



TRABAJO FIN DE GRADO
INGENIERÍA INFORMÁTICA

Estimación de la distancia cámara-sujeto en fotografías faciales mediante técnicas de aprendizaje profundo

Autor

Iván Salinas López

Directores

Pablo Mesejo Santiago

Enrique Bermejo Nievas



ESCUELA TÉCNICA SUPERIOR DE INGENIERÍAS INFORMÁTICA Y DE
TELECOMUNICACIÓN

Granada, Junio de 2024

Estimación de distancia cámara-sujeto en fotografías faciales usando aprendizaje profundo

Iván Salinas López

Palabras clave: palabra_clave1, palabra_clave2, palabra_clave3,

Resumen

Poner aquí el resumen.

Camera-subject distance estimation in facial photographs using deep learning

Iván Salinas López

Keywords: Keyword1, Keyword2, Keyword3,

Abstract

Write here the abstract in English.

Yo, **Iván Salinas López**, alumno de la titulación **TITULACIÓN** de la **Escuela Técnica Superior de Ingenierías Informática y de Telecomunicación de la Universidad de Granada**, con DNI 78026145W, autorizo la ubicación de la siguiente copia de mi Trabajo Fin de Grado en la biblioteca del centro para que pueda ser consultada por las personas que lo deseen.

Fdo: Iván Salinas López

Granada a X de Julio de 2023

D. **Nombre Apellido1 Apellido2 (tutor1)**, Profesor del Área de **XXXX** del Departamento **YYYY** de la Universidad de Granada.

D. **Nombre Apellido1 Apellido2 (tutor2)**, Profesor del Área de **XXXX** del Departamento **YYYY** de la Universidad de Granada.



Informan:

Que el presente trabajo, titulado ***Título del proyecto, Subtítulo del proyecto***, ha sido realizado bajo su supervisión por **Nombre Apellido1 Apellido2 (alumno)**, y autorizamos la defensa de dicho trabajo ante el tribunal que corresponda.

Y para que conste, expiden y firman el presente informe en Granada a X de mes de 201 .

Los directores:

Nombre Apellido1 Apellido2 (tutor1) **Nombre Apellido1 Apellido2 (tutor2)**

Agradecimientos

Poner aquí agradecimientos...

Índice general

| | |
|---|----------|
| 1. Introducción | 1 |
| 1.1. Definición del problema | 1 |
| 1.2. Motivación | 3 |
| 1.3. Objetivos | 4 |
| 1.4. Planificación del proyecto | 5 |

Índice de figuras

| | |
|--|---|
| 1.1. Efectos de la distorsión de perspectiva en características faciales de fotografías realizadas a diferentes SCD: 0.5 m, 1 m y 3 m. Estos efectos varían en relación a la distancia y son independientes de la longitud focal [1] | 2 |
| 1.2. Ejemplo de superposición craneofacial [18]. | 3 |
| 1.3. A la izquierda la fotografía distorsionada, a la derecha la fotografía sin distorsión [24]. | 4 |



Índice de cuadros

| | |
|---|---|
| 1.1. Planificación inicial del proyecto | 6 |
|---|---|

Capítulo 1

Introducción

1.1. Definición del problema

Durante la última década, la identificación facial ha ganado una relevancia significativa gracias a la revolución del aprendizaje profundo y los sistemas automáticos de reconocimiento facial. Este avance ha conducido a una expansión de las posibles aplicaciones en el mercado actual desde los campos del cumplimiento de la ley o la ciencia forense, hasta áreas del sector privado como el comercio minorista, las aplicaciones multimedia o la seguridad. Además, el desarrollo tecnológico en el ámbito de la imagen ha mejorado tanto la calidad como la disponibilidad de datos fotográficos, lo cual facilita un análisis más exhaustivo de los factores que influyen en las imágenes fotográficas.

Las técnicas de identificación facial normalmente son realizadas por expertos con o sin la ayuda de sistemas automáticos. Los expertos analizan los datos y evalúan las características anatómicas de un individuo desconocido para compararlas con las de uno o varios individuos conocidos. Actualmente existen cuatro métodos de comparación facial reconocidos: análisis morfológico, superposición, fotogrametría y comparación holística [8].

