

IMAGEN Y PODER EN LA ERA DE LA VISIÓN ARTIFICIAL: UNA INTERPRETACIÓN A PARTIR DE GILBERT SIMONDON Y GILLES DELEUZE*

IMAGE AND POWER IN THE AGE OF COMPUTER VISION:
AN INTERPRETATION USING GILBERT SIMONDON AND
GILLES DELEUZE

CLAUDIO CELIS BUENO**

RESUMEN: El presente artículo utiliza las nociones de información y modulación en Gilbert Simondon y en Gilles Deleuze para pensar el concepto de imagen en el nuevo contexto de visión artificial. El desarrollo de algoritmos de aprendizaje automático ha posibilitado la automatización cada vez mayor de sistemas de percepción visual en los cuales las imágenes son producidas y consumidas por máquinas. En este contexto, pareciera ser que las categorías tradicionales para pensar la relación entre imagen y poder quedan obsoletas. Ante esto, este artículo explora en qué medida las nociones simondonianas de información y modulación, en particular a través de su recepción en la obra de Gilles Deleuze, ofrecen un punto de vista alternativo para pensar una noción no representacional y no hilemórfica de imagen, así como su vínculo con nuevas formas de poder (sociedades de control).

PALABRAS CLAVE: Gilbert Simondon, algoritmos, Gilles Deleuze, sociedades de control, visión artificial.

ABSTRACT: This article uses the notions of modulation and information in order to re-define the notion of image in the new context of artificial vision. Machine learning algorithms have enable an ever greater automation of visual perception where images are produced and consumed by machines. It seems that traditional categories to think the relationship image-power become obsolete in light of these new technologies. Thus, this article explores how Gilbert Simondon's notions of modulation and information particularly as received in Gilles Deleuze's work, offer an alternative approach to a post-representational concept of image and its relation to new power regimes (societies of control).

KEYWORDS: Gilbert Simondon, Algorithms, Gilles Deleuze, Control Societies, Artificial Vision.

Recibido: 10.04.2019. Aceptado: 01.07.2020

* El presente artículo ha sido escrito en el marco del proyecto Fondecyt de iniciación N° 11170065, CONICYT, Chile.

** Doctor en Critical and Cultural Theory. Académico e investigador del Instituto de Humanidades, Universidad Academia de Humanismo Cristiano, Santiago, Chile. Correo electrónico: ccelis@academia.cl. Orcid: <https://orcid.org/0000-0001-9976-100X>

1. INTRODUCCIÓN

EN LOS ÚLTIMOS AÑOS, los sistemas de visión artificial han evolucionado de manera considerable. Como sugiere Adam Greenfield (2017, p. 214) esto ha sido posible gracias a la aplicación de algoritmos de aprendizaje automático [*machine-learning algorithms*] al reconocimiento de imágenes. La visión artificial se ha comenzado a utilizar de manera cada vez más presente en procesos productivos (para coordinar líneas de producción altamente automatizadas), en dispositivos de seguridad (para el reconocimiento facial o de incidentes a través de cámaras de vigilancia), en vehículos de conducción autónoma (como es el caso del *Self-Driving Car* de Google o el *Auto-pilot* de Tesla), y en tecnologías militares (como drones o misiles guiados autónomamente).

El artista norteamericano Trevor Paglen (2019) utiliza la noción “imágenes invisibles” para definir este nuevo régimen de la percepción visual en el cual una progresiva cantidad de imágenes es producida por y para las máquinas, excluyendo gradualmente al ojo humano del acto perceptivo. Este nuevo contexto técnico, sugiere Paglen, exige nuevas herramientas conceptuales para pensar la compleja relación entre imagen y poder. Por relación entre imagen y poder nos referimos a capacidad de las imágenes para reflejar, reproducir o interrumpir determinadas relaciones de poder. Tradicionalmente, conceptos como espectáculo, ideología, mimesis, representación, identificación, contemplación, fetichismo, etc., han intentado dar cuenta de la relación entre imagen y poder desde la perspectiva del espectador humano, sujeto y agente de la percepción visual (Paglen, 2019). Una característica central de esta relación remite al hecho de que el vínculo entre la representación y lo representado es flexible y ambivalente. Esto implica que el hiato entre la representación y lo representado se presenta como un territorio de disputa política (Rancière, 1995; Baudrillard, 1979). En el caso de la visión artificial, en cambio, esta ambivalencia pierde su carácter constitutivo y comienza a ser reemplazada por la performatividad y efectividad de un sistema técnico (Paglen, 2019).

Con el objetivo de desplazar el análisis desde la dimensión ambivalente de la imagen visible hacia la dimensión operativa de la imagen invisible y con ello pensar el estatuto político de la imagen en la era de la visión artificial, este artículo utiliza el marco conceptual abierto por los filósofos franceses Gilbert Simondon (2015 y 2016) y Gilles Deleuze. Específicamente, este artículo propone como hipótesis principal que las categorías de información y modulación (propuestas por Simondon y elaboradas posterior-

mente por Deleuze) ofrecen un importante punto de partida para definir un nuevo marco teórico desde donde pensar la relación entre imagen y poder en el nuevo contexto de visión artificial.

2. EL CONCEPTO DE MODULACIÓN SEGÚN GILBERT SIMONDON

La filosofía de Gilbert Simondon (2015) busca redefinir la noción de individuación. Para Simondon, el pensamiento occidental ha pensado el ser a partir de un principio de individuación y no en tanto proceso de individuación. Esto se debe a que el pensamiento filosófico solo ha conocido un modo de pensar el equilibrio: el equilibrio estable (p. 11). Este modo de equilibrio “excluye el devenir”, ya que representa un sistema en el cual “todas las transformaciones posibles” fueron ya realizadas (p. 11). Como sostiene Yuk Hui (2015, p. 80), el pensamiento occidental es incapaz de pensar la individuación en términos procesuales (el proceso de individuación precede al principio de individuación), relacionales (los seres que entran en relación no preceden al proceso de individuación que los constituye), e incompletos (en todo proceso de individuación persiste un remanente de energía potencial que posibilita futuras individuaciones).

