

IDENTIFICACIÓN DE PROCESOS FOTOGRAFICOS MEDIANTE TÉCNICAS DE APRENDIZAJE DE MÁQUINA

Gustavo Lozano San Juan

Archivo Fotográfico Manuel Toussaint.

Instituto de Investigaciones Estéticas.

Universidad Nacional Autónoma de México

RESUM

En aquesta ponència es presenten les experiències i els resultats de dos projectes basats en l'aplicació d'eines d'intel·ligència artificial per a la identificació dels processos fotogràfics en el context del treball arxivístic. Al primer es va utilitzar un arbre de decisió basat en l'algoritme C4.5 que opera sobre les metadades relacionades amb les característiques físiques, i en el segon es va utilitzar la xarxa neuronal convolucional VGG19, que opera sobre la imatge de l'objecte en qüestió.

Es descriuen les particularitats del problema a resoldre, les consideracions tècniques de cadascuna de les estratègies, els desafiaments per a la interpretabilitat dels resultats i del procés i es conclou amb una sèrie de recomanacions per a treballs futurs.

SUMMARY

This conference presents the experiences and results of two projects based on the application of artificial intelligence tools for the identification of photographic processes in the context of archival work. In the first one, a decision tree based on the C4.5 algorithm that operates on feature-related metadata physical, and in the second the convolutional neural network VGG19 was used, which operates on the image of the object in question.

The particularities of the problem to be solved are described, the technical considerations of each of the strategies, the challenges for the interpretability of the results and the process and concludes with a series of recommendations for future work.

Este trabajo describe dos experimentos con técnicas de aprendizaje de máquina en los que se aborda el problema de la identificación del proceso fotográfico en el contexto del trabajo de documentación, digitalización y conservación que se realiza al interior de los archivos fotográficos históricos. A lo largo del texto se analizan las particularidades del problema que se busca resolver, se describen los aspectos técnicos de cada experimento, se evalúan los resultados y se plantean recomendaciones para investigaciones a futuro.

Herramienta web para identificar procesos fotográficos

POLARIDAD

¿Cuál es la polaridad de su fotografía?

Todas las fotografías se pueden dividir en dos categorías, Positivos y Negativos. A esta característica física se le conoce como Polaridad y se refiere a la relación que existe entre los tonos de la escena original y los de la fotografía, cuando los tonos de la fotografía son iguales a los de la escena original decimos que se trata de un Positivo y cuando los tonos de la fotografía son opuestos a los de la escena original se trata de un Negativo.

Observa los ejemplos de abajo y selecciona una de las dos opciones.

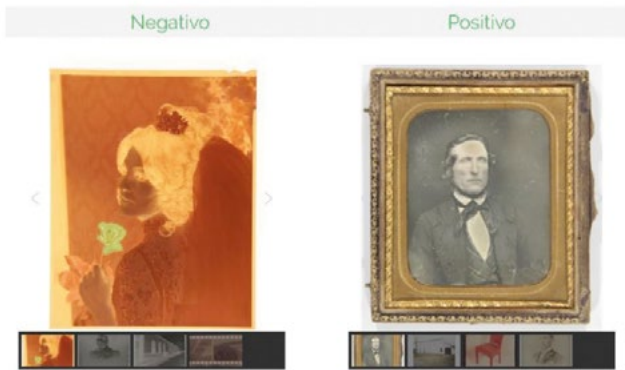


Ilustración. 1. Propuesta de interfaz de usuario para el árbol de decisiones.

En el primer proyecto se utilizó un árbol de decisión, una herramienta de apoyo para la toma de decisiones que permite clasificar entidades con base en sus atributos. Una de las mayores virtudes de los árboles es que se pueden aplicar de forma manual, esto es posible gracias a que visualmente son muy claros y explícitos, por lo que su utilización y comprensión resulta muy intuitiva para los usuarios.

Se eligió el algoritmo C4.5 y el software Weka, se ingresaron como datos de entrada los metadatos que describen las características físicas de las fotografías, como el tipo de soporte, la polaridad, el tono, la tonalidad, entre otros; el software cuantificó la contribución de cada uno de los metadatos en términos de información, iterando hasta producir el árbol más eficiente en brindar la respuesta correcta. El resultado de este proceso fue un árbol de 26 nodos, 25 ramas y 18 hojas, que tiene un porcentaje de instancias clasificadas correctamente de 85.41%.

A futuro se podría delimitar el problema general de identificación a un subconjunto de más acotado de procesos de una misma tipología, tono o polaridad; también sería recomendable evaluar el uso del árbol con usuarios que puedan aportar su experiencia y conocimiento al modelo; convendría también trasladar el árbol a un esquema que permita a los usuarios ubicar su posición y visualizar la ruta que deben seguir dentro del proceso, así como diseñar una interfaz gráfica amigable que explique los conceptos a identificar y ofrezca ejemplos de las respuestas posibles.

En el segundo proyecto se reentrenó la red neuronal convolucional Imagenet para identificar el proceso fotográfico a partir de una imagen digital del objeto original. Las redes neuronales convolucionales (RNC) son dispositivos para el análisis de información visual que han tenido un desarrollo muy acelerado en los últimos diez años. Algunas sus principales características son la complejidad de su operación interna (relacionado con los elementos que los conforman, como capas, neuronas, pesos, sesgos, convoluciones, filtros y funciones de activación), así como la amplia gama de herramientas disponibles que buscan simplificar su implementación.

La operación de convolución que da nombre a estas redes es aquella en la que el dígito que corresponden a cada pixel de la imagen se multiplica por un filtro o kernel, el resultado es una versión de la imagen en la que se destacan algunas de sus características y se simplifican otras. La convolución se repite en diferentes capas de la red, con lo cual progresivamente se va obteniendo una versión más reducida, pero a la vez más destilada de la imagen, llamada mapa de características, con ésta es posible alimentar la última capa conocida como clasificador y de esta manera la red identifica el contenido de una imagen en alguna de las categorías utilizadas para su entrenamiento.

Las redes más comunes han sido entrenadas con Imagenet, un conjunto de datos en el que se muestran personas, acciones, objetos, lugares, animales y vegetales clasificados en 20,000 categorías diferentes. En este proyecto se utilizó la red Inception y fue reentrenada mediante el proceso llamado transferencia de aprendizaje para diferenciar el proceso fotográfico en imágenes de históricas. Para ello se utilizaron 600 ejemplos de fotografías divididas en tres procesos fotográficos: ambrotipo, daguerrotipo y ferrotipo.