Para que este análisis sea confiable y concluyente, los datos (fotografías faciales), deben estar en unas condiciones adecuadas (calidad, resolución, enfoque o iluminación) y la escena (punto de vista de la cámara, pose de la cabeza, expresión facial) debe ser lo más neutral y representativa posible [7, 17]. Estos requisitos nos aseguran que los rasgos faciales sean lo más fieles posible a las características anatómicas del individuo, y por tanto, permiten que las técnicas de identificación sean más robustas.

Muchos estudios han identificado limitaciones en los actuales métodos de reconocimiento facial automático. Los principales factores desafiantes son la pose, la iluminación, la expresión y la variación en la edad [12, 13]. Sin

embargo, también existen otros factores importantes como la **oclusión**, el **género o la etnia** [3, 10].

Uno de los factores más importantes en las fotografías faciales es la distorsión de perspectiva, la cual puede provocar deformaciones en los rasgos faciales, como en las orejas, la nariz o la forma general del rostro, especialmente cuando la cámara está muy cerca del sujeto al momento de tomar la fotografía [14] (ver Figura 1.1). Esta alteración en la perspectiva repercute negativamente en los sistemas de reconocimiento automático [6, 16, 20], lo cual puede obstaculizar la identificación precisa de individuos.



Figura 1.1: Efectos de la distorsión de perspectiva en características faciales de fotografías realizadas a diferentes **SCD**: 0.5 m, 1 m y 3 m. Estos efectos varían en relación a la distancia y son independientes de la longitud focal [1]

La distorsión de perspectiva está estrechamente relacionada con la distancia cámara-sujeto (**subject-to-camera distance**, SCD en adelante), de hecho, la relación es de decremento logarítmico, es decir, valores pequeños de la SCD corresponden con una mayor distorsión, mientras que la distorsión disminuye conforme la **SCD** aumenta [18]. Conocer la SCD en fotografías faciales permite cuantificar la cantidad de distorsión presente en la imagen, así como las diferencias en la distorsión entre dos conjuntos de imágenes. Esta información es útil para reproducir con fiabilidad las condiciones originales de una escena, ~~especialmente en el ámbito forense~~, y para facilitar el desarrollo de técnicas que permitan la corrección precisa de dicha distorsión.

La SCD, a diferencia de otros parámetros de la cámara como la longitud focal o el tamaño del sensor, no puede obtenerse directamente desde los metadatos de la fotografía [21]. Por tanto, se necesita un método preciso para su estimación.

En los últimos años se han utilizado varios métodos que combinan técnicas manuales y automatizadas basadas en puntos de referencia o en características anatómicas de la cara [9, 4]. Sin embargo, no se han obtenido resultados favorables debido a la dificultad para obtener estimaciones precisas en largas distancias por la diversa fisonomía de la cara, y a los problemas

relacionados con los parámetros de la cámara, como el recorte de imágenes o la combinación de diferentes longitudes focales en el mismo conjunto de datos.

El único método totalmente automatizado para estimar la SCD en fotografías faciales, hasta la fecha, se conoce como FacialSCDnet [1]. Este método utiliza una arquitectura basada en aprendizaje profundo para procesar fotografías faciales y estimar la SCD con precisión. Sin embargo, las soluciones basadas en aprendizaje profundo tienen sesgos sobre los datos y las muestras deben ser lo más representativas posible.

Considerando todos estos aspectos, el presente Trabajo de Fin de Grado (TFG) consiste en mejorar el método actual del estado del arte en la estimación automática de la distancia cámara-sujeto en fotografías faciales mediante el uso de técnicas de aprendizaje profundo.

1.2. Motivación

En el ámbito forense, el principal foco recae sobre la determinación de la identidad humana cuando se dispone de información esquelética [2]. En las últimas décadas, los antropólogos han centrado su atención en mejorar las técnicas para realizar una identificación más precisa. En este contexto, la estimación de la SCD juega un papel crucial, ya que, si estimamos la SCD con fiabilidad, podemos recrear la imagen con los restos esqueléticos (aplicando el valor de la SCD) [19]. A continuación, realizamos las comparaciones anatómicas mediante la superposición craneofacial [5] para identificar si se corresponde con la misma persona (ver Figura 1.2). Si los parámetros de adquisición entre ambas imágenes (normal y esquelética) son distintos entonces dificultaría el análisis morfológico [11].

