

## Viabilidad de Búsqueda de Parentesco en Problema Inverso Familiar.

Julieta Goria<sup>1{0009-0001-7087-6871}</sup>, Francisco Sandalinas<sup>1</sup>, Pablo Negri<sup>1,2{0000-0003-0250-5208}</sup>, y Daniel Acevedo<sup>1,2{0000-0002-2170-2533}</sup>.

<sup>1</sup> Departamento de Computación, FCEN-UBA

<sup>2</sup> Instituto de Investigación en Ciencias de la Computación (ICC), UBA-CONICET

[jgoria@dc.uba.ar](mailto:jgoria@dc.uba.ar), [fsandalinas@dc.uba.ar](mailto:fsandalinas@dc.uba.ar), [pnegri@dc.uba.ar](mailto:pnegri@dc.uba.ar),  
[davecedo@dc.uba.ar](mailto:davecedo@dc.uba.ar)

**Resumen.** El reconocimiento familiar a partir de imágenes faciales es una tarea desafiante que amplía la verificación facial tradicional al incorporar variaciones genéticas y generacionales. Mientras que los enfoques existentes, como los explorados en el desafío Recognizing Families In the Wild (RFIW), se centran en verificar relaciones familiares en una dirección temporal lineal (comparando mayormente padres con hijos jóvenes), nuestra investigación aborda un problema inverso: la identificación de hijos mediante la comparación de adultos con versiones más jóvenes de sus padres.

El objetivo de este trabajo es determinar si esta formulación inversa constituye un problema distinto al reconocimiento de parentesco tradicional o si puede resolverse con los mismos enfoques. Para ello, utilizamos ArcFace [Deng et al., 2019] para la alineación de rostros y extracción de *embeddings* faciales. Desarrollamos además un conjunto de datos que refleja esta inversión temporal, basado en imágenes extraídas del sitio IMDb (Internet Movie Database). Evaluamos el comportamiento del modelo comparando las distribuciones de similitud del coseno en ambos datasets: el tradicional (FIW) y el nuevo dataset inverso.

**Palabras Clave:** ArcFace, Reconocimiento Familiar, Reconocimiento Facial.

**Abstract.** Kinship recognition from facial images is a challenging task that broadens traditional facial verification techniques by incorporating genetic and generational variations. While preexisting approaches, such as those explored in the Recognizing Families In the Wild (RFIW), focus on verifying familial relationships in a standard temporal direction (mostly comparing parents to their younger children), our research addresses the inverse problem: identifying children by comparing adults with younger versions of their parents.

This work's goal is to determine whether this inverse formulation constitutes a different problem from traditional kinship recognition, or whether it can be addressed using the same approaches. To this end, we used ArcFace [Deng et al., 2019] for facial alignment and embedding extraction. We also developed a new dataset reflecting this temporal inversion, based on images extracted from IMDb (Internet Movie Database). We evaluated the model's performance by comparing the distribution of the cosine similarity in both

datasets: the traditional approach (FIW) and our newly proposed inverse dataset.

**Palabras Clave:** ArcFace, Kinship Recognition, Facial Recognition.

## 1 Introducción

El reconocimiento de parentesco a partir de imágenes faciales es una línea de investigación reciente con aplicaciones relevantes en diversas áreas como el análisis forense, reuniones de familias desplazadas, genealogía y redes sociales. En particular, tiene un potencial significativo para contribuir a la identificación de personas desaparecidas, como en el caso de los nietos apropiados durante la última dictadura militar en Argentina.

Esta tarea, conocida como *kinship verification*, fue formalizada a gran escala por el desafío “Recognizing Families In the Wild” (RFIW) [Robinson et al., 2017]. RFIW propone tareas como la verificación binaria de parentesco, la verificación padre/madre-hijo/a y la búsqueda de familiares en una galería, utilizando un dataset de más de 500.000 pares de imágenes con siete tipos de relaciones familiares.

Estas tareas han impulsado el desarrollo de modelos basados en redes neuronales profundas [Robinson et al., 2022], como SphereFace [Liu et al., 2017] y ArcFace [Deng et al., 2019], que establecieron líneas sólidas para evaluar el parecido familiar en contextos no controlados. El equipo que logró el mejor rendimiento en el desafío RFIW, TeamCNU, utilizó una arquitectura basada en las características de ArcFace y una función de triplet loss [Zhang et al., 2021] para la verificación de parentesco. ArcFace, en particular, ha sido clave para mejorar la precisión en la discriminación de identidades al incorporar una pérdida angular en el espacio de características, lo que permite maximizar las diferencias inter-clase y minimizar las variaciones intra-clase. Este enfoque ha demostrado ser especialmente útil para tareas de verificación de parentesco, brindando una mayor robustez en condiciones no controladas como las presentes en el reconocimiento de familiares a partir de imágenes históricas. Estos avances sientan las bases para adaptar y aplicar estas tecnologías a problemáticas locales.

