

# Detección del ruido ISO por medio de técnicas de Datamining y Redes Neuronales

Pablo Niklas Facultad de Ingeniería  
 Universidad de Palermo  
 Ciudad Autónoma de Buenos Aires  
 Email: pablo.niklas@gmail.com

**Resumen**—En este trabajo se presenta la problemática actual del ruido ISO. Comenzando por un recorrido por la historia de la imagen y sus implicancias sociales. A continuación se presentan los algoritmos de datamining más utilizado y se aplican dos algoritmos a la experimentación. Ante el fracaso de la misma, se recurre a las redes neuronales, donde se explica el concepto de éstas y luego se experimenta entrenándolas y probando la fiabilidad del método intentando detectar el ruido ISO.

**Index Terms**—Fotografía, Ruido ISO, Datamining, Redes Neuronales, Sistemas Expertos.

## I. INTRODUCCIÓN

“Una imagen vale mas que mil palabras”, proverbio chino que invita a reflexionar sobre el valor de la imagen. Ésta capta a tal punto la atención consciente del observador que procura atracción o rechazo, amor u odio, es adictiva, viral, sugestiva, inspiradora de sensaciones. Pero, ¿Qué pasa si el mensaje es interpretado erróneamente por una distorsión involuntaria de la imagen?

En concordancia al proverbio chino, Santiago Ruiz y Luciana Tuquell en su publicación “*Imágenes y palabras en la lucha por imposición de sentidos: la imagen generador del relato*” [1] afirman sobre la necesidad de estudiar las funciones de la imagen. Para los autores se produce una lucha social por imponer unos relatos sobre otros y en esa disputa por el sentido, las imágenes actúan como mediadoras, como generadoras de relatos argumentativos, narrativos o explicativos; y aquí nuevamente la pregunta: ¿Se puede confiar en la imagen pura como medio conciliador de relato o ésta puede verse alterada por *intervenciones mecánicas autónomas*, que nada tienen que ver con la finalidad buscada por su autor?

El deseo del Hombre por retratar momentos considerados importantes, se remonta al comienzo del arte, con su variante rupestre en cuevas de la India, hace ya unos 200.000 o 300.000 años. [2] En ese entonces no había problemas de “intervenciones indeseables”, ya que se pintaba sobre la roca, en cuevas. La imagen ha sido y seguirá siendo un medio comunicacional a través del tiempo.

Durante décadas, lejos de pensar en esas alteraciones como lo que hoy se denomina ruido<sup>1</sup>; era aceptado que la nitidez y calidad de una imagen pudiera verse afectada por una infinidad de razones, a las que no se les otorgaba mayor sentido. En un esfuerzo imaginativo audaz, cada interprete completaba

los datos distorsionados, con aquellos que su propia memoria consciente le otorgaba, suponiendo que los mismos cumplían con la finalidad del autor. Ahora bien, el sentido originario generalmente distaba de una representación certera y a lo largo de los años se volvió un problema.

En la actualidad el Ser Humano, reconoce aquello que afecta la fotografía, sabe diferenciar la imagen real del ruido y hasta generó un sistema de clasificación; en su camino hacia la verdad natural de las cosas procura eliminarlo, sin embargo, para ello necesita un proceso informático que logre su detección certera.

En la sección II, “Historia”, se comenta la historia de la imagen como canal de información, su evolución en la fotografía actual y la influencia negativa del ruido ISO en la misma. En la sección III, “Ruido”, se trata su definición desde un aspecto mas técnico, y a pesar de ser algo “no deseado”, logró una simbiosis en la fotografía “analógica”. En la sección IV, “En Fotografía Digital”, se da una definición mas exacta del ruido que se quiere detectar, sus causas y los distintos tipos. En la sección V, “Algoritmos de Datamining”, se define el termino “Datamining” y presentan los distintos algoritmos del mencionado campo. En la sección VI, “Experimentación con técnicas de datamining”, se presentan los pasos que se utilizan en la experimentación, como se procesa el archivo fotográfico, y la selección de la herramienta de datamining respectiva para su procesamiento, y para finalizar el resultado de la experimentación. En la sección VII, “Conclusión de la hipótesis inicial”, se deja asentado que la experimentación por medio de técnicas de datamining, fracasaron o no fueron concluyentes, ya que la densidad de ruido, hace que su detección por medio de los algoritmos utilizados, no fueron satisfactorias. En la sección VIII, “Redes Neuronales”, se presentan el concepto de ANS o redes neuronales artificiales, sus diferencias con las redes neuronales naturales, y la lógica difusa. En la sección IX, “Experimentación con redes neuronales”, se describe la experimentación de detección por medio de ANS, y se presenta la ANS utilizada, la CNN, red neural convulacional. La utilización del lenguaje Python y la biblioteca Keras para entrenar una red neural y luego poder verificar que la red detecta las imágenes que tienen ruido ISO. Se describen los distintos pasos de la experimentación. En la sección X, “Conclusión”, se realiza la conclusión final del trabajo, donde queda plasmado que con ninguno de los métodos utilizados, es posible la detección del ruido ISO.

<sup>1</sup>En semiología, se define al ruido como “interferencia que afecta a un proceso de comunicación”.

## II. HISTORIA

Como se mencionó en la sección anterior, la necesidad de generar y almacenar imágenes, se remonta a tiempos muy lejanos y con el correr de los milenios su evolución deriva en la fotografía. Su nombre del griego foto (luz) y grafos (escritura), hace que muchos la llamen “el arte de escribir o pintar con luz”. Su historia se remonta a más de dos siglos atrás, en 1839 mediante el nacimiento del daguerrotipo [3], como ejemplo se encuentra la figura 1. Este procedimiento fotográfico, de origen Francés, satisfizo la necesidad de retratar personas o eventos para la posteridad.



Figura 1: Daguerrotipo del Almirante Dn. Guillermo Brown.

Este método, utilizaba las sales de plata, que recubrían una placa de cobre, como material fotosensible; al exponerla a la luz, la fotografía se revelaba con vapores de mercurio que daban como resultado una imagen perfecta, la cual debía resguardarse de la abrasión del aire, mediante la protección de un cristal. Los haluros de plata<sup>2</sup> eran la mínima unidad de información que se disponía sobre la placa fotosensible, equivalente al *pixel*<sup>3</sup> de la fotografía digital actual. Sin embargo el daguerrotipo acabo desapareciendo por su impracticabilidad explícita: producía una única imagen sin la posibilidad de sacar copias sin destruir el original, el tiempo de exposición era muy largo, la placa de cobre emulsionada con sales de plata se craqueaba fácilmente en contacto con el agua o el aire, lo que la volvía extremadamente frágil, era muy pesado, de difícil manipulación y altamente toxico.

Unps años más tarde del advenimiento del daguerrotipo, se presenta a la sociedad el calotipo, procedimiento fotográfico

<sup>2</sup>Se denomina haluro de plata a un compuesto químico formado entre la plata y uno de los halógenos — bromuro de plata (AgBr), cloruro de plata (AgCl), yoduro de plata (AgI), y tres formas de fluoruros de plata. A menudo como grupo son referidos como los haluros de plata, y suele utilizarse la notación pseudo-química AgX.

<sup>3</sup>Un píxel o pixel, plural píxeles (acrónimo del inglés *picture element*, “elemento de imagen”), es la menor unidad homogénea en color que forma parte de una imagen digital.

perfeccionado por William Henry Fox Talbot, en la figura 2, que consistía en utilizar papel negativo, el cuál permitía realizar un número ilimitado de copias a partir de un único negativo; una vez que se concluía el proceso de revelado, la imagen lograda se sumergía en hiposulfito sódico para fijarla y hacerla permanente con un tiempo de exposición muy breve, alrededor de 30 segundos.



Figura 2: William Henry Fox Talbot, ca. 1864

Alrededor de los años 50, el calotipo da paso al colodión<sup>4</sup>, figura 3 y con él, la posibilidad de una imagen instantánea, que abre la puerta a la fotografía callejera y una nueva actividad comercial. El colodión, conocido también como algodón-pólvora, es una clase de explosivo cuya base es la celulosa nítrica. Estos negativos debían ser expuestos y revelados mientras estaban húmedos, por tal razón los fotógrafos necesitaban de un cuarto oscuro cercano; si bien fue un paso importante para acercarse a la fotografía instantánea, todavía distaba mucho de ser aquella que actualmente se conoce como tal. [3]



Figura 3: Ejemplo de colodión húmedo

Sin bien, muchos reconocen que en la actualidad todavía se utiliza el colodión como expresión artística, en 1882,

<sup>4</sup>También conocido como colodión húmedo



Figura 4: Fotografía con sensibilidad ISO 1600.



Figura 5: Fotografía con sensibilidad ISO 25.

cede su espacio protagónico a la imagen con gelatina de bromuro; de esta forma y con la vigencia revolucionaria del colodión y su posibilidad de sacar copias, la fotografía fue evolucionando constantemente, muchos fueron los ensayos y vastas las pruebas hasta llegar a nuestros días.

