

ESTO NO ES UNA FOTO.

Por: Hito Steyerl. Revista Anfibia. 20/07/2018

La tecnología de las cámaras de los smartphones es diferente a la de las tradicionales: las lentes son diminutas y básicamente, una basura. Casi la mitad de los datos captados por el sensor de la cámara son ruido. El truco está en el algoritmo que limpia todo y distingue la imagen detrás del ruido. Un ensayo político sobre la imagen de la investigadora Hito Steyerl, que forma parte de Arte Duty Free, recientemente publicado por Caja Negra.

Hace un tiempo conocí a un programador extremadamente interesante, que estaba trabajando en una tecnología para cámaras de *smartphones*. Tradicionalmente, se piensa que la fotografía representa lo que está allí afuera por medio de la tecnología, idealmente a través de un vínculo indicativo. ¿Pero hoy en día esto es realmente cierto? El desarrollador me explicó que la tecnología de las cámaras de los teléfonos contemporáneos es bastante diferente a la de las cámaras tradicionales: las lentes son diminutas y básicamente son una basura, lo que significa que casi la mitad de los datos captados por el sensor de la cámara son, en realidad, ruido. El truco, entonces, está en el algoritmo que limpia el ruido o, más bien, que distingue la imagen detrás del ruido.

¿Pero cómo puede saber la cámara cómo hacerlo? Muy simple: escanea en su totalidad las otras imágenes guardadas en el teléfono o en nuestras redes sociales y examina a nuestros contactos. Analiza las fotos que ya tomamos, o aquellas asociadas a nosotros, e intenta identificar rostros y formas para volver a vincularlos con nosotros. Al comparar lo que nosotros y nuestra red ya fotografiamos, el algoritmo predice lo que podríamos haber querido fotografiar esta vez. Crea la imagen presente sobre la base de imágenes anteriores, en nuestra/su memoria. Este nuevo paradigma se llama fotografía computacional.^[1]

El resultado puede ser una imagen de algo que nunca existió, pero que el algoritmo piensa que nos podría gustar ver. Este tipo de fotografía es especulativa y relacional. Es un juego de probabilidades que apuesta a la inercia. Hace que ver

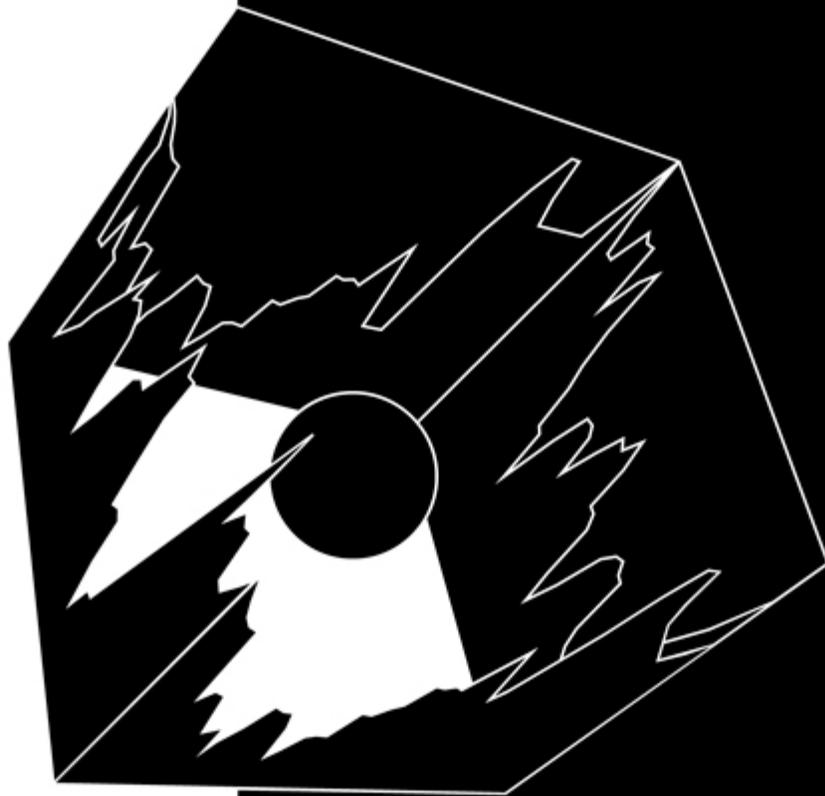
cosas imprevistas sea más difícil. A medida que aumenta la cantidad de ruido, aumenta también la cantidad de interpretaciones aleatorias.

