

<https://artnodes.uoc.edu>

ARTÍCULO

NODO «IA, ARTE Y DISEÑO: CUESTIONANDO EL APRENDIZAJE AUTOMÁTICO»

Vistas panorámicas sobre el patrimonio visual colectivo a través de redes neuronales convolucionales

Las exposiciones *Revolutionary Arkive* y *Mnemosyne 2.0* de Pilar Rosado

Pilar Rosado Rodrigo

Universidad de Barcelona

Ferran Reverter Comes

Universidad de Barcelona

Fecha de presentación: marzo de 2020

Fecha de aceptación: junio de 2020

Fecha de publicación: julio de 2020

Cita recomendada

Rosado Rodrigo, Pilar y Ferran Reverter Comes. 2020. «Vistas panorámicas sobre el patrimonio visual colectivo a través de redes neuronales convolucionales. Las exposiciones *Revolutionary Arkive* y *Mnemosyne 2.0* de Pilar Rosado». En: Andrés Burbano; Ruth West (coord.) «IA, arte y diseño: Cuestionando el aprendizaje automático». *Artnodes*, N.º 26: 1-12. UOC. [Consulta: dd/mm/aa]. <http://doi.org/10.7238/a.v0i25.3354>



Los textos publicados en esta revista están sujetos –si no se indica lo contrario– a una licencia de Reconocimiento 4.0 Internacional de Creative Commons. La licencia completa se puede consultar en https://creativecommons.org/licenses/by/4.0/deed.es_ES.

Resumen

En el contexto de una sociedad saturada de imágenes, las redes neuronales convolucionales (*convolutional neural networks* - *CNN*), preentrenadas a partir de la información visual contenida en miles y miles de imágenes, constituyen una herramienta de gran utilidad para ayudarnos a ordenar el patrimonio visual, ofreciéndonos así un punto de acceso que de otra manera sería imposible.

Una de las responsabilidades del artista contemporáneo es adoptar posiciones que ayuden a dar significado, a proyectar sentido sobre la acumulación visual a la que nos enfrentamos. En este artículo pasamos a describir dos exposiciones de la artista Pilar Rosado: *Revolutionary Arkive* y *Mnemosyne 2.0*. En ellas se ha utilizado la red neuronal artificial *ResNet-50* para extraer las características visuales de grandes conjuntos de imágenes y los descriptores de imagen obtenidos se han usado como entrada para el algoritmo *t-SNE*. Así, se han elaborado grandes mapas visuales formados por miles de imágenes ordenadas siguiendo criterios de similitud formal en los que se ponen de manifiesto los patrones visuales de los arquetipos de una determinada categoría semántica.

La manera de archivar y recuperar nuestra memoria colectiva debe tener una correlación con los avances tecnológicos y científicos de nuestra época para que se nos vayan descubriendo progresivamente nuevos horizontes de conocimiento. En su práctica artística, Rosado explora, en el ámbito de la cocreación humano-máquina, cuestiones políticas que se pueden abordar desde la imagen y que implican a las tecnologías de aprendizaje automático, como la gestión de la información en los archivos visuales del futuro o la revisión de la memoria visual colectiva.

Palabras clave

Deep learning, postfotografía, *computer vision*, *t-SNE*, *convolutional neural networks* (*CNN*), Aby Warburg.s.

Panoramic views on the collective visual heritage through convolutional neural networks. The exhibitions Revolutionary Arkive and Mnemosyne 2.0 by Pilar Rosado

Abstract

In the context of a society saturated in images, convolutional neural networks (CNNs), pre-trained from the visual information contained in thousands and thousands of images, constitute a tool which is of great use for helping us to order visual heritage, in this way offering us a point of access which would otherwise be impossible.

One of the responsibilities of the contemporary artist is to adopt positions which can help to provide sense, to cast meaning onto the accumulation that we face. In this article we describe two exhibitions by the artist Pilar Rosado: Revolutionary Arkive and Mnemosyne 2.0. The artificial neuronal network ResNet-50 has been used in these exhibitions in order to extract the visual characteristics of large sets of images, and the image descriptors obtained have been used as an entry point for the algorithm t-SNE. In this way, large visual maps have been produced which are composed of thousands of images and are arranged by following criteria of formal similitude in which the visual patterns of the archetypes of a certain semantic category are displayed.

The method of archiving and recovering our collective memory must have a correlation with the technological and scientific advances of our age so that we can progressively discover new horizons of knowledge. In her artistic practice Rosado explores, in the sphere of human-machine co-creation, political issues which can be approached from the perspective of image and which involve the machine learning technologies, such as the management of information in the visual files of the future or the review of the collective visual memory.

Keywords

Deep Learning, Post-photography, Computer Vision, t-SNE, Convolutional Neural Networks (CNNs), Aby Warburg.

Torrente visual

«...organizarse ya es, en cierto modo, tener ojos»
(Saramago, 2003: 330).

Cada minuto transcurrido, millones de imágenes son tomadas y vertidas a la red, como si de un río de infinito caudal se tratase. Con ellas abocamos también el rastro de nuestros gestos, muchas veces compartidos. La dificultad de rescatarlas de este torrente y de acceder a su visualización de una manera inteligible puede convertirlas en invisibles. Estos enormes archivos visuales pueden volver a algún tipo de comprensibilidad gracias al uso de tecnologías de *deep learning*.

La posibilidad de poder detener este incesante y vertiginoso flujo de información creando una instantánea que nos ayude a reflexionar, desde lo visual, sobre lo que hemos producido, especialmente sobre la «forma» de las imágenes que hemos producido, se vislumbra como una poderosa herramienta para pensar. La belleza de este tipo de visualización de datos estará implícita en los rastros de verdad que puedan destilarse de su análisis.