En 1923, aparece en el mercado una cámara de 35 mm, ligera, versátil, denominada Leica. Esta y sus competidoras en el mercado, hicieron de la fotografía, un arte accesible para todos. En 1936 se implementa el uso de la película de color y finalmente en 1947 con la conocida cámara Polaroid Land, se hace realidad la fotografía instantánea. Ya de la mano de la fotografía tradicional, mucho no paso para llegar a la fotografía digital, pero esa es una historia que ya se conoce.

### III. RUIDO

En la electrónica, el ruido es “una señal no deseada”. No es posible eliminarlo completamente, aunque gracias al avance tecnológico y a procesos actualizados de diseño de circuitos, se logra mitigar. Con el correr del tiempo, el ruido se ha manifestado en diferentes intensidades y ha evolucionado al ritmo de la tecnología. Tal es así, que paradójicamente existen sensores, que facilitan su “creación” y otros que hasta ayudan a detectarlo y permiten configuraciones que logran minimizarlo.

Poco se habla sobre el ruido en la imagen *analógica*, pero si nos remontamos un peldaño hacia atrás en el tiempo, podemos observar que la fotografía tradicional ideó una especie de simbiosis con el ruido. Si bien, estamos haciendo mención a un proceso fotográfico contemporáneo, de conocimiento masivo, ha derivado para nuestra época en obsoleto ya que lo digital prima sobre cualquier procedimiento que se le oponga. Esta clase de fotografía analógica-tradicional, requiere de granos más grandes de material fotosensible a los efectos de asegurar la alta sensibilidad de la misma. Cuando el grano se vuelve altamente visible, estamos ante la presencia de ruido. La ventaja de este tipo de fotografía, contrariamente a la digital, es que el mayor tamaño de gránulo de material fotosensible, queda estéticamente bien. En la figura 5 se pueden ver en mayor detalle los haluros, y la figura 4, se pueden ver los

haluros en menor tamaño, pero tomando en cuenta que el acercamiento distorsiona un poco, vale tener en cuenta que los haluros en este caso, son de mayor tamaño, debido a la mayor sensibilidad.

En la fotografía digital, el paradigma es más complejo y el ruido adquiere en ella su propia evolución, si bien la primera imagen puramente digital, data de 1957 con Russell Kirsch<sup>5</sup>, la trascendencia del ruido digital obtiene preponderancia años más tarde y su detección sigue siendo hoy, una asignatura pendiente.

### IV. EN FOTOGRAFÍA DIGITAL

El ruido digital es la variación aleatoria (que no se corresponde con la realidad) del brillo o el color en las imágenes digitales producido por el dispositivo de entrada (la cámara digital en este caso).

Está universalmente aceptado que, así como, en la imagen analógica el grano es tolerable e incluso estético, en la fotografía digital el ruido es anti-estético e indeseable como se aprecia en el ejemplo de la figura 6.

Salvando las distancias, los dos conceptos son similares: en la fotografía analógica el grano aparece cuando los puntos que forman la imagen (granos de material fotosensible) adquieren tanto tamaño que son apreciables a simple vista. Las variables que influyen en la aparición de grano son la marca de la película, la sensibilidad (a más sensibilidad, más grano) y el proceso de revelado. Evidentemente la cantidad de ampliación de la copia también aumenta el tamaño del grano. En la fotografía digital, el ruido es ese efecto “indeseable” que consiste en la aparición aleatoria de señales ajenas a la imagen original, especialmente apreciable en las zonas de sombra de la imagen. [4]. Muchas veces el ruido digital refleja una mala captura, un mal equipo o un mal procesado.

#### IV-A. Causas

Para comprender los motivos que originan el ruido digital se debe fijar la atención en el funcionamiento del sensor

<sup>5</sup>Russell A. Kirsch (nacido en 1929) es un ex ingeniero estadounidense de la Oficina Nacional de Estándares que desarrolló el primer escáner de imágenes digitales.

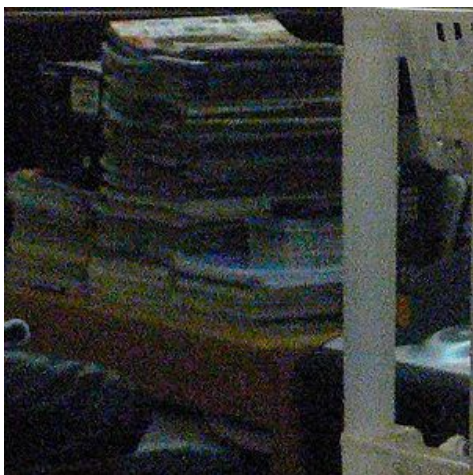


Figura 6: Ejemplo de ruido en la fotografía digital.

de imagen, que hace posible la captura de la foto (CCD<sup>6</sup>). Un CCD, mostrado en la figura 7, está compuesto por una matriz de celdas fotosensibles, encargadas de recibir la imagen formada por las lentes que conforman la óptica de la cámara. Cada una de esas celdas contendrá uno o más fotodiodos con capacidad de convertir la luz en una señal eléctrica, para ser procesada como una imagen (color), es decir, será convertida a datos numéricos que se almacenarán en forma binaria en la memoria de la cámara para dar origen a un pixel. Lo cierto es que esa actividad eléctrica, en sí misma, generará una cierta señal aún en ausencia de luz, que oscilará en relación con la temperatura, generando datos al azar, que originarán ruido.

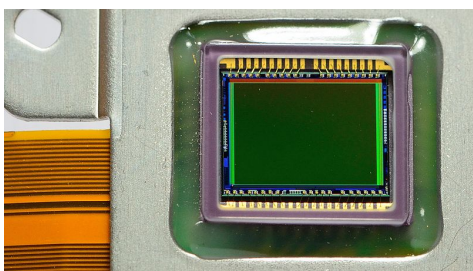


Figura 7: Sensor CCD Canon SX30

Si se realiza una definición comparativa, se puede decir que el término ruido, es apropiado para definir esos puntos de color dispersos en las fotografías digitales, similares al “soplido de fondo” que se escucha en un equipo de música cuando se eleva el volumen y no se pone ninguna fuente de sonido. El soplido que aumenta según se incrementa el volumen es consecuencia del propio circuito electrónico (agitación de los electrones en las junturas de los transistores, debido a la polarización de éstas y a la temperatura). Aplicado a la fotografía digital, cuando se incrementa la sensibilidad del CCD, aumentado el valor de ISO se estará haciendo algo parecido a subir el volumen, incrementado la actividad eléctrica, forzando el sensor de imagen, con el consiguiente incremento de ruido

<sup>6</sup>Un dispositivo de carga acoplada (en inglés charge-coupled device) es un circuito integrado que contiene un número determinado de condensadores enlazados o acoplados.

en las fotografías. De hecho, la sensibilidad de cada uno de los pixeles del CCD es fija, correspondiéndose al valor más bajo de ISO que ofrece la cámara (generalmente 100 ISO), aumentar la sensibilidad (200, 400, 800 ISO o más) no se logra por el incremento de la sensibilidad de los fotones, sino a través de la amplificación posterior de la señal que estos emiten.

A vista de lo expuesto, a la hora de capturar una señal luminosa débil (sombras, fotos nocturnas, etc), en el caso de amplificarla al hacer uso de un valor alto de ISO, estamos ampliando igualmente los datos aleatorios (puntos de color dispersos por la toma) fruto de la actividad eléctrica del sensor; es decir, aumentaremos el ruido de la fotografía, que al mismo tiempo se verá incrementado con los tiempos largos de exposición (más de 2 segundos) y altas temperaturas.

Otro factor a tener en cuenta es que el ruido también aumenta cuando el tamaño del pixel disminuye. La dificultad de interpretación, consecuencia de la reducción del tamaño del pixel, por la interacción de los propios píxeles y su extrema cercanía, generará igualmente datos aleatorios que se verán incrementados enormemente en el caso de amplificar la señal. Esto explica porque los modelos de cámaras réflex, al contar con un sensor de pixeles de gran tamaño (sensiblemente mayores que los pixeles encontrados en las cámaras compactas u hogareñas), generen un nivel de ruido mucho menor que el obtenido en modelos compactos, soportando igualmente valores de ISO muy elevados (800, 1600...) con un nivel de ruido muy bajo en relación con el que generaría una cámara hogareña.

#### IV-B. Tipos de Ruido

**IV-B1. Ruido Impulsional o “Sal y Pimienta”:** Este tipo de ruido conocido como ruido sal y pimienta, se caracteriza porque los píxeles de la imagen son muy diferentes en color o intensidad a los píxeles circundantes. El hecho que define este tipo de ruido es que el pixel ruidoso propiamente dicho, no tiene relación alguna con los píxeles circundantes. Generalmente, este tipo de ruido, afectará a una pequeña cantidad de píxeles de la imagen. Al ver la imagen, se encontrarán puntos blancos sobre puntos negros o puntos negros sobre puntos blancos, de ahí el término sal y pimienta. Defectos que contribuyen a este tipo de ruido son las manchas de polvo en la óptica de la cámara, o bien un CCD defectuoso, que realizará una captura errónea.

