

Ninguno de estos embriones existe en la vida real. Generando imágenes de embriones con Inteligencia Artificial.

None of these embryos exist in real life. Generating images of embryos

Dr. Francho Meléndez, Sergio González.
Fenomatch SL.

RESUMEN

En este artículo mostramos cómo la inteligencia artificial (IA) y el procesamiento de imágenes puede usarse en la generación de datos artificiales en el campo de la fertilidad. El acceso a datos es, en general, problemático para investigadores y académicos en todas las áreas, pero cuando tratamos con datos médicos el desafío es incluso mayor llegando a bloquear desarrollos e innovaciones en el campo. La generación de datos de forma artificial está ganando importancia dado que permite tener una alta accesibilidad y control sobre los datos. En este artículo exploramos modelos de redes neuronales generativas antagónicas (GAN) para crear imágenes de embriones emulando aquellas tomadas en una incubadora, los cuales pueden ayudar a crear bases de datos de entrenamiento para algoritmos e incluso para especialistas.

Palabras clave: *Análisis facial, Inteligencia Artificial, Generative Adversarial Networks, Selección embrionaria, Time-lapse, Reproducción Asistida, Fenomatch.*

Correspondencia: Luis José Arenaz Villalba
luis.arenaz@fenomatch.com
SOLICITUD REIMPRESIÓN: Email: fertilidad@editorialmedica.com

SUMMARY

This paper illustrates how artificial intelligence (AI) and image processing can be used to generate artificial data in the field of fertility. Access to data is in general problematic for researchers and academics in all areas, but when dealing with medical data, this becomes even more challenging, sometimes blocking developments and innovations in the field. Generating artificial data for training algorithms has gained importance since it allows for large availability and control over the data. In this paper we explore GANs to generate images of embryos, emulating those taken from an incubator which can help to generate training data for algorithms, and even for specialists.

Key words: *Facial analysis, Artificial Intelligence, Generative Adversarial Networks, Embryo selection, Time-lapse, Assisted Reproduction, Fenomatch.*

1. INTRODUCCIÓN

Las técnicas modernas de aprendizaje automático y en particular el aprendizaje profundo, basan su efectividad en bases de datos extensas con gran variabilidad. De esta forma, los algoritmos pueden optimizar modelos que sean capaces de encontrar resultados correctos en diferentes escenarios, en vez de solo en los casos presentes en los datos de entrenamiento, lo que se denomina sobreajuste (*overfitting*).

El auge de la IA en el campo de la fertilidad y otras disciplinas médicas requiere por tanto de grandes bases de datos que por su naturaleza son menos accesibles que en otros ámbitos. Los datos médicos tienen mayores niveles de protección, y por tanto su accesibilidad es más compleja. A su vez, por la naturaleza no técnica en cuanto al manejo de datos de los profesionales médicos, en muchos casos dichos datos no se encuentran apropiadamente documentados, o no han sido gestionados de forma que resulte sencilla su extracción y análisis. En algunos casos, la dificultad de extraer estos datos puede provenir de algo tan sencillo como adquirir el permiso por parte de los pacientes, ya sea por imposibilidad de contactarlos, por la negativa del paciente, o simplemente por la falta de colaboración de la entidad que gestiona los datos.

Incluso una vez conseguidos los datos, en el caso de aprendizaje supervisado, se necesita conocimiento experto para etiquetar los datos de entrenamiento. Imaginemos que tenemos un algoritmo para clasificar el estadio y la morfología de un embrión a partir de imágenes de una incubadora. Una colección de imágenes para un entrenamiento de tamaño mediano puede oscilar entre 5000 y 10000 imágenes, las cuales han de ser previamente clasificadas y etiquetadas en su correcto estadio y morfología por un embriólogo. Este trabajo es arduo y puede requerir el desarrollo de herramientas software específicas.

Vista esta problemática, es fácil entender que tanto la generación artificial como el etiquetado automático de datos son

de gran utilidad para la creación de bases de datos para el entrenamiento de algoritmos de aprendizaje automático. Con una cantidad de datos relativamente pequeña, podemos generar una cantidad ilimitada de los mismos, y en algunos casos que incluso dichos datos estén ya etiquetados y listos para el entrenamiento de otros algoritmos.

En este estudio hemos presentamos una prueba de concepto donde generamos imágenes de embriones utilizando un algoritmo generativo, valiéndonos de nuestra experiencia generando imágenes artificiales de caras para entrenar nuestros algoritmos de análisis facial de Fenomatch. El resultado es un modelo que es capaz de generar imágenes fotorealistas de embriones similares a las capturadas por las cámaras de las incubadoras. Este tipo de imágenes podrían utilizarse para entrenar algoritmos de reconocimiento, clasificación, o detección de elementos del embrión. Este tipo de técnicas podrían usarse también en el ámbito de la formación de futuros embriólogos.

