

[Centro de Información de COVID \(CIC\): Charlas científicas relámpago](#)

Transcripción de una presentación de George K. Thiruvathukal (Loyola University Chicago), 22 de septiembre de 2021



Título: [RAPID: Aprovechando nuevas fuentes de datos para analizar el riesgo de COVID-19 en lugares concurridos](#)

Subvención de La Fundación Nacional de Ciencias (NSF, por sus siglas en inglés) # [2027524](#)

[Grabación de YouTube con diapositivas](#)

[Septiembre 2021 Información del seminario web del CIC](#)

Editora de la transcripción: Macy Moujabber

Traductora: Isabella Graham Martínez

Transcripción

Diapositiva 1

Muy bien, muchas gracias. En primer lugar, cuando te presentan como un orador fantástico, es algo difícil estar a la altura, pero gracias por esa presentación tan amable a todos nosotros, Lauren. Y voy a zambullirme en esta charla en el espíritu de las charlas relámpago. Por lo tanto, voy a hablar de "Observando la movilidad humana durante COVID-19."

Diapositiva 2

Muy bien y nuestro trabajo es acerca de ser capaz de aprovechar lo que llamamos cámaras públicas, cámaras de red pública, y analizar esta gran- esta gran cantidad de datos visuales capturados de estas cámaras para ver cómo la gente está siguiendo el distanciamiento social durante varias etapas de los cierres en varios lugares. En realidad, hemos analizado en este estudio piloto en sólo cinco países y tres estados para analizar la eficacia de estas políticas de bloqueo para ver si podríamos- si esto puede ser una herramienta, en la caja de herramientas, para ayudar a reducir la propagación de COVID-19. Y el principal desafío del estudio es que tenemos millones de imágenes, de hecho, al día de hoy tenemos probablemente cerca de 7 millones de imágenes. Y luego lo que queremos hacer es ser capaces de tomar esas imágenes, y queremos ser capaces de ver dónde se encuentran estas imágenes y básicamente mirar las políticas que están en esas ubicaciones y ser capaces de ver si hay algún patrón en los datos. Por lo tanto, este es un gran proyecto de datos y quiero decir que normalmente le digo a la gente: no intente esto en su computadora portátil en casa. Es probable que no funcione, y usted puede sobrecalentar su computadora dentro de unos minutos del estudio.

Diapositiva 3

Así que, sin más preámbulos. Así que, básicamente, esta diapositiva muestra sólo algunas muestras del tipo de imágenes que tenemos disponibles. Todo esto viene de nuevo de las cámaras públicas que están disponibles en la red, básicamente, típicamente, como los gobiernos

y los individuos u otras organizaciones como los parques nacionales tienen estas cámaras y las incrustan en sus páginas web. Quieren ponerlos a disposición de otros, no necesariamente para investigación, pero es público- son cámaras públicas. Así que, a diferencia de la red masiva de cámaras de red cerradas que existen en muchos municipios, por ejemplo, estas son fuentes que pueden ser legítimamente utilizadas para el estudio sin plantear problemas significativos de privacidad y seguridad. Muy bien, así que en nuestro proyecto de investigación global, que es el proyecto Cam Squared, que había mencionado en la diapositiva del título, pero no dije mucho sobre ello por el interés de mantener el tiempo, en realidad tenemos un proceso para descubrir estas cámaras. Tenemos papeles sobre ese tema. Y el estudio que hemos hecho es básicamente donde capturamos datos continuamente de todas las cámaras de nuestra red de cámaras, así que unas 36.000 cámaras que hemos descubierto a través de este proceso automático que creamos y lo que— recopilamos datos entre abril de 2020 y marzo de 2021. Al igual que muchas personas que trabajan en COVID pensamos al principio, tal vez va a ser un corto episodio y luego estamos, al igual que, bueno, esto parece seguir así que tal vez deberíamos empezar algunos trabajos de supercomputación sólo para recoger datos de toda nuestra red de cámaras y ver qué pasa. Quizás lo analicemos en algún momento y eso es lo que hicimos. Entonces, esto es usar una gran supercomputadora en el Laboratorio Nacional Argonne donde tengo una cita de visita, y tenemos acceso a algunos recursos increíbles allí, pero nuestra red de almacenamiento es capaz de almacenar petabytes de datos. Acaba de ser actualizado a 200 petabytes. Eso es más de lo que la mayoría de la gente puede imaginar qué hacer, pero fuimos capaces de recopilar unos 70 terabytes de datos sólo para las cámaras que estábamos viendo aquí. Vale, de nuevo, es un poco difícil poner eso en tu portátil. Entonces, lo que hicimos fue básicamente tener un trabajo que corría todos los días, normalmente unas cinco o seis veces al día, y luego tomar una captura de todas estas cámaras simultáneamente usando un trabajo de computación de clústeres. Se tarda unos 30 minutos para recoger datos simultáneamente de todas estas cámaras, casi simultáneamente es una mejor manera de decirlo. Pero eso es algo bastante extraordinario en sí mismo. Es sólo haber demostrado cómo podemos recopilar tanta información visual en las súper computadoras. Luego, lo que hicimos fue examinar el número de seres humanos y vehículos a lo largo del tiempo porque queríamos analizar la movilidad humana de manera más general, no sólo los peatones, sino también las personas que se suben a los coches de nuevo, las personas que no están en los coches y otros vehículos. Y luego también vimos cómo esto se correlaciona con el Índice de Rigor de Oxford, que es un índice que mide lo que está pasando con la política en varios lugares. El remate de nuestro estudio es que creemos que los datos visuales, especialmente en futuras pandemias, serán un método que se utilizará porque ya se está demostrando que es efectivo para poder entender las políticas que les mostraré en algunos gráficos hacia el final.

