



Transformar Datos de Observación de la Tierra en Conjuntos de Datos de Infraestructura Construida para la Modelación del Riesgo de Desastres
3 - 10 de octubre 2023

2^{nda} Sesión: Preguntas y Respuestas

Por favor escriba sus preguntas en el cuadro para preguntas. Si tiene preguntas adicionales, por favor comuníquese con cualquiera de los siguientes instructores: Marina Mendoza (mtm@imagecatinc.com), Brock Blevins (brock.blevins@nasa.gov)

Pregunta 1: Profe y que diferencia hay con el Google Earth o algo asi donde puedo ver desde arriba las calles y edificios? ¿Qué diferencia tienen los datos presentadosn (NYS FIDSS) con los datos de huella de edificio Google?

Professor, what's the difference with Google Earth or something like that where I can see the streets and buildings from above? How is the data presented (NYS FIDSS) different from Google building footprint data?

Respuesta 1: El beneficio de usar Google fue el use del Google Street view que nos dejo ver la fachada lateral del edificio que nos dio mas informacion sobre la condicion y la altura del edificio, tambien como los desagües de agua (alcantarillas). NYS FIDSS incluye informacion sobre infraestructura critica y cataloga todos los edificios. El proceso de Mechanical Turking es el mismo para los dos, pero diferente informacion se collecciono para los dos casos. En NYS fue el estado completo, y el trabajo con Google empezo con puntos aleatorios para encontrar los ejemplos y formular un analysis de precision.

Pregunta 2: (Transformar Datos de Observación de la Tierra en Conjuntos de Datos de Infraestructura Construida) Que porcentaje de precisión tiene esa base de datos de huellas de edificaciones.

What percentage of precision does that database of building footprints have?

Respuesta 2: This is a great question. There are two components of the building footprint extraction workflow: the CNN-model, and the VVM post-processing that removes false positives and performs regularization. Our CNN-model yields precision-recall metrics around 85-90% for evaluations at pixel-level. The VVM



operates around 96% accuracy: it removes 96% of false positives, while very minimally reducing recall (true positives).

[Translation] Esta es una excelente pregunta. Hay dos componentes del flujo de trabajo de extracción de la huella de un edificio: el modelo CNN y el posprocesamiento VVM que elimina los falsos positivos y realiza la regularización. Nuestro modelo CNN produce métricas de recuperación de precisión de alrededor del 85-90% para evaluaciones a nivel de píxel. El VVM opera con una precisión de alrededor del 96%: elimina el 96% de los falsos positivos, al tiempo que reduce mínimamente el recuerdo (verdaderos positivos).

Pregunta 3: Que significa los puntos naranjas?

What do the orange dots mean?

podemos usarlos para actividad volcanica calentamiento de océanos?

Can we use them for volcanic activity, ocean warming?

Respuesta 3: Por favor aclare la pregunta. ¿Tal vez sea el mapa de colores de probabilidad Verdadero Positivo, que va del rojo al azul?)

El flujo de trabajo que presentamos solo está destinado al análisis de edificios.

Please clarify the question. Maybe it's the True Positive probability colormap, that ranges from red to blue?)

The workflow we presented is only intended for building analysis

Pregunta 4: Buenas tardes, en que programa se utiliza el mechanical turk ?

Good afternoon, in which program is mechanical turk used?

Respuesta 4: El Turk mecánico abierto (localturk) se ejecuta dentro de Node.js (<https://nodejs.org/>). Si instala node.js, puede agregar turk local con el comando en la página del repositorio de GitHub (<https://github.com/danvk/localturk>) y luego ejecutar el proceso con uno de los comandos que se muestran. Amazon Turking se ejecuta en su plataforma, debe configurar una cuenta (<https://www.mturk.com/>) y Amazon ejecuta el software; usted lo configura todo a través de su navegador.



The open mechanical Turk (localturk) is run within Node.js (<https://nodejs.org/>). If you install node.js you can add local turk with the command on the GitHub repo page (<https://github.com/danvk/localturk>) and then run the process with one of the commands shown. Amazon turking is run on their platform, you need to set up an account (<https://www.mturk.com/>) and Amazon runs the software, you do configure it all via your browser.

Pregunta 5: **Buenas tardes, que porcentaje de validación in situ es necesaria para evitar falsos positivos en la realización de los inventarios urbanos?**

Respuesta 5: Es muy difícil eliminar los falsos positivos. Si desea limitar el error a un cierto nivel, es mejor calcular un tamaño de muestra basado en una fórmula relativamente simple:

$$n = Z^2 p (1 - p) / m^2$$

n: tamaño de la muestra

Z: nivel de confianza, también conocido como valor z

p: desviación estándar o proporción poblacional (generalmente 0,5 si no se conoce)

m: margen de error (+/-)

Hay calculadoras en línea que simplifican esto: simplemente puede elegir el nivel de confianza deseado (por ejemplo, 95 %) y el margen de error para la medición (por ejemplo, +/- 5 %) y arrojará un tamaño de muestra para una población que es ilimitado. Ver: <https://www.calculator.net/sample-size-calculator.html>

It's very hard to eliminate false positives. If you want to limit error to a certain level, it's best to calculate a sample size that's based on a relatively simple formula:

$$n = Z^2 p (1 - p) / m^2$$

- n: sample size
- Z: confidence level, aka z-value
- p: standard deviation or population proportion (generally 0.5 if not known)
- m: margin of error (+/-)



Transformar Datos de Observación de la Tierra en Conjuntos de Datos de Infraestructura Construida para la Modelación del Riesgo de Desastres
3 - 10 de octubre 2023

There are online calculators that simplify this, you can just choose a desired confidence level (e.g., 95%) and margin of error for the measurement (e.g., +/- 5%) and it will spit out a sample size for a population that is unlimited. See:

<https://www.calculator.net/sample-size-calculator.html>