Para desarrollar su noción de individuación en tanto proceso y no en tanto principio, Simondon incorpora la noción de “equilibrio metaestable” (descubierta por la física termodinámica en el siglo XIX). La metaestabilidad de un sistema refiere a la energía potencial en reposo relativo que posibilita futuras transformaciones. Desde esta perspectiva, la individuación no consiste en el mero tránsito del devenir al equilibrio (como ha sido pensada tradicionalmente por la filosofía), sino en la “resolución de un sistema metaestable” (2015, p. 12). Esto quiere decir que una determinada forma no agota la energía potencial de un sistema, sino simplemente “mantiene su nivel energético y conserva sus potenciales” (p. 25). Más aún, sostiene Simondon, para que un estudio de la individuación sea capaz de incorporar el punto de vista de la metaestabilidad, la noción de forma debe ser reemplazada por la de información (p. 25).

Es importante mencionar que la noción de información adquiere en Simondon una definición particular y no debe ser reducida a “las señales o soportes o vehículos de información, como tiende a hacerlo la teoría tecnológica de la información” (p. 25). Para Simondon (2016), la información no debe ser pensada como una cosa, sino como “la operación de una cosa que llega a un sistema y que produce allí una transformación” (p. 139). La infor-

mación, por ende, “no puede definirse más allá de este acto de incidencia transformadora y de la operación de recepción” (Simondon, 2016, p. 139). La función de información, por lo tanto, es la modificación de una realidad local (un cambio de estado energético en un sistema metaestable) a partir de una señal incidente (Simondon, 2016, p. 140). Asimismo, un receptor de información es virtualmente “toda realidad que no posee enteramente por sí misma la determinación del curso de su devenir” (Simondon, 2016, p. 140). Dicho de otro modo, todo sistema metaestable es un potencial receptor de información, es decir, un receptor de una señal incidente que gatilla una transformación de estado. Por su parte, toda señal incidente que gatille una transformación energética será considerada información.

Ahora bien, para comprender mejor el concepto de información en Simondon, la noción de modulación es de vital importancia (Iliadis, 2013, p. 94). Simondon introduce el concepto de modulación en *La Individuación a la luz de las nociones de forma y de información* (2015), en el marco de su crítica al hilemorfismo (que comprende la individuación a partir del encuentro entre una materia pasiva y una forma activa). Luego de argumentar que la comprensión de la producción de un ladrillo no se puede limitar a la ecuación materia pasiva (arcilla) y forma activa (molde), ya que el encuentro entre ambos polos representa solo un momento de dos semi-cadenas de procesos técnicos, Simondon postula que una mejor comprensión del proceso de individuación puede tomar como modelo el funcionamiento del “relé electrónico” (2015, p. 38). Es a partir del análisis de este dispositivo técnico que Simondon establecerá una distinción entre el molde (ladrillo) y la modulación (relé electrónico), sosteniendo que este último término refiere a un “molde temporal continuo” (2015, p. 39). Moldear y modular constituyen los dos polos opuestos de un proceso de individuación: “moldear es modular de manera definitiva; modular es moldear de manera continua y perpetuamente variable” (2015, p. 39). En el primer caso, la individuación ocurre de una vez, es decir, una vez efectuado el proceso de individuación se alcanza un estado de equilibrio que persiste más allá del acto de moldeado. En el segundo caso, la individuación es temporal: en cuanto deja de ocurrir el proceso de individuación, la forma individuada vuelve a su estado pre-individual de energía potencial (metaestable).

Esta distinción entre molde y modulación vuelve a aparecer en una conferencia de 1962, titulada *La amplificación en los procesos de información*, en la cual Simondon (2016) nuevamente vincula la operación de modulación al análisis del relé electrónico. El relé eléctrico (antecesor del relé electrónico y del transistor informático) utiliza una bobina electromagnética

para activar o desactivar un segundo sistema eléctrico acoplado. Como se puede observar en la figura 1, la bobina electromagnética (B) es activada por una señal de baja energía (A) que, al cargarse, moviliza un interruptor (C) que activa una energía considerablemente mayor en el segundo sistema (D). Con ello, es posible controlar, a través de una muy baja señal, una serie de interruptores de un sistema de gran carga energética. Se podría sostener que la invención del relé eléctrico ha jugado un rol central en la historia de la automatización industrial y posteriormente de los sistemas informáticos¹. Por su parte, para Simondon, el relé eléctrico permite ejemplificar las nociones de información y modulación que están en la base de su teoría de la individuación.

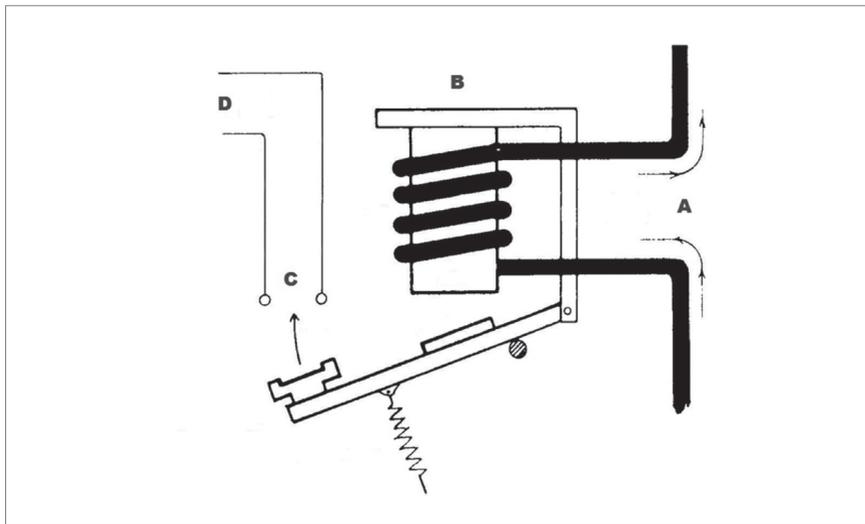


Figura 1. Esquema de un relé eléctrico, elaboración propia.