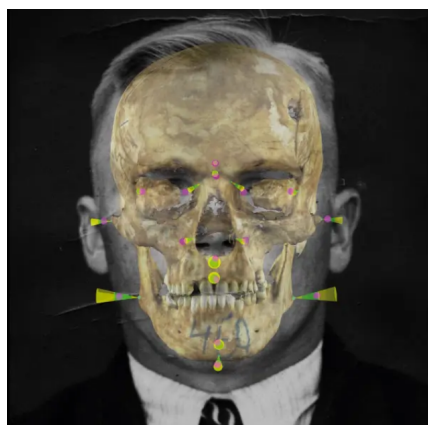


Figura 1.2: Ejemplo de superposición craneofacial [18].



En el ámbito de la biometría y el reconocimiento facial, conocer la SCD es fundamental para comprender y corregir las distorsiones de perspectiva que pueden afectar a la precisión de la identificación. Esta distancia influye directamente en la representación facial capturada, afectando la geometría y proporciones del rostro (ver Figura 1.3). Un entendimiento preciso de esta medida permite aplicar técnicas de corrección de distorsión [23, 24] para garantizar una representación fiel y coherente de las características faciales, mejorando de esta manera la precisión y fiabilidad de los sistemas biométricos y de reconocimiento facial.



Figura 1.3: A la izquierda la fotografía distorsionada, a la derecha la fotografía sin distorsión [24].



En el campo del aprendizaje profundo, es común encontrar sesgos [22] en los datos que pueden afectar a la precisión y generalización de las soluciones. Dichos sesgos pueden surgir de diversos factores, como la falta de diversidad o el ruido excesivo en el conjunto de datos utilizado para entrenar los modelos. Esto también ocurre en los métodos actuales para la estimación automática de la SCD. Es por ello que, una forma de mejorar el conjunto de datos podría ser incluir una variedad más amplia de grupos étnicos, edades, géneros, expresiones faciales, condiciones de iluminación y fondos, así como modelos tanto faciales como de cuerpo completo. Al hacerlo, se busca mitigar los sesgos inherentes a los datos y desarrollar modelos más robustos y equitativos. Esta mejora implica una mayor capacidad de generalización y adaptación a una amplia gama de situaciones y contextos.

Por todo ello, la estimación de la SCD de forma precisa y generalizable tiene el potencial de impulsar avances significativos en diversos campos de investigación.

1.3. Objetivos



El objetivo general de este TFG consiste en desarrollar un modelo de aprendizaje profundo adecuado para mejorar la estimación de la distancia cámara-sujeto en fotografías faciales. Para el desarrollo del proyecto, dividi-

remos el objetivo general en una serie de objetivos parciales:

1. Realizar un análisis exhaustivo del estado del arte para la estimación de la SCD en fotografías faciales.
2. Generar un conjunto de datos sintético de mayor calidad (utilizando modelos 3D más completos de cuerpo entero y poses distintas, con fondos e iluminación más realistas).
3. Realizar un estudio experimental que permita validar los enfoques propuestos y extraer conclusiones sobre su aplicabilidad al problema.
4. Desarrollar y entrenar con nuevas arquitecturas que mejoren los resultados.
5. Usar tecnologías más recientes que mejoren los tiempos de aprendizaje y los resultados obtenidos.

1.4. Planificación del proyecto

Para abordar el desarrollo de este proyecto, es esencial considerar que el TFG tiene asignados 12 créditos ECTS, lo que equivale a aproximadamente 300 horas de trabajo. Dada la distribución temporal del segundo cuatrimestre, con unas 20 semanas disponibles, se estima que se requerirá dedicar al TFG unas 20 horas semanales, equivalentes a 4 horas diarias durante 5 días a la semana. Se reservan así 4 semanas como margen para posibles retrasos o imprevistos que puedan surgir durante el desarrollo del proyecto.