En el contexto de esta sociedad de la sobreinformación y de la enorme producción fotográfica, una de las responsabilidades del artista es adoptar posiciones que ayuden a dar significado, a proyectar sentido sobre la acumulación visual a la que nos enfrentamos. Acerca de este tema, Joan Fontcuberta escribe: «El valor de creación más determinante no consiste en fabricar imágenes nuevas, sino en saber gestionar su función, sean nuevas o viejas» (Fontcuberta, 2016: 39).

Situamos nuestra propuesta en el ámbito de lo postfotográfico y proponemos que, aplicando modelos de *deep learning* para la categorización automática de grandes cantidades de imágenes de archivo, podemos darles un orden, una belleza que nos muestre cosas nunca vistas en una de «las múltiples reencarnaciones de la imagen», como las llama Fontcuberta.

Antecedentes

La genealogía de estos intereses se remonta al historiador del arte y la cultura Aby Warburg (1866-1929) quien, con su *Atlas Mnemosyne*, formado por una colección de imágenes y escaso texto, pretendía narrar la historia de la memoria de la civilización europea. Con una visión premonitory Warburg considera que las imágenes por sí mismas y en su relación mutua generan un espacio de pensamiento (Warburg, 2010).

También podemos citar a André Malraux, ministro francés de cultura y escritor, que en 1947 lanza la idea de crear un museo imaginario basado en reproducciones fotográficas, sin paredes, sin muros, permitiendo la posibilidad a cada persona de formar su propio museo (Malraux, 1949). A pesar de limitaciones insalvables como la reproducción de la textura, el volumen o la escala de una obra de

arte, y teniendo claro que una experiencia artística válida nunca podrá substituir la vivencia en el lugar, la ventaja de este tipo de museo imaginario es que cada espectador puede crear su propio museo en función de su sensibilidad.

Podemos, así mismo, conectar con la llamada *fotografía aumentada* de autores como Lev Manovich con sus análisis visuales de obras literarias, videojuegos o las portadas de la revista *Time*. En la obra *TimeLine*, de Jeremy Douglass y Lev Manovich de 2009, vemos en una sola imagen las portadas de todos los números de la revista *Time* publicados entre 1923 y el verano de 2009, un total de 4.535. Las imágenes aparecen dispuestas de acuerdo con el orden de publicación (de izquierda a derecha y de arriba a abajo) (Manovich, 2009). También resulta adecuado mencionar el proyecto de Lev Manovich y William Huber de 2010 que permite visualizar la estética general del mundo del juego, además de los cambios visuales y narrativos dentro de la progresión de un solo juego, ordenando 22.500 *frames* tomados del videojuego japonés *Kingdom Hearts II* cada 6 segundos de juego. La obra nos muestra 62,5 horas de juego representadas en una sola imagen (Manovich, 2012).

En este marco presentamos dos proyectos de creación concebidos por Pilar Rosado con el soporte estadístico de Ferran Reverter, con quien colabora activamente en proyectos de investigación desde el año 2012. En primer lugar, se presenta la exposición *Revolutionary Arkive: The eye-machine is watching*, que tuvo lugar entre el 27 y el 30 de septiembre de 2018 durante el Festival Panoràmic en Granollers (Barcelona), donde se volvieron a revivir las revueltas del 1968 a través del cine y de la fotografía. En segundo lugar, se muestra el proyecto expositivo *Mnemosyne 2.0: cartografías computacionales de la memoria visual*, que se presentó entre el 9 de mayo y el 9 de junio de 2019 en la Sala 0 de Roca Umbert Fábrica de las Artes de Granollers (Barcelona) y que fue el resultado de la residencia artística que Pilar Rosado desarrolló durante el año 2019 en ese contexto.

En los próximos apartados pasaremos a describir estas dos exhibiciones artísticas, para cuya realización se utilizaron miles de fotografías recuperadas de internet clasificadas según los modelos de *deep learning* que explicaremos a continuación.

Métodos computacionales

Para recuperar las miles de imágenes utilizadas, Rosado realizó múltiples búsquedas textuales en Google y, para descargar las imágenes, accedió a los repositorios de internet utilizando *software* de *web scraping* incorporado como una extensión en el navegador.

No podemos obviar que los algoritmos de búsqueda de Google se diseñaron para ayudarnos a encontrar lo que buscamos entre millones de páginas web en una fracción de segundo. Estos sistemas de clasificación se basan no en uno, sino en toda una serie de algoritmos que tienen en cuenta muchos factores como las palabras

de la consulta, la relevancia y usabilidad de las páginas, el grado de especialización de las fuentes, la ubicación y la configuración. El peso que se da a cada factor cambia dependiendo de la naturaleza de la consulta (según sitio web Búsqueda de Google).

Desde la visión artificial se están haciendo muchos esfuerzos para que las máquinas emulen la habilidad que tenemos las personas de comprender una imagen. En la actualidad estamos viviendo un gran avance en el uso del *deep learning* en tareas de catalogación automática de imágenes.

El *deep learning* se fundamenta en la idea del aprendizaje mediante ejemplos dados. El ordenador, a partir de ellos, construye un modelo capaz de predecir reglas subyacentes. El algoritmo modifica el modelo cuando se producen errores de predicción, así, de manera iterativa, estos modelos pueden aprender de forma extremadamente precisa, gracias a que el sistema es capaz de extraer patrones relevantes para la tarea planteada.

La potencia del *deep learning* reside en la capacidad de estas metodologías de descubrir por sí mismas los descriptores de las imágenes sin necesidad de usar unos predefinidos. Manovich apuntó en 2019 que este aspecto del aprendizaje automático es, de hecho, algo nuevo en la historia del arte computacional; un ordenador que por sí solo puede aprender la estructura del mundo visual.