En particular, podemos mencionar cuatro consecuencias conceptuales que Simondon extrae a partir del análisis del relé eléctrico. En primer lugar, el relé eléctrico ilustra una noción de información que se define estricta-

¹ En términos generales, un transistor es un relé eléctrico que reemplaza la bobina electromagnética por un circuito de silicio, reduciendo tanto el umbral de activación como la velocidad de respuesta.

mente por su capacidad para gatillar una transformación energética (y no por una singularidad contenida en el mensaje transmitido). Como hemos dicho, para Simondon la información no es una cosa, sino la operación que gatilla una señal en un sistema metaestable. Desde este punto de vista, una información puede ser generada intencionalmente o no. Lo que determinará su condición de información será su recepción en tanto señal de activación de una transformación energética. En el caso del relé eléctrico, la señal de baja intensidad funciona como información gracias a su capacidad para activar la bobina electromagnética y con ello activar la descarga eléctrica del segundo sistema².

En segundo lugar, Simondon destaca que el relé eléctrico posibilita que “una señal incidente de débil energía” (información) controle “una energía local más considerable” (Simondon, 2016, p. 152). La modulación es amplificadora porque una señal débil modifica un estado energético considerablemente mayor. De hecho, en la década de 1930, Edwin Armstrong descubre que el relé eléctrico puede ser utilizado para amplificar una onda (Fig. 2): al introducir una onda débil en la señal de entrada del relé eléctrico (A), la señal de salida (B) generará la misma onda pero con una carga energética mayor. Este punto es central para comprender la diferencia entre molde y modulación: mientras el primero solo puede individualizar una forma fijada de antemano (p. ej.: un ladrillo), el segundo puede modular cualquier información que ingrese al amplificador (p. ej.: onda acústica). De allí la definición del modulador como “molde temporal continuo”.

² Este punto es importante para comprender la crítica de Simondon a la “teoría matemática de la información”. Para Simondon, el problema del concepto de información propuesto por Shannon (1948) y por Wiener (1948) es que se limita al análisis del mensaje y su transmisión, dejando de lado la problemática de su recepción (2015, p. 281). Según Simondon, la información es una señal que “no es solamente emitida y transmitida por modulación de energía”, sino que “es también recibida por un dispositivo que posee su propio funcionamiento y que debe integrar la señal de información dentro de su funcionamiento, haciéndole jugar un rol de información eficaz” (2015, p. 281). Por ello es que la señal de información no debe ser comprendida solamente como aquello transmitido “sin deterioro causado por el ruido de fondo y de los demás aspectos del azar y de degradación de energía”, sino como aquello “que debe ser recibido, es decir, lo que debe adoptar una significación, tener eficacia para un conjunto que posee un funcionamiento propio” (2015, p. 281). Para un análisis de la relación entre el concepto de información de Gilbert Simondon y la teoría matemática de la comunicación de Shannon y Wiener, véase Terranova (2004).

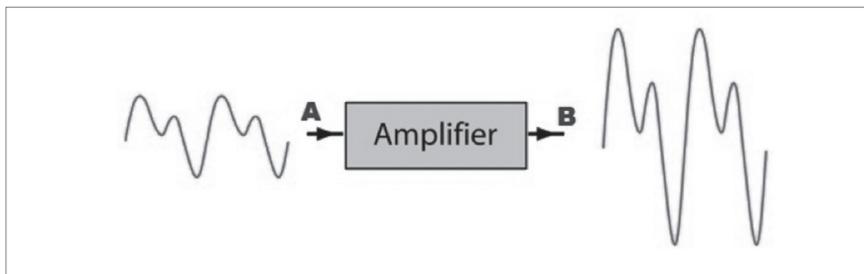


Figura 2. Esquema de amplificación de onda, elaboración propia.

En tercer lugar, el sistema de activación (señal/información) del relé eléctrico está separado energéticamente de la energía potencial del receptor. Esta es una de las diferencias fundamentales respecto de la amplificación transductiva en la cual el propio proceso de transformación actúa como información de futuras transformaciones (Simondon, 2016, p. 144). Esto significa que en la operación de modulación el sistema receptor es autónomo energéticamente (posee una energía potencial que lo mantiene en equilibrio metaestable), pero heterónimo “desde el punto de vista de la causa que desencadena las transformaciones” (2016, p. 140). Para que ello sea posible, el sistema receptor debe poseer un mecanismo de “aislamiento interno” que evite el desgaste de la totalidad de su energía potencial. La operación de modulación implica que la señal de información (“energía débil”) no entre “directamente en trato con la energía potencial del receptor” (es decir, que no se añada o se sustraiga a ella), sino que modifique la “estructura de aislamiento interno del receptor” y con ello provoque su transformación energética o modulación (Simondon, 2016, p. 140). En el caso del relé eléctrico presentado en la *Figura 1*, la energía incidente (A) no ingresa en trato directo con la energía local (D), sino que opera a través de la modificación del sistema de aislamiento interno (C) gracias a la modificación de su estructura (B).