Y ni siquiera mencionamos todavía las interferencias externas que pueden afectar lo que nuestros teléfonos están registrando. Existen todo tipo de sistemas que pueden prender o apagar nuestras cámaras remotamente: las empresas, los Gobiernos, los militares. Pueden ser desactivadas en ciertos lugares: se podrían, por ejemplo, bloquear sus funciones de registro cerca de protestas o, al contrario, transmitir todo lo que ven en ese entorno. De modo similar, un dispositivo podría ser programado para autopixelar, borrar o bloquear contenidos secretos, sexuales o con copyright. Podría combinarse con los así llamados *dick algorithm* [algoritmos de penes] para eliminar el contenido nsfw (*Not Safe/Suitable for Work*: no seguro/apropiado para el trabajo), modificar automáticamente el vello púbico, amoldar u omitir cuerpos, intercambiar o combinar contextos, o insertar anuncios geolocalizados, ventanas *pop-up* o *feeds* en vivo. Podría denunciarnos a nosotros o a alguien de nuestra red a la policía, a las agencias de relaciones públicas o a los generadores de *spam*. Podría señalar nuestras deudas, jugar nuestros juegos, transmitir nuestra frecuencia cardíaca. La fotografía computacional se ha expandido hasta cubrir todas estas posibilidades.

HITO STEYERL

ARTE DUTY FREE

El arte en la era de la
guerra civil planetaria



Este tipo de fotografía vincula a la robótica del control con el reconocimiento de objetos y las tecnologías del *machine learning*. De modo que, si tomamos una foto en un *smartphone*, el resultado no es tanto premeditado como premediado. La imagen quizás pueda mostrar algo inesperado, porque puede haber hecho referencias cruzadas entre muchas bases de datos diferentes: controles de tráfico, bases de datos médicas, galerías de fotos de nuestros *frenemies* en Facebook, datos de tarjetas de crédito, mapas y cualquier otra cosa que quiera.