Con el objetivo de clasificar las imágenes, se ha utilizado la red neuronal convolucional *ResNet-50* (He et al., 2016) que fue preentrenada en un subconjunto de la base de datos *ImageNet* (Russakovsky et al., 2015). Entrenada con más de un millón de imágenes, tiene 177 capas convolucionales en total, asociadas a una red residual de 50 capas, y puede clasificar imágenes en 1000 categorías de objetos.

Para clasificar las imágenes utilizadas en las exposiciones que comentaremos más adelante, se usó esta red *ResNet-50* como un extractor de características visuales, tomando las activaciones en la capa de salida (fc1000) como descriptores artificiales de las imágenes de la colección, obteniéndose una representación inducida por la red del contenido visual de cada una de las imágenes. Así, podemos decir que esta etapa del proceso creativo está condicionada por el entrenamiento supervisado de la red *ResNet-50*.

Posteriormente, se emplearon estos descriptores artificiales obtenidos como entrada para el algoritmo *t-SNE* (*stochastic neighbor embedding*) (Maaten y Hinton, 2008). *t-SNE* representó los vectores 1000-dimensionales correspondientes a las imágenes en un espacio 2-dimensional (plano de representación) preservando óptimamente las relaciones de similitud de contenido visual de las imágenes (ver figura 1). Esta etapa del proceso creativo es no supervisada, ya que depende exclusivamente de las similitudes entre las imágenes que se están analizando.

Al trabajar con grandes conjuntos de fotografías, el problema es que, cuando se aplica *t-SNE* a las imágenes, tal como hemos comentado, resulta que una gran cantidad de las localizaciones de las imágenes en el plano de representación corresponden a elementos superpuestos que a menudo impiden inspeccionar una

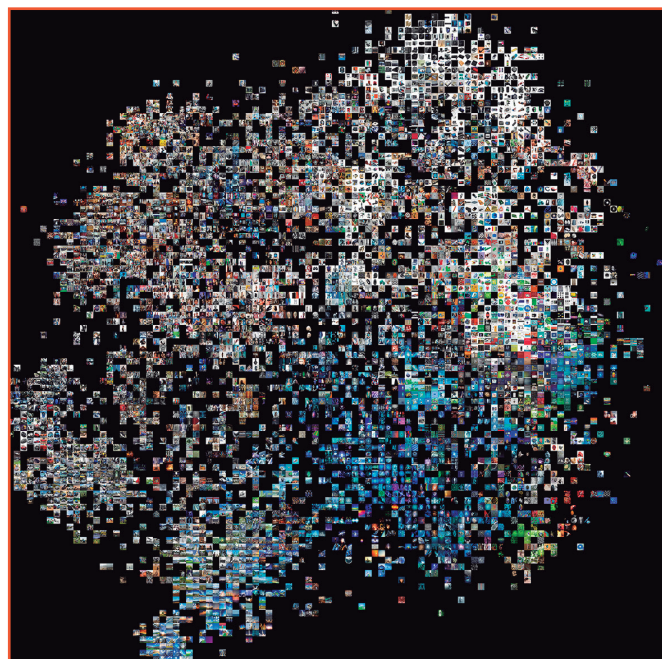


Figura 1. 10.000 imágenes ordenadas por una red neuronal artificial *Res-Net50* y el algoritmo *t-SNE* en función de sus similitudes gráficas. Imagen digital de 2 x 2 metros. © Pilar Rosado

sola imagen. Para solventar esta problemática se utilizó el código de Karpathy (2017), que reasigna las localizaciones sobre una cuadrícula regular, pero sin destruir las relaciones de vecindad que surgieron del agrupamiento por similitud basado en *t-SNE* (ver figura 2).

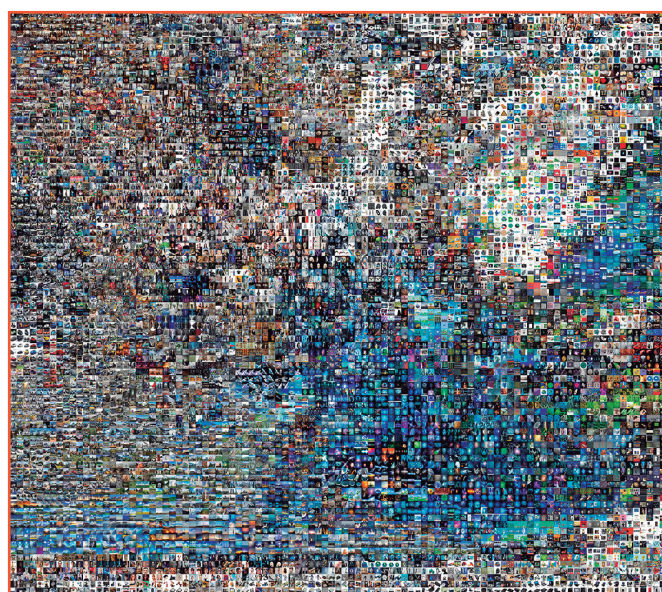


Figura 2. Se utiliza el código de Karpathy (Karpathy 2017) que reasigna las localizaciones sobre una cuadrícula regular, pero sin destruir las relaciones de vecindad que surgieron del agrupamiento por similitud basado en *t-SNE*. Imagen digital de 2 x 2 metros. © Pilar Rosado.

En el proceso de elaboración de ambas exposiciones, la toma de decisiones hasta llegar al resultado final es compartida entre el humano y la máquina, configurando un proceso de cocreación híbrido que alienta posibilidades novedosas para la obtención de resultados.