El ejemplo del relé eléctrico, además, permite comprender que todo proceso de información (y por ende de individuación) ocurre en un sistema de disparidad energética. El sistema metaestable (D) se encuentra en una permanente disparidad energética que se presenta como condición de posibilidad de futuras transformaciones. La información operará entonces como señal transformadora no porque agregue o sustraiga energía al sistema, sino porque “resuelve” la disparidad ya presente en el sistema metaes-

table a través de una modificación del sistema de “aislamiento interno” (C). Esta resolución, sin embargo, no agota la disparidad, sino que la mantiene (Simondon ofrece como ejemplo la visión binocular, la cual requiere de la disparidad entre ambos ojos para producir una visión en profundidad).

Estas cuatro conclusiones implican una consecuencia más general del modelo del relé eléctrico para una teoría de la individuación en Simondon: el relé eléctrico no presupone el principio de individuación (a diferencia del modelo del ladrillo para la tradición hilemórfica que presupone la forma del ladrillo como condición de posibilidad de la individuación), sino que evidencia el proceso de adquisición de forma (en tanto señal de información que gatilla un cambio de estado energético en un sistema metaestable). A partir de esto, se puede decir que la diferencia central entre la noción de forma y la de información es que la primera presupone un molde fijo (principio de individuación) mientras que la segunda opera mediante un modulador, es decir, un molde temporal variable (proceso de individuación).

3. LA VISIÓN ARTIFICIAL COMO MODULACIÓN

Como hemos mencionado en la introducción, la visión artificial ha experimentado una evolución considerable en el último tiempo gracias al desarrollo de algoritmos de aprendizaje automático. Hannah Fry (2018, p. 10) explica que existen dos tipos de algoritmos: algoritmos basados en reglas y algoritmos de aprendizaje automático. En el primer caso, la operación de un algoritmo está pre-definida de antemano por las reglas que ha establecido el programador. Esto significa que los algoritmos basados en reglas sirven solo para resolver problemas para los cuales el programador conoce de antemano los pasos necesarios para su resolución (Fry, 2018, p. 10). Los algoritmos de aprendizaje automático, en cambio, son sistemas que permiten, a través de un proceso de entrenamiento, optimizar un algoritmo para alcanzar un resultado para el cual no se conocen las reglas de antemano. Este tipo de algoritmos ha sido posible técnicamente por un encuentro de dos tecnologías paralelas: el desarrollo del *Big Data* (la gran acumulación de datos que ahora puede ser utilizada para el entrenamiento de los algoritmos de aprendizaje automático) y la evolución de los procesadores gráficos (los cuales posibilitan una mayor cantidad de poder de procesamiento paralelo de datos y acelera así considerablemente el proceso de entrenamiento de los algoritmos de aprendizaje automático).

Para algo tan variable y complejo como la percepción visual, los algo-

ritmos basados en reglas resultan extremadamente limitados. La emergencia de algoritmos de aprendizaje maquina, en cambio, ha posibilitado el mejoramiento de los sistemas de reconocimiento de imágenes gracias a su capacidad de entrenamiento y permanente corrección. Uno de los inventos fundamentales para el desarrollo de estos algoritmos de aprendizaje maquina (y por ende para el desarrollo de las tecnologías de visión artificial) fue el algoritmo de clasificación binaria *Perceptron* (Rosenblatt, 1958). Basado en la concepción que en 1958 se tenía del funcionamiento de una neurona, el *Perceptron* (Fig. 3) es un algoritmo en el cual una señal de entrada, al sobrepasar determinado umbral de activación, gatilla una nueva señal de salida. Tal como el relé electrónico, el *Perceptron* se compone de cuatro partes: un valor de entrada (X); un operador que realiza una determinada operación sobre el valor de entrada (W); una operación de sumatoria que combina los resultados de la operación sobre los distintos valores de entrada (Σ) y un umbral de activación (A) que define si la señal de salida será 0 o 1 (Y). El *Perceptron* es el fundamento de los algoritmos de aprendizaje maquina (Greenfield, 2017, pp. 214-215). A través del proceso de aprendizaje que introduce un set de datos de entrenamiento (X), el algoritmo va ajustando autónomamente el valor del operador (W) y del umbral de activación (A) con el objetivo de extraer el resultado que reduzca al máximo el margen de error de clasificación (Y).

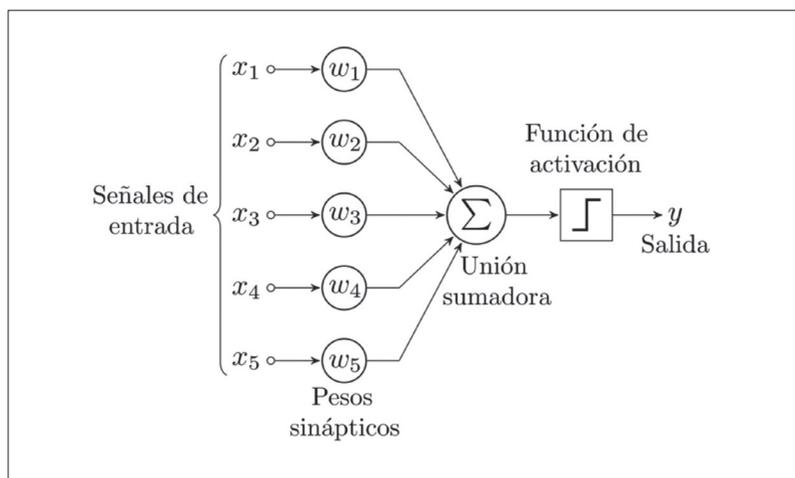


Figura 3. Modelo del algoritmo *Perceptron* (Rosenblatt, 1958).