Fotografía relacional

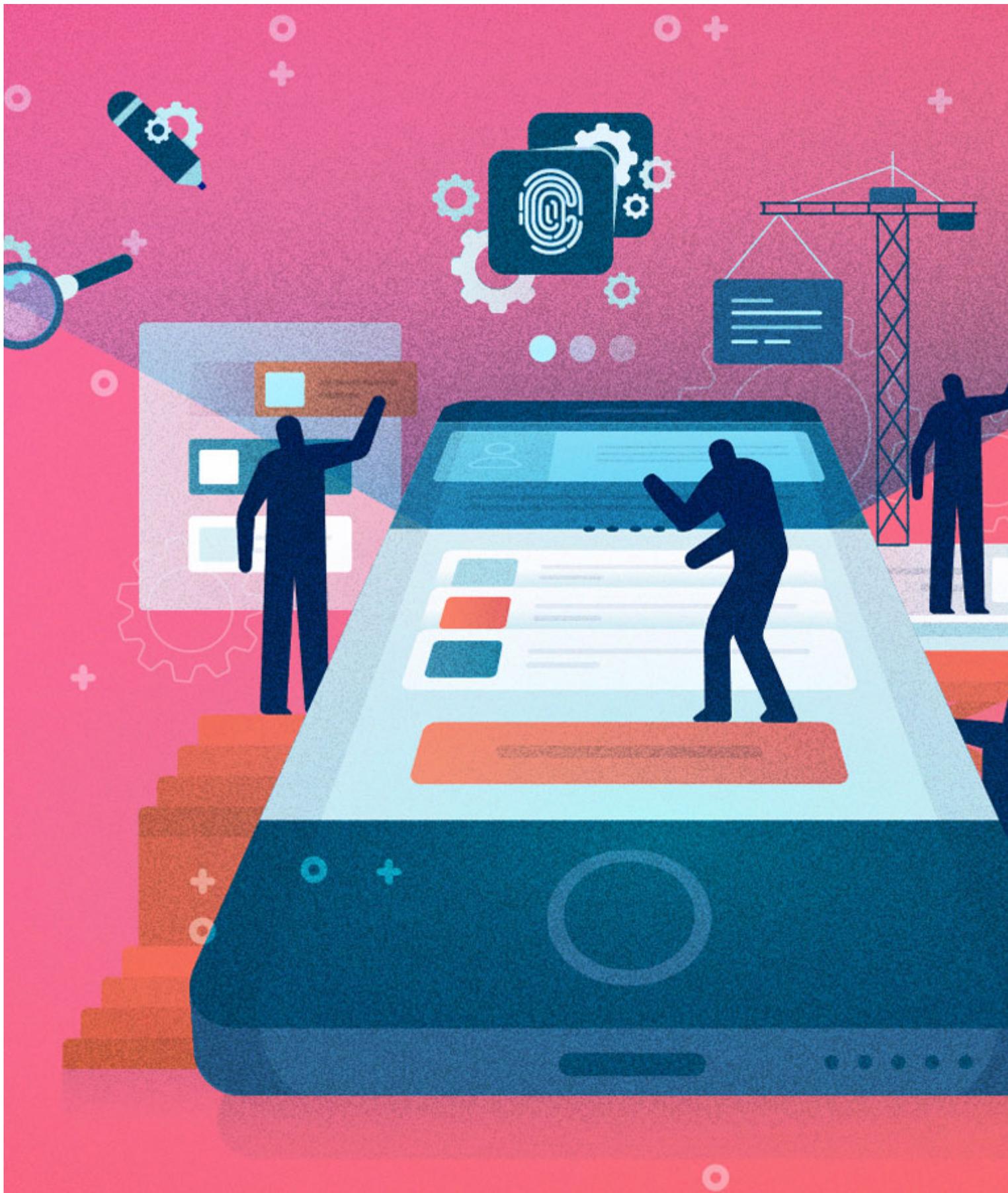
La fotografía computacional es por lo tanto intrínsecamente política, no en su contenido pero sí en su forma. No solo es relacional, sino que es realmente social: potencialmente, incontables sistemas y personas interfieren las imágenes antes incluso de que se vuelvan visibles.^[ii] Y por supuesto que esta red no es neutral. Tiene reglas y normas grabadas en sus plataformas, que representan una mezcla de parámetros y efectos jurídicos, morales, estéticos, tecnológicos y comerciales que son ocultados sin rodeos. Podríamos ser retocados, redireccionados, grabados, borrados, remodelados o reemplazados en nuestra propia imagen. La cámara se transforma en un proyector social más que en una grabadora. Nos muestra una superposición de cómo piensa que a nosotros nos gustaría vernos junto a lo que otros piensan que deberíamos comprar o ser. Pero la tecnología raramente hace cosas por su cuenta. La tecnología está programada por muchas entidades con metas opuestas, y la política es un problema de definir cómo separar en ellas el ruido de la información.^[iii]

¿Cuáles son entonces las políticas que ya existen para definir la separación entre ruido e información, o que incluso definen el ruido y la información como tales en primer lugar? ¿Quién o qué decide qué es lo que la cámara *verá*? ¿Cómo se realiza ese proceso? ¿Por quién o por qué? E, incluso, ¿por qué esto es importante?