Exposición *Mnemosyne 2.0*: Cartografías computacionales de la memoria visual

Las obras que duermen en los archivos como documentos pueden volver a divertir (amuser) y excitar gracias a un análisis que nos las haga ver con ojos nuevos; un placer que estaría entre la plástica y el estudio riguroso (Pérez-Hita, 2019: 2).

En 1924 el historiador Aby Warburg inicia su *Atlas Mnemosyne*. Rastreando temas y patrones visuales recurrentes a través del tiempo, pasando por la antigüedad, el Renacimiento y llegando a la cultura contemporánea, Warburg pretendía elaborar algo así como una base de datos de imágenes de todos los tiempos y culturas. Es la suma del trabajo de toda su vida y quedó inacabado a su muerte en 1929. Las placas originales estaban hechas de tablas de madera de 1,7 x 1,4 metros cubiertas por lienzo negro en las que Warburg fijaba todo tipo de imágenes, como fotografías, reproducciones de pinturas, dibujos, ilustraciones extraídas de libros y material de periódicos. Cada tabla se dedicó a un tema específico que reflejaba sus intereses de investigación, estaban numeradas y eran fotografiadas en el gran salón de la biblioteca de Warburg. El autor no conservó las tablas originales, sino que constantemente reorganizaba las imágenes y agregaba nuevos tableros y temas (The Warburg Institute, 2018). Para Warburg las tablas, más que un medio de presentación, eran un medio de investigación y comprensión.

El proyecto expositivo *Mnemosyne 2.0* de Pilar Rosado es un intento de actualizar, mediante las nuevas tecnologías, la idea que abordó Aby Warburg con su *Atlas Mnemosyne*. De una forma análoga a como Warburg, sentado en su mesa, debía trabajar para establecer relaciones visuales entre sus imágenes, la artista trabajó frente a su ordenador utilizando el recurso de búsqueda de información de Google. Para elegir sus temas de interés, Rosado estableció un paralelismo con la magnífica exposición de 2010 en el MNCARS titulada «Atlas. ¿Cómo llevar el mundo a cuestras?», comisariada por Didi-Huberman, que se conformaba en cuatro grandes apartados titulados: «Conocer por imágenes», «Recomponer el orden de las cosas», «Recomponer el orden de los lugares» y «Recomponer el orden de los tiempos» (Didi-Huberman, 2010). La artista, recogiendo el testigo de Didi-Huberman, descargó imágenes de internet explorando conceptos y procesos tecnológicos actuales que han revolucionado los hábitos y espacios sociales, con la intención de poder revisar de forma crítica este patrimonio visual compartido. Para ello estructuró sus estrategias de búsqueda a partir de los criterios que describimos a continuación.

Para «recomponer el orden de las cosas», buscó imágenes de objetos producidos en los últimos tiempos por las nuevas tecnologías, realizando búsquedas de términos como «tecnología portable», «nanotecnología», «impresión 3D», «ciberseguridad», «robótica»...

Para «recomponer el orden de las personas», escribió en el buscador de Google el nombre de actividades realizadas por personas a raíz del uso de nuevas tecnologías como «hacktivistas», «mineros del coltán», «ciborgs», «hacedores de selfies», «liquidadores de Chernóbil»...

Para «recomponer el orden de los lugares», buscó espacios del planeta afectados por la «producción de tecnología» como «centrales nucleares», «cambio climático», «grandes basureros», «gentrificación»...

En el mundo contemporáneo, a pesar del espectacular desarrollo de los nuevos medios, la población tiene la necesidad de seguir manifestándose en la calle para propiciar cambios sociales y hacer valer sus derechos. Así que Rosado, con la intención de «recomponer el orden de los tiempos», buscó imágenes de «manifestaciones por los derechos humanos», «manifestaciones por el orgullo gay», «manifestaciones contra el cambio climático»...

En la exposición se mostraban también de forma relevante los textos de estas búsquedas (ver figura 3), que se iniciaron de una forma deliberada siguiendo los intereses de la artista, pero que pronto se vieron condicionados por los resultados de las propias búsquedas, en un proceso de serendipia que favoreció el descubrimiento o hallazgo afortunado a partir de una imagen que sugería una próxima búsqueda inesperada, pero valiosa. Se constituyó este método de creación cooperativo entre la artista y la computadora en un proceso estimulante de aprendizaje visual, de descubrimiento dirigido por la voluntad, pero también por la casualidad que deja abierta una puerta a la habilidad del sujeto que busca para reconocer que ha hecho un descubrimiento importante, aunque inesperado.



Figura 3. Imagen de todas las frases que la artista utilizó para recuperar en Google las 40.000 imágenes que muestra la exposición. Junto a cada frase aparece el número de imágenes que se obtuvo en esa búsqueda. © Pilar Rosado

Rosado comparte con Warburg la intención de crear una cartografía personal estableciendo relaciones y analogías que después

permitan un proceso abierto e infinito de relecturas, pero se diferencia metodológicamente en el hecho de que establece un diálogo de cocreación en el que permite que sus criterios de búsqueda se vean estimulados e hibridados con los criterios de la computadora.



Figura 4. Exposición *Mnemosyne 2.0* en la Sala 0 de Roca Umbert Fábrica de las Artes de Granollers (Barcelona) en marzo de 2019. © Pilar Rosado

Una vez descargadas 10.000 imágenes para cada grupo de conceptos (cosas, personas, lugares y manifestaciones), se aplicaron sobre ellas los métodos de *deep learning* descritos en el apartado «Métodos computacionales» para conseguir tener todas las imágenes ordenadas en una gran fotografía de 2 x 2 metros. La intención de la artista era llegar a cristalizar en grandes mapas visuales sus pensamientos e intereses relacionados con la tecnología, mediante este deambular por la información visual contenida en este conjunto de imágenes extraído de internet.