Cuando Frank Rosenblatt diseñó el algoritmo *Perceptron* en 1958, sin embargo, no existía ni la capacidad de procesamiento ni la cantidad de datos para explotar su potencial. Por ello es que este algoritmo ha permanecido olvidado por casi 50 años (Kurenkov, 2015; Greenfield, 2017, p. 214). Con el desarrollo del *Big Data* y el avance en procesadores gráficos, el *Perceptron* de Rosenblatt ha sido recuperado y se ha convertido en el fundamento de los sistemas de aprendizaje automático y de reconocimiento de imágenes. En el caso particular de estos últimos, los algoritmos de aprendizaje automático son entrenados a través de bancos de imágenes que poseen de antemano una serie de categorías [*labelled datasets*]³. Por ejemplo, un algoritmo puede ser entrenado para reconocer imágenes de perros. Para ello se le entrega al algoritmo una serie de miles o incluso millones de imágenes (unas de perros y otras no, pero todas ellas ya categorizadas), y el *Perceptron* irá ajustando, a través de un proceso de retro-propagación, sus valores (W y A) para optimizar el índice de precisión y reducir la tasa de error de la información de salida (Y): perro (1) o no-perro (0). Según Google, su aplicación de reconocimiento de imágenes puede alcanzar tasas de precisión de hasta un 99%. La escala y velocidad que hacen posible este nivel de precisión, sin embargo, implica que los programadores de la arquitectura inicial del algoritmo no puedan comprender las reglas de reconocimiento de imágenes que este proceso ha generado (Fry, 2018, p. 11). Investigadores japoneses (Su & Sakurai, 2019), por ejemplo, han mostrado que en ciertos casos la modificación de un solo pixel genera una categoría completamente nueva de reconocimiento de imágenes, evidenciando la distancia radical entre los modos de percepción humana y el desconocido procedimiento de los algoritmos de visión artificial. Por esta misma razón, programadores de Google (Mordvintsev y Tyka, 2015) han creado la aplicación *Deep Dream*, que invierte el algoritmo de reconocimiento de imágenes para intentar visualizar su funcionamiento interno. En vez de identificar una categoría en una imagen dada (por ejemplo, un perro), *Deep Dream* proyecta en una imagen cualquiera la categoría buscada, generando así una visualización del modo de percepción oculto del algoritmo (Fig. 4).

³ Cada vez que operamos un test CAPTCHA para acceder a un sitio en línea estamos colaborando con la producción de estos sets de datos categorizados [*labelled data*] que son luego utilizados para el entrenamiento de algoritmos de reconocimiento de imágenes. Para un análisis detallado del entrenamiento de los algoritmos de reconocimiento de imágenes véase Crawford y Paglen (2019).



Figura 4. Ejemplo de imagen generada por el algoritmo *Deep Dream* de Google.

Ahora bien, al igual que en el caso del relé electrónico analizado por Simondon, el *Perceptron* opera como un mecanismo de “amplificación moduladora” que permite que una señal de entrada (información) active o no una determinada energía potencial. Desde esta perspectiva, el funcionamiento del *Perceptron* (como el de una neurona) es análogo al de un relé electrónico en el cual una señal débil que supera un determinado umbral de activación gatilla la transformación de un sistema metaestable. La diferencia fundamental, sin embargo, es que el *Perceptron* permite actualizar y modificar la estructura del modulador, de allí su aporte fundamental a los sistemas de aprendizaje maquínico. Es decir, la modulación realizada por el *Perceptron* no solo amplifica una determinada señal a partir de la modulación de una energía potencial en estado metaestable, sino que además permite, gracias al proceso de retro-propagación, modular gradualmente la estructura de modulación. Si el ladrillo es un molde y el relé electrónico un modulador, el *Perceptron* es un algoritmo de aprendizaje maquínico que posibilita una modulación auto-modulante.

El fundamento técnico de la visión artificial es un algoritmo de aprendizaje maquínico en el cual la capacidad de modulación del *Perceptron* modula no solo la energía potencial del sistema, sino que modula también la estructura de modulación. Mientras la modulación descrita por Simondon refiere a un “molde temporal continuo”, la modulación que está en juego en

los sistemas de aprendizaje maquínico y que hace posible la visión artificial constituye una modulación continua de la estructura modulante. En este sentido, los algoritmos de aprendizaje maquínico podrían ser leídos como un ejemplo concreto de ese “molde auto-deformante que cambia constantemente y a cada instante” descrito por Gilles Deleuze (1999, p. 280) para definir el poder en las sociedades de control. Dicho de otro modo, sería posible sostener que el concepto de modulación de Gilbert Simondon sugiere un vínculo interno entre visión artificial y aquello que Deleuze (1999) llamó sociedades de control. Este vínculo, como insiste Deleuze, no presupone que una máquina técnica explique una máquina social, sino, por el contrario, que toda máquina técnica está siempre inscrita en una máquina social que la pone en funcionamiento (1999, p. 274).

4. MODULACIÓN Y PODER

En una conversación con Antonio Negri, Gilles Deleuze dice:

Es evidente que puede buscarse siempre la correspondencia entre un tipo de sociedad y un tipo de máquina: las máquinas simples o dinámicas de las sociedades de soberanía, las máquinas energéticas de las sociedades disciplinares, las máquinas cibernéticas y los ordenadores de las sociedades de control. Pero las máquinas no explican nada, es preciso analizar los dispositivos colectivos de enunciación de los cuales las máquinas no son más que una parte. (1999, p. 274)

Esta observación es de gran importancia para el análisis de la visión artificial. Por un lado, establece una relación entre las sociedades de control y la creciente “algoritmización de la vida” (Rouvroy & Berns, 2013). Por el otro, sin embargo, nos advierte que una máquina técnica no puede ser separada del régimen de poder en el cual dicha máquina técnica opera. En este sentido, el análisis de la visión artificial en tanto dispositivo técnico no puede ser separado del análisis de la visión artificial en tanto dispositivo de poder. Ahora bien, las nociones de información y modulación de Gilbert Simondon nos ofrecen una pista significativa para explorar la intersección entre la dimensión técnica y el régimen de poder propios de esta visión artificial.