**IV-B2. Ruido Gaussiano:** En el ruido de tipo gaussiano<sup>7</sup>, todos y cada uno de los píxeles que componen la imagen cambian su valor, de acuerdo con una distribución normal o gaussiana, tal cual se aprecia en la figura 9. Se podría aplicar otro tipo de distribución, pero la distribución de Gauss, se toma como el modelo que más se aproxima, debido al teorema

<sup>7</sup>Johann Carl Friedrich Gauß (Brunswick; 30 de abril de 1777 - Gotinga; 23 de febrero de 1855) fue un matemático, astrónomo, geodesta y físico alemán que contribuyó significativamente en muchos campos, incluida la teoría de números, el análisis matemático, la geometría diferencial, la estadística, el álgebra, la geodesia, el magnetismo y la óptica.





Figura 8: Imagen con ruido tipo “Sal y Pimienta”.

central del límite<sup>8</sup>, que dice que la suma de los diferentes ruidos tiende a aproximarse a una distribución normal o gaussiana. [5] [6]

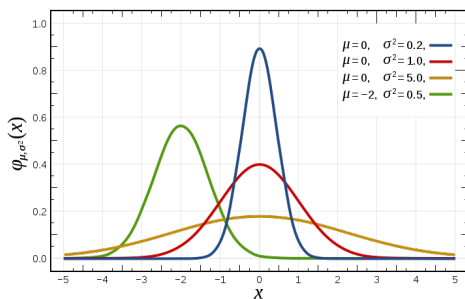


Figura 9: Curvas de distribución normal o gaussiana

**IV-B3. Ruido de disparo:** Es un tipo de ruido eléctrico que tiene lugar cuando el número finito de partículas que transportan energía, tales como los fotones en un dispositivo óptico, es suficientemente pequeño como para generar la aparición de fluctuaciones estadísticas apreciables en una medición. Se sabe que cuanto mayor sea el valor promedio de la intensidad luminosa, mayor será el nivel de este ruido, figura 10. Sin embargo, el nivel de señal crece con mayor rapidez cuanto mayor es su nivel promedio; aún así, el ruido de disparo sólo supone un problema cuando se trabaja con intensidades luminosas bajas.

**IV-B4. Ruido de luminancia:** En Fotometría, la luminancia se define como la densidad angular, rectangular y superficial

<sup>8</sup>El teorema del límite central o teorema central del límite indica que, en condiciones muy generales, si  $S_n$  es la suma de  $n$  variables aleatorias independientes y de varianzas no nula pero finita, entonces la función de distribución de  $S_n$  “se aproxima bien” a una distribución normal. Así pues, el teorema asegura que esto ocurre cuando la suma de estas variables aleatorias e independientes es lo suficientemente grande.

de flujo luminoso que incide, atraviesa o emerge de una superficie siguiendo una dirección determinada. Alternativamente, también se puede definir como la densidad superficial de intensidad luminosa en una dirección dada.

El ruido de luminancia aparece como variaciones de luminosidad en la imagen, se nota sobre todo en superficies que deberían tener un color uniforme.

**IV-B5. Ruido de crominancia:** La crominancia, es la señal que en los sistemas de vídeo transporta la información de color de la imagen, separadamente de la señal luma o señal  $Y$ . La crominancia es generalmente representada por dos componentes de diferencia de color,  $B - Y$  (azul-luma) y  $R - Y$  (rojo-luma). Cada una de estas señales puede tener factores de escala o corrimientos, como se especifica en las normas de vídeo aplicables. Los sistemas digitales de vídeo y fotografía a veces usan una descomposición luma-crominancia para una mejor compresión. Por ejemplo, cuando una imagen digital común RGB se comprime mediante la norma JPEG, el espacio de color RGB se convierte primero, mediante una matriz de rotación, en un espacio de color  $YCbCr$ , debido a que los tres componentes en el espacio tienen menos redundancia de correlación y porque las componentes de crominancia pueden ser submuestreadas en un factor de 2 o 4 para comprimir aún más la imagen. En la descompresión, el espacio  $Y'CbCr$  se reconvierte a RGB.

El ruido de crominancia aparece como pequeños puntos de colores de manera aleatoria.

Es el ruido más fácil de eliminar con un software de filtrado anti-ruido, pero se pierde fidelidad en los colores.

**IV-B6. Ruido Termal o de corriente oscura:** Es el producido por el calor de la cámara (al ser radiación electromagnética puede mostrarse en el sensor). De fácil detección, todo aquello que no es parte de la escena, es reconocido como ruido. Se puede reducir enfriando el sensor o limitando el tiempo de exposición. El ruido termal genera ruido blanco, figura 11.

**IV-B7. Ruido de lectura:** Se denomina así, al ruido eléctrico generado por los circuitos que leen los valores del sensor de píxeles.

**IV-B8. Ruido de cuantificación:** Esta clase de ruido esta dada por los denominados errores de redondeo, los cuales tienen lugar cuando la señal analógica es convertida mediante un ajuste finito de valores discretos digitales. No se aprecia comúnmente, se puede reducir usando un sensor que almacene mas bits por pixel.

#### IV-C. Reducción del ruido

**IV-C1. Compromiso:** Cuando se elige un algoritmo de reducción de ruido, se deben sopesar varios factores: el de la capacidad de cómputo disponible y el tiempo a partir del cual no resulta rentable su tratamiento: una cámara digital debe aplicar la reducción del ruido en fracciones de segundos, utilizando para ello una pequeña CPU empotrada en la misma; mientras que un ordenador de escritorio tiene mucha mayor capacidad de cómputo y tiempo. También hay que tener en cuenta si el sacrificio de algunos detalles de la imagen real es aceptable ante la posibilidad de eliminar mayor cantidad de ruido (la agresividad con que decidir si las variaciones en la imagen son el ruido o no).

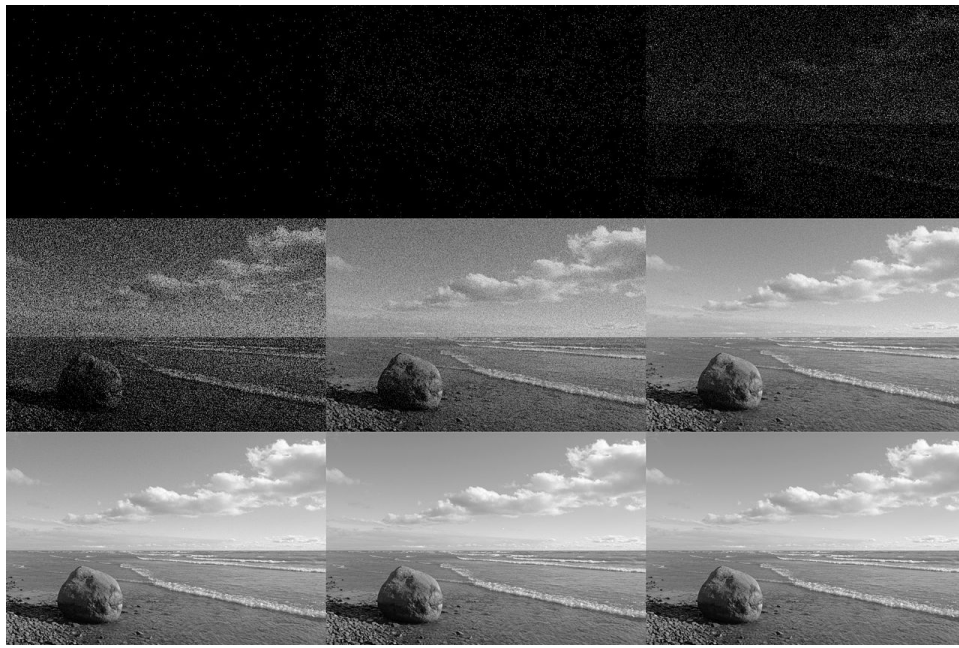


Figura 10: La cantidad de fotones por pixel aumenta de izquierda a derecha y de arriba hacia abajo.

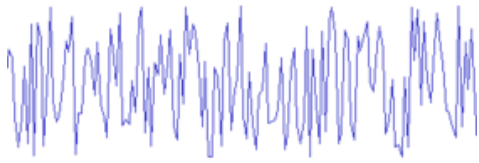


Figura 11: Gráfica del ruido blanco en función del tiempo.

**IV-C2. Ruido de crominancia y ruido de luminancia:** En las fotografías del mundo real, el mayor nivel de detalle consiste principalmente en las variaciones del brillo (luminancia), más que en las variaciones de tono (crominancia).

Dado que cualquier algoritmo de reducción debe tratar de eliminar el ruido sin sacrificar los detalles de la escena fotografiada, se corre el riesgo de una mayor pérdida de detalle, ante la reducción de ruido de luminancia que en la reducción de ruido de crominancia.