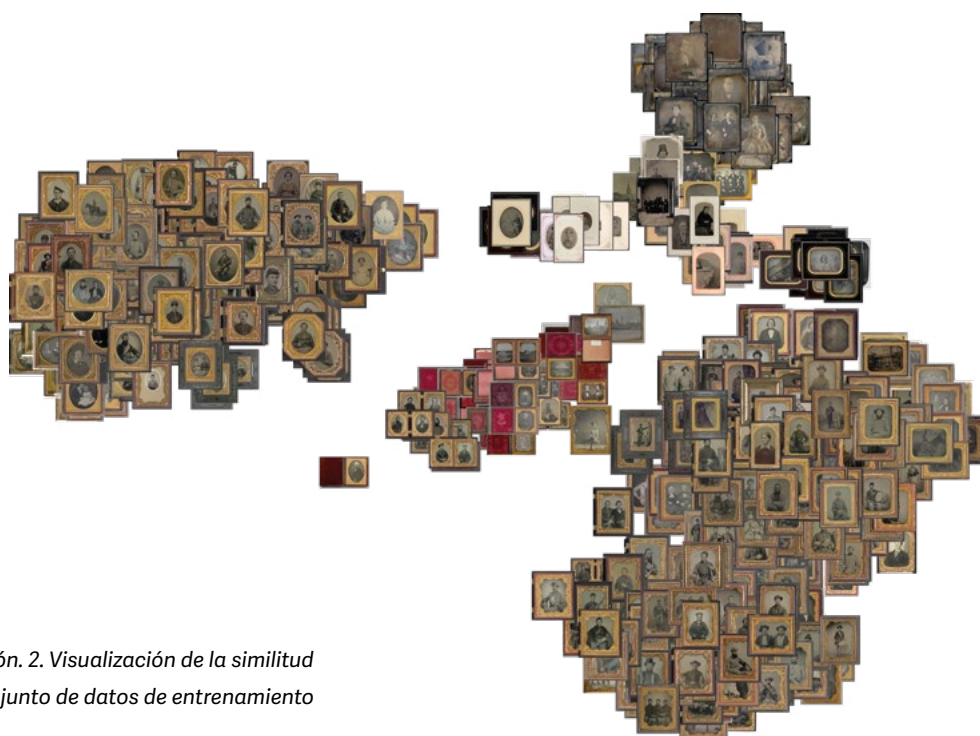


Ilustración 2. Visualización de la similitud en el conjunto de datos de entrenamiento

El resultado de este proceso fue una red con un porcentaje de 85 % de precisión en la fase de entrenamiento y de 75 % en la de validación. Una técnica de visualización basada en la reducción multidimensional permitió identificar que la identificación de las imágenes se basa inicialmente en aspectos como la forma de montaje y presentación (estuche, marco o tarjeta) y solo de forma secundaria en las características físicas relacionadas con el proceso fotográfico, a futuro será necesario intentar que la red enfatice este tipo de característica ya que son las que más interesan en este caso.

Los dos experimentos realizados han dado resultados alentadores que ameritan una investigación más detallada y completa, de momento se pueden considerar como un punto de partida para un proyecto formal a futuro. El principal recurso que habría de sumarse es la participación de un especialista en inteligencia artificial que analice críticamente los planteamientos iniciales para corregirlos, refinarlos y robustecerlos.

Identificación de Procesos Fotográficos

Mi interés, por reflexionar y experimentar con diferentes herramientas físicas y metodológicas para la identificación de procesos fotográficos, surge de la oportunidad que he tenido de impartir una gran cantidad de talleres sobre este tema. La mayoría de ellos han sido breves y solo algunos han tenido una duración de más de cuatro horas; sin embargo, casi siempre el tiempo resulta insuficiente para abordar el tema a profundidad.

Generalmente, son talleres prácticos en los que participan colegas que trabajan en archivos y que comparten diversas responsabilidades simultáneas como la conservación, la documentación, la digitalización, el servicio de consulta, etcétera. Ellos y ellas cuentan con una amplia experiencia en el trabajo diario, poseen un gran compromiso por su labor y sienten respeto y afecto por el patrimonio que resguardan –aun cuando su formación profesional no está vinculada con el ámbito de la archivística, la conservación o la fotografía-. Y es, a través de este tipo de talleres, que buscan desarrollarse profesionalmente para cumplir de mejor manera con su encargo.

En el contexto del trabajo de un archivo fotográfico, cuando nos acercamos por primera vez a una fotografía que no conocemos, surgen en nuestra mente algunas preguntas básicas como las siguientes: ¿Qué representa la imagen? y ¿de qué año es?, ¿Quién es el autor o autora?, ¿Cuál es su tamaño, material de soporte, color y técnica con la que fue creada? Personas con una diferente formación profesional pueden sentir un mayor interés por una o por otra de estas preguntas. Por ejemplo, los historiadores pueden tener una curiosidad por conocer el origen, función y trayectoria de la imagen; mientras que un fotógrafo podría estar más interesado en los aspectos técnicos de la toma, como el encuadre, la composición, el rango tonal, el género fotográfico, entre otros aspectos.

Sin embargo, en este primer contacto siempre existe una inquietud por conocer los aspectos esenciales que definen la materialidad del objeto fotográfico. De

hecho, es frecuente que la respuesta a las preguntas que se plantean sea una visión unificada en la que se cruzan la función original, la circulación de la imagen; así como las huellas físicas que adquiere el objeto a lo largo del tiempo como testimonio de su trayectoria y valoración. La importancia de esta información es que nos permite contextualizar la fotografía, que conforma parte de la base sobre la que posteriormente se pueden construir indagaciones más amplias y complejas.



Ilustración 1. Diferentes procesos fotograficos

Tras este primer acercamiento, se necesita hacer una diversidad de tareas dentro del archivo y requieren que contemos con información sobre el proceso fotográfico de una determinada fotografía, serie o fondo; para conseguir un objetivo específico, entre las que destacan:

Para hacer un uso eficiente del espacio, las guardas y los contenedores de almacenamiento, al momento de elaborar la propuesta de ordenación.

Para asentar, en los instrumentos de consulta, los datos relacionados con las características físicas. Para saber qué tan rara o común es la fotografía, serie o fondo, y así poder realizar una valuación monetaria. Para saber qué técnicas y equipos utilizar en su digitalización. Para saber cómo debe manipularse duran-

te la consulta y evitar causar daños. Para redactar la cédula que acompañe la imagen en una exposición o el pie de foto en una publicación. Para comprender el estado físico en el que se encuentra y anticipar los deterioros que puede sufrir. Para establecer las medidas de conservación que requiere y saber que tratamientos de restauración podemos aplicar.