Pese a que son los primeros pasos que nuestro equipo da en esta dirección, discutimos también aquí algunas direcciones para el trabajo futuro, que consideramos pueden hacer que estas técnicas generativas tengan un impacto importante en el mundo de la fertilidad.

2. ESTUDIOS PREVIOS

Las redes neuronales han probado ser superiores a los métodos anteriores en la generación de imágenes sintéticas. Los dos modelos más comunes para esta tarea son *Autoencoders* y Redes Generativas Antagónicas (*Generative adversary networks*, GANs). Nos centraremos en este estudio en dos algoritmos basados en GANs, dado que a día de hoy son considerados superiores en la generación de imágenes sintéticas.

Las GANs han recibido mucha atención desde que Ian Goodfellow presentara el concepto en su artículo (1). La idea de este modelo es que una red neuronal genera imágenes sintéticas, y otra red juzga dichas imágenes y decide si son

reales o sintéticas. Ambas redes van aprendiendo a hacer su función conforme son expuestas a más datos. La red generadora va mejorando sus creaciones hasta que consigue “engañar” a la otra red. Una vez entrenado el algoritmo, podemos usar la red generadora para crear imágenes fotorealistas similares a las presentes en la base de datos original.

Más recientemente, Karras et al. (2) presentaron su trabajo StyleGAN en 2019 con unos resultados sin precedentes en la generación de caras fotorealistas en alta resolución. Para nuestros experimentos utilizamos una variación de su trabajo realizado en 2020 (3) que simplifica su entrenamiento utilizando la aumentación de datos (*data-augmentation*) (4) lo cual permite un entrenamiento más eficiente con un número de imágenes de entrenamiento menor.

Muy recientemente, se ha diseñado un algoritmo con una arquitectura basada en GANs, pero que tan solo requiere una imagen para su entrenamiento (5). Las imágenes generadas con este método mantienen ciertas estructuras con la original, pero presentan variaciones en la distribución y cantidad de los elementos.

En el caso de generación de imágenes de embriones, encontramos un trabajo con espíritu similar al nuestro donde Dirvanauskas y su equipo generan imágenes en los primeros tres estadios del embrión (5). Nuestros experimentos han generado imágenes con más cantidad de detalle y hemos sido capaces de generar imágenes en todas las fases de desarrollo.

3. MATERIALES Y MÉTODOS: GENERANDO IMÁGENES DE EMBRIONES.

3.1. Style-based GAN

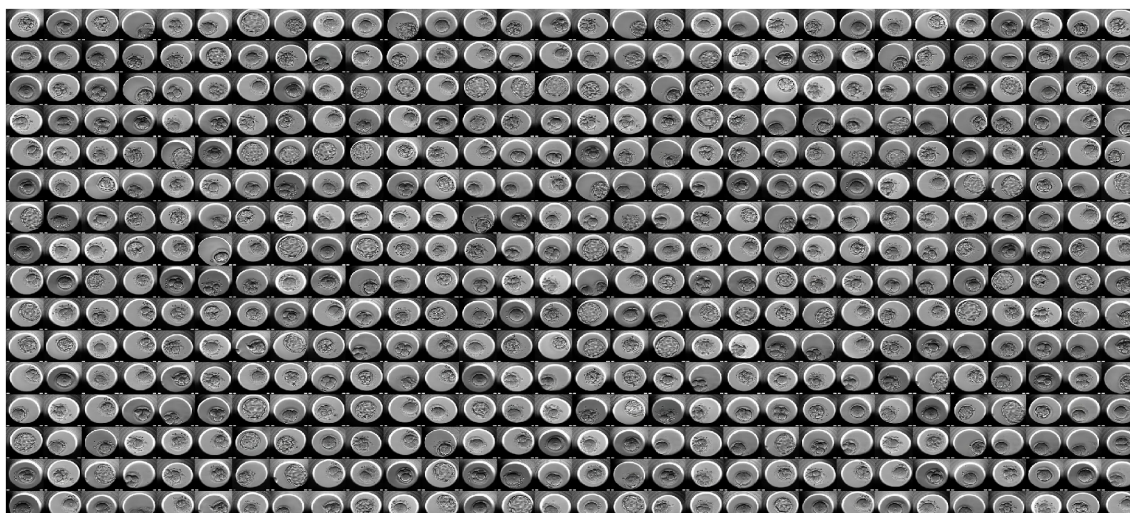
En este trabajo proponemos el uso de las arquitecturas presentadas en los estudios de Karras y colaboradores (3) y Dirvanaskus y colaboradores (5) para la generación de imágenes de embriones similares a las capturadas durante un time-lapse en una incubadora. Con este experimento, queremos comprobar la capacidad de esta arquitectura para generar imágenes con la estructura y morfología correctas.