Más aún, si se toma en consideración que desde una perspectiva óptica, el ruido cromático en las imágenes, resulta más desagradable que el ruido de luminancia; es doble entender que se traten de forma separada. Las manchas de color se consideran como algo artificial, en desmedro de la imagen real, en comparación con la apariencia granulada de ruido de luminancia, el cuál resulta comparable con el grano de una película. Por estas dos razones, la mayoría de los algoritmos de reducción de ruido fotográfico, dividen la imagen en las componentes de crominancia y luminancia.

La mayoría de programas informáticos dedicados de reducción de ruido permiten al usuario controlar la reducción del ruido de crominancia y de luminancia por separado.

**IV-C3. Filtros lineales:** Un método para eliminar el ruido es por convolución<sup>9</sup> de la imagen original con una máscara que representa un filtro de paso bajo o la operación de suavizado. Por ejemplo, la máscara de Gauss está compuesta por elementos determinados por una función de Gauss. Esta convolución trae el valor de cada píxel más parecido en valor con sus vecinos. En general, en un filtro de suavizado de píxeles cada pixel establece cual es el valor promedio o la media ponderada de sí mismo y sus vecinos cercanos. Los filtros de suavizado, tienden a difuminar la imagen, porque los valores de los píxeles que son significativamente más altos o más bajos, respecto al valor de sus vecinos, tienden a “manchar” toda la zona, provocando el difuminado de los bordes, por lo que tienen que ser utilizados con cierta precaución.

Otro filtro utilizado es el denominado, de media, el mismo trabaja de la siguiente forma: dada una imagen  $f(i, j)$ , se procede a generar una nueva imagen  $g(i, j)$  cuya intensidad para cada píxel se obtiene promediando los valores de intensidad de los píxeles  $f(i, j)$  incluidos en un entorno de vecindad predefinido.

**IV-C4. Filtros no lineales:** Un ejemplo por excelencia de filtros no lineales es el filtro de mediana<sup>10</sup> el cuál resulta muy bueno para conservar detalles de la imagen. Para aplicar un filtro de mediana se ha de considerar cada píxel de la imagen, ordenar los píxeles vecinos de acuerdo a sus intensidades y sustituir el valor original del píxel por la mediana del valor de la lista.

<sup>9</sup>En matemáticas y, en particular, análisis funcional, una convolución es un operador matemático que transforma dos funciones  $f$  y  $g$  en una tercera función que en cierto sentido representa la magnitud en la que se superponen  $f$  y una versión trasladada e invertida de  $g$ .

<sup>10</sup>En el ámbito de la estadística, la mediana representa el valor de la variable de posición central en un conjunto de datos ordenados.

*IV-C5. Filtro gaussiano:* En esta clase de filtro, el valor de cada punto es el resultado de promediar con distintos pesos los valores vecinos a ambos lados de dicho punto. Este tipo de filtro también tiene el problema del difuminado de los bordes, pero no es tan acusado como el caso de la media simple. Este tipo de filtro reduce especialmente el ruido tipo gaussiano.

## V. ALGORITMOS DE DATAMINING

Datamining (minería de datos) es un término acuñado por el economista *Michael Lovell* en 1983. Si bien había expresiones léxicas similares precisadas 20 años antes [7], se puede definir el vocablo datamining como el campo interdisciplinario en la intersección de la inteligencia artificial<sup>11</sup>, aprendizaje automático o *machine learning*<sup>12</sup>, estadística y los sistemas de base de datos [8].

Una de las herramientas, comúnmente utilizadas en la minería de datos, son los sistemas que construyen clasificadores. Tales sistemas toman para análisis un conjunto de casos, cada uno perteneciente a una clase, la cual a su vez, integra de un pequeño número de clases y que se describe por sus valores, para un conjunto fijo de atributos. La salida de un clasificador podría predecir con precisión la clase a la que pertenecería un nuevo caso.

A continuación se lista una muestra de los algoritmos más conocidos: [9]

### V-A. Árboles de decisión

Un árbol de decisión es una forma gráfica y analítica de representar todos los eventos (sucesos) que pueden surgir a partir de una decisión asumida en cierto momento. Ayudan a tomar la decisión más “acertada”, desde un punto de vista probabilístico, ante un abanico de posibles decisiones. Estos árboles permiten examinar los resultados y determinar visualmente cómo fluye el modelo. Los resultados visuales ayudan a buscar subgrupos específicos y relaciones que tal vez se encontrarían con estadísticas más tradicionales. Los árboles de decisión son una técnica estadística para la segmentación, la estratificación, la predicción, la reducción de datos y el filtrado de variables, la identificación de interacciones, la fusión de categorías y la discretización de variables continuas. [10]

### V-B. Clasificadores por conjunto de reglas

Las reglas de asociación se utilizan para descubrir hechos que ocurren en comúnmente dentro de un determinado conjunto de datos.

### V-C. KMeans

K-means es un método de agrupamiento, que tiene como objetivo la partición de un conjunto de  $n$  observaciones en  $k$  grupos en el que cada observación pertenece al grupo cuyo valor medio es más cercano. El algoritmo requiere que se le indique la cantidad de grupos o *clusters*, figura 12.

<sup>11</sup>Es una rama de las ciencias computacionales encargada de estudiar modelos de cómputo capaces de realizar actividades propias de los seres humanos en base a dos de sus características primordiales: el razonamiento y la conducta.

<sup>12</sup>Es el subcampo de las ciencias de la computación y una rama de la inteligencia artificial, cuyo objetivo es desarrollar técnicas que permitan que las computadoras aprendan.

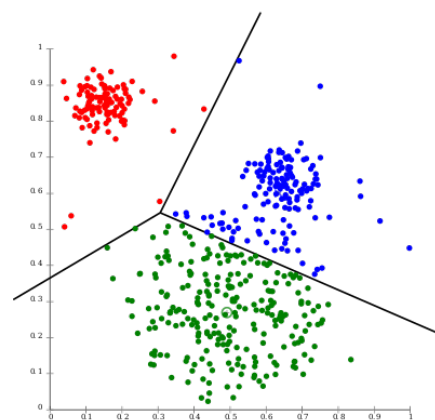


Figura 12: Kmeans con  $k = 3$

### V-D. Máquinas de vectores de soporte

Dado un conjunto de puntos, subconjunto de un conjunto mayor (espacio), en el que cada uno de ellos pertenece a una de dos posibles categorías, un algoritmo basado en SVM construye un modelo capaz de predecir si un punto nuevo (cuya categoría desconocemos) pertenece a una categoría o a la otra. Es similar a los árboles de decisión en la funcionalidad provista.

### V-E. EM

El algoritmo esperanza-maximización o algoritmo EM se usa en estadística para encontrar estimadores de máxima verosimilitud de parámetros en modelos probabilísticos que dependen de variables no observables.

## VI. EXPERIMENTACIÓN CON TÉCNICAS DE DATAMINING

### VI-A. Introducción

Para verificar la fiabilidad de las técnicas de datamining para la reducción del ruido ISO, se debe contar con un conjunto de datos para ser procesado. La foto, en la figura 13, tiene que estar en formato BMP<sup>13</sup>, es decir sin compresión alguna, ya que no solo distorsiona la información, sino que también poseen campos EXIF<sup>14</sup> agrega metadatos propios del formato de la imagen (geolocalización, apertura y velocidad de diafragma, fecha y hora de la toma, sensibilidad ISO, entre otros).

<sup>13</sup>Los archivos de mapas de bits se componen de direcciones asociadas a códigos de color, uno para cada cuadro en una matriz de píxeles tal como se esquematizaría un dibujo de “colorea los cuadros” para niños pequeños. Normalmente, se caracterizan por ser muy poco eficientes en su uso de espacio en disco, pero pueden mostrar un buen nivel de calidad. A diferencia de los gráficos vectoriales al ser reescalados a un tamaño mayor, pierden calidad. Los archivos BMP no son utilizados en páginas web debido a su gran tamaño en relación a su resolución.

<sup>14</sup>Es una especificación para formatos de archivos de imagen usado por las cámaras digitales. Fue creado por la Japan Electronic Industry Development Association (JEIDA). La especificación usa los formatos de archivos existentes como JPEG, TIFF Rev. 6.0, y RIFF el formato de archivo de audio WAVE, a los que se agrega tags específicos de metadatos. No está soportado en JPEG 2000 o PNG.





Figura 13: Foto usada en la experimentación

Cuadro I: Valores RGB de los píxeles

Xcoor	Ycoor	Red	Green	Blue
0	0	201	172	116
0	1	197	161	139
0	2	209	165	138
0	3	202	174	134
0	4	240	180	126
0	5	198	162	74
0	6	192	163	121
0	7	185	164	147
0	8	185	125	124

## VI-B. Generación del conjunto de datos

La generación de datos se realiza con un pequeño programa, cuyo código puede leerse en el Listado 1, hecho en lenguaje Python<sup>15</sup>, que leyendo el archivo en formato BMP, genera un archivo de texto en formato CSV<sup>16</sup> con cinco columnas, las primeras dos consisten en las coordenadas ( $x, y$ ) y las siguientes tres, el color del pixel en codificación RGB<sup>17</sup> ( $r, g, b$ ). Estas últimas columnas almacenarán valores sin signo de 8 bits<sup>18</sup>.