Sin embargo, no todas estas tareas requieren del mismo nivel de precisión al momento de identificar el proceso, adicionalmente, no siempre es posible llegar a una respuesta de la que se tenga certeza absoluta. Quizá, el caso más delicado sea el de la restauración, porque es bien conocido que algunos tratamientos y sustancias tienen efectos negativos y pueden causar daños irreversibles. Como el lavado por inmersión acuosa, en el caso de las impresiones a la albúmina o la limpieza con acetona en el caso de los negativos e impresiones con aglutinante de colodión. Este es un tema muy interesante que requiere de mayor atención, aunque no profundizaré en él en este momento; considero necesario enumerar los siguientes cuestionamientos:

¿Qué tareas requieren un mayor nivel de precisión y confianza y cuáles no? ¿Cómo comunicamos un mayor o menor grado de incertidumbre en la identificación del proceso? ¿Cómo operamos alrededor de esta incertidumbre?

Este marco nos lleva al cuestionamiento principal del presente trabajo: ¿En qué consiste el problema de identificación del proceso fotográfico? Desde mi punto de vista, se trata de un problema en el que buscamos comprender; y no simplemente asignar a una categoría la ubicación de una fotografía dentro de un espacio complejo con varias dimensiones o capas, como son la histórica, la tecnológica o la de la cultura visual.

En torno a estos diferentes aspectos y a través de observaciones puntuales, es que construimos un perfil de lo que nuestra fotografía es y también de lo que no es. Por ejemplo, reconocer que tengo ante mí una fotografía con soporte de vidrio, me permite ubicarla cronológicamente en una temporalidad cercana al siglo XIX o a los inicios del S. XX; centrar la atención en tipologías concretas como los negativos o las imágenes de cámara; y excluir simultáneamente a casi todos los procesos policromos. En este sentido, se trata de una indagación que se realiza en comparativa con otras imágenes y no de manera aislada, por lo que se

exige, tradicionalmente como requisito indispensable, conocer el universo completo de posibilidades para posteriormente ir eliminando las probabilidades.

En su aspecto performativo, la identificación del proceso fotográfico, consta de dos componentes que se accionan simultáneamente y se retroalimentan entre sí. Por un lado, se encuentra la manipulación del objeto fotográfico y la percepción multisensorial de sus principales características materiales; entre las que destacan las visuales pero que también incluye las táctiles, olfativas e incluso sonoras, según los planteamientos más esotéricos. Entre estas características se encuentran: el brillo, textura, color, tono, grosor, tamaño, rango tonal, nitidez y el contraste. Por otro lado, se ubica la interpretación y vinculación de aquello que está siendo observado en un vasto cuerpo de información, donde confluyen temas de diversas disciplinas -la historia, los estudios de cultura visual, la fotografía y la conservación- que incluyen aspectos que a continuación se describen.

Dentro de la historia, se debe considerar la cronología general de la invención, uso y obsolescencia de los procesos fotográficos, así como la ubicación temporal de los principales hitos tecnológicos, como el uso de los diferentes soportes plásticos, la introducción de la fotografía a color, la evolución de los formatos y su relación con la democratización de la producción fotográfica.

En los estudios de cultura visual: el dominio de estilos, y estética en la vestimenta, el peinado y el diseño en general durante diferentes contextos. Así como los distintos géneros y funciones fotográficas -vistas mexicanas, retratos de estudio, fotografía aérea, reprografía-. Los formatos de presentación y el estilo de elementos decorativos en impresiones, soportes secundarios, estuches, marcos y paspartús.

En la fotografía: los conceptos y mecanismos de formación de la imagen: imagen latente, revelado, fijado y virado. La distinción funcional de las diferentes tipologías fotográficas: negativo, impresión, diapositiva e imagen de cámara. Los rasgos expresivos de los materiales fotográficos como: el brillo, la textura, el grosor, y el color del papel y de la imagen en impresiones.

Por último, en la conservación: las técnicas de observación detallada y de manipulación cuidadosa, los materiales constitutivos y las técnicas artesanales e

industriales de fabricación, la estratigrafía del objeto fotográfico -soporte, aglutinante, partícula formadora de la imagen- y los posibles agentes, mecanismos y efectos de deterioro.

Comúnmente, las metodologías de los talleres, publicaciones y páginas web, que tratan sobre la identificación de procesos fotográficos plantean que las personas que quieran aprender, deberán dominar previamente este vasto cuerpo de información. Lo anterior, implica una curva de aprendizaje demasiado extensa que demanda un esfuerzo sostenido a lo largo de un tiempo prolongado para poder acceder a las habilidades y conocimientos antes descritos. Esto, limita la diseminación y el acceso a dicho conocimiento y, en última instancia, reduce las posibilidades de éxito para personas interesadas en el tema.

Por todo lo anterior, este trabajo explora la factibilidad de dos herramientas digitales basadas en la inteligencia artificial, que ayuden a los profesionales de los archivos a llevar a cabo la identificación de procesos fotográficos de manera más ágil y efectiva. La idea es tomar como punto de partida las necesidades, motivaciones, experiencias y el contexto de estos usuarios; que permita diseñar un instrumento que no los sustituya, sino que los guíe y los apoye en la realización de la tarea. De tal manera que, progresivamente, se familiaricen con el problema que se busca resolver y, eventualmente, construyan la red de conocimientos que les permita completar con calidad y confianza las funciones que se han enumerado previamente.

Inteligencia Artificial y Aprendizaje de Máquina

Los dos proyectos que aquí se exploran están basados en una técnica computacional derivada de una rama de la inteligencia artificial, denominada: Aprendizaje de Máquina (AM). Donde se enfatiza el aprender a realizar una operación o tarea a partir de un conjunto de ejemplos del resultado que se busca obtener: a diferencia del paradigma instruccional, en el que la tarea se descompone en una secuencia de pasos individuales que se programan explícitamente para ser ejecutados por la computadora.

Esta tecnología se encuentra vinculada con las recomendaciones de servicios de transmisión de videos y música en línea; la interacción de los asistentes personales de

voz, la identificación de candidatos para un puesto de trabajo, entre otras aplicaciones. En el ámbito de la imagen, las aplicaciones más conocidas son el reconocimiento de señales de tráfico para el sistema de navegación en vehículos autónomos y en el reconocimiento facial, usado tanto en sistemas de vigilancia como en redes sociales.

Para generar un modelo de aprendizaje a partir de un conjunto de datos, a grandes rasgos se llevan a cabo los siguientes pasos: obtención, limpieza y preparación de datos; ejecución de la heurística o algoritmo y evaluación de los resultados. A continuación, se describen dos propuestas que siguen esta ruta: la primera está basada en Árboles de Decisiones (AD) y la segunda en Redes Neuronales Convolucionales (RNC).