En cuanto a la metodología de desarrollo, se ha optado por seguir un enfoque basado en el ciclo de vida en cascada [15], aunque con una variante que permite retroalimentación. Aunque el proyecto presenta requisitos y objetivos claros, se reconoce la posibilidad de ajustes menores durante su desarrollo, especialmente a medida que se obtenga más información sobre el problema y los métodos. Esta flexibilidad se considera crucial para adaptarse a posibles cambios en el contexto o los requisitos del proyecto.

Las fases del ciclo de vida del proyecto son las siguientes:


A continuación se describen las fases del ciclo de vida del proyecto:

- **Análisis de Requisitos:** Consiste en las reuniones iniciales con los clientes, en este caso los directores del TFG. Se realiza un análisis del problema y un estudio detallado de la bibliografía existente
- **Diseño:** Consiste en la exploración y selección de los métodos apropiados basados en el análisis previo, tanto para la resolución como para la



validación de la solución propuesta. Además, se llevarán a cabo pruebas preliminares y se elaborará el diseño del software experimental.

- Implementación: Consiste en la adaptación del código de los modelos investigados, la implementación de nuevas funcionalidades y la generación de un conjunto de datos sintético junto con su posterior pre-procesado.
- Pruebas: Consiste en la realización de diversos experimentos para validar el funcionamiento del software desarrollado, utilizando los modelos y datos previamente definidos.



| Tarea | Semanas - Horas | Febrero | | | | Marzo | | | | Abril | | | | | Mayo | | | | Junio | | |
|------------------------|--------------------|---------|----|----|----|-------|----|----|----|-------|---|----|----|----|------|----|----|----|-------|----|----|
| | | 5 | 12 | 19 | 26 | 4 | 11 | 18 | 25 | 1 | 8 | 15 | 22 | 29 | 6 | 13 | 20 | 27 | 3 | 10 | 17 |
| Análisis de Requisitos | 3 - 60 | | | | | | | | | | | | | | | | | | | | |
| Diseño | 3 - 60 | | | | | | | | | | | | | | | | | | | | |
| Implementación | 5 - 90 | | | | | | | | | | | | | | | | | | | | |
| Pruebas | 5 - 90 | | | | | | | | | | | | | | | | | | | | |

Tabla 1.1: Planificación inicial del proyecto



Bibliografía

- [1] Enrique Bermejo et al. «FacialSCDnet: A deep learning approach for the estimation of subject-to-camera distance in facial photographs». En: *Expert Systems with Applications* 210 (2022), pág. 118457. ISSN: 0957-4174.
- [2] Soren Blau y Luis Fondebrider. «Human Skeletal Remains: Identification of Individuals». En: *Encyclopedia of Global Archaeology*. 2014, págs. 3556-3558.
- [3] Joy Buolamwini y Timnit Gebru. «Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification». En: *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*. Vol. 81. 2018, págs. 77-91.
- [4] Xavier P. Burgos-Artizzu, Matteo Ruggero Ronchi y Pietro Perona. «Distance Estimation of an Unknown Person from a Portrait». En: *Computer Vision – ECCV 2014*. Vol. 8689. 2014, págs. 313-327.
- [5] Sergio Damas et al. «Forensic identification by computer-aided craniofacial superimposition: A survey». En: *ACM Computing Surveys* 43 (2011), págs. 1-27. ISSN: 0360-0300.
- [6] Naser Damer et al. «Deep Learning-based Face Recognition and the Robustness to Perspective Distortion». En: *24th International Conference on Pattern Recognition*. 2018, págs. 3445-3450.
- [7] Gary Edmond et al. «Law’s Looking Glass: Expert Identification Evidence Derived from Photographic and Video Images». En: *Current Issues in Criminal Justice* 20 (2009), págs. 337-377.
- [8] FISWG. *Facial Comparison Overview and Methodology Guidelines V2.0*. Accedido el 7 de Febrero de 2024. 2022. URL: https://fiswg.org/fiswg_facial_comparison_overview_and_methodology_guidelines_V2.0_2022.11.04.pdf.
- [9] Arturo Flores et al. «Camera Distance from Face Images». En: *Advances in Visual Computing*. Vol. 8034. 2013, págs. 513-522.