El problema del pene

Examinemos un ejemplo: la separación entre la cara [*face*] y el culo [*butt*], o entre partes del cuerpo “aceptables” e “inaceptables”. No es coincidencia que Facebook se llame Facebook y no Buttbook, porque no se pueden publicar culos en Facebook. ¿Pero cómo se erradican entonces los culos? Un documento filtrado por un *freelancer*

indignado revela las precisas instrucciones para construir y mantener la cara de Facebook, y nos muestra lo que es bien sabido: que la desnudez y los contenidos sexuales están estrictamente prohibidos, excepto los desnudos artísticos y las tetillas masculinas; pero también que sus políticas contra la violencia son mucho más laxas, siendo aceptables incluso las decapitaciones y las grandes cantidades de sangre.^[iv] “Las cabezas, extremidades, etc., reventadas están permitidas mientras que no se muestren las entrañas”, se lee en una de las pautas. “Está permitido mostrar heridas profundas en la piel; está permitido mostrar sangre en exceso”. Estas reglas se hallan todavía controladas por humanos, o más precisamente, por una fuerza de trabajo global subcontratada en Turquía, Filipinas, Marruecos, México e India, personas que trabajan desde su casa y ganan alrededor de cuatro dólares por hora. Estos trabajadores son contratados para distinguir entre partes del cuerpo aceptables (la cara) e inaceptables (el culo). En principio, no hay nada malo en tener reglas para las imágenes que se encuentran disponibles públicamente. Algún tipo de proceso de filtrado tiene que ser implementado en las plataformas online; nadie quiere recibir en su casilla porno vengativo^[v] o atrocidades, independientemente de si existen mercados para ese tipo de imágenes. La cuestión concierne adónde y cómo se traza la línea, tanto como quién la dibuja, y de parte de quién. ¿Quién decide entre señal versus ruido?



Volvamos a la eliminación de contenidos sexuales. ¿Existe un algoritmo para ejecutarla, así como lo hay para el reconocimiento facial? Esta pregunta surgió públicamente por primera vez en el así llamado dilema de Chatroulette. Chatroulette era un servicio ruso de videos online que permitía que las personas se conocieran en la Web. Rápidamente se volvió famoso por su botón de “siguiente”, para el que el término “botón de disgusto” [*unlike button*] sería demasiado respetuoso. Al comienzo, la audiencia del sitio explotó hasta alcanzar 1,6 millones de usuarios mensuales en 2010. Pero luego emergió lo que se llamó “el problema del pene”, en referencia a las muchas personas desnudas que usaban el servicio para conocer a otras personas.^[vi] El ganador de un concurso online convocado para “solucionar” el problema ingeniosamente sugirió utilizar un reconocimiento facial veloz o un escaneo de seguimiento de ojos en los *feeds* de video: si no había ningún rostro discernible, deduciría que debía ser un pene.^[vii]

Exactamente el mismo flujo de trabajo fue utilizado también por el Servicio Secreto Británico, en su programa de espionaje Optic Nerve, que en secreto realizó una extracción masiva de capturas de cámaras web de usuarios. Los *feeds* de video de casi 1,8 millones de usuarios de Yahoo fueron interceptados para desarrollar tecnologías de reconocimiento facial y de iris. Pero, de un modo que no sorprende, resultó que alrededor del 7% del contenido no mostraba ningún rostro. Entonces –tal como fue sugerido en el caso Chatroulette– utilizaron escaneos con reconocimiento facial en todo el material y trataron de excluir los penes por no ser rostros. Pero no funcionó muy bien. En un documento filtrado, el Cuartel General de Comunicaciones del Gobierno admitió la derrota: “No existe una capacidad perfecta para censurar materiales que puedan resultar ofensivos”.^[viii]

Las soluciones subsiguientes fueron un poco más sofisticadas. La detección probabilística de pornografía calcula la cantidad de píxeles de tonos oscuros en ciertas regiones de la imagen, produciendo complicadas fórmulas taxonómicas como la siguiente:

*Si el porcentaje de píxeles color piel en relación con el tamaño total de la imagen es menor al 15%, entonces la imagen no es un desnudo. En caso contrario, siga con el próximo paso.