Árboles de decisión

En tanto que este proyecto tiene una trayectoria de varios años, los aspectos fundamentales se pueden consultar en la charla: "Propuesta metodológica para identificar el proceso fotográfico en fotografías históricas", presentada en marzo de 2019 dentro del *III Coloquio de Archivística: Los archivos filmicos y fotográficos*, organizado por la Escuela Nacional de Estudios Superiores Morelia de la UNAM¹. Posteriormente en 2020 publiqué, junto a mi colega Rodrigo Colín, el texto: *Propuesta metodológica y de análisis computacional para identificar el proceso fotográfico en fotografías históricas del siglo XIX y XX*², texto en el que se describe con mayor detalle la metodología que se retoma en este trabajo de manera general.

Un árbol de decisión es una herramienta de apoyo para la toma de decisiones. Como su nombre lo indica, se basa en un modelo con forma de árbol en el que se representan de manera gráfica los puntos en los que se requiere tomar una decisión, siendo las bifurcaciones consecuencia de éstas. Visto de otra forma los árboles de decisión también permiten clasificar entidades (representadas por las hojas del árbol) con base en sus atributos (representados por las ramas)³.

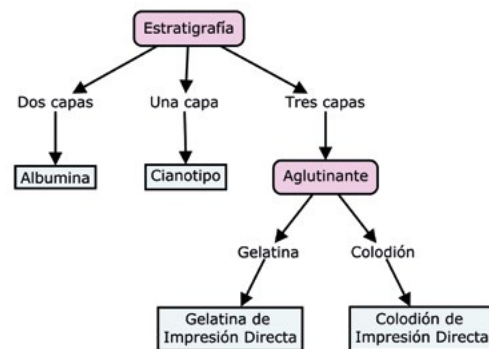


Ilustración 2. Ejemplo básico de un árbol de decisión

Hay dos fases esenciales en la construcción del árbol: la generación y la aplicación. En la generación se utiliza un conjunto de datos confiables, que deben ser normalizados y depurados: posteriormente se ingresan en un programa especializado en minería de datos, que debe configurarse con los parámetros que se detallarán más adelante. En este caso se eligió el programa WEKA⁴ del cual se obtuvo un árbol que se puede utilizar para tomar una decisión que nos guíe a identificar o clasificar una entidad determinada, dentro de un universo con diversas alternativas.

Es importante especificar que se utilizan los atributos en formato textual para obtener como resultado la identificación del proceso fotográfico de manera textual. En nuestro caso, clasificar quiere decir identificar correctamente el proceso fotográfico de una fotografía de archivo; mientras que los atributos corresponden a las diferentes características físicas que presentan.

La aplicación del AD en un archivo fotográfico puede tomar dos rutas. La primera, por medio de un algoritmo el cual opere automáticamente y tome como datos de entrada la información de los campos que documentan las características físicas, como parte de un catálogo o algún otro instrumento de control intelectual o de consulta, así como datos de salida se obtendría el proceso fotográfico para todo un fondo o colección simultánea-

1 - III Coloquio de archivística, 2019, <https://www.youtube.com/watch?v=Ogl37lkmuDQ>.

2 - Gustavo Lozano y Rodrigo Colín, "Propuesta metodológica y de análisis computacional para identificar el proceso fotográfico en fotografías históricas del siglo XIX y XX", en *Inteligencia artificial y datos masivos en archivos sonoros y audiovisuales* (México: Instituto de Investigaciones Bibliotecológicas y de la Información. UNAM, 2020).

3 - "Decision Tree Learning", en Wikipedia, el 31 de diciembre de 2021, https://en.wikipedia.org/w/index.php?title=Decision_tree_learning&oldid=1063019167.

4 - "Home - Weka Wiki", consultado el 26 de julio de 2022, <https://waikato.github.io/weka-wiki/>.

mente. La segunda ruta, que considero más adecuada, es utilizar el AD como una herramienta metodológica interactiva que los profesionales de los archivos puedan seguir paso a paso. Aprovechando las características del AD para estructurar, optimizar e ilustrar un proceso, que de otra manera resultaría muy complejo.

Los árboles de decisión se componen de tres elementos: la raíz o punto de partida; los nodos intermedios, los cuales representan a los atributos que describen a las entidades y, en este caso, corresponden a las características físicas de las fotografías; y finalmente las hojas, estas representan a cada uno de los procesos fotográficos que es posible encontrar en el archivo.

Algunas ventajas de los árboles de decisión son: se pueden generar sin necesidad de una computadora, sin embargo, mientras mayor sea el número de atributos esto será más complejo.

Aunque la toma de decisión se puede automatizar, una de sus mayores virtudes es que los árboles se pueden aplicar de manera manual. Esto, es posible gracias a que visualmente son claros y explícitos, y su uso y comprensión es muy intuitivo. A diferencia de otros modelos, como las redes neuronales; considerados modelos de caja negra porque usualmente es difícil de explicar en términos simples por qué y cómo se generan sus predicciones o clasificaciones.

Para el proceso de generación del árbol se analizaron dos algoritmos. El algoritmo ID3 (*Iterative Dichotomiser 3*) fue inventado por Ross Quinlan 'investigador en ciencias de la computación en el área de minería de datos y teoría de la decisión' para crear árboles a partir de bases de datos. Al calcular la entropía -unidad de medida del desorden en un conjunto de datos, según la teoría de la información- para cada atributo del conjunto de datos, permite dividir en subconjuntos basados en el valor de mínima entropía, y construir de manera recursiva cada nodo del árbol de decisión. En conjunción, ID3 utiliza también el método de ganancia de información, la medida de las diferencias de entropía antes y después de que un atributo se divide en subconjuntos de datos, para determinar el nodo raíz en cada llamada recursiva del algoritmo.

Quinlan efectuó mejoras en su algoritmo original y creó el algoritmo C4.5, que de igual manera emplea el método de ganancia de información, pero hace más evidente que el árbol resultante pueda ser usado como modelo de clasificación. Entre las mejoras más notables, respecto a ID3, se encuentra la habilidad de trabajar con atributos continuos, como valores numéricos o fechas. También permite trabajar con valores de atributos vacíos o "huecos" en la base de datos, sin que esto afecte la ganancia de información de un atributo durante la construcción del árbol. Y por último, el árbol creado con C4.5 es recortado después de su creación, es decir, algunas ramificaciones del árbol son modificadas por nodos terminales cuando contribuyen a simplificar el árbol⁵.