Diapositiva 4

Vale, esto es solo un poco sobre el flujo de trabajo que usamos para esto. Ya mencioné el proceso de descubrimiento de la cámara, Y luego lo que hicimos fue aplicar un par de filtros sólo para asegurarnos de que realmente nos estamos enfocando en cámaras que realmente van a ayudar en el estudio porque hay muchas cámaras a las que tenemos acceso que son poco interesantes. Este es un ejemplo de una cámara sin interés. Habíamos encontrado una cámara que en realidad se utiliza para monitorear como una señal de cierre de carretera para asegurarse de que las luces siguen parpadeando. Eso no nos va a decir mucho sobre la movilidad humana. Y luego, así que miramos- hicimos algunos pre-procesamiento sólo para averiguar cuál de estas cámaras nos están dando cualquier tipo de datos relacionados con la movilidad, así que esto podría ser- vemos a los seres humanos allí, vemos vehículos allí, etc. Muy bien, y luego después de encontrar las cámaras que realmente cumplen con estos requisitos, entonces realmente pasamos y hacemos el análisis aquí. Zoom está bloqueando mi última parte del diagrama, pero las partes clave son, en primer lugar, encontrar los modelos relevantes. Tenemos dos modelos diferentes que estamos

usando. Uno de ellos se llama Pedestron, que es un modelo probado en el tiempo para poder mirar el tráfico de tipo peatonal y otro tráfico humano, tal vez no siempre un peatón caminando como en un paso de peatones, es como en varios entornos. Y luego tenemos como YOLOv3 que es un detector general de objetos. Lo usamos para- y lo hemos entrenado para poder mirar vehículos. Y luego hay algo de inteligencia incorporada en la agregación de los datos por ubicación. Básicamente usamos un número de servicios de geolocalización para poder hacer eso, y así es como pudimos llegar a estos análisis por país o por estado. Bien, por supuesto que el resultado final es que queremos hacer unos gráficos bonitos, y espero que sean lo suficientemente bonitos para esta presentación, todavía estamos finalizando eso.

Diapositiva 5

De todos modos, lo único que hacemos es, tenemos cámaras en muchos lugares. Así que, en muchos casos podemos tomar las coordenadas geográficas que tenemos para estas cámaras y luego podemos corroborar eso con lo que Google Street View nos está diciendo y hay alguna participación humana en este proceso. Así que, en realidad teníamos que saber que para nuestras cámaras representativas que miramos en realidad, correlacionarlos con Google Street View y como se puede ver aquí como una de estas cámaras se muestra- sí se muestra el nuestro bien y otro muestra de Google. Bueno, y vemos algunos de los mismos signos- no son las mismas palabras en los signos, pero definitivamente vemos la misma colocación de los signos allí y tenemos mucha confianza en que esa cámara está donde creemos que está.

Diapositiva 6

La otra cosa que- para hacer este tipo de trabajo requiere que usted establezca un conjunto de datos de validación que se va a utilizar para ayudarnos a hacer un poco mejor análisis con ambos de los modelos que le mencioné. Así que, lo que hicimos fue tomar algunas de nuestras, un subconjunto de nuestras imágenes y sólo las usamos para proporcionar un poco mejor etiquetado de - ya saben, que esto es un peatón, esto es un coche, y así sucesivamente. Y uno de los desafíos que enfrentamos en nuestro conjunto de datos es que debido a que estas son cámaras públicas, son como ángulos aleatorios, están como colocados en todas partes como están, algunos de ellos están colocados lejos, algunos están cerca. Y por lo tanto, en realidad requiere que mires diferentes escenas y hagas algo entrenando allí para asegurarte de que cuando realmente miras cualquier imagen en cualquiera de nuestros conjuntos de datos, tienes una mayor probabilidad de ser capaz de clasificar con precisión.