*Si el número de píxeles color piel en la región más grande de piel es menor al 35%

del total de piel contabilizada, el número de píxeles color piel en la segunda región más grande de piel es menor al 30% del total de piel contabilizada y el número de píxeles color piel en la tercera región más grande de piel es menor al 30% del total de piel contabilizada, la imagen no es un desnudo.

*Si el número de píxeles color piel en la región más grande de piel es menor al 45% del total de piel contabilizada, la imagen no es un desnudo.

*Si el total de piel contabilizada es menor al 30% del número total de píxeles en la imagen y el número de píxeles de piel dentro del polígono limitado es menor al 55% del tamaño del polígono, la imagen no es un desnudo.

*Si el número de regiones de piel es mayor a 60 y la intensidad promedio dentro del polígono es menor a 0,25, la imagen no es un desnudo.

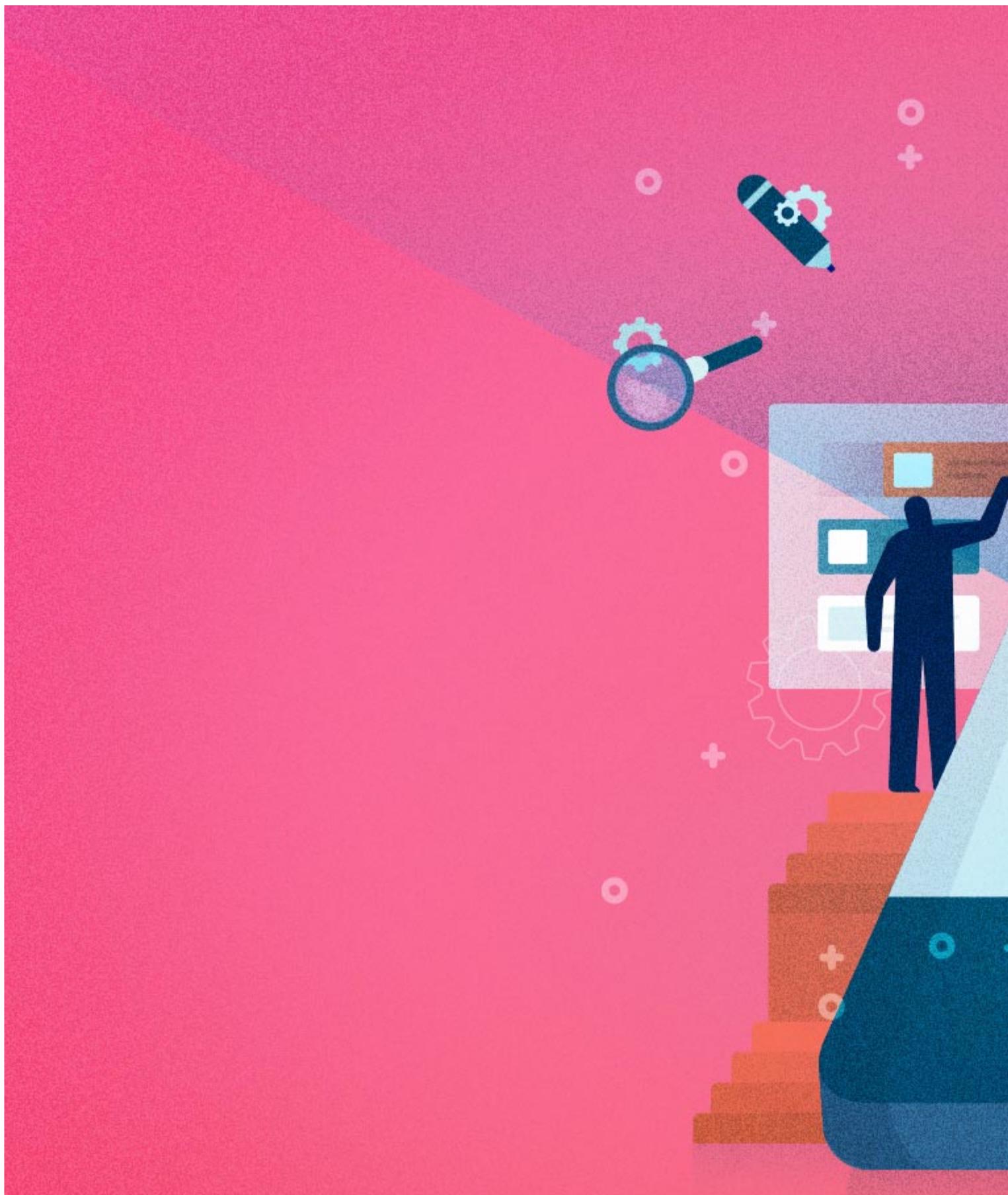
*En caso contrario, la imagen es un desnudo.[\[ix\]](#)