Una vista del contenido del archivo generado puede verse en el cuadro II.

<sup>15</sup>Python es un lenguaje de programación interpretado cuya filosofía hace hincapié en una sintaxis que favorezca un código legible.

<sup>16</sup>CSV (sigla en inglés de Comma Separated Values, en español “Valores separados por comas”). Es un formato tradicional y bastante conocido para el intercambio de información.

<sup>17</sup>RGB (sigla en inglés de red, green, blue, en español “rojo, verde y azul”) es la composición del color en términos de la intensidad de los colores primarios de la luz.

<sup>18</sup>Bit es el acrónimo de Binary digIT (o sea “dígito binario”, en español denominado como bit, y en menor medida como bitio). Es un dígito del sistema de numeración binario

Listado 1: Listado del generador del conjunto de datos

```
#!/usr/bin/env python

from __future__ import with_statement
from PIL import Image

im = Image.open('foto.bmp')

pix = im.load()

width, height = im.size

with open('foto.csv', 'w+') as f:
    f.write('Xcoor,Ycoor,Red,Green,Blue\n')

    for x in range(width):
        for y in range(height):
            r = pix[x, y][0]
            g = pix[x, y][1]
            b = pix[x, y][2]

            f.write('{0},{1},{2},{3},{4}\n'.format(
                x, y, r, g, b))
```

## VI-C. Comienzo del trabajo

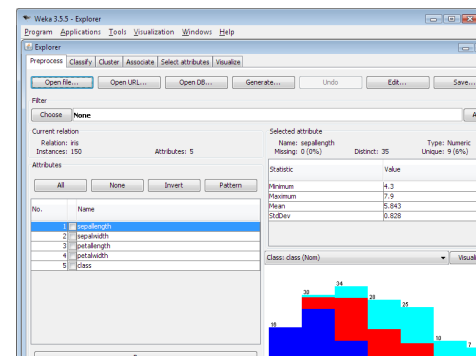


Figura 14: Pantalla de WEKA mostrando clases de atributos.

## VI-D. Hipótesis inicial de trabajo

La hipótesis de este trabajo consiste en aprovechar los algoritmos de datamining que son utilizados para detectar *outliners*, es decir, valores fuera del promedio.

Habiendo ya explicado en que consiste el ruido ISO, el patrón en común para los casos listados, radica en un cambio drástico de color (RGB) del pixel del ruido, respecto al resto circundante.

En consecuencia, el enfoque es utilizar un sistema de datamining para la detección de los *outliners*.





Figura 15: Canon 7D Mark II

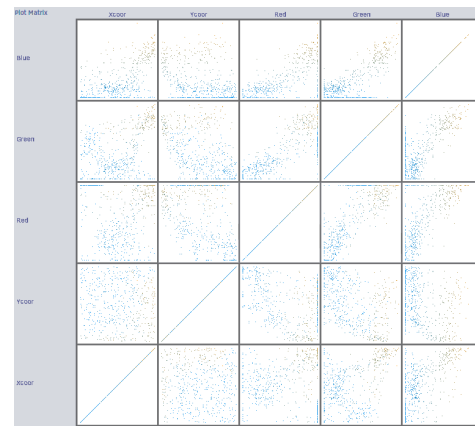


Figura 16: Matriz de los atributos.

#### VI-E. Selección de Herramientas

Para la realización de los experimentos se seleccionaron las siguientes herramientas:

- Herramienta Datamining: W.E.K.A.<sup>19</sup> [11]. Figura 14.
- Cámara de fotos: Canon 7D Mark II. Figura 15. La cámara utilizada, del tipo APS-C<sup>20</sup>, se seleccionó debido por ser un modelo de la gama media alta de la marca Canon<sup>TM</sup>, lo que permite que las imágenes obtenidas sean aptas para la realización del experimento.

El criterio de selección del software, fue en orden:

- **Costo cero:** En base a una metodología de planeamiento y presupuesto, se seleccionaron los recursos económicos necesarios, para lograr los resultados deseados, sin inversiones significativas.
- **Usabilidad:** Se buscó un conjunto de herramientas de fácil accesibilidad, caracterizados por la sencillez operacional necesaria para el tiempo de experimentación.
- **Capacidad de multiprocesamiento**<sup>21</sup> Se buscaron algoritmos con capacidad operativa alta, que permitieran la ejecución de varios procesos en simultáneo, a fin de optimizar los tiempos de prueba.

<sup>19</sup>Wakato Environment for Knowledge Analysis. es una plataforma de software para el aprendizaje automático y la minería de datos escrito en Java y desarrollado en la Universidad de Waikato. WEKA es software libre distribuido bajo la licencia GNU-GPL.

<sup>20</sup>En fotografía digital, APS-C (Advanced Photo System type-C, "Sistema Avanzado de fotografía tipo C") hace referencia a un formato de sensor de imagen.

<sup>21</sup>En el caso de WEKA, el multiprocesamiento es posible por medio de la instalación del paquete *wekaServer*.

#### VI-F. Pasos de la experimentación

Una vez realizada la selección de herramientas, se procede a realizar los siguientes pasos:

1. A los efectos de mantener la información de la imagen tal cual fue captada por el sensor CCD de la cámara, la toma se realiza en formato RAW<sup>22</sup>. La foto utilizada en el experimento, figura 13, es de una resolución de 927x1088 píxeles.
2. Para el revelado digital, se utiliza el software Darktable<sup>23</sup>.
3. Luego del procesamiento, se exporta a formatos TIFF, debido a que Darktable, no permite exportar a BMP.
4. Con el programa GIMP<sup>24</sup>, se abre el TIFF y se convierte a BMP.
5. Se procesa el BMP, con el programa del listado 1, para obtener el archivo CSV con la información RGB de cada uno de los píxeles de la foto.
6. Este archivo CSV es importado por WEKA y grabado como archivo ARFF (Attribute-Relation File Format). Una vez importado con WEKA se puede visualizar la matriz de los atributos. Figura 16.
7. Para la validación, se compara la información generada por el algoritmo de datamining, contra la foto aumentada al 9000 %. Figura 17. De esta forma se puede ubicar fácilmente el píxel informado por WEKA y decrementando el zoom, validar el resultado.

<sup>22</sup>En el formato RAW la cámara no realizará postprocesado alguno, ya que se limitará a almacenar la información en la memoria. Luego de la toma fotográfica, el formato RAW debe ser "revelado digitalmente".

<sup>23</sup>Es un programa de procesamiento fotográfico en formato raw de código abierto. En vez de ser un editor de gráficos rasterizados como Adobe Photoshop, trabaja con un conjunto de herramientas específicamente enfocadas al procesamiento y postproducción de imágenes raw no destructivo y está principalmente enfocado a mejorar el proceso de trabajo del fotógrafo facilitando la gestión de grandes cantidades de imágenes. Está disponible libremente en versiones para las principales distribuciones Linux, Windows, OS X y Solaris bajo la licencia GPL versión 3 o posterior.

<sup>24</sup>GIMP (GNU Image Manipulation Program) es un programa de edición de imágenes digitales en forma de mapa de bits, tanto dibujos como fotografías. Es un programa libre y gratuito. Forma parte del proyecto GNU y está disponible bajo la Licencia pública general de GNU y GNU Lesser General Public License

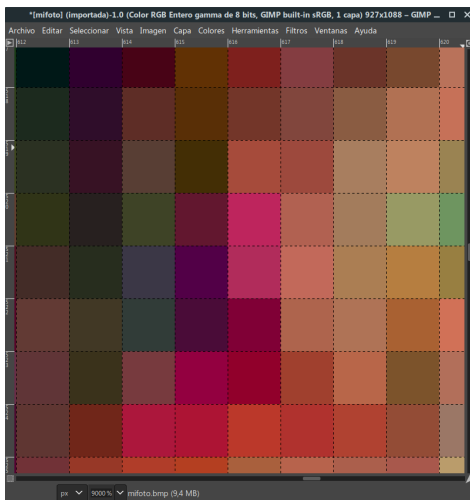


Figura 17: Zoom al 9000 % para la validación.

#### VI-G. KMean

La visualización de los 5 clusters, se ve en la figura 18. Nótese la similitud con la foto original, aunque invertida verticalmente.

