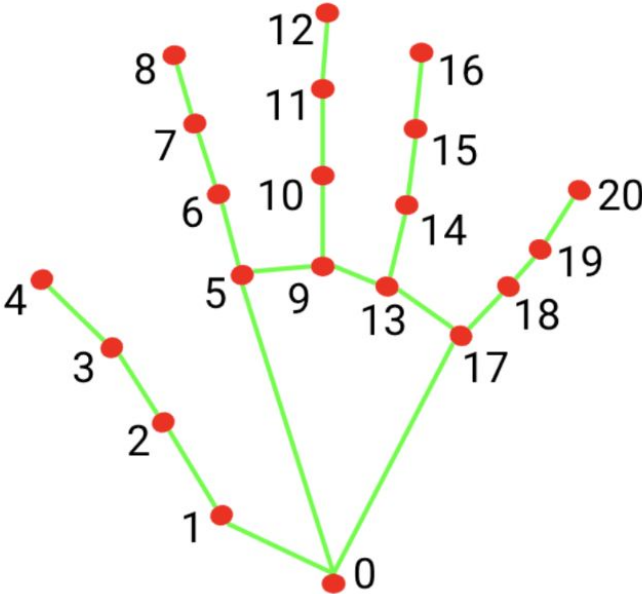
<https://github.com/tensorflow/tfjs-models/tree/master/hand-pose-detection>

- Modelo de red neuronal que se ejecuta en imágenes RGB para detectar y predecir la ubicación de las articulaciones de las manos humanas en tiempo real
- Este modelo es parte de la solución de **MediaPipe** desarrollada por Google. Es altamente optimizado y eficiente, diseñado para funcionar bien en dispositivos en tiempo real.
- Puede ejecutarse a **>30FPS** en la mayoría de los portátiles modernos y dispositivos móviles
- Más adecuado para detectar una o dos manos que están a una distancia de entre **0.5 y 2 metros** de la cámara
- Detecta las manos más visibles en el centro de la imagen y predice **21 puntos clave** de cada mano, incluyendo las posiciones de las articulaciones y las puntas de los dedos
- Proporciona coordenadas 3D de los puntos clave en metros, lo que permite calcular la orientación y la posición en el espacio

hand-pose-detection

- hand-pose-detection contiene un modelo de detección de palma y un modelo de detección de puntos de referencia de la mano
- El modelo de detección de palma ubica las manos dentro de la imagen de entrada
- El modelo de detección de puntos de referencia de la mano identifica puntos de referencia específicos de la mano en la imagen recortada de la mano definida por el modelo de detección de palma

Diagrama de los punto clave

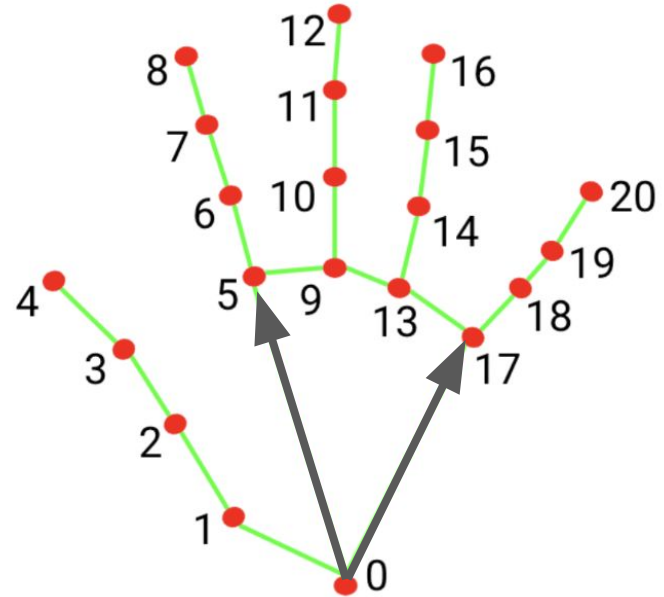


Sigla	Término anatómico	Traducción
CMC	CarpoMetaCarpal joint	Articulación carpometacarpiana
MCP	MetaCarpoPhalangeal joint	Articulación metacarpofalángica
PIP	Proximal InterPhalangeal joint	Articulación interfalángica proximal
DIP	Distal InterPhalangeal joint	Articulación interfalángica distal
TIP	fingerTIP	Punta del dedo

- 0. WRIST
- 1. THUMB_CMC
- 2. THUMB_MCP
- 3. THUMB_IP
- 4. THUMB_TIP
- 5. INDEX_FINGER_MCP
- 6. INDEX_FINGER_PIP
- 7. INDEX_FINGER_DIP
- 8. INDEX_FINGER_TIP
- 9. MIDDLE_FINGER_MCP
- 10. MIDDLE_FINGER_PIP
- 11. MIDDLE_FINGER_DIP
- 12. MIDDLE_FINGER_TIP
- 13. RING_FINGER_MCP
- 14. RING_FINGER_PIP
- 15. RING_FINGER_DIP
- 16. RING_FINGER_TIP
- 17. PINKY_MCP
- 18. PINKY_PIP
- 19. PINKY_DIP
- 20. PINKY_TIP