Posteriormente, será el espectador quien realice sus propios recorridos visuales a través de estos mapas (ver figura 5), que constituyen un gran atlas visual o, como el subtítulo de la exposición sugiere, «cartografías computacionales de la memoria visual» contenida en internet acerca de conceptos relacionados con la tecnología.

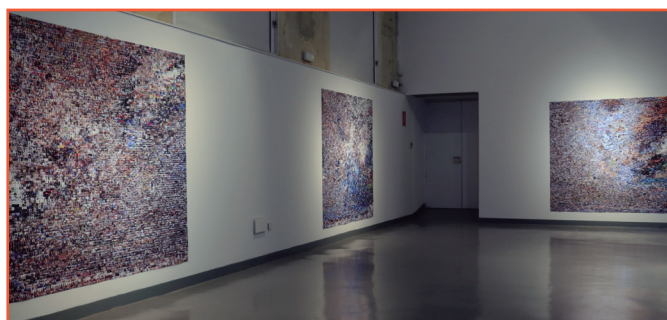


Figura 5. Exposición *Mnemosyne 2.0* en la Sala 0 de Roca Umbert Fábrica de las Artes de Granollers (Barcelona) en marzo de 2019. © Pilar Rosado

Un atlas constituye una forma visual del saber. No es un diccionario, ni un manual, ni un catálogo. Proporciona un saber abierto:

ante las imágenes, sobran las palabras. El atlas desvela una pasión por las morfologías, pero se aleja de la categorización estricta. En un atlas el lector puede recorrer con la mirada, deambular erráticamente para encontrar o no, tiene la posibilidad de construir sus propias cartografías (Didi-Huberman, 2010).

«Entiendo la fotografía como una herramienta para pensar la realidad y en *Mnemosyne 2.0* propongo poner un poco de orden en el maremágnum fotográfico que nos rodea aplicando estrategias de catalogación automática para construir un atlas formado por imágenes hechas de imágenes, cartografías de nuestra memoria visual que nos ayuden a comprender un poco mejor los rastros, los patrones, los arquetipos que construimos como sociedad en nuestro quehacer fotográfico. Los resultados obtenidos nos permitirán visitar la iconografía que constituirá en un futuro los archivos de nuestra memoria, desde un nuevo punto de vista más libre de preconcepciones e ideologías» (Rosado, 2019: 10).



Figura 6. Detalle ampliado del gran mapa de lugares afectados por la tecnología. Imagen digital de 2 x 2 metros. © Pilar Rosado

Exposición *Revolutionary Arkive: The eye-machine is watching*

La exposición *Revolutionary Arkive: The eye-machine is watching* tuvo lugar en el contexto del Festival Panoràmic. Bajo el lema *The whole world is watching*, la segunda edición del festival proyectó 7 largometrajes, 7 cortometrajes, organizó actividades, debates y conferencias y acogió las propuestas artísticas de 46 creadores (Festival Panoràmic, 2018).

Revolución, mutación, procesos de cambio para modificar el orden establecido, movimientos para desobedecer los dictados del poder. *Revolutionary Arkive*, ante la paradoja de un archivo que mantenga en orden historias de desórdenes, opta por transformar el archivo de imágenes de manifestaciones, en una colección viva, mutable, donde la visión artificial interrelaciona información visual sobre la revuelta.

«A Pilar Rosado parece interesarle la posibilidad de representar el alma, la personal y la colectiva. En su obra, *Revolutionary Arkive*, reúne imágenes de revueltas y revoluciones buscando e invitando al espectador a buscar analogías, diferencias, hilos de conexión o denominadores comunes en las imágenes y los gestos de levantamientos de multitudes reivindicativas» (Pérez-Hita, 2019: 1).

Las revueltas o revoluciones nacen cuando nos quieren imponer una serie de estructuras y de límites que decidimos romper para liberarnos. Podríamos decir que los conceptos se nos quedan pequeños y necesitamos ampliarlos. Estos desórdenes son en realidad mutaciones. La vida, llena de cambios está, por tanto, llena de revoluciones. Como apunta Fontcuberta, «Las imágenes son construcciones ideológicas, un campo de batalla político» (Fontcuberta, 2019: 177).

La exposición presentó, por un lado, el vídeo titulado *The eye-machine is watching*, que consistía en una serie de grabaciones de manifestaciones reales de los trabajadores de Granollers a lo largo de los años exigiendo sus derechos. Las películas fueron analizadas mediante un algoritmo de detección de rostros que localizaba rápidamente las caras de los manifestantes, pero las etiquetas de categorización habían sido manipuladas por la artista de manera que, en lugar de identificar «rostros», identificaban a los manifestantes con etiquetas subversivas como «agitador», «revolucionario», «rebelde», «insurgente», «disidente», etc., anunciando un hipotético futuro en el que las computadoras sean capaces de identificar este tipo de cualidades personales, simplemente al detectar nuestro rostro (ver figura 7).



Figura 7. Exposición *Revolutionary Arkive* en el contexto de Festival Panoràmic de Granollers en 2018. © Camila Marinone.

«*Revolutionary arkive: The eye-machine is watching* es un estudio revolucionario que demuestra que la tecnología es capaz de modular la manera como hasta ahora habíamos ordenado la información. Así como 1968 supuso un antes y un después en el método de recepción de las imágenes, actualmente los avances tecnológicos y la inteligencia artificial han alterado nuestra interpretación del mundo. (...) si las máquinas están preparadas para categorizar la producción fotográfica, también pueden categorizar a las personas

de, por ejemplo, subversivas a través de algoritmos que detecten rostros. El mundo distópico orwelliano vigilado por un Gran Hermano ya no queda tan lejos» (Company, 2018).

