



Reporte de avances

# Estudios para la sistematización y valoración de información de infraestructura

---

JUNIO 2026

COORDINADORES

José Alejandro Ruiz Sánchez

Elio Atenógenes Villaseñor García

COLABORADORES

María Fernanda Benítez Banegas

Alejandra Figueroa Martínez

Irving Gibran Cabrera Zamora

Ranyart Rodrigo Suárez Ponce de León



# Antecedentes

Proyecto de investigación que contribuye a la sistematización de la información contenida en la Plataforma de Infraestructura Pública Nacional y su posterior aprovechamiento para valoraciones económicas de la infraestructura.

Específicamente, el proyecto:

- 1) Apoya procesos de validación e identificación de direcciones.
- 2) Apoya la validación y sistematización de información compartida por Unidades del Estado.
- 3) Identifica información relevante para las estimaciones económicas de infraestructura de autopistas.
- 4) Identificación de plantas de tratamiento y subestaciones eléctricas a través de la clasificación de objetos en imágenes satelitales.

## TEMA 1

---

# Generación de herramientas para el procesamiento, validación e integración de información



DETECTAR VIGENTES

VALIDAR COORDENADAS



TRANSFORMAR ATRIBUTOS

POST /DETECTAR-VIGENTES

## ARCHIVOS DE ENTRADA



Archivo A

Referencia principal  
(Base)



Archivo B

Archivo a comparar  
(Nuevo)

## PARÁMETROS OPCIONALES

ID\_COLUMN\_A

Ej: RFI

Columna ID del Archivo A (auto si vacío)

ID\_COLUMN\_B

Ej: RFI

Columna ID del Archivo B (auto si vacío)

STATUS\_COLUMN

ESTATUS\_RFI

Columna de estatus a generar (use el valor por defecto si se deja vacío)

Procesar archivos →

Procesando archivos, por favor espera...

# Integración Bases de datos

Consiste en generación de herramientas que apoyan en la integración de bases de datos. Con el fin de unificar la información de manera ágil.

DETECTAR VIGENTES

VALIDAR COORDENADAS



TRANSFORMAR  
ATRIBUTOS

POST /VALIDAR-COORDENADAS

ARCHIVO DE ENTRADA



Excel

Archivo con columnas de latitud y longitud

PARÁMETROS OPCIONALES

LAT\_COLUMN

Ej: Latitud

Columna de latitud (auto si vacío)

LOK\_COLUMN

Ej: Longitud

Columna de longitud (auto si vacío)

ENTIDAD\_COLUMN

Ej: ENTIDAD

Columna de entidad federativa (opcional)

MUNICIPIO\_COLUMN

Ej: MUNICIPIO

Columna de municipio (opcional)

BUFFER\_KM

10

Buffer en los paises validación de país (por defecto: 10 km)

Validar coordenadas →

Validando coordenadas, por favor espera...

## Validación y geocodificación

Se desarrolla una metodología para la validación de geolocalización contrastando la información recibida y el Marco Geoestadístico. De esta manera ratificar la información existente. A su vez, permite conseguir coordenadas a partir de direcciones y viceversa

## TEMA 2

---

# Metodología para la estimación de costos de carreteras



# Regresión econométrica

La metodología consiste en una regresión con base en la información de documentos de la Secretaría de Infraestructura, Comunicaciones y Transporte (SICT) que brinda **los costos paramétricos para la construcción de los distintos tipos de caminos del país.**

Para los años: 2020, 2021, 2023, 2024, 2025 y 2026

$$\hat{C} = \beta_0 + \beta_1 \text{tipo} + \beta_2 \text{carpeta} + \beta_3 \text{cuerpo} + \beta_4 \text{region} + \beta_5 \text{orografía} + \beta_6 \text{año} + \epsilon$$

Donde:

$\hat{C}$  : costo estimado de tramo carretero

*tipo*: clasificación de carretera

*carpeta*: carpeta de tipo asfáltica o hidráulica