Para nuestro experimento contamos con una base de datos propia consistente de 4560 imágenes, pertenecientes a 16 time-lapse de ciclos completos de desarrollo embrionario. Las imágenes fueron recortadas para tener una proporción cuadrada de 512x512 píxeles y después reducidas a 256x256 píxeles para reducir el tiempo y el consumo de memoria en el entrenamiento. El modelo utilizado requiere preprocesar las imágenes para crear versiones en diferentes resoluciones. La generación también es de forma progresiva, empezando por una imagen de baja resolución y añadiendo detalle en las siguientes iteraciones. El modelo estuvo entrenando por un total de 100 horas en una GPU con 8 GB de RAM

La gran mayoría de las imágenes generadas tienen la morfología correcta cubriendo todo el espectro en las fases de desarrollo del embrión. En la figura 1, podemos observar

FIGURA 1

Uso de redes generativas adversarias (GAN) para producir nuevas imágenes de embriones diferentes a las usadas para el entrenamiento del algoritmo



una muestra aleatoria de 480 imágenes muestra de la diversidad de las imágenes generadas. La figura 2 contiene una pequeña parte de las imágenes de la figura 1 ampliadas para su mejor evaluación.

En la figura 3 podemos comparar imágenes pertenecientes a la base de datos original e imágenes generadas con nuestro modelo. El algoritmo es capaz de generar imágenes realistas con la estructura correcta y variabilidad semejante a la de los datos originales.

En algunas de ellas aparecen artefactos o la imagen es incompleta como puede verse en la figura 4. Algunos de los artefactos detectados son estructuras morfológicas que no aparecen bien definidas, la forma del pocillo no es completamente circular, o se mezclan elementos enfocados con elementos desenfocados. Dichos artefactos deberían reducirse al mínimo entrenando el modelo durante más tiempo para mejorar su convergencia.

La conclusión del experimento es que este tipo de modelos

FIGURA 2

Una muestra ampliada de las imágenes generadas en la Figura 1

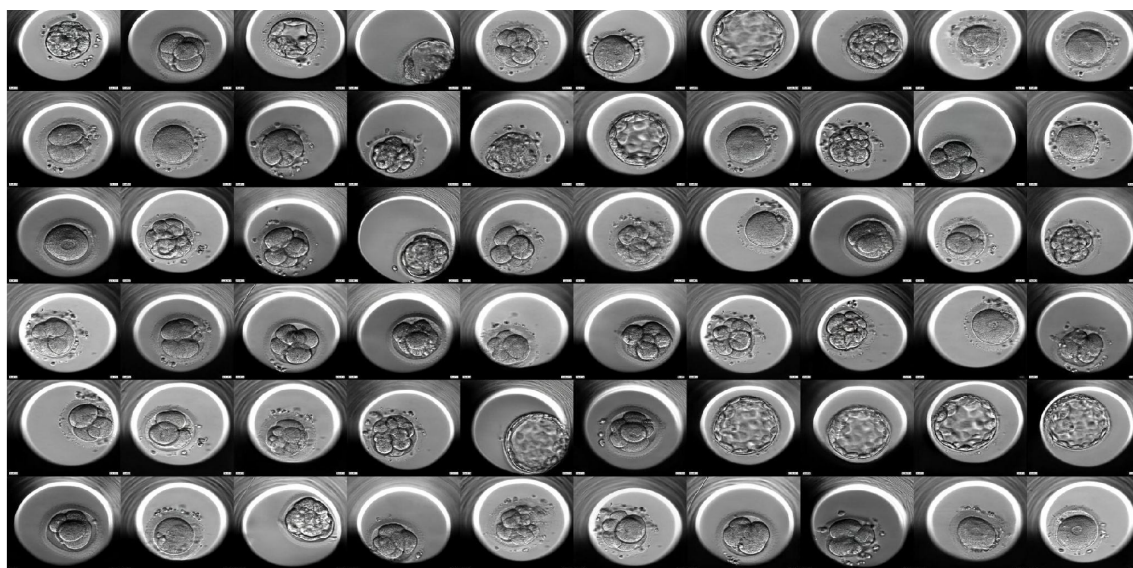


FIGURA 3

Izq. 6 fotografías de embriones de la base de datos propia. Dch. 6 imágenes generadas durante el experimento usando una arquitectura GAN [5].