Pero este método quedó en ridículo bastante rápido porque produjo una gran cantidad de falsos positivos, incluyendo, en algunos casos, albóndigas envueltas, tanques y ametralladoras. Las aplicaciones más recientes de detección de pornografía utilizan tecnologías autodidactas basadas en redes neuronales, teorías verbales de computación y computación cognitiva. No intentan suponer la imagen estadísticamente, sino que más bien intentan entenderla por medio de la identificación de los objetos a través de sus relaciones.[\[x\]](#)

Según la descripción del desarrollador Tao Yang, hay todo un nuevo campo de estudios cognitivos de la visión que se basan en la cuantificación de la cognición en cuanto tal, para poder hacerla mensurable y computable.[\[xi\]](#) Si bien todavía existen dificultades técnicas considerables, este intento representa un nuevo nivel de formalización; un nuevo orden de imágenes, una gramática de las imágenes, un sistema algorítmico de la sexualidad, la vigilancia, la productividad, la reputación y la computación que se vincula con la gramaticalización de las relaciones sociales por parte de las corporaciones y los gobiernos.

¿Cómo funciona entonces? Para poder aprender a reconocer las partes objetables, el sistema de detección de pornografía de Yang debe observar una masa considerable de estas e inferir sus relaciones. Así que, básicamente, se comienza cargando una gran cantidad de fotos que contienen las partes del cuerpo que se

busca eliminar de la computadora. La base de datos consiste en carpetas llenas de estas partes del cuerpo, listas para iniciar el relacionamiento formal. No solo vaginas, pezones y *fellatios* sino también anos/solos y anos/mezclados_con_vaginas. Sobre la base de este catálogo, un gran espectro de detectores se preparan para el trabajo: el detector de pechos, el detector de vaginas, el detector de vello púbico, el detector de *cunnilingus*, el detector de mamadas, el detector de anos, el detector de manos-que-tocan-vaginas. Estos identifican posiciones sexuales fascinantes como las técnicas del Bostezo y el Pulpo, la Persuasión del Deudor, el Tocar el Cello y el Mirar el Partido.



Esta gramática, así como el catálogo de los objetos parciales, evocan la noción de “gramática pornográfica” de Roland Barthes, cuando describe los escritos del marqués de Sade como un sistema de posiciones y partes corporales listas para transformarse en todas las combinaciones posibles.^[xii] Sin embargo, este sistema marginal y abiertamente opresivo podría ser visto como el reflejo de una gramática del conocimiento más general desarrollada durante el así llamado Iluminismo. Tanto Michel Foucault como Theodor W. Adorno y Max Horkheimer han comparado los sistemas sexuales de Sade con los sistemas convencionales de clasificación. Ambos se articulan a partir del recuento y el ordenamiento, creando taxonomías tediosas, pedantes y exhaustivas. De modo similar, el entusiasmo del señor Yang por formalizar las partes del cuerpo y sus relaciones recíprocas da cuenta de un enorme esfuerzo por hacer que el conocimiento, las imágenes y el comportamiento sean cada vez más cuantificables y mensurables, para un sistema de valor de cambio basado en datos.