Para la generación de árboles de decisión se empleó el software de código abierto WEKA (<https://www.cs.waikato.ac.nz/ml/weka>). El cual, cuenta con una colección de algoritmos y herramientas de pre procesamiento de datos que permiten trabajar de manera flexible con conjuntos de datos; asimismo facilita la experimentación, obtención de resultados estadísticos y visualización tanto de los datos de entrada como de los resultados obtenidos.

Los datos que se emplearon en este trabajo son una acotada selección de diversas características fotográficas de la colección de estudio, perteneciente al área de conservación del Archivo Fotográfico Manuel Toussaint del Instituto de Investigaciones Estéticas de la UNAM. Todas las características físicas, de los procesos fotográficos a identificar, se registraron en una hoja de cálculo que sirve de insumo inicial en esta etapa del proceso.

En la etapa de limpieza y preparación de los datos, de las 27 columnas originales de la hoja de cálculo, se mantuvieron únicamente 15 columnas o atributos que representan conjuntos de información, que se acotaron a un rango de valores bien identificados. Las razones principales para descartar algunas de las columnas son la ausencia de información en campos opcionales (por ejemplo, subtipo de soporte) y la dispersión de datos, en el caso de las columnas donde se escriben comentarios o anotaciones que suelen ser únicos (por ejemplo, particularidades del objeto fotográfico).

5 - Lozano y Colín, "Propuesta metodológica y de análisis computacional para identificar el proceso fotográfico en fotografías históricas del siglo XIX y XX".

Clasificaciones		Atributos comunes				Atributos particulares									
Proceso	Tipología	Soporte primario	Iluminación	Polaridad	Tono	Fecha	Estratigrafía	Magnificación	Tonalidad	Brillo	Superficie	Particularidades Objeto	Texto	Deterioro	
Daguerstipo	Imagen de cámara	Metal	Reflexión	Positivo	Monocromático	1839 - 1860			Neutro	Muy brillante	Positivo-negativo			Delaminación de plata, corrosión de cobre	
Ambrótipo	Imagen de cámara	Vidrio	Reflexión	Positivo	Monocromático	1851 - 1865			Café	Brillante	Luces lechosas			Corrosión, faltantes, craqueladuras	
Ferrotipo	Imagen de cámara	Metal	Reflexión	Positivo	Monocromático	1855 - 1890			Café	Semi mate					
Cianotipo	Impresión	Papel	Reflexión	Positivo	Monocromático	1840 - 1920	Una	Fibras visibles	Con	Mate					
Albumina	Impresión	Papel	Reflexión	Positivo	Monocromático	1851 - 1890	Das	Fibras visibles	Amarillo, Café rojizo	Sem mate	Textura del pap	Soporte primario delgado, Soporte secundario grueso		Casquedatos: amarillamiento, pérdida de densidad en las luces y sombras	
Carbón	Impresión	Papel	Reflexión	Positivo	Monocromático	1860 - 1940	Das		Otro		Relieve en sombras				
Colodión de impresión directa	Impresión	Papel	Reflexión	Positivo	Monocromático	1885 - 1910	Tres	No se ven fibras	Purpura_Rojo	Brillante	Textura lisa	Soporte primario grueso, Soporte secundario grueso, Indiscencia		Abrasiones, pérdida de densidad en las luces	
Plata gelatina de impresión directa	Impresión	Papel	Reflexión	Positivo	Monocromático	1885 - 1910	Tres	No se ven fibras	Amarillo	Brillante	Textura lisa	Soporte primario grueso, Soporte secundario grueso, Indiscencia		pérdida de densidad en las luces	
Goma bicromatada	Impresión	Papel	Reflexión	Positivo	Monocromático	1890 - 1930	Das		Otro						
Platinotipo	Impresión	Papel	Reflexión	Positivo	Monocromático	1890 - 1930	Una	Fibras visibles	Neutro					Ghosting	
Plata gelatina	Impresión	Papel	Reflexión	Positivo	Monocromático	1890 - 2018	Tres	No se ven fibras	Neutro						
Colodión mate de impresión directa	Impresión	Papel	Reflexión	Positivo	Monocromático	1895 - 1910	Tres	No se ven fibras	Neutro, Café (platino), púrpura (oro)	Semi mate	Texturizado	Soporte primario grueso, Soporte secundario grueso y de color,		Ghosting, No hay pérdida de densidad	
Difusión de plata	Impresión	Papel	Reflexión	Positivo	Monocromático	1942 - 2018	Tres		Neutro		Lisa	Pestañas, borde irregular o perforado, superficie con restos de adhesivo,	Polaroid	Revelado irregular, amarillamiento, pérdida de balance de color y desajuste de colorantes	
Cromógeno	Impresión	Papel	Reflexión	Positivo	Policromático	1940 - 2018	Tres				Brillante	Lisa o texturizada	Papel de fibra, RC, acetato pigmentado	< 1980 amarillamiento, pérdida de balance de color y desajuste de colorantes	
Difusión de colorantes por transferencia	Impresión	Papel	Reflexión	Positivo	Policromático	1963 - 2018	Tres				Brillante	Lisa	Pestañas, borde irregular o perforado, superficie con restos de adhesivo,	Polaroid	Revelado irregular,
Banqueo de colorantes	Impresión	Papel	Reflexión	Positivo	Policromático	1963 - 2018	Tres				Brillante	Lisa,	Marco blanco con contenedor de químicos de procesamiento	Polaroid	Revelado irregular, craqueladuras, migración de colorantes en áreas blancas
Difusión de colorantes por transferencia interna	Impresión	Papel	Reflexión	Positivo	Policromático	1972 - 2018	Tres				Brillante	Lisa,			
Colodión húmedo	Negativo	Vidrio	Transmisión	Negativo	Monocromático	1851 - 1885			Café		Aglutinante irregular, Barniz	Vidrio grueso e irregular		Abrasión	
Gelatina seca	Negativo	Vidrio	Transmisión	Negativo	Monocromático	1880 - 1925			Neutro		Aglutinante irregular	Vidrio delgado y recto		Especo de plata	
Plata gelatina filtro de celulosa	Negativo	Plástico	Transmisión	Negativo	Monocromático	1889 - 1950						Muscas "V"	Nitrato	Amarillamiento del soporte, deformación,	
Plata gelatina Acetato de celulosa	Negativo	Plástico	Transmisión	Negativo	Monocromático	1935 - 2018						Muscas "U"	Safety	Canalías, burbujas, deformaciones, olor a vinagre	
Plata gelatina Poliéster	Negativo	Plástico	Transmisión	Negativo	Monocromático	1955 - 2018							Safety		
Cromógeno Acetato de celulosa	Negativo	Plástico	Transmisión	Negativo	Policromático	1942 - 2018							Safety		
Cromógeno Poliéster	Negativo	Plástico	Transmisión	Negativo	Policromático	1955 - 2018							Safety		
Plata gelatina Vidrio	Transparencia	Vidrio	Transmisión	Positivo	Monocromático	1880 - 1940									
Plata gelatina Acetato de celulosa	Transparencia	Plástico	Transmisión	Positivo	Monocromático	1935 - 2018									
Plata gelatina poliéster	Transparencia	Plástico	Transmisión	Positivo	Monocromático	1955 - 2018									
Procesos aditivos	Transparencia	Vidrio	Transmisión	Positivo	Policromático	1907 - 1935		Retícula				Lineas rectas paralelas	Bajo brillo y saturación	Delaminación, puntos verdes	
Cromógeno Acetato de celulosa	Transparencia	Plástico	Transmisión	Positivo	Policromático	1935 - 2018								Safety	
Cromógeno Poliéster	Transparencia	Plástico	Transmisión	Positivo	Policromático	1955 - 2018								Safety	