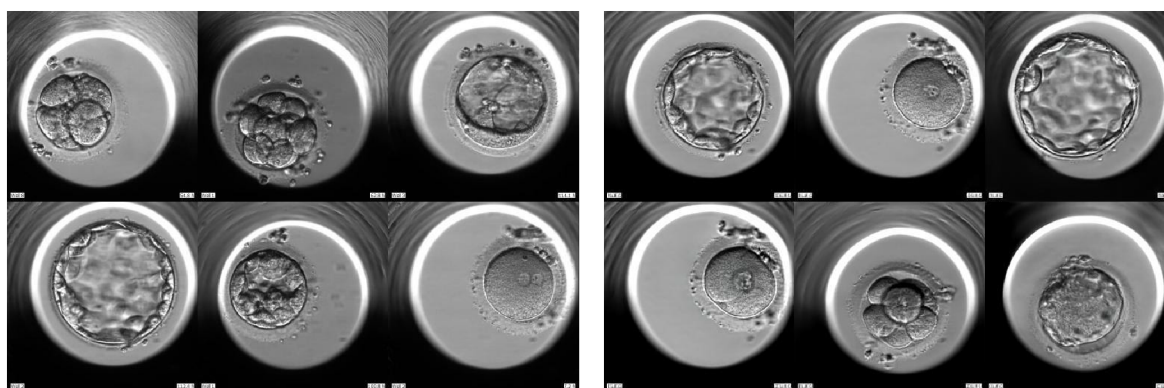
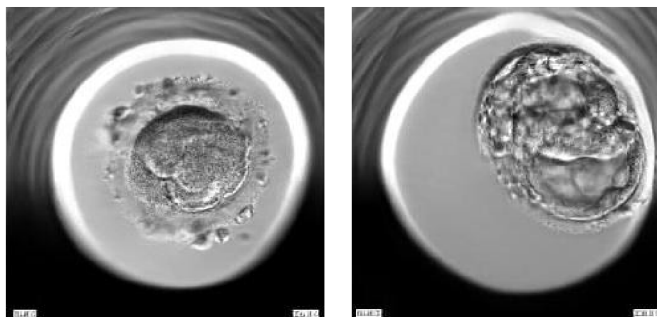


FIGURA 4

Imágenes con artefactos generadas con nuestro modelo



es capaz de generar imágenes realistas de embriones con suficiente variabilidad como para generar time-lapses de ciclos completos. Los modelos GAN basados en estilos confirman aprender una estructura de los embriones suficientemente restringida como para generar diferentes estadios con un nivel de detalle equivalente a las originales.

3.2. Single-Image GAN

En nuestro segundo experimento, utilizamos el modelo presentado por Dirvanaskus y colaboradores (5), el cual solo necesita una imagen de entrenamiento. El modelo necesita alrededor de 3 horas de entrenamiento, y una vez entrenado permite la aplicación de dicho modelo a diferentes algoritmos de procesamiento de imagen, siendo la generación de imágenes una de ellas. En la figura 5 podemos ver la imagen inicial utilizada para el entrenamiento y una de las imágenes generadas por el algoritmo. Podemos observar que la imagen es muy similar a la original, y que la estructura no es

del todo correcta, se pierden los contornos de la célula y el interior de la misma está borroso al igual que algunas regiones del exterior.

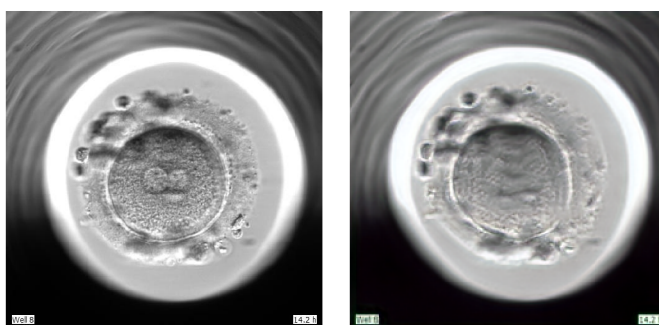
Podemos ver que este modelo es muy limitado para este caso de uso. El tipo de variaciones que es capaz de generar rompen la estructura general de la imagen, creando imágenes incorrectas o demasiado parecidas al original.