Por consiguiente, las partes del cuerpo no deseadas se transforman en elementos de una nueva gramática, legible por medio de máquinas y basada en la imagen, que usualmente opera en paralelo a las redes de reputación y control, pero que también puede vincularse con estas en cualquier momento. Su estructura podría ser el reflejo de los modos contemporáneos de recolección, acumulación y financiación del “conocimiento” basado en datos y producido en masa por una cacofonía de algoritmos parcialmente sociales incrustados en la tecnología.

El ruido y la información

Pero volvamos a la pregunta con la que comenzamos: ¿cuáles son los algoritmos políticos y sociales que separan a la información del ruido? El énfasis, nuevamente, está puesto en la política, no en los algoritmos. Jacques Rancière ha mostrado de un modo muy bello que esta división corresponde a una fórmula social mucho más antigua: la distinción entre el ruido y el habla, establecida con el fin de dividir a la masa entre la turba y los ciudadanos.^[xiii] Si quisiéramos no tomarnos a alguien seriamente, o limitar sus derechos o su estatus, actuaríamos como si su discurso fuera tan solo ruido, un quejido incoherente o un llanto, como si esa persona careciera de razón y por lo tanto estuviera exceptuada de ser un sujeto, mucho menos poseedora de derechos. En otras palabras, esta política descansa en un acto de decodificación consciente, que separa el “ruido” de la “información”, el “habla” del

“quejido” o la “cara” del “culo”, y que a partir de allí organiza cuidadosamente los resultados en jerarquías y clases verticales.^[xiv] Los algoritmos incluidos hoy en la tecnología de las cámaras de los *smartphones* para definir a la imagen antes de su aparición actúan de un modo similar.

A la luz de las proposiciones de Rancière, quizás todavía estemos lidiando con una idea más tradicional de la política como representación.^[xv] Si todos somos representados auráticamente (o visualmente), y nadie es descartado como ruido, entonces la igualdad podría estar cerca. Sin embargo, las redes han cambiado tan drásticamente que casi todos los parámetros de la política representativa se han desplazado. En la actualidad, muchas más personas de las que hubo nunca en la Web pueden subir un número casi ilimitado de tales representaciones. Pero, mientras tanto, el nivel de participación política a través de la democracia parlamentaria parece haber disminuido. Al tiempo que las imágenes flotan cuantiosamente, las elites contraen y centralizan el poder.

Como si esto fuera poco, nuestras caras están desconectándose, no solo de nuestros culos, sino también de nuestras voces y cuerpos. Nuestras caras son ahora un elemento: una cara/mezclada_con_el_teléfono, lista para combinarse con cualquier otro ítem del catálogo. Si es necesario se agregan leyendas o texturas. Se toman impresiones de las caras. Una imagen se vuelve menos una representación que una delegación [*proxy*], una mercenaria de la apariencia, una textura-superficie-mercancía flotante. Las personas son montadas, mezcladas, ensambladas, incorporadas. Los humanos y las cosas se entremezclan en constelaciones siempre nuevas para transformarse en bots o cyborgs.^[xvi] Mientras que los humanos proveen de afectos, pensamientos y sociabilidad a los algoritmos, estos a su vez les proveen de lo que solía llamarse subjetividad. Este desplazamiento es lo que dio lugar a una política posrepresentacional a la deriva en el espacio de la información.^[xvii]

^[i] Daniel Rubinstein y Katrina Sluis, “Notes on the Margins of Metadata: Concerning the Undecidability of the Digital Image”, en *Photographies*, vol. 6, nro. 1, 2013, pp. 151-158. Ver también los escritos y entrevistas de Katrina Sluis sobre esta noción.

^[ii] Sobre la política inherente a las nociones de ruido e información, ver Tiziana Terranova: “La política cultural de la información implica un regreso a las condiciones mínimas de la comunicación (la relación de la señal con el ruido y el

problema de hacer contacto)” (*Network Culture: Politics for the Information Age*, Londres, Pluto, 2004, p. 10).