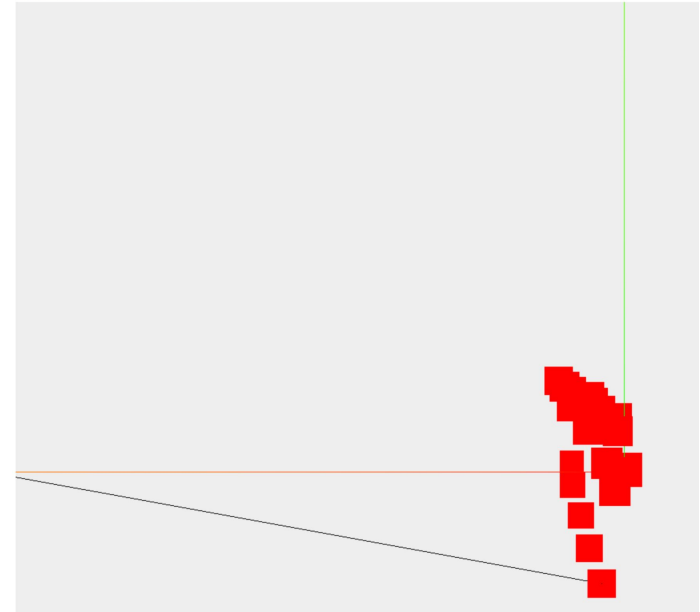
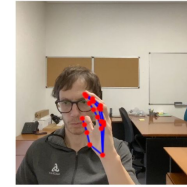
Ejemplo: calcular la dirección de la mano en 3D

- MediaPipe Hand Pose Detection utiliza un sistema de coordenadas 3D para diestros:
 - Eje x: dirección horizontal (de izquierda a derecha desde la perspectiva de la cámara)
 - Eje y: dirección vertical (de arriba a abajo desde la perspectiva de la cámara)
 - Eje z: profundidad (valores negativos más cercanos a la cámara)
- La normal a un plano se calcula con el producto vectorial
- El vector normal es una dirección perpendicular a la palma (hacia fuera) y ayuda a describir su orientación en el mundo
- Para la mano derecha, cuando la palma está orientada hacia la izquierda:
 - Los puntos de referencia INDEX_FINGER_MCP, PINKY_MCP y WRIST definen los vectores utilizados para calcular la normal de la palma



Ejemplo: calcular la dirección de la mano en 3D

- Interpretación de valores que nos da la dirección de los componentes **(para mano derecha!)**
 - $x > 0$: la mano está orientada más hacia la izquierda (**desde la perspectiva de la mano**)
 - $x < 0$: la mano está orientada más hacia la derecha
 - $y > 0$: la palma está orientada hacia arriba
 - $y < 0$: la palma está orientada hacia abajo
 - $z > 0$: la palma está orientada hacia la cámara
 - $z < 0$: la palma está orientada en dirección contraria a la cámara (palma mirando hacia mí)
- Combinación de componentes: al combinar estos componentes, puede describir completamente la orientación de la palma
 - Ejemplo: $(x = 0,5, y = -0,2, z = -0,6)$ indica una palma ligeramente inclinada hacia la izquierda, orientada hacia arriba y hacia la cámara



Pasos para calcular el vector normal

El vector normal representa la dirección perpendicular al plano de la palma

1. Normalizar los puntos
 - a. Calcular la magnitud máxima
 - i. Para cada punto 3D, se calcula la distancia desde el origen (0, 0, 0) utilizando la fórmula de la magnitud: $\text{magnitud} = \sqrt{x^2 + y^2 + z^2}$
 - ii. El mayor valor de estas magnitudes se utiliza como factor de escala
 - b. Escalar los puntos
 - i. Cada punto se normaliza dividiendo sus coordenadas (x, y, z) por este factor de escala, asegurando que todos los puntos tengan valores relativos entre 0 y 1
2. Seleccionar los puntos clave
 - a. wrist: Punto en la muñeca (keypoints3D[0])
 - b. indexMCP: Base del dedo índice (keypoints3D[5])
 - c. pinkyMCP: Base del dedo meñique (keypoints3D[17])
3. Se construyen dos vectores que definen el plano de la palma
 - a. vector1: Desde wrist hacia indexMCP
 - b. vector2: Desde wrist hacia pinkyMCP
4. Calcular el vector normal
 - a. El vector normal se calcula como el **producto vectorial** de los dos vectores

Tutorial 4: Detección de puntos clave de las manos con hand-pose-detection

<https://docs.google.com/document/d/1oDMLipi-8F0bT0EfMS6z71zChpj8g5t9EnF-Wh1U-3c>

Aprender a:

- Cargar el modelo hand-pose-detection en TensorFlow.js
- Ejecutar el modelo prefabricado hand-pose-detection para realizar estimaciones sobre fotogramas de la cámara web
- Visualizar los puntos clave de las manos detectadas y las conexiones entre articulaciones

2. Segunda práctica

Resumen

- Existen dos tipos de modelos pre-entrenados brutos (o sin procesar): (1) Modelos de capas (Layers Model), (2) Modelo de grafo (Graph Model):
- TFJS proporciona modelos pre-entrenados brutos para detectar la ubicación de las articulaciones humanas de una sola persona: MoveNet
- Un tensor es un objeto de la clase `tf.Tensor`
- TFJS proporciona métodos para realizar operaciones con tensores