4. DISCUSIÓN

La generación de imágenes fotorealistas es sin duda una de las áreas del procesamiento de imagen que ha dado un salto cualitativo con el desarrollo de las redes neuronales. Su aplicación al ámbito de la medicina y más particularmente al campo de la fertilidad, está todavía en su infancia. Sin embargo, como podemos ver en estos experimentos, estas tecnologías son perfectamente aplicables y pueden tener aplicaciones muy interesantes en su desarrollo en el ámbito de la fertilidad.

FIGURA 5

Izq. Fotografía original. Dch. Imagen generada durante el experimento usando una arquitectura GAN de una sola imagen [5].



En nuestros experimentos, ambos métodos han sido capaces de generar imágenes de embriones, pero presentando diferentes beneficios y limitaciones. El segundo método, *Single-image* GAN, requiere un tiempo de entrenamiento realmente pequeño, de unas pocas horas, y el mínimo dataset posible. Sin embargo, la calidad y la variabilidad de las imágenes creadas no es todavía suficiente para nuestro caso de uso. Con algunas mejoras en la calidad de las imágenes generadas, dado su tiempo de entrenamiento tan bajo, podríamos entrenar el modelo con imágenes de diferentes estadios de desarrollo embrionario para conseguir producir imágenes de las diferentes fases.

En contraste, el primer experimento, genera imágenes de gran calidad y variabilidad, con el coste de una base de datos de entrenamiento más grande. El artículo original de generación basada en estilos (2) utilizaba más de 90.000 imágenes mientras que la versión actual (3) parece funcionar bien con un número más limitado como hemos demostrado. El tiempo de entrenamiento es por supuesto mucho mayor, entre uno y dos órdenes de magnitud superior, pero las posibilidades generativas del algoritmo son mucho más altas. Las redes generativas basadas en estilos, una vez entrenadas, crean la representación numérica del modelo (espacio latente) que permite otra serie de aplicaciones, como la creación de animaciones entre imágenes generadas. Esto ha sido explorado en la literatura en la generación de imágenes de caras, siendo posible crear una cara donde se modifica el aspecto del pelo o el color de piel, pero manteniendo los rasgos faciales. Este tipo de exploración del espacio latente es sin duda una línea de investigación muy interesante para nuestro tipo de datos.

En ambos experimentos presentados aquí, los resultados generados son puramente aleatorios. Un mayor control en el proceso generativo es deseable para conseguir un rango de aplicaciones más amplio de estos algoritmos, y es actualmente un área de investigación muy activa. Ejemplos de

este tipo de control incluyen: poder elegir el momento del desarrollo embrionario de la imagen que queremos generar; poder seleccionar diferentes áreas de la imagen donde generar diferentes partes del embrión, por ejemplo, la posición y tamaño de una de las células, o el área ocupada por la zona pelúcida, el trofoblasto o la masa celular interna; crear animaciones de ciclos completos.

5. CONCLUSIÓN

Hemos presentado dos experimentos para la generación de imágenes de embriones similares a las generadas por una incubadora time-lapse utilizando redes neuronales antagónicas. Los resultados muestran que la arquitectura basada en estilos GAN es capaz de generar imágenes realistas de embriones con un alto nivel de detalle y una estructura morfológica correcta en la mayoría de los casos.

6. BIBLIOGRAFÍA

1. **Generative Adversarial Networks**, I. Goodfellow and Jean Pouget-Abadie and Mehdi Mirza and Bing Xu and David Warde-Farley and Sherjil Ozair and Aaron C. Courville and Yoshua Bengio, *Advances in Neural Information Processing Systems*, 2014.
2. **A style-based generator architecture for generative adversarial networks**, Karras, Tero and Laine, Samuli and Aila, Timo, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019.
3. **Training Generative Adversarial Networks with Limited Data**, Tero Karras and Miika Aittala and Janne Hellsten and Samuli Laine and Jaakko Lehtinen and Timo Aila, *Advances in Neural Information Processing Systems*, 2020.
4. **A survey on Image Data Augmentation for Deep Learning**, Connor Shorten and T. Khoshgoftaar, *Journal of Big Data*, 2019.
5. **SinGAN: Learning a Generative Model From a Single Natural Image**, Shaham, Tamar Rott, and Dekel, Tali and Michaeli, Tomer, 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019.
6. **Dirvanauskas D, Maskeliūnas R, Raudonis V, Damaševičius R, Scherer R**. HEMIGEN: Human Embryo Image Generator Based on Generative Adversarial Networks. *Sensors (Basel)*. 2019;19(16):3578. Published 2019 Aug 16. doi:10.3390/s19163578