La exposición presentaba también una serie de grandes paneles con conjuntos de fotografías descargadas de internet al realizar búsquedas en Google sobre manifestaciones de países de todo el mundo (ver figura 8). La intención era realizar el proceso de categorización automática descrito en el apartado de «métodos computacionales» para extraer arquetipos visuales de revuelta, iconos de rebelión, insubmisión, agitación política o reivindicación social y, al mismo tiempo, para hacer visibles las diferencias o particularidades que tienen las fotografías que capturamos en cada país o región.

«La artista trabaja con redes neuronales artificiales, algoritmos de *deep learning* que, si tienen al alcance la suficiente cantidad de datos, pueden extraer los elementos formales que se repiten de manera estadísticamente significativa para constituir los rasgos de la representación arquetípica de una determinada categoría semántica» [Traducción propia]¹¹ (Fontcuberta, 2019: 178).



Figura 8. Frame del vídeo *The eye-machine is watching*, proyectado en la exposición *Revolutionary Arkive* en el que la artista manipula un algoritmo de detección de rostros para asignar etiquetas relacionadas con la revuelta como; insurgencia, agitación, disidencia, rebeldía, etc. © Pilar Rosado.

Para poder obtener una representación visual de cada continente, se seleccionaron imágenes a partir del concepto «manifestación», añadiendo el criterio de lugar: «manifestaciones en Cataluña», «manifestaciones en Rusia», «manifestaciones en China», «manifestaciones en Nicaragua», «manifestaciones en Sudáfrica» o «manifestaciones

1. En palabras de Joan Fontcuberta: «L'artista treballa amb xarxes neuronals artificials, algorismes de *Deep Learning* que, si tenen a l'abast la suficient quantitat de dades, poden extreure els elements formals que es repeteixen de manera estadísticament significativa per constituir els trets de la representació arquetípica d'una determinada categoria semàntica».

en Egipto»; y se ordenaron en diferentes paneles, a la manera de Aby Warburg. De este modo, la artista creó fotomosaicos donde las imágenes de cada conflicto aparecen ordenadas por afinidades gráficas (ver figura 9).



Figura 9. A la izquierda imágenes de manifestaciones en Nicaragua y a la derecha imágenes de manifestaciones en Rusia, en ambos casos ordenadas por la red neuronal como se describe en el apartado de Métodos. 2018. Imagen digital de 1 x 0,7 metros. © Pilar Rosado.

Una vez posicionadas las imágenes según su contenido visual, se puede apreciar qué relevancia tienen los tanques o los elementos bélicos, la mortalidad, las masas de gente unida, las banderas, las

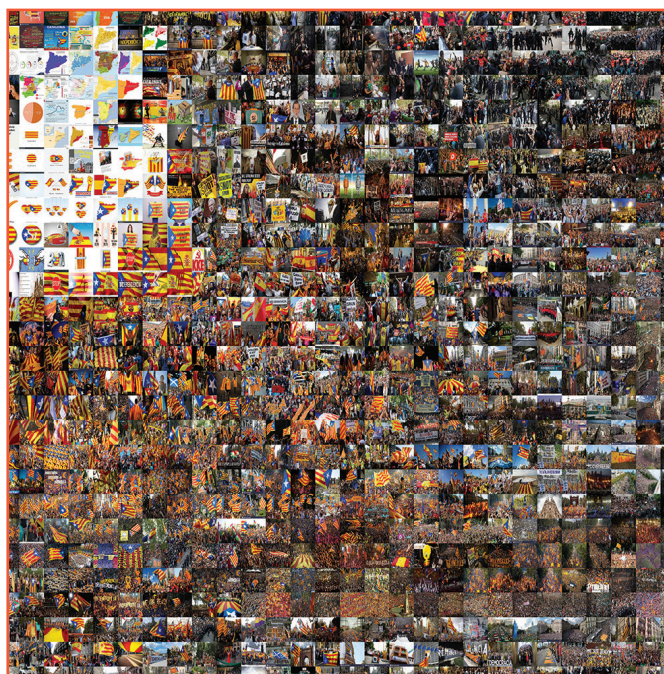


Figura 10. Imagen con 800 fotografías sobre manifestaciones en Catalunya, ordenadas mediante la red neuronal convolucional ResNet-50. Imagen digital de 1x1 metros. © Pilar Rosado

pancartas o la acción policial. Evidentemente, tratándose de manifestaciones, las imágenes con gente o policías son las que más abundan en cada mapa o, en el caso de Cataluña, por ejemplo, las banderas también tienen mucho peso (ver figura 10), pero claramente se observaron diferencias en función del lugar. Sin embargo, los mapas visuales presentados no son categóricos y solo muestran una realidad en la que el análisis de cada persona puede profundizar.

El objetivo no es únicamente identificar un patrón para definir las revoluciones, sino poner de manifiesto que las nuevas tecnologías dan un punto de vista alternativo al humano sobre las imágenes, que aportan mucha información y que permiten catalogarlas desde una perspectiva que hasta ahora no se había explorado.

Notas finales

La diferencia fundamental entre los mapas visuales de las dos exposiciones presentadas radica en los criterios de búsqueda de imágenes: en *Mnemosyne 2.0* la artista utiliza en sus búsquedas múltiples conceptos muy variados que obedecen a sus intereses personales relacionados con la «tecnología». Por el contrario, en *Revolutionary arkive* todas las búsquedas de imágenes parten de un único concepto, el de «manifestación», pero en diferentes países del mundo, con la intención de extraer los arquetipos visuales de la colectividad. Hay un juego entre lo personal y lo colectivo en ambas propuestas y las dos se ven enriquecidas por la cooperación de los algoritmos.