Ilustración 5. Conjunto de datos de entrenamiento. Características físicas descritas textualmente

Es importante resaltar la necesidad que los datos estén normalizados, ya que, si existen errores ortográficos u otras inconsistencias, el funcionamiento del algoritmo no será el correcto. De hecho, para mantener consistencia e integridad de los datos, se puede definir un vocabulario.

Para utilizar el algoritmo C4.5 en WEKA, se hace uso de la versión de código abierto llamada J4.8 -versión ligeramente mejorada del algoritmo C4.5 revisión 8 que está implementado en el lenguaje de programación Java-.

El procedimiento es bastante directo, el primer paso consiste en seleccionar el archivo de datos desde el explorador de WEKA para apreciar el conjunto de datos en su totalidad. Enseguida se selecciona la pestaña de clasificación y se selecciona de la lista de opciones el algoritmo J48. Se asignan los parámetros correspondientes, se selecciona el atributo o colum-

na que representa las clases -en este caso, el proceso fotográfico-, se ejecuta el algoritmo y se observan los resultados que se obtienen⁶.

Es crucial determinar los parámetros adecuados para obtener el mejor árbol de decisión. Si bien, la cantidad de parámetros es acotada, son suficientes para generar diversidad de resultados, que solamente pueden ser considerados mejores o peores, en comparación con el porcentaje de registros clasificados correctamente y según un conjunto de pruebas (el cual puede ser el mismo conjunto de datos). Los parámetros más significativos, para lograr un árbol de decisión suficientemente compacto pero que logre clasificar correctamente la mayor cantidad de elementos. no siempre es evidente. Y requiere de experimentación y múltiples ejecuciones del algoritmo, en este caso, la configuración utilizada fue la que se muestra en la siguiente ilustración:

Parámetro	Descripción	Valor usado
Unpruned tree (unpruned)	Usar árbol de decisión sin poda o recortes (genera árboles más grandes).	False
Reduced-error pruning (reducedErrorPruning)	Determina si se utiliza un método alternativo de poda (no necesariamente con mejores resultados).	False
Pruning confidence (confidenceFactor)	Factor que se utiliza para podar o recortar el árbol; a menor valor mayor poda.	0.25
Minimum number of instances (minNumObj)	Establece el mínimo número de instancias por hoja o el número mínimo de ramificaciones.	2
MDL-correction (useMDLcorrection)	Ajuste basado en MDL (mínima longitud de descripción) para calcular ganancia de información.	False

Ilustración. 4. Parámetros de configuración del programa

El resultado de este proceso fue un árbol de 26 nodos, 25 ramas y 18 son nodos terminales u hojas, que permiten clasificar fotografías según los 15 atributos que se definieron en la base de datos. Se tiene un porcentaje de instancias clasificadas correctamente del 85.41%, que es el porcentaje más alto que se alcanzó con los parámetros que ya se mencionaron. Esto no representa un árbol ideal -que clasifique correctamente el 100% de las muestras- pero evidencia que, al recorrer el árbol desde la raíz hasta una hoja, hay características de la fotografía que tienen mayor peso para lograr una rápida clasificación.

Como resultado de este proceso, se obtuvo un modelo funcional con un nivel de certeza aceptable, que ofrece una base sólida sobre la cual se puede continuar investigando y construyendo. A futuro, se podrían introducir mejoras como: limitar el problema de identificación a un subconjunto de procesos como aquellos que correspondan a una misma tipología, tono o polaridad; esto ayudaría a separar el universo de posibilidades en bloques más manejables de información. Por otro lado, sería recomendable evaluar la usabilidad del AD con usuarios auténticos, que puedan aportar su experiencia y conocimiento para hacer recomendaciones que se integren al modelo. Convendría trasladar el AD a un esquema visual que resulte atractivo a los usuarios, que les permita ubicar su posición y comprender la ruta que deben seguir dentro del proceso. Finalmente considero que el algoritmo se debería acompañar de un cuestionario donde, además de plantear las preguntas, también explique los conceptos que se busca identificar y ofrezca ejemplos gráficos de las diferentes respuestas posibles.

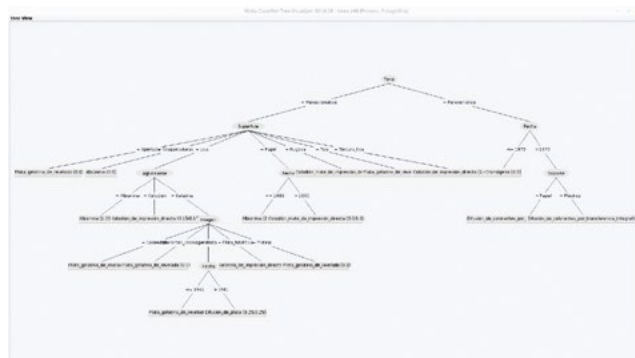


Ilustración. 5. Árbol de decisión resultante



Ilustración. 6. Propuesta de interfaz de usuario para el árbol de decisiones.

Redes Neuronales Convolucionales

Las redes neuronales convolucionales (RNC) son un tipo de sistemas computacionales para el análisis de información visual. Estas, han tenido un desarrollo muy acelerado en los últimos diez años; por lo que, actualmente hay proyectos académicos para evaluar su efectividad y conveniencia en una amplia variedad de aplicaciones, por ejemplo: el diagnóstico de padecimientos mediante el reconocimiento de imágenes médicas como radiografías y ultrasonidos; en vigilancia y seguridad, por medio del reconocimiento facial y la detección de objetos en imágenes de cámaras de seguridad; y en el área de humanidades digitales y patrimonio cultural, para hacer posible la búsqueda por imágenes y el reconocimiento de características técnicas, estilísticas y de contenido en cerámicas, pinturas, periódicos, y otros objetos.