En el ensayo sobre las sociedades de control, Deleuze (1999) distingue entre el molde propio de las sociedades disciplinares y la modulación característica de las sociedades de control: “Los encierros son moldes [*mou-*

lages] o moldeados diferentes, mientras que los controles constituyen una modulación [*modulation*], como una suerte de molde autodeformante que cambia constantemente y a cada instante, como un tamiz cuya malla varía en cada punto” (1999, p. 280). Para ejemplificar esta diferencia entre molde y modulación, Deleuze distingue entre la fábrica Fordista y la empresa post-Fordista: en la primera, la estandarización del modelo productivo (molde) posibilita la aceleración de la producción, la reducción de costos y el aumento de la plusvalía; en la segunda, la informatización de la producción (modulación) permite flexibilizar el proceso productivo y con ello optimizar el proceso de valorización del capital.

Para Yuk Hui (2015, pp. 77-79), si bien Deleuze no refiere directamente a Simondon en el ensayo sobre las sociedades de control, la oposición entre molde y modulación es claramente una apropiación de las categorías simondonianas. Además del texto sobre las sociedades de control, Hui identifica en la obra de Deleuze otras dos referencias significativas a la noción de modulación de Simondon: el libro *El Pliegue* (2005) y el curso *El Diagrama* (2014). Estas tres referencias identificadas por Hui permiten proponer ciertas hipótesis en torno a la relación entre imagen y poder en el contexto de la visión artificial, es decir, establecer la relación entre la máquina técnica y la máquina social.

En *El Pliegue* (2005, p. 30), Deleuze utiliza la distinción entre molde y modulación para definir una noción de objeto desde una perspectiva no-hilemórfica. En la filosofía barroca de Leibniz, nos dice Deleuze, el objeto no se relaciona con “un molde especial, es decir, con una relación materia-forma, sino con una modulación temporal que implica tanto una puesta en variación continua de la materia como un desarrollo continuo de la forma” (2005, p. 30). Esta noción de objeto permite cuestionar el carácter hilemórfico de la imaginación que se encuentra en el centro del proceso de objetivación en la filosofía Kantiana. Para Kant (2003, p. 245), es la facultad de la imaginación la que posibilita la producción de un objeto en tanto “objeto=X”, es decir, como la subsunción de una multiplicidad dada a la sensibilidad bajo la unidad de un concepto del entendimiento. En esta definición, la imaginación funciona como un molde hilemórfico en el cual una forma activa (concepto del entendimiento) individualiza una materia pasiva (la materia sensible dada a la sensibilidad). En contra de esta noción hilemórfica, la noción de objeto propuesta por Deleuze (2005) se define no como molde sino como modulación, es decir, como un proceso de formación permanente de la materia y de un desarrollo continuo de la forma (pp. 30-31). Esta oposición entre el objeto en la filosofía de Leibniz y en la

filosofía de Kant permite establecer una reflexión acerca del estatuto de la facultad de la imaginación en la era de la visión artificial. En el caso de esta tecnología, lo que está en juego es la capacidad del algoritmo para vincular una determinada información visual con una determinada categoría. A diferencia del modelo hilemórfico de la imaginación de Kant, sin embargo, la visión artificial no impone un concepto unitario para ordenar la materia sensible, sino que modula una cierta información en vistas a generar en cada caso el resultado que reduzca al mínimo un determinado margen de error: es una formación permanente de la materia y un desarrollo continuo de la forma. Tomemos el ejemplo referido más arriba del sistema de reconocimiento de imágenes que cambia la categoría de reconocimiento a partir de la modificación de un pixel (Su & Sakurai, 2019). Lo que este experimento demuestra es la distancia radical entre el modo de percepción visual humano (para el cual el cambio de un pixel es completamente imperceptible y no modifica en nada el proceso de reconocimiento de un objeto) y el modo de percepción de la visión artificial (en el cual no se trata de la subsunción de una multiplicidad sensible bajo el molde de una categoría del entendimiento, sino del cálculo estadístico que relaciona entre sí a una serie de miles o millones de imágenes)⁴. Más aún, el aprendizaje maquínico permite la modulación permanente de las diferentes estructuras modulantes. A diferencia del esquematismo kantiano en el cual el molde es inmutable (categorías trascendentales del entendimiento), en la visión artificial opera un tipo de esquematismo en el cual el molde se modula de manera continua. La noción de modulación permitiría pensar una “imaginación algorítmica” que explique el proceso de objetivación (objeto=X) en los sistemas de visión artificial más allá de toda lógica hilemórfica de la imagen.