- [10] Nicholas Furl, P.Jonathon Phillips y Alice J O'Toole. «Face recognition algorithms and the other-race effect: computational mechanisms for a developmental contact hypothesis». En: *Cognitive Science* 26 (2002), págs. 797-815. ISSN: 0364-0213.
- [11] Sean Healy y Carl Stephan. «Perspective distortion tolerances and skull-face registration in craniofacial superimposition: an analytical review». En: *International Journal of Legal Medicine* 137 (2023).
- [12] Anil K. Jain, Brendan Klare y Unsang Park. «Face Matching and Retrieval in Forensics Applications». En: *IEEE MultiMedia* 19 (2012), págs. 20-20.
- [13] Zhifeng Li, Unsang Park y Anil K. Jain. «A Discriminative Model for Age Invariant Face Recognition». En: *IEEE Transactions on Information Forensics and Security* 6 (2011), págs. 1028-1037.
- [14] Bo Peng et al. «Position Determines Perspective: Investigating Perspective Distortion for Image Forensics of Faces». En: *IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2017, págs. 1813-1821.
- [15] R.S. Pressman. *Software Engineering: A Practitioner's Approach*. 2005. ISBN: 9780073019338.
- [16] Zahid Riaz. y Michael Beetz. «On the effect of perspective distortions in face recognition». En: *Proceedings of the International Conference on Computer Vision Theory and Applications*. Vol. 1. 2012, págs. 718-722.
- [17] Nicole Spaun. «Facial Comparisons by Subject Matter Experts: Their Role in Biometrics and Their Training». En: *International Conference on Biometrics*. 2009, págs. 161-168.
- [18] Carl N. Stephan. «Perspective distortion in craniofacial superimposition: Logarithmic decay curves mapped mathematically and by practical experiment». En: *Forensic Science International* 257 (2015), 520.e1-520.e8. ISSN: 0379-0738.
- [19] Carl N. Stephan. «Estimating the Skull-to-Camera Distance from Facial Photographs for Craniofacial Superimposition». En: *Journal of Forensic Sciences* 62 (2017), págs. 850-860.
- [20] Joachim Valente y Stefano Soatto. «Perspective distortion modeling, learning and compensation». En: *IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2015, págs. 9-16.
- [21] Andrea Valsecchi. *Comprensión de los parámetros de la cámara*. Accedido el 17 de Febrero de 2024. 2019. URL: <https://skeleton-id.com/investigaciones/comprension-de-los-parametros-de-la-camara>.

-
- [22] Gal Vardi. «On the Implicit Bias in Deep-Learning Algorithms». En: *Communications of the ACM* 66 (2022), págs. 86-93.
 - [23] Zhixiang Wang et al. «DisCO: Portrait Distortion Correction with Perspective-Aware 3D GANs». En: *arXiv preprint arXiv:* (2023).
 - [24] Yajie Zhao et al. «Learning Perspective Undistortion of Portraits». En: *17th IEEE/CVF International Conference on Computer Vision*. 2019, págs. 7848-7858.

Búsquedas en Scopus

Estimación de la distancia en fotografías faciales (430)

TITLE-ABS-KEY ((distance OR depth) AND estimation AND (photographs OR images) AND facial) AND (LIMIT-TO (SUBJAREA , “COMP”) OR LIMIT-TO (SUBJAREA , “ENGI”))

Estimación de la distancia en fotografías faciales mediante IA (218)

TITLE-ABS-KEY ((deep AND learning) OR (machine AND learning) OR (artificial AND intelligence) OR (computer AND vision) OR (soft AND computing) AND ((distance OR depth) AND estimation AND (photographs OR images) AND facial)) AND (LIMIT-TO (SUBJAREA , “COMP”) OR LIMIT-TO (SUBJAREA , “ENGI”))

Estimación de la distancia en fotografías faciales mediante deep learning (126)

TITLE-ABS-KEY ((deep AND learning) AND ((distance OR depth) AND estimation AND (photographs OR images) AND facial)) AND (LIMIT-TO (SUBJAREA , “COMP”) OR LIMIT-TO (SUBJAREA , “ENGI”))