Algunas de las principales características de las RNC son la complejidad de su operación interna -relacionado con los elementos que los conforman, como capas, neuronas, pesos, sesgos, convoluciones, filtros y funciones de activación-; así como la amplia gama de herramientas disponibles que buscan simplificar su implementación: lenguajes, bibliotecas, *frameworks*, conjuntos de datos y fuentes de información, entre los que se pueden mencionar TensorFlow, Keras, Google Colab, ImageNet, Kaggle, etcétera-.

Las redes más comúnmente utilizadas como Inception, Mobilenet o VGG 19 han sido entrenadas con el conjunto de datos ImageNet. El cual está formado por 1.4 millones de fotografías tomadas de internet, en las que se muestran personas, acciones, objetos, lugares, animales y vegetales clasificados en 20,000 categorías diferentes⁷. Debido a esta característica no sería capaz de diferenciar el proceso fotográfico en imágenes de épocas pasadas. No obstante, es posible adaptarla para que realice esta función mediante un proceso llamado transferencia de aprendizaje, que consiste en reentrenar a la red con ejemplos de las nuevas categorías que deseamos que aprenda a identificar.

La principal ventaja de las redes neuronales es su velocidad y precisión en el reconocimiento de imágenes dentro de un conjunto de categorías delimitadas. Sin embargo, para alcanzar un funcionamiento adecuado requieren de una gran cantidad de imágenes durante su entrenamiento y, dependiendo de la aplicación que se busque abordar, esto puede ser muy costoso o simplemente imposible de realizar. Una crítica frecuente a estos sistemas es la poca transparencia que tienen, tanto en su operación interna como en sus resultados, frecuentemente son llamados "cajas negras" debido a la dificultad que existe para su interpretabilidad. Lo que nos puede llevar a cuestionar su utilidad en los ámbitos de las humanidades y el patrimonio cultural; ámbitos en que la comprensión y el cuestionamiento de los métodos de trabajo son tan importantes como la obtención de resultados.

En un sentido básico, una imagen digital es un conjunto de dígitos (un arreglo) que indican el color de cada uno de los píxeles que forman la imagen. Este

arreglo numérico es el que sirve como dato de entrada para la red neuronal convolucional, compuesta por cientos de neuronas artificiales que están agrupadas en capas con diferentes funciones (convolución, pooling, relu, clasificación). Las cuales, como ocurre en el sistema nervioso humano, reciben información de entrada, la procesan y transforman para después transmitirla como información de salida hacia la capa siguiente.

La operación de convolución, que da nombre a estas redes, es aquella en la que el dígito que corresponden a cada pixel de la imagen se multiplica por un filtro o kernel. Por lo que el resultado es una versión de la imagen en la que se destacan algunas de sus características y se simplifican otras, según el filtro del que se trate. La convolución se repite en diferentes capas de la red, con lo cual progresivamente se va obteniendo una versión más reducida, pero a la vez más destilada de la imagen. La imagen de salida alimenta a la siguiente capa y el proceso se repite varias veces; como resultado final se obtiene un mapa de características que es una especie de huella numérica de la imagen. La cual contiene la suma de las características más importantes que han sido destacadas por los filtros. De esta manera, con el mapa de características, es posible alimentar una última capa conocida como "clasificador", que compara el mapa de características de la imagen de entrada con el mapa de las fotografías con las que fue entrenada la red neuronal. Es de esta manera como la red identifica el contenido visual dentro de las categorías usadas para su entrenamiento.

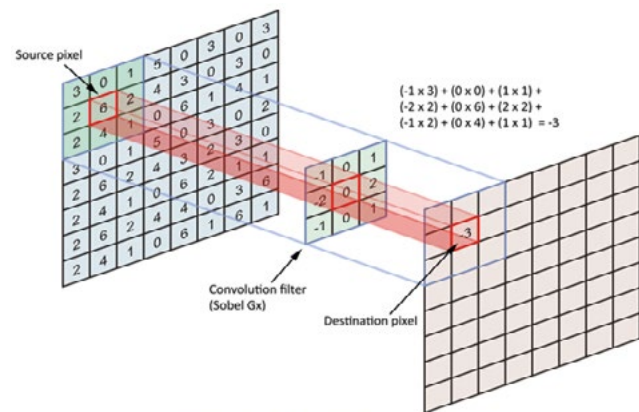


Ilustración 7. Operación de convolución

7 - Melvin Wevers y Thomas Smits, "The Visual Digital Turn: Using Neural Networks to Study Historical Images", *Digital Scholarship in the Humanities* 35, núm. 1 (el 1 de abril de 2020): 194–207, <https://doi.org/10.1093/llc/fqy085>.

Ahora, si se quiere usar la red para identificar categorías diferentes se pueden re-aprovechar la mayoría de las capas que contiene. Ya que, únicamente se debe sustituir la última capa, que es la encargada de clasificar. Esto es posible porque los mapas de características producidos por la red en sus capas internas no representan objetos dentro de la imagen; sino características formales como colores, formas y texturas que se combinan para formar un gran número de variaciones y generalmente se relacionan con el aspecto semántico o de contenido de las imágenes -objetos, lugares y rostros de personas-. Pero estos colores, formas y texturas pueden ser también elementos útiles para identificar el proceso fotográfico de una imagen.

El proceso de re entrenamiento o transferencia de aprendizaje se realizó siguiendo las instrucciones del repositorio Tensorflow for poets⁸ en la versión adaptada por Tim Sherrat en sus colección Glam Workbench⁹. En mi repositorio de Github¹⁰ se puede descargar el conjunto de datos, el cuaderno Jupyter con el código y otras indicaciones para reproducir los resultados que aquí se presentan.

En esencia lo que se hizo fue sustituir el “clasificador” original con uno entrenado con las categorías que nos interesan, utilizando para ello la mayor cantidad posible de ejemplos. En este caso se usaron 600 ejemplos de fotografías divididas en tres procesos fotográficos: ambrotip¹¹, daguerrotipo¹² y ferrotipo¹³. Estas imágenes fueron descargadas de las páginas web de varios archivos bibliotecas y museos y se asume que están correctamente identificadas.