En este trabajo se propone una metodología para analizar la viabilidad del reconocimiento de parentesco en un escenario inverso al tradicional, donde se busca identificar posibles hijos comparando sus rostros actuales con imágenes de sus padres a una menor edad relativa a la suya. Nuestra hipótesis plantea que la distribución de similitudes faciales, medidas a través de distancias del coseno entre los *embeddings* generados por ArcFace, debería mantener la noción de separabilidad entre el problema tradicional y su versión inversa. Es decir, si los rasgos faciales heredables que permiten reconocer el parentesco son suficientemente robustos, la comparación hijo adulto con padre joven debería mostrar una distribución análoga o similar a la comparación padre adulto con hijo joven.

El resto del artículo se estructura de la siguiente manera: la siguiente sección presenta la metodología de recolección y armado del dataset, luego en la sección

3 se detallan las métricas y métodos de evaluación y se realiza una comparación contra el dataset FIW, y en la sección 4 se esbozan posibles líneas futuras de investigación.

## 2 Metodología

En esta sección se presenta el diseño de un *pipeline* que combina técnicas modernas de detección y reconocimiento facial con un conjunto de datos específicamente creado para representar la problemática planteada anteriormente.

### 2.1 Dataset IMDB inverso

A partir de imágenes disponibles públicamente en IMDb (figura 1), se construyó un conjunto de datos orientado específicamente al problema de inversión temporal. El mismo incluye imágenes donde para cada par hijo/(padre/madre), el rango etario del progenitor es de entre 15 a 40 años, y de cada hijo de 20 años en adelante.



Figura 1: Ejemplo de pares familiares del dataset IMDb. La fila superior muestra las capturas de los hijos, y la fila inferior las capturas de los respectivos padres.

Se realizó un proceso de *scrapping* a partir del dataset abierto de IMDb. El mismo asocia individuos (actores, asistentes de producción, etc.) con su ID único en IMDb, junto con algunos datos personales básicos como nombre y año de nacimiento. El protocolo de búsqueda, descarga y procesamiento se dividió en las siguientes etapas, ilustrado en la figura 2:

1. A partir de un *id* de IMDb, se accedió al perfil del individuo para determinar sus relaciones de parentesco, a las cuales es posible acceder a través de un hipervínculo y a su vez obtener los *ids* correspondientes.

2. Se descargaron las imágenes de los individuos asociados a cada familia que cumplan con los requisitos etarios y se categorizaron como tal.
3. Se realizó un proceso de alineación y detección de rostros con InsightFace, óptimo para el posterior procesamiento con ArcFace. Luego las imágenes se recortaron en un parche de tamaño  $112 \times 112$  pixels.
4. Fue necesario como último paso un proceso de curado automático para eliminar intrusos en las imágenes. Para el mismo, se obtuvieron los *embeddings* de rostros de cada imagen del paso anterior, y se agruparon según la distancia del coseno. Se preservaron los rostros con menor distancia entre una imagen de referencia del individuo en cuestión, la cual presenta un único rostro para asegurar la correcta identificación del individuo.

Esto produjo un dataset de 275 imágenes, los cuales produjeron 2200 pares positivos. Estas imágenes se organizaron por familia y persona, siguiendo una jerarquía análoga al dataset FIW.



Figura 2: Pasos para la descarga e identificación de las fotografías de cada celebridad (Imagen tomada de Di Ielsi and Negri, 2021).

## 2.2 Evaluación

Con el objetivo de analizar el comportamiento de los datos en el escenario tradicional y en el inverso, se compararon las distribuciones de similitud del coseno normalizadas (entre 0 y 1) para pares de imágenes faciales en ambos conjuntos de datos: FIW y el nuevo dataset IMDb inverso.

Para esta evaluación, se calcularon las distancias del coseno entre pares de *embeddings* generados con ArcFace, distinguiendo dos categorías:

- Pares positivos: personas con vínculo familiar directo (padre-hijo).
- Pares negativos: personas sin parentesco.

La idea central fue observar si las distancias del coseno permiten una separación efectiva entre pares positivos y negativos en ambos contextos. No se aplicaron técnicas de balanceo de clases ni se realizó un control exhaustivo de condiciones intra-clase (como pose, edad o iluminación). Sin embargo, para tareas posteriores (como un posible fine-tuning), será necesario un proceso de curado más riguroso que garantice la representatividad y calidad del conjunto.