[iii] Este es el problema que dio lugar a la teoría de la información en un artículo seminal de Claude Shannon publicado en 1948. Que por supuesto incluye la intención de resolver cómo vincular en redes y modular estos parámetros a través de muchas plataformas diferentes. Ver C.E. Shannon, “A Mathematical Theory of Communication”, en *Bell System Technical Journal*, vol. 27, nro. 3, julio de 1948, pp. 379-423, y vol. 27, nro. 4, octubre de 1948, pp. 623-656.

[iv] Adrian Chen, “Inside Facebook’s outsourced anti-porn and gore brigade, where ‘camel toes’ are more offensive than ‘crushed heads’”, disponible online en *gawker.com*, 16 de febrero de 2012.

[v] *Revenge porn*: llamado en español “porno vengativo” o “pornovenganza”, se trata de la publicación de imágenes pornográficas sin el consentimiento de la persona involucrada, con el objeto de perjudicarla. [N. del T.]

[vi] Brad Stone, “In airtime video chat reboot, nudists need not apply”, disponible online en *www.bloomberg.com*, 5 de junio de 2012.

[vii] Nicholas Carlson, “Here’s THE solution to Chatroulette’s penis problem”, disponible online en *businessinsider.com*, 8 de abril de 2010 [última consulta: marzo de 2018].

[viii] Spencer Ackerman y James Ball, “Optic Nerve: millions of Yahoo webcam images intercepted by gchq”, disponible online en *The Guardian*, 28 de febrero de 2014.

[ix] Rigan Ap-apid, “An Algorithm for Nudity Detection”, en *wenku.baidu.com*.

[x] Software para la detección de pornografía en videos e imágenes del Yang’s Scientific Research Institute, Estados Unidos, disponible en *yangsky.com*.

[xi] Tao Yang, “Applications of Computational Verbs to Effective and Realtime Image Understanding”, en *International Journal of Computational Cognition*, vol. 4, nro. 1, 2006.

[xii] “El sistema de Sade (según Barthes), como un lenguaje, tiene su propia gramática (‘una gramática pornográfica’), que consiste en algunos elementos básicos. La postura sexual es el principal, y los otros son: sexo, masculino o femenino; posición social; ubicación, por ejemplo, convento, calabozo, ¡incluso habitación!, etc. Sade combina estos elementos en todo tipo de transformaciones *exhaustivas*, para elaborar un conjunto de posibilidades por completo sustancioso” (Girish Shambu sobre *Sade Fourier Loyola*, de Barthes, en *girishshambu.blogspot.de*).

[xiii] Jacques Rancière, “Ten Theses on Politics”, en *Theory & Event* vol. 5, nro. 3, 2001. [Edición en español: “Diez tesis sobre política”, en *Política, policía, democracia*, Santiago de Chile, LOM, 2006].

[xiv] Y también todo tipo de otras jerarquías, obviamente.

[xv] Rancière articuló esta idea por primera vez en *La mesentente*, París, Galilée, 1995. Desde entonces, las políticas del sonido y la imagen se han transformado bastante dramáticamente a causa el desarrollo de los medios sociales basados en la Web.

[xvi] En la legendaria descripción de Donna Haraway, un cyborg es un organismo cibernético, un híbrido de máquina y organismo, una criatura tanto de la realidad social como de la ficción. Ver su “Manifiesto para ciborgs: ciencia, tecnología y feminismo socialista a finales del siglo xx”, en *Ciencia, ciborgs y mujeres. La reinvencción de la naturaleza*, Madrid, Cátedra, 1991, y editado también de forma independiente como *Manifiesto para cyborgs*, Mar del Plata, Puente Aéreo, 2014.

[xvii] Terranova distingue en *Network Culture* entre los espacios representacional e informacional.

[LEER EL ARTÍCULO ORIGINAL PULSANDO AQUÍ.](#)

Fotografía: Revista Anfibia

Fecha de creación

2018/07/20