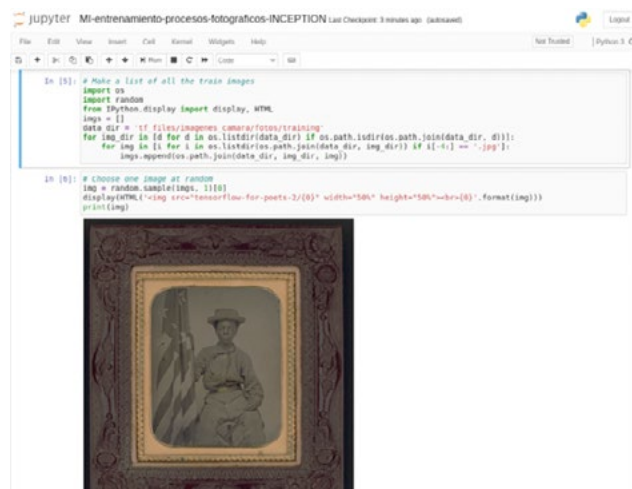


Ilustración 8. Cuaderno Jupyter basado en el tutorial Tensorflow for poets

El resultado de este proceso fue una red con un porcentaje de 85 % de precisión en la fase de entrenamiento y de 75 % en la fase de validación.

Una técnica de visualización de las similitudes entre imágenes basada en la reducción multidimensional¹⁴ nos permite explorar e interpretar la operación de la red. Vemos dos grandes grupos y tres o cuatro conjuntos más pequeños, al parecer los rasgos visuales predominantes en la categorización de las imágenes son los relacionados con la forma de montaje y presentación, así tenemos zonas en las que se ubican las imágenes con estuche, pero sin tapa; imágenes con estuche y tapa con el clásico terciopelo rojo; imágenes montadas en tarjetas; placas desnudas y por último imágenes con marcos negros y un borde dorado.

8 - “googlecodelabs/tensorflow-for-poets-2”, consultado el 26 de julio de 2022, <https://github.com/googlecodelabs/tensorflow-for-poets-2>.

9 - Tim Sherrat, “GLAM-Workbench/Image-Recognition” (Zenodo, el 21 de noviembre de 2019), <https://doi.org/10.5281/zenodo.3549628>.

10 - Gustavo Lozano, “Image Recognition with Tensorflow”, Jupyter Notebook, el 31 de mayo de 2020, <https://github.com/gustavolsj/image-recognition>.

11 - “Ambrotype | The Historic New Orleans Collection”, consultado el 26 de julio de 2022, <https://www.hnoc.org/virtual/daguerreotype-digital/ambrotype>.

12 - “[Portrait of a Seated Woman in Matron Cap] (The J. Paul Getty Museum Collection)”, The J. Paul Getty Museum Collection, consultado el 26 de julio de 2022, <https://www.getty.edu/art/collection/object/104F7Q>.

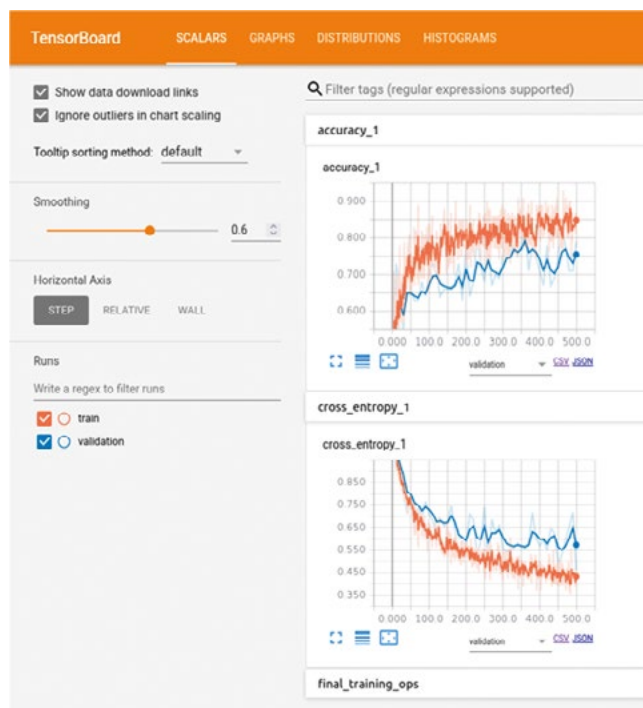
13 - “Tintype | The Historic New Orleans Collection”, consultado el 28 de julio de 2022, <https://www.hnoc.org/virtual/daguerreotype-digital/tintype>.

14 - “Machine Learning for Artists”, Python (2016; repr., ml4a, el 3 de marzo de 2022), https://github.com/ml4a/ml4a/blob/d71b61a99c417b9ace3404420b37d22f6da06153/examples/info_retrieval/image-search.ipynb.



Il·lustració 9. Gràfics de precisió i entropia de la red

El caso más interesante es el de los dos grupos principales ya que parecen contener imágenes muy similares, por lo que no es clara la razón del porqué están divididas, una hipótesis que explicaría este fenómeno es que la red sí considera aspectos visuales más sutiles que corresponden a características físicas relacionadas con el proceso fotográfico ya que si se observa el conjunto de la derecha en él se encuentran varias imágenes que presentan faltantes de la imagen que aparecen como lagunas de color negro, mientras que en el otro grupo no. Este aspecto deberá explorarse a futuro para intentar que la red enfatice este tipo de característica para su clasificación y minimice el impacto de los rasgos visuales relacionados con otros aspectos de las imágenes.



Il·lustració 10. Gràfics de precisió i entropia de la red

Conclusiones

Las dos diferentes técnicas de aprendizaje de máquina que se exploran en este trabajo han dado resultados alentadores que ameritan una investigación más detallada y completa: se trata de un punto de partida sobre el cual es posible plantear un proyecto interdisciplinario a futuro. Para ello, el recurso más importante que habría de sumar es la participación de un especialista en inteligencia artificial, que pueda analizar críticamente los planteamientos iniciales para corregirlos, refinarlos y robustecerlos. La aportación más relevante de este trabajo es el análisis del problema que se busca resolver; ya que en éste se profundiza en aspectos como la motivación, experiencia

y habilidades que caracterizan a los profesionales de los archivos, potenciales usuarios de esta aplicación. Y también se describen las particularidades del problema de la identificación del proceso fotográfico, tanto en lo relacionado con sus aspectos performativos como teóricos.

Se debe considerar que, una eventual propuesta metodológica resultante de este proyecto, deberá guiar y no sustituir al usuario, orientándolo sobre la forma correcta de llevar a cabo la manipulación y observación del objeto fotográfico y ofreciéndole acceso oportuno, pertinente y contextualizado al cuerpo de información necesario en cada etapa de la identificación