Diapositiva 7

Una cosa, creo que no sé qué pasó aquí. No mencioné una cosa; mencioné las escenas seleccionadas, pero no tuve la oportunidad de mencionar brevemente que se puede ver aquí en la parte inferior del conjunto de imágenes. Estos son a los que estamos aplicando nuestros detectores de objetos y nosotros básicamente- estos detectores de objetos nos van a dar típicamente- el verde es el falso- son los positivos y luego tenemos el azul que está mostrando algunos que son falsos positivos, y el rojo son los negativos. Porque pueden ver que como estos detectores de objetos son muy buenos en la clasificación. Esto es sólo mostrando a los peatones aquí, pero hacen un muy buen trabajo para identificar a los peatones, incluso como objetos distantes, lograron ser clasificados con bastante precisión como peatones- bien, o personas.

Diapositiva 8

Muy bien, así que de todos modos, ¿cuáles son los hallazgos clave? Esto es lo que quiero pasar un poco de tiempo hablando-bien. Voy a- Acabo de dejar caer algo en el suelo. Bueno, los hallazgos clave son que esto funciona muy bien. Así que, quería mostrarles un poco sobre cómo se organizan estos gráficos aquí. Así que básicamente, tenemos una apertura como esta cuando hay una apertura sucediendo. Aquí es cuando hay un bloqueo así que por desgracia hay algo de

pixelación sucediendo aquí en cuando fui de . pdf a . png pero esta es una apertura en Francia tan O de Francia, bloqueo en Francia, y luego otra apertura en Francia. Y está claro que, y especialmente en lugares donde había políticas consistentes sobre abrir y cerrar, definitivamente vemos que conoces el aumento en la actividad que ocurre tanto en el tráfico vehicular como en el peatonal y las cámaras que tenemos en esos lugares, Está bien, realmente hace cosquillas. Por supuesto, a medida que las malas noticias comienzan a marcar, empezamos a ver a la gente incluso retroceder antes del encierro que tiene lugar, por ejemplo, en Francia.

Bien, todos sabemos que los EE.UU. es una historia interesante y una de las cosas que es bastante fascinante y todo lo que— especialmente nuestros gráficos de lugares como Georgia, podemos ver que como porque el índice de rigor es bastante plano aquí, bueno, lo siento. He perdido un poco de mi, así que puedes ver como hay un cierre y una apertura en Georgia y luego esa apertura simplemente continúa para siempre, y por supuesto, la actividad que es en realidad vemos como especialmente como la actividad vehicular sigue aumentando como, probablemente, como la economía se está abriendo, pero que lo ves de otra manera un tipo de historia inconclusa allí. Pero definitivamente cuando se trata de lugares con políticas consistentes, y por supuesto, Europa ha sido realmente bastante buena en este sentido, vemos que los patrones de aperturas, cierres, y si hay como una segunda apertura que sucedió, Bueno, no tuvimos una reapertura a tiempo para Alemania en el momento de este estudio. Sí, obtenemos todo esto de lo que se ha publicado sobre la apertura de los cierres especialmente. Pero sí, los países europeos están bastante bien. Vemos que Australia y luego los estados de EE.UU. son un poco menos consistentes. Hawaii en realidad todavía estamos tratando de entender más acerca de lo que está pasando allí, pero una cosa que sabemos es que en Hawaii había un número de personas que viajan a Hawaii y estaban básicamente siguiendo diferentes reglas que la mayoría de las personas que vivían en realidad en Hawaii. El patrón se parece mucho a Georgia pero tal vez un poco mejor que Georgia en términos de lo que realmente está pasando. Por lo tanto, la conclusión es que al menos cuando vemos que, bueno, política coherente, somos capaces de ver una conexión bastante fuerte al analizar los datos visuales de lo que realmente está pasando tanto con el tráfico peatonal y vehicular. Ciertamente, la esperanza es que en el futuro habrá un poco más de consistencia cuando se trata de las pólizas de COVID-19 y- esta es una de las razones por las que emprendimos nuestro estudio, en primer lugar es que queríamos tener una idea de cómo las personas están realmente respondiendo a las pólizas que están presentes. Y por supuesto, sé que voy a tener una pregunta sobre si hemos analizado otros estados. Sí, lo hemos hecho, pero también queríamos asegurarnos de que teníamos suficientes datos, especialmente datos visuales, para las ubicaciones que estábamos analizando porque algunos estados, por ejemplo, en los Estados Unidos no tenemos tantas cámaras como otros.

Diapositiva 9

De todos modos, sólo quiero decir un poco sobre que tenemos un equipo de investigación muy grande. Este es un equipo que comprende probablemente- hemos tenido un tamaño variable. Tenemos cerca de 15 estudiantes universitarios, hemos tenido tres estudiantes de posgrado involucrados, y por supuesto tenemos cuatro profesores involucrados: yo y Yung-Hsiang Lu fuimos los líderes de la facultad para este trabajo. Y tenemos a David Shoham, quien también ha colaborado con nosotros del Departamento de Salud Pública. Estuvo en Loyola y se ha mudado a la Universidad Estatal de Tennessee Oriental, y Wei Zakharov de Ciencias de la Información. Así que no sería posible hacer un estudio como este sin estos estudiantes y profesores de gran talento, así que sólo quiero asegurarme de que los reconozco.