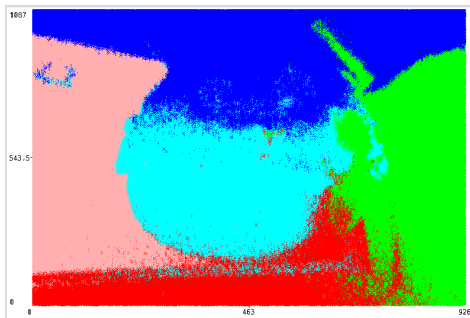


Figura 18: Visualización de los 5 clusters

#### VI-H. EM

Este algoritmo es similar a KMeans, con la única diferencia que el número de clústers se detecta automáticamente en función de las características de la población. Figura 19. El proceso tomó 9579.89 segundos.

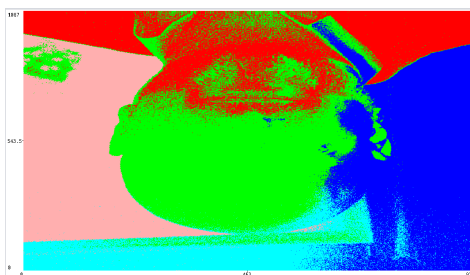


Figura 19: Clusters detectados con EM

Cuadro II: Comparación entre ANS y el cerebro [13]

Item	Cerebro	Computadora
Velocidad de proceso	$10^{-2}$ seg. (10Hz)	$10^{-9}$ seg. (1000Hz)
Estilo de procesamiento	paralelo	secuencial
Número de cores	$10^{11} - 10^{12}$	pocos
Conexiones	10.000 por procesador	pocas
Almacenamiento	distribuido	direcciones fijas
Tolerancia a fallos	amplia	nula
Tipo de control de proceso	auto-organizado	centralizado

## VII. CONCLUSIÓN DE LA HIPÓTESIS INICIAL

No se pudo comprobar, *a prima facie*<sup>25</sup> la validez de la hipótesis inicial, ya que la alta densidad de ruido ISO en la foto del experimento, hace que la aproximación por medio de técnicas de datamining no sea exacta. La visualización de los clusters generados por la aplicación de dos algoritmos distintos (Figuras 19 y 18) dio como resultado, la similitud con la foto de origen, esto es debido a que claramente los pixeles que componen la foto, **en promedio**, son los que componen las distintas partes de la misma. El ruido ISO existente en la foto, por sus valores RGB promedio y con los algoritmos de datamining probados, no puede distinguirse del valor RGB promedio del sector correspondiente de la foto.

## VIII. REDES NEURONALES

Las redes neuronales artificiales (o ANS<sup>26</sup>) no son más que otra forma de emular ciertas características propias de los humanos, como la capacidad de memorizar y de asociar hechos. Si se examinan con atención aquellos problemas que no pueden expresarse a través de un algoritmo, se observará que todos ellos tienen una característica en común: la experiencia.

El ser humano es capaz de resolver estas situaciones complejas acudiendo a la experiencia acumulada. Así, parece claro que una forma de aproximarse al problema consista en la construcción de sistemas que sean capaces de reproducir esta característica humana. En definitiva, las redes neuronales no son más que un modelo artificial y simplificado del cerebro humano, que es el ejemplo más perfecto de sistema conocido, con la particularidad de desarrollar la capacidad de adquirir conocimiento a través de la experiencia. Una red neuronal es “un nuevo sistema para el tratamiento de la información, cuya unidad básica de procesamiento está inspirada en la célula fundamental del sistema nervioso humano: la neurona”. [12] Las ANS emulan el hardware cerebral, mientras que los sistemas difusos o borrosos, emulan el software. Algunas comparaciones entre el cerebro humano y ANS se pueden ver en la tabla II.

### VIII-A. Sistemas borrosos o difusos

Como se mencionó anteriormente, los sistemas borrosos o difusos, emulan el software cerebral, tratando de reproducir los razonamientos de mas alto nivel, especialmente, el de razonamiento aproximado. [13] La lógica borrosa, asigna términos lingüísticos a propiedades físicas, como por ejemplo

<sup>25</sup>Locución latina en ablativo absoluto que significa “A primera vista”

<sup>26</sup>Artificial Neural Systems. Si bien también existen otras definiciones como Redes Neuronales Artificiales, Neurocomputación, o Sistemas Conexionistas.

temperatura *templada*, y así proporciona un marco de trabajo para manejar esos conceptos, ver figura 20.

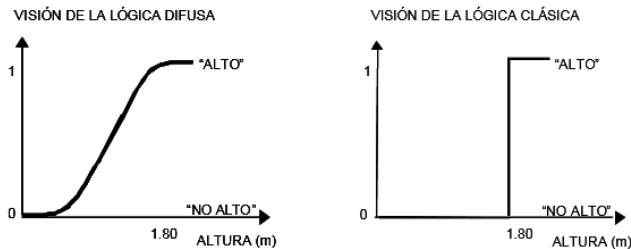


Figura 20: Lógica difusa vs clásica

En resumen:

- Las redes neuronales artificiales, son capaces de descubrir automáticamente relaciones entrada/salida, en función de datos empíricos, merced a su capacidad de aprendizaje a partir de ejemplos.
- Los sistemas borrosos permiten emplear el conocimiento disponible por los expertos para el desarrollo de sistemas inteligentes.
- Es fundamental la integración de diversas técnicas, existiendo en particular, estrechas relaciones entre redes neuronales y sistemas borrosos.

## IX. EXPERIMENTACIÓN CON REDES NEURONALES

### IX-A. Introducción

El proceso que se realiza consiste en la generación de información para entrenar al sistema experto y luego verificar su fiabilidad para la detección de ruido ISO.

### IX-B. Generación del conjunto de datos para entrenamiento

Por medio de los buscadores de internet google<sup>27</sup> y bing<sup>28</sup>, se obtienen imágenes de pequeños gatos domésticos, filtrando las mismas por derecho de uso (libre de derechos) y tamaño mayor a 640 pixeles de ancho por 480 de alto. Luego a través del complemento del navegador chrome<sup>29</sup>, denominado *ImageAssistant Batch Image Downloader*<sup>30</sup>, se procede a realizar una bajada masiva<sup>31</sup> renombrando los archivos con un nombre normalizado, ya que al ser obtenidos de diferentes fuentes, los mismos tienen diferentes nombres. Los archivos duplicados son detectados por medio del proceso hash<sup>32</sup> MD5<sup>33</sup>. Para lograr esto, se utiliza la siguiente pieza de código bash<sup>34</sup> `md5sum * | sort | cut -d ' ' -f1`

<sup>27</sup><http://www.google.com>

<sup>28</sup><http://www.bing.com>

<sup>29</sup><https://www.google.com.ar/chrome/>

<sup>30</sup><https://chrome.google.com/webstore/detail/imageassistant-batch-image-downloader>

<sup>31</sup>También denominada *bulk*.

<sup>32</sup>Función matemática que dada una entrada de cualquier tamaño, devuelve una salida siempre del mismo tamaño y cuyo contenido corresponde “unívocamente” al contenido procesado.

<sup>33</sup>En criptografía, MD5 es un algoritmo de reducción criptográfico de 128 bits ampliamente usado. Uno de sus usos es el de comprobar que algún archivo no haya sido modificado. [14]

<sup>34</sup>Intérprete de órdenes o *shell* de la mayoría de los sistemas operativos GNU/Linux.

| uniq -d. Habiendo depurado la fuente de archivos para el entrenamiento, se procede a generarle el ruido correspondiente. En primera instancia se prueba una aproximación *gaussiana* dando como resultado la inexistencia de ruido suficiente. En función de este resultado, se genera ruido “Impulse”. Este tipo de ruido, ocurre a menudo en proceso de adquisición o transmisión de imágenes digitales como resultado de fallos del sensor fotoeléctrico o por errores en los distintos canales de bits, esto es debido en gran parte a interferencia electromagnética [15]. El código fuente de *gnoise.sh*, que realiza la generación de ruido, es el listado 2. Luego para el procedimiento de la totalidad de los archivos, se invocará *gnoise.sh* para cada uno de los mismos. A dichos efectos, se utiliza el siguiente código bash: `for F in training_denoise/*; do echo $F; ./gnoise.sh $F; done.`

Listado 2: Generador de ruido

```
#!/bin/bash

filename=$(basename -- "$1")
extension="${filename##*.*}"
filename="${filename%.*}"

convert $1 -size 1024x768 +noise Impulse
        training\_noise/\${filename}\_impulse
        .\${extension} &
convert $1 -size 1024x768 +noise
        Gaussian training\_noise/\${filename}\
        _gaussian.\${extension}
```

Luego se reducen las imágenes a una definición menor, mediante el siguiente código: `for f in training_noise/*; do convert $f -size 800x600 training_size/\$(basename $f); done`

### IX-C. Red neural convulacional

La Red Convolutiva CNN o ConvNet (Convolutional neural network) se denomina así por el concepto matemático “convolución”, una transformación lineal<sup>35</sup> de dos funciones en una tercera que puede representar la magnitud, de valor real, en la que se superponen las anteriores. [16].