Por su parte, en el curso sobre pintura, Deleuze (2014) utiliza la distinción entre molde y modulación para definir el concepto de diagrama en vistas a comprender el acto de pintar. “Pintar es modular”, nos dice Deleuze (2014, p. 144). La pintura es modulación de luz o de color (o de ambos), y el diagrama es la matriz de esta modulación (2014, p. 145). El uso del concepto de modulación en el curso de pintura permite pensar una imagen asignificante (como propone Deleuze, una “imagen sin semejanza”, una “figura no figurativa”). Desde aquí es posible extrapolar una definición de

⁴ Sobre este punto resulta interesante la tesis de Luciana Parisi (2016) sobre cómo los sistemas de aprendizaje maquínico ponen en crisis la lógica deductiva y requieren de nuevos marcos conceptuales para pensar una lógica “inferencial”.

imagen no-representacional que permita una nueva conceptualización de la imagen en la era de la visión artificial. Esta definición no referiría ya a la dimensión de significación y representación de las imágenes visibles que circulan en el terreno de los espectadores humanos, sino que considera primordialmente la condición de información de las imágenes invisibles, es decir, su capacidad para modular procesos de transformación energética. Como propone Trevor Paglen (2019), las imágenes invisibles actúan en nuestra vida cotidiana, cambiando su función representacional por una serie de activaciones, ejecuciones y operaciones. En el caso del reconocimiento facial algorítmico (Fig. 5), por ejemplo, las imágenes del rostro no operan como significantes de una subjetividad más profunda, sino como datos de información que, a través de un proceso relacional que involucra miles o millones de imágenes faciales, busca la reducción al mínimo de su tasa de error. El vínculo entre una imagen facial y un individuo particular (un potencial “terrorista”, “criminal”, “consumidor”, o “usuario”), por lo tanto, será solo un efecto secundario (de superficie) del proceso de individuación relacional en el cual no operan entidades ya individualizadas sino solo procesos de información y posterior transformación energética. Desde este punto de vista sería posible sostener que las imágenes invisibles propias de la visión artificial no operan al nivel de una significación (que implica una subjetividad receptora), sino al nivel asignificante de relaciones informacionales. Más aún, se ha mencionado que la noción de información en Simondon no se refiere al mensaje que comunica un determinado contenido, sino a la señal que activa una determinada transformación energética. Esta noción de información, junto a la noción de imagen como modulación propuesta por Deleuze, enfatiza la importancia de establecer un giro desde una concepción representacional de la imagen hacia una concepción pragmática capaz de identificar no lo que una imagen quiere decir sino aquello que una imagen puede hacer. En el caso del reconocimiento facial algorítmico, la imagen no significa, sino que ejecuta una acción al interior de un marco social y político determinado, movilizandando relaciones de poder concretas⁵.

⁵ Para un análisis más detallado del reconocimiento facial algorítmico, véase Celis (2020).



Figura 5. *Machine Readable Hito*, Trevor Paglen, 2016.

A partir de las nociones de información y modulación en Simondon es posible definir la imagen como un proceso de individuación en el cual entran en relación un medio (un potencial energético) y una señal (o información). Desde esta perspectiva, una imagen no representa ni significa, sino que modula. La definición de información en el análisis de la ampliación moduladora de Simondon permite pensar la imagen no como el soporte de un mensaje o significación, sino como una señal incidente que activa un proceso de transformación energética en un sistema metaestable. Este giro puede ser leído desde la perspectiva enfatizada por Félix Guattari (2017, p. 216) de un análisis pragmatista de los signos en el capitalismo contemporáneo que integra tanto una semiótica significativa como una semiótica asignificante. La imagen opera como un “acto de imagen” que no comunica un mensaje a un receptor sino que gatilla una operación de cambio energético (que involucra un entramado de máquinas sociales, biológicas y técnicas). Además, el concepto de modulación utilizado por Deleuze en su análisis de la pintura y del diagrama ofrece pistas para una definición no-representacional de la imagen. Desde la perspectiva abierta por Deleuze (2014), la imagen opera como un dispositivo modulador que es anterior a la figuración y la semejanza (p. 144). En el ejemplo de *Deep Dream* de Google (ver Fig. 4), la figuración aparece solamente como un efecto de los procesos de modulación informática, una inferencia estadística que pone en relación la totalidad de las imágenes de perros utilizadas en el entrenamiento del algoritmo (que a su vez entran en relación estadística

con los píxeles que componen la imagen utilizada como base, en este caso, la imagen de las Torres Gemelas el 11 de septiembre de 2001). Se trata de una definición de imagen que no presupone necesariamente un ojo humano (presuposición que guía las definiciones fenomenológica, semiológica y hermenéutica de la imagen). Mientras en la perspectiva de la imagen como representación la política de la imagen dependía de su carácter ambivalente y flexible (Paglen, 2019), en la noción de imagen como modulación su dimensión política pasa por su capacidad para resolver una situación de disparidad y asimetría energética.

Por último, el texto ya referido sobre las sociedades de control utiliza la noción de modulación para definir una forma de poder más allá de la categoría de sujeto (Deleuze, 1999). El control contemporáneo no operaría entonces a través de un molde que se impone como norma o normalización de un cuerpo individual (como en el régimen disciplinar), sino a través de una modulación temporal que extrae de una población sus curvas de normalidad y anticipa sus posibles trayectorias (vectores). En el caso del régimen disciplinar, el poder define un molde y luego lo impone a cada cuerpo en vistas a generar un proceso de subjetivación que normaliza a cada individuo al interior de la masa. En el caso de las sociedades de control, no se genera un molde dado de antemano, sino que se utilizan técnicas estadísticas para modular los flujos sociales e identificar patrones que permitan predecir comportamientos. Mientras en el primer caso la multiplicidad se administra a través del binomio masa/individuo, en el segundo caso se trata de la producción de “dividuos” y “bancos de datos” cuyo efecto no es moldear cuerpos sino modular comportamientos predecibles (Deleuze, 1999, p. 280). Esto exige pensar el poder no desde la perspectiva del individuo moderno, sino desde nuevas formas de aquello que Antoinette Rouvroy ha llamado “gubernamentalidad algorítmica” (Rouvroy & Berns, 2013). Para Michel Foucault (2009), la gubernamentalidad consiste en el arte de dirigir fuerzas sociales para, de este modo, conseguir un determinado fin sin imponer sobre estas fuerzas un molde predefinido (como en la lógica disciplinar), sino a través del cálculo estadístico y de la reorientación de flujos sociales. Dicho en términos de Simondon, disciplinar es imponer moldes, mientras que gobernar es modular determinadas fuerzas sociales con el objetivo de conseguir una amplificación energética. De igual modo podemos plantear que la gubernamentalidad algorítmica consiste en el uso de tecnologías de aprendizaje maquínico para así efectuar una determinada modulación de energías sociales. Según Rouvroy & Berns (2013, pp. 169-172), esto se logra a través de un triple proceso: captura de datos, proce-