Los grandes repositorios de imágenes constituyen una parte del recuerdo común de la humanidad y es creciente el interés hacia el concepto de «memoria colectiva». Con respecto a la historia y la teoría del arte y la cultura, Malraux y Warburg estaban interesados en el papel de las imágenes como representaciones y vías de transmisión de la memoria psicológica de las sociedades, pero también en la metamorfosis de formas y modos de expresión en la memoria social global (Grebe, 2010: 8).

En las exposiciones presentadas, la posibilidad de abarcar el contenido visual total de una colección nos permite considerar una información a la que no tendríamos acceso analizando solo las partes, las propiedades de la totalidad no resultan de los elementos constituyentes, sino que emergen de las relaciones espacio-temporales del todo (Köhler, 1947; Koffka, 1967). Estas metodologías nos abren nuevas vías para el diálogo con el pasado y prometen iluminar muchos aspectos de la historia y evolución de las imágenes.

«Artistas y científicos tienen que habérselas con esos datos e imágenes, y no solo investigar y criticar lo que se está haciendo con ellos desde arriba, sino proponer otras vías, otros juegos y preguntas que apunten a algo que no sea vender o venderse»

(Pérez-Hita, 2019: 2).

Desde hace años Margaret Boden estudia la forma en que la inteligencia artificial puede ayudarnos a explicar la creatividad humana

y, aunque reconocer la creatividad en máquinas o en humanos no es nada sencillo, Boden aclara que se pueden distinguir tres tipos: creatividad combinatoria, exploratoria y transformacional. En la creatividad combinatoria las ideas conocidas se mezclan de maneras desconocidas, se construyen analogías a partir de similitudes estructurales. La creatividad exploratoria se apoya en reglas y estructuras que ya tienen valor cultural como, por ejemplo, estilos pictóricos o musicales, para generar nuevas propuestas (Boden, 2017: 73). El proceso creativo seguido por Rosado en ambas exposiciones pone en juego la creatividad combinatoria y la exploratoria. La combinatoria porque se propicia el encuentro entre las analogías visuales que establecen los algoritmos informáticos y los criterios de búsqueda que decide la artista. En lo que respecta a la creatividad exploratoria, en el proceso se solapan los arquetipos visuales de la autora con los patrones visuales que detecta el algoritmo en los grupos de fotografías analizados.

Podríamos decir que en este proceso creativo se superponen las preconcepciones visuales del humano y las de la máquina. Esta es una diferencia sustancial a nivel metodológico con las estrategias para establecer analogías visuales utilizadas por Aby Warburg, ya que él solo disponía de su propio criterio. En ambos casos la estrategia tiene en común el deseo de establecer un diálogo con las imágenes para pensar.

Desde este punto de vista, en las obras de Rosado descritas, se pone de manifiesto la forma en que la conjugación humano-máquina puede favorecer los procesos de cocreación. Esta colaboración depende en gran medida del juicio humano, quien al final es el responsable del éxito del sistema; sin embargo, los procesos de cálculo aumentado del computador no hacen más que estimular este proceso creativo proporcionando rapidez a unas búsquedas que serían imposibles de otro modo. Sin olvidar que los resultados de las búsquedas utilizadas para obtener el conjunto de imágenes están mediatizados por los algoritmos que utiliza Google, no cabe duda de que la interacción con el computador alimenta encuentros fortuitos e inesperados, además de los previsible, y que en este proceso de búsqueda sirven para activar la imaginación de la artista.

Para finalizar, hay que puntualizar que el proceso de decisión del tamaño y número de imágenes descargadas condiciona la medida final de los mapas y esto se decide a partir de la obtención de grandes fotografías que no superen la escala humana, para que sea posible abarcar con la mirada el todo y la parte. Así se nos desvelan a un tiempo los detalles de las pequeñas imágenes y las configuraciones de los arquetipos de la totalidad. En la distancia, las grandes fotografías revelan unas valorables cualidades pictóricas de arte abstracto, las pequeñas fotografías se asemejan a las pinceladas, se reducen las representaciones objetuales contenidas en ellas hasta quedar únicamente representada la tensión visual que se establece por la gradación formal en la que se disponen. La información se hace arte y su belleza pretende hacernos despertar.

Agradecimientos

Nuestro sincero agradecimiento a los revisores anónimos que con sus comentarios ayudaron a mejorar el manuscrito original y a Roca Umberto Fábrica de las Artes de Granollers (Barcelona), Festival Panoràmic, Archivo Municipal de Granollers y ADAG.