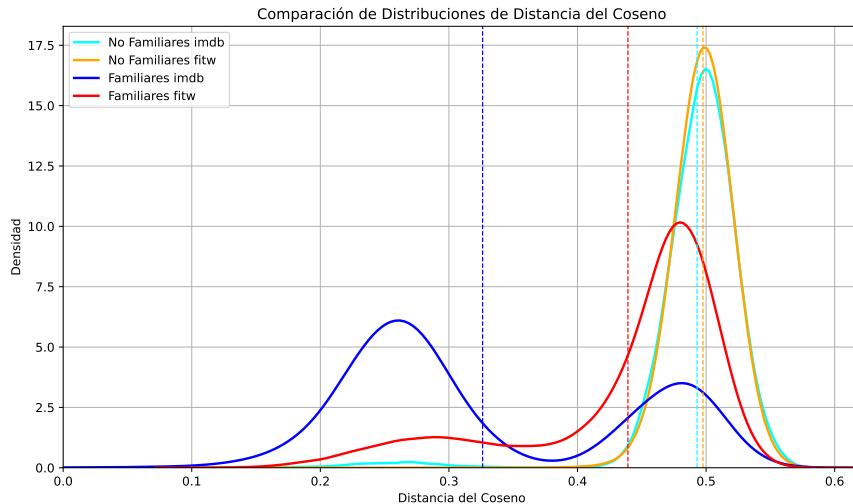


Figura 3: Imagen comparativa de la distribución de distancias del coseno según dataset y clasificación familiar.

### 3 Resultados.

Se observa de la figura 3 una marcada diferencia entre las distancias del coseno normalizadas para los pares familiares y no familiares del dataset IMDb. Esto sugiere una coherencia visual fuerte entre los integrantes de una misma familia. Por su parte, los pares no familiares presentan una distribución centrada en valores más altos (alrededor de 0.5), lo que permite una separación clara entre ambas clases. Este resultado se ve respaldado por la presencia de algunos casos atípicos en cada grupo, ubicados principalmente en la zona correspondiente a la clase opuesta. Esto podría indicar errores de etiquetado o similitudes visuales.

En el dataset FIW, si bien los pares familiares también tienden a concentrarse en distancias más bajas, la diferencia con los no familiares es menos pronunciada. Las distribuciones de ambas clases se superponen considerablemente en torno a 0.5, lo que dificulta una separación clara. Este comportamiento puede deberse a una mayor diversidad en los vínculos familiares representados, diferencias de

edad más marcadas, menor calidad de imagen o la inclusión de relaciones más distantes. La baja separabilidad en FIW sugiere un escenario más complejo y realista, lo que refuerza su valor como benchmark para evaluar la capacidad de generalización de los modelos de reconocimiento de parentesco.

## 4 Conclusiones

En conjunto, estos resultados muestran que la distribución de distancias del coseno no solo refleja la similitud facial, sino también la complejidad inherente del dataset. IMDb presenta un caso de alta separabilidad, lo que indica su viabilidad como problema resolvable.

Como pasos a seguir, proponemos:

- Explorar cómo varían las distancias en función del tipo de relación (padre-hijo, hermanos, abuelos-nietos).
- Realizar fine-tuning de modelos de *embeddings* en el problema específico de reconocimiento de parentesco.
- Investigar la posibilidad de mejorar la calidad del dataset IMDb eliminando ejemplos ambiguos o incorrectamente etiquetados, que podrían estar influyendo en los outliers observados.

Estas direcciones permitirán profundizar en el análisis de la viabilidad del reconocimiento de parentesco automático y mejorar la comprensión de las limitaciones actuales de los modelos en escenarios realistas.

## Reconocimientos.

Este trabajo fue financiado por el proyecto UBACyT 20020220400306BA.

## Bibliografía.

- Deng, J., Guo, J., Xue, N., & Zafeiriou, S. (2019). Arcface: Additive angular margin loss for deep face recognition. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 4690–4699. <https://doi.org/10.1109/CVPR.2019.00482>
- Di Ielsi, C., & Negri, P. (2021). Base de imágenes latinoamericana para reconocimiento facial. *II Simposio Argentino de Imágenes y Visión (SAIV 2021)-JAIO 50*.
- Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B., & Song, L. (2017). Spheredface: Deep hypersphere embedding for face recognition. *CVPR*.
- Robinson, J. P., Shao, M., & Fu, Y. (2022). Survey on the analysis and modeling of visual kinship: A decade in the making. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44 (08), 4432–4453. <https://doi.org/10.1109/TPAMI.2021.3063078>

- Robinson, J. P., Shao, M., Zhao, H., Wu, Y., Gillis, T., & Fu, Y. (2017). Recognizing families in the wild (rfiw): Data challenge workshop in conjunction with acm mm 2017. *Proceedings of the 2017 Workshop on Recognizing Families In the Wild*, 5–12. <https://doi.org/10.1145/3134421.3134424>
- Zhang, X., Min, X., Zhou, X., & Guo, G. (2021). Supervised contrastive learning for facial kinship recognition. *2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021)*, 01–05.