samiento de datos, y ejecución de una acción a partir de las predicciones generadas. Si volvemos al ejemplo del reconocimiento facial algorítmico (ver Fig. 5), el proceso de individuación (en este caso, la selección de una determinada emoción expresada por el rostro) opera gracias a un proceso de captura de datos y de su posterior procesamiento. La predicción generada no refiere a una singularidad propia del rostro individual, sino que es el resultado de un proceso relacional que involucra las imágenes faciales de miles o incluso millones de individuos.

La singularidad de los aparatos de aprendizaje maquínico en tanto aparatos de modulación es su capacidad para, en cada caso, actualizar no solo una energía potencial sino también su propia estructura de actualización. El modulador algorítmico está él mismo en constante proceso de modulación. Desde esta perspectiva es posible sugerir una doble correspondencia. Por un lado, entre régimen disciplinar, individuo y noción hilemórfica de imagen: la imagen en tanto significación apelaría a un sujeto que es individualizado a través de técnicas disciplinares. Por otro lado, la imagen en tanto información y modulación se corresponde con el régimen de control que opera en dimensiones pre- y supra-individuales. Estas imágenes no apelan ya a la significación para un sujeto, sino que dependen de una semiótica asignificante capaz de gatillar transformaciones energéticas específicas que articulan de manera indiferente cuerpos, afectos, máquinas técnicas, flujos naturales, y flujos informacionales (Lazzarato, 2014, pp. 27-28). Creemos que esta doble correspondencia permite establecer, a su vez, las bases conceptuales para comenzar a redefinir la relación entre imagen y poder en el contexto de la visión artificial.

REFERENCIAS

- Baudrillard, J. (1979). Hacia una crítica de la economía política del signo. *Crítica de la economía política del signo* (pp. 166-193). Buenos Aires: Siglo XXI.
- Celis, C. (2020). The Face Revisited: Using Deleuze and Guattari to Explore the Politics of Algorithmic Face Recognition. *Theory, Culture & Society* 37(1), 73-91.
- Crawford, K., & Paglen, T. (2019). *Excavating AI: The Politics of Training Sets for Machine Learning*. <https://www.excavating.ai/>
- Deleuze, G. (1999). *Conversaciones: 1972-1990*. Valencia: Pre-Textos.
- Deleuze, G. (2005). *El pliegue: Leibniz y el barroco*. Buenos Aires: Paidós.
- Deleuze, G. (2014). *Pintura, el concepto de diagrama*. Buenos Aires: Editorial Cactus.

- Foucault, M. (2009). *Seguridad, territorio, población*. Buenos Aires: Fondo de Cultura Económica.
- Fry, H. (2018). *Hello World: How to be Human in the Age of the Machine*. New York: W. W. Norton & Company.
- Greenfield, A. (2017). *Radical Technologies: The Design of Everyday Life*. London: Verso.
- Guattari, F. (2017). *La Revolución Molecular*. Madrid: Errata Naturae.
- Hui, Y. (2015). Modulation after Control. *New Formations*, 84/85, 74-91.
- Iliadis, A. (2013). A New Individuation: Deleuze's Simondon Connection. *Media Tropes* 4(1), 83-100.
- Kant, I. (2003). *Crítica de la razón pura*. Buenos Aires: Losada.
- Kurenkov, A. (2015). A 'Brief' History of Neural Nets and Deep Learning. <http://www.andreykurenkov.com/writing/ai/a-brief-history-of-neural-nets-and-deep-learning/>
- Lazzarato, M. (2014). *Signs and Machines: Capitalism and the production of subjectivity*. California: Semiotext(e).
- Mordvintsev, A. & Tyka, M. (2015). Inceptionism: Going Deeper into Neural Networks. *Google AI Blog*. <https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html>
- Paglen, T. (2019). Imágenes invisibles: tus fotografías te miran. *Revista La Fuga*, 22. <http://www.lafuga.cl/imagenes-invisibles/944>
- Parisi, L. (2016). Automated Thinking and the Limits of Reason. *Cultural Studies ↔ Critical Methodologies* 16(5), 471-481.
- Rancière, J. (1995). *El desacuerdo: Política y filosofía*. Buenos Aires: Nueva Visión.
- Rosenblatt, F. (1958). The Perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review* 65(6), 386-408.
- Rouvroy, A. & Berns, T. (2013). Algorithmic Governmentality and the Prospects of Emancipation. *Reseaux* 177(1), 163-196.
- Shannon, C. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal* 27, 379-423.
- Simondon, G. (2015). *La individuación a la luz de las nociones de forma y de información*. Buenos Aires: Editorial Cactus.
- Simondon, G. (2016). La amplificación en los procesos de información. *Comunicación e Información* (pp. 139-162). Buenos Aires: Editorial Cactus.
- Su, J. & Sakurai, K. (2019). One Pixel Attack for Fooling Deep Neural Networks. *arXiv*. <https://arxiv.org/pdf/1710.08864.pdf>
- Terranova, T. (2004). Communication beyond Meaning: On the Cultural Politics of Information. *Social Text* 80, 22(3), 51-73.
- Wiener, N. (1948). *Cybernetics, or control and communication in the animal and the machine*. Cambridge (MA): The MIT Press.