Referencias bibliográficas

- Boden, Margaret Ann. 2017. *Inteligencia Artificial*. Madrid: Turner Noema.
- Company, Oriol. 2018. «Revolutionary Arkive. Ordenar el desordre». *Núvol. El Digital de la Cultura*, October 10. <https://www.nuvol.com/pantalles/revolutionary-arkive-ordenar-el-desordre-55800>.
- D'Oultremont, Maria. 2019. «Pilar Rosado i la panoràmica perduda de les imatges». *Núvol. El Digital de la Cultura*. May 17. <https://www.nuvol.com/art/pilar-rosado-i-la-panoramica-perduda-de-les-imatges-60640>.
- Didi-Huberman, George. 2010. *Atlas. ¿Cómo llevar el mundo a cuestas?*. Madrid: TF Editores.
- Festival Panoràmic. 2018. «The whole world is watching». <http://panoramicanollers.cat/revolutionary-arkive/>.
- Fontcuberta, Joan. 2016. *La furia de las imágenes. Notas sobre la postfotografía*. Barcelona: Galaxia Gutenberg.
- Fontcuberta, Joan. 2019. «El gran riu. Per a una teoria general de la revolta». *El món d'Ahir. Història d'autor*, 13: 176-189.
- Google. «Cómo funcionan los algoritmos de Google». May 22, 2020. https://www.google.com/intl/es_es/search/howsearchworks/algorithms/.
- Grebe, Anja. 2010. «Museum and Mnemosyne. Aby Warburg, André Malraux and the re-/construction of art history as social history». *“Le Musée Imaginaire” and Temptations of the Orient and Japan*, ed. Tanaka Hidemichi, 55-61. Tokyo: Akita.
- He, Kaiming; Xiangyu Zhang, Shaoqing Ren y Jian Sun. 2016. «Deep residual learning for image recognition». *In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 770-778. http://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html. <https://doi.org/10.1109/CVPR.2016.90>.
- Karpathy, Andrej. 2017. «t-SNE visualization of CNN codes», accessed January 12. <https://cs.stanford.edu/people/karpathy/cnnembed/>.
- Koffka, Kurt. 1967. *Principles of Gestalt Psychology*, New York: Mimesis International.
- Köhler, Wolfgang. 1947. *Gestalt psychology: an introduction to new concepts in modern psychology*. New York: Liveright.
- Malraux, André. 1949. *Le Musée imaginaire*. Ginebra: Skira.
- Manovich, Lev. 2009. «TimeLine», accessed January 10. <http://manovich.net/index.php/exhibitions/timeline>.

- Manovich, Lev. 2012. «¿Cómo ver 1000000 de imágenes? *Deforma cultura online*, 3, 1-11, Recuperado el 10 de Agosto de 2018 de http://www.deforma.info/es/product.php?id_product=24.
- Manovich, Lev. 2019. «Defining AI Arts: Three Proposals». *AI and Dialog of Cultures*, Exhibition Catalog. Saint-Petersburg, Russia: Hermitage Museum.
- Pérez-Hita, Félix. 2019. «Pilar Rosado y la resurrección de los archivos». *Mnemosyne 2.0. Cartografías computacionales de la memoria visual*, 1-2. Granollers: Roca Umbert Fàbrica de les Arts.
- Rosado, Pilar. 2019. *Mnemosyne 2.0. Cartografías computacionales de la memoria visual*. Granollers: Roca Umbert Fàbrica de les Arts.
- Russakovsky*, Olga, Jia Deng*, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, y Li Fei-Fei (* = equal contribution). 2015. «ImageNet Large Scale Visual Recognition Challenge». *International Journal of Computer Vision (IJCV)*, 115: 211–252. <https://doi.org/10.1007/s11263-015-0816-y>.
- Saramago, José. 2003. *Ensayo sobre la ceguera*. Madrid: Alfaguara.
- Maaten, Laurens van der y Geoffrey Hinton. 2008. «Visualizing data using t-SNE». *Journal of Machine Learning Research*, 9 (Nov): 2579-2605.
- The Warburg Institute. 2018. «The Mnemosyne Atlas, October 1929», accessed 2020 February 20. <https://warburg.sas.ac.uk/collections/warburg-institute-archive/bilderatlas-mnemosyne/mnemosyne-atlas-october-1929>.
- Warburg, Aby. 2010. *Atlas Mnemosyne*, traducido por Joaquín Chamorro Mielke. Madrid: Akal.

CV



Pilar Rosado Rodrigo

Universitat de Barcelona

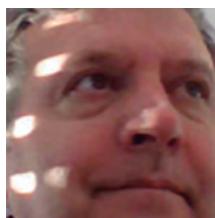
pilarrosado@ub.edu

Pilar Rosado es una artista e investigadora cuya práctica explora el uso de algoritmos de aprendizaje automático en el ámbito de la creación. Trabajando con grandes colecciones de imágenes de archivo y estas nuevas tecnologías pretende proporcionar puntos de vista alternativos para la reflexión y que cuestionen las convenciones de nuestra mirada. Doctora en Bellas Artes y Licenciada en Biología, entre 2012 y 2015 fue becada por la Generalitat de Catalunya para llevar a cabo su investigación titulada «Formas latentes: protocolos de visión artificial para la detección de analogías aplicados a la catalogación y creación artística», cuyos resultados han sido publicados en la prestigiosa revista Leonardo del MIT Press en junio de 2019. Combina su carrera artística con la investigación y la docencia como profesora en la Facultad de Bellas Artes de la Universidad de Barcelona, siendo miembro de varios proyectos de I+D.

ORCID: <https://orcid.org/0000-0001-7142-5047>

www.pilarrosado.eu

CV



Ferran Reverter Comes

Universitat de Barcelona

freverter@ub.edu

Ferran Reverter es profesor agregado de Estadística del Departamento de Genética, Microbiología y Estadística de la Universidad de Barcelona. Profesor del Máster de Estadística e Investigación Operativa UB-UPC en la asignatura de Statistical Learning. Ha realizado estancias posdoctorales en el Laboratoire de Statistique et Probabilités de la Université Paul Sabatier (Toulouse, 2008) y en el Centre de Regulació Genòmica del Parc de Recerca Biomèdica (Barcelona, 2013-2017). En la actualidad centra su investigación en el desarrollo e implementación de herramientas estadísticas para el análisis de datos en biociencias, la mejora de métodos basados en funciones de núcleo (kernel methods) para aumentar la interpretabilidad y visualización, y en la implementación de herramientas de análisis computacional de imágenes en el ámbito biomédico y el artístico.

ORCID: <https://orcid.org/0000-0002-9489-3350>.