En 1962 David Hunter Hubel y Torsten Nils Wiesel publican “*Receptive fields, binocular interaction and functional architecture in the cat's visual cortex.*”, sugieren un nuevo modelo de cómo los mamíferos perciben el mundo visualmente. Mostraron que las cortezas visuales de gato y mono incluyen neuronas que responden exclusivamente a su ambiente directo. Esta es la inspiración de las redes convolucionales. [17]

Una CNN es una red multicapa diseñada inicialmente para reconocer patrones visuales, lo que permite mejorar ciertas propiedades respecto de otras arquitecturas, por ejemplo: la

<sup>35</sup>Es toda función cuyo *dominio* y *codominio* son espacios vectoriales que cumplen unas propiedades determinadas.



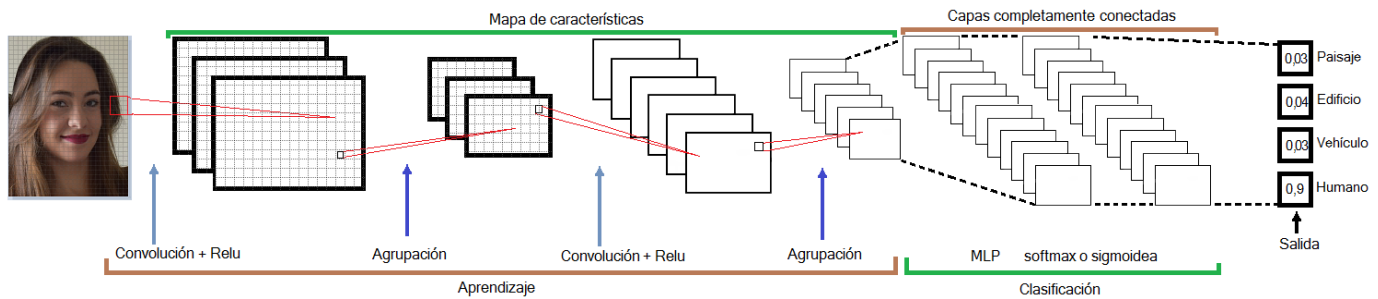


Figura 21: Esquema de una red convolucional de tres capas.

función de retrocesión<sup>36</sup> o *backpropagation* es más eficiente de implementar y reduce enormemente la cantidad de parámetros en la red.

La importancia de este proceso consiste en que, a medida que se entrena la red, las neuronas de las capas intermedias se organizan a sí mismas de tal modo que las distintas neuronas aprenden a reconocer distintas características del espacio total de entrada. Después del entrenamiento, cuando se les presente un patrón arbitrario de entrada que contenga ruido o que esté incompleto, las neuronas de la capa oculta de la red responderán con una salida activa si la nueva entrada contiene un patrón que se asemeje a aquella característica que las neuronas individuales hayan aprendido a reconocer durante su entrenamiento.

La red toma el concepto de capas con especialización, cada neurona de una capa no recibe conexiones entrantes de todas las neuronas de la capa anterior, sino de algunas. Esto hace que una neurona se especialice en un dominio de la capa anterior reduciendo así el número de operaciones. La **reducción** es el proceso relevante.

Otras redes son más eficientes para trabajar con matrices con pocas características, pero cuando esas matrices llegan a ser muy grandes su rendimiento cae, esto ocurre cuando trabajamos con imágenes. Para solventar este inconveniente, las primeras redes convolucionales extraían las peculiaridades más significativas y reducían así las matrices de características, para ello utilizaban una imagen de entrada de tamaño fijo, pero este es un requisito artificial y puede reducir la precisión al reconocer imágenes de un tamaño arbitrario. Los procedimientos que se realizaban para transformar las imágenes de entrada y ponerlas en el formato requerido eran: recortar la imagen (solo seleccionando una parte del tamaño correcto de la imagen) y deformarla (cambiando la proporción).

Las redes CNN solventan la necesidad de transformar las imágenes y reducen el número de operaciones a realizar. La estructura general de una red convolucional se construye con tres tipos de capas distintas, como se aprecia en la figura 21.

Dado a la naturaleza de las convoluciones dentro de las redes neuronales convolucionales, estas son aptas para poder

aprender a clasificar todo tipo de datos donde estos estén distribuidos de una forma continua a lo largo del mapa de entrada, y a su vez sean estadísticamente similares en cualquier lugar del mapa de entrada. Por esta razón, son especialmente eficaces para clasificar imágenes, por ejemplo para el auto-etiquetado de imágenes.

Sin embargo, las redes neuronales convolucionales también pueden ser aplicadas para la clasificación de series de tiempo o señales de audio utilizando convoluciones en 1D, así como para la clasificación de datos volumétricos usando convoluciones en 3D.

#### IX-D. Entrenar la red con Keras

*Keras* es un conjunto de funciones de redes neuronales de alto nivel escrita en Python. El motor defacto de *Keras* es *Tensorflow*<sup>37</sup>, pero también usa *CNTK*<sup>38</sup> [19] y *Theano*<sup>39</sup>. [21].

Estas bibliotecas pueden utilizar el poder computacional de las GPU<sup>40</sup>, sin embargo, en la mayoría de las notebooks no es una decisión acertada, debido a que por cuestiones propias de la ingeniería del mencionado hardware (tamaño, disipación de calor, consumo de energía, entre otros) es muy posible que siga siendo mas performante utilizar los núcleos del CPU de la computadora, en vez de los GPU.

La biblioteca *Keras*, tal cual como viene empaquetada, utiliza un conjunto genérico de instrucciones de CPU a los efectos de mantener la compatibilidad entre CPUs. Para esta investigación, la versión de biblioteca *Keras*, no utilizaba los

<sup>36</sup>Es un método de cálculo del gradiente utilizado en algoritmos de aprendizaje supervisado utilizados para entrenar redes neuronales artificiales. El método emplea un ciclo propagación – adaptación de dos fases. Una vez que se ha aplicado un patrón a la entrada de la red como estímulo, este se propaga desde la primera capa a través de las capas siguientes de la red, hasta generar una salida. La señal de salida se compara con la salida deseada y se calcula una señal de error para cada una de las salidas.

<sup>37</sup>TensorFlow™ es una biblioteca de software libre que se utiliza para realizar cálculos numéricos mediante diagramas de flujo de datos. Los nodos de los diagramas representan operaciones matemáticas y las aristas reflejan las matrices de datos multidimensionales (tensores) comunicadas entre ellas. [18]

<sup>38</sup>Microsoft Cognitive Toolkit (CNTK) es un conjunto de herramientas de código abierto para el aprendizaje profundo distribuido de calidad comercial. Describe las redes neuronales como una serie de pasos computacionales a través de un gráfico dirigido. CNTK implementa el aprendizaje de descenso de gradiente estocástico con diferenciación automática y paralelización en múltiples GPU y servidores.

<sup>39</sup>Theano es una biblioteca de Python que permite definir, optimizar y evaluar expresiones matemáticas, especialmente aquellas con matrices multidimensionales (*numpy.ndarray*) [20]. También puede superar C en una CPU en muchos órdenes de magnitud aprovechando las GPU recientes.

<sup>40</sup>Procesador principal de las placas de video.



Figura 22: Consumo de memoria y CPU de la ejecución del aprendizaje

conjuntos de instrucciones AVX2<sup>41</sup> y FMA<sup>42</sup>, al descargar de internet una versión compilada con dicho soporte, se logra en cada *epoch*<sup>43</sup> una reducción de 6 horas a 4 horas.

El entrenamiento fue realizado utilizando una red neuronal convolucional con imágenes de 64x64 pixels en formato RGB. Las pruebas determinaron que a pesar de distintos parámetros ingresados, la red no fue capaz de distinguir una foto con ruido ISO de una sin ruido apreciable.

Se aumentó a 1024x768, no siendo posible terminar el proceso, debido al consumo excesivo de memoria RAM. 16 GB no alcanzaron, dando el aviso `tensorflow/core/framework/allocator.cc:122: Allocation of 1249132544 exceeds 10% of system memory`. Si a pesar del aviso, se continúa, el proceso insiste constantemente en pedir memoria, haciendo que el sistema operativo se vea forzado a utilizar la partición de *swapping*<sup>44</sup>, hasta que llega a un punto donde la ejecución es inviable, y el kernel<sup>45</sup> toma acción, eliminando el proceso de aprendizaje.

Habiendo reiniciando la computadora para que se depure la memoria y reestablezcan los parámetros de intercambio, sucesivos ensayos determinaron que se podía ejecutar sin errores de falta de memoria, cuando la imagen tenía un tamaño de: 640x480. Figura 22

<sup>41</sup>Intel®Advanced Vector Extensions es una nueva extensión de conjunto de instrucciones de 256 bits para Intel®SSE y está diseñada para aplicaciones que son intensivas en punto flotante (FP). [22]

<sup>42</sup>Una suma combinada fusionada (también conocida como *fmadd*) es una operación de suma multiplicada en coma flotante realizada en un solo paso, con un solo redondeo. Es decir, cuando una suma múltiple no fusionada calcularía el producto  $b.c$ , lo redondearía a  $N$  bits significativos, agregaría el resultado  $a$  y redondearía a  $N$  bits significativos, una suma múltiple fusionada fusionaría la expresión completa  $a + b.c$  a su máxima precisión antes de redondear el resultado final a  $N$  bits significativos.

<sup>43</sup>En Machine Learning, un *epoch* es el procesamiento completo de un conjunto de datos de entrenamiento por el algoritmo de aprendizaje.

<sup>44</sup>En informática, el espacio de intercambio es una zona del disco (un fichero o partición) que se usa para guardar las imágenes de los procesos que no han de mantenerse en memoria física. A este espacio se le suele llamar swap, del inglés “intercambiar”.

<sup>45</sup>Es el principal responsable de facilitar a los distintos programas acceso seguro al hardware de la computadora o en forma básica, es el encargado de gestionar recursos, a través de servicios de llamada al sistema.

El proceso de aprendizaje fue posible ejecutarlo únicamente utilizando el CPU, y no el GPU, debido a que cuando se quiso implementar GPU en esta nueva configuración, el mismo canceló dando el error: `tensorflow.python.framework.errors_impl.ResourceExhaustedError: OOM when allocating tensor with shape[596608,128] and type float on /job:localhost/replica:0/task:0/device:GPU:0 by allocator GPU_0_bfc`. El entrenamiento finalizó satisfactoriamente luego de 32 horas de entrenamiento utilizando el CPU únicamente. Sin embargo éste no generó información suficiente como para predecir con exactitud que foto posee ruido ISO o cual no.

## X. CONCLUSIÓN

Si bien se experimenta con algunos de los algoritmos de datamining presentados y con una red neuronal especializada en el reconocimiento de imágenes, no es posible el reconocimiento de ruido ISO en las imágenes.

Para el caso de Datamining, la densidad del ruido es un factor determinante, debido a que el algoritmo no puede discernir entre el ruido y sus píxeles vecinos.

Para el caso de Redes Neuronales, hubo que reducir las imágenes a una resolución de 640x480 lo que hace que el ruido aumente al ser mayor su densidad. Esto sumado al entrenamiento que la red tuvo acotado por el hardware existente al momento de la experimentación, no genera suficiente conocimiento para una detección fiel de ruido ISO en las imágenes presentadas.

## AGRADECIMIENTOS

Se lo dedico a mi familia, mi hija Delfina y mi esposa Analía, por su apoyo incondicional. A los docentes, que fueron una fuente de inspiración y consulta inagotables.

## REFERENCIAS

- [1] S. Ruiz and X. Triquell, “Imágenes y palabras en la lucha por imposición de sentidos: la imagen como generadora de relatos,” *Cuadernos del Centro de Estudios de Diseño y Comunicación* N° 61. [Online]. Available: [http://fido.palermo.edu/servicios\\_dyc/publicacionesdc/vista/detalle\\_articulo.php?id\\_articulo=12434&id\\_libro=596](http://fido.palermo.edu/servicios_dyc/publicacionesdc/vista/detalle_articulo.php?id_articulo=12434&id_libro=596)
- [2] R. G. Bednarik, “Los primeros testimonios del espíritu creador,” *El correo de la Unesco*.
- [3] M. Brady. El colodión húmedo. [Online]. Available: [https://fido.palermo.edu/servicios\\_dyc/blog/docentes/trabajos/42301\\_160424.pdf](https://fido.palermo.edu/servicios_dyc/blog/docentes/trabajos/42301_160424.pdf)
- [4] J. S. Mohd Awais Farooque, “Survey on various noises and techniques for denoising the color image,” *International Journal of Application or Innovation in Engineering and Management (IJAIE)*, vol. 2, no. 11, 2013. [Online]. Available: <http://www.ijaiem.org/volume2issue11/IJAIE-2013-11-24-070.pdf>
- [5] D. P. Cattin. Image restoration: Introduction to signal and image processing. [Online]. Available: [https://miac.unibas.ch/SIP/06-Restoration.html#\(1\)](https://miac.unibas.ch/SIP/06-Restoration.html#(1))
- [6] J. Nakamura, *Image Restoration: Introduction to Signal and Image Processing*, C. Press, Ed.
- [7] M. C. Lovell, “Data mining,” *The Review of Economics and Statistics*, vol. 65, no. 1, 1983. [Online]. Available: [https://www.jstor.org/stable/1924403?origin=crossref&seq=1#page\\_scan\\_tab\\_contents](https://www.jstor.org/stable/1924403?origin=crossref&seq=1#page_scan_tab_contents)
- [8] Curriculum design philosophy. [Online]. Available: <http://www.kdd.org/curriculum/view/curriculum-design-philosophy>
- [9] IEEE International Conference on Data Mining. [Online]. Available: <http://www.cs.uvm.edu/~icdm/algorithms/10Algorithms-08.pdf>

- [10] V. B. M. J. R. H. R. V. Baños. Cómo aplicar árboles de decisión en SPSS. [Online]. Available: <http://diposit.ub.edu/dspace/handle/2445/43762>
- [11] WEKA. [Online]. Available: <https://www.cs.waikato.ac.nz/ml/weka/>
- [12] D. J. Matich. Redes neuronales: Conceptos básicos y aplicaciones. [Online]. Available: <ftp://decsai.ugr.es/pub/usuarios/castro/Material-Redes-Neuronales/Libros/matich-redesneuronales.pdf>
- [13] B. M. del Brio; Alfredo Sanz Molina, *Redes Neuronales y Sistemas Difusos*.
- [14] R. Rivest. The md5 message-digest algorithm. [Online]. Available: <http://www.ietf.org/rfc/rfc1321.txt>
- [15] P.VENKATESAN and G.NAGARAJAN, "Removal of gaussian and impulse noise in the colour image progression with fuzzy filters," *International Journal of Electronics Signals and Systems(IJESS)*, vol. 3, no. 1, 2013. [Online]. Available: <https://pdfs.semanticscholar.org/a622/921a28a71210e2401e7c447ec5dedd8634b4.pdf>
- [16] Cnn-rnn convulacional. [Online]. Available: <http://numerentur.org/convolucionales/>
- [17] Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1359523/>
- [18] An open source machine learning library for research and production. [Online]. Available: <https://www.tensorflow.org/>
- [19] The microsoft cognitive toolkit. [Online]. Available: <https://docs.microsoft.com/en-us/cognitive-toolkit/>
- [20] Theano at a glance. [Online]. Available: <http://deeplearning.net/software/theano/introduction.html>
- [21] Keras - biblioteca de aprendizaje profundo para python. [Online]. Available: <https://keras.io/>
- [22] Isa extensions. [Online]. Available: <https://software.intel.com/en-us/isa-extensions/intel-avx>



**Pablo Niklas** Estudiante avanzado de la Licenciatura de Tecnología de la Información correspondiente a la Facultad de Ingeniería de la Universidad de Palermo. Tiene como hobby coleccionar computadoras de los 80's y la fotografía<sup>46</sup>, se lo puede contactar al mail [pablo.niklas@gmail.com](mailto:pablo.niklas@gmail.com).

#### ÍNDICE DE FIGURAS

1.	Daguerrotipo del Almirante Dn. Guillermo Brown.	2
2.	William Henry Fox Talbot, ca. 1864 . . . . .	2
3.	Ejemplo de colodión húmedo . . . . .	2
4.	Fotografía con sensibilidad ISO 1600. . . . .	3
5.	Fotografía con sensibilidad ISO 25. . . . .	3
6.	Ejemplo de ruido en la fotografía digital. . . . .	4
7.	Sensor CCD Canon SX30 . . . . .	4
8.	Imagen con ruido tipo "Sal y Pimienta". . . . .	5
9.	Curvas de distribución normal o gaussiana . . . . .	5
10.	La cantidad de fotones por pixel aumenta de izquierda a derecha y de arriba hacia abajo. . . . .	6
11.	Gráfica del ruido blanco en función del tiempo. . . . .	6
12.	Kmeans con $k = 3$ . . . . .	7
13.	Foto usada en la experimentación . . . . .	8
14.	Pantalla de WEKA mostrando clases de atributos. . . . .	8
15.	Canon 7D Mark II . . . . .	9
16.	Matriz de los atributos. . . . .	9
17.	Zoom al 9000 % para la validación. . . . .	10
18.	Visualización de los 5 clusters . . . . .	10
19.	Clusters detectados con EM . . . . .	10
20.	Lógica difusa vs clásica . . . . .	11
21.	Esquema de una red convulacional de tres capas. . . . .	12
22.	Consumo de memoria y CPU de la ejecución del aprendizaje . . . . .	13

#### ÍNDICE DE CUADROS

I.	Valores RGB de los píxeles . . . . .	8
II.	Comparación entre ANS y el cerebro [13] . . . . .	10

<sup>46</sup>[www.pabloniklas.com](http://www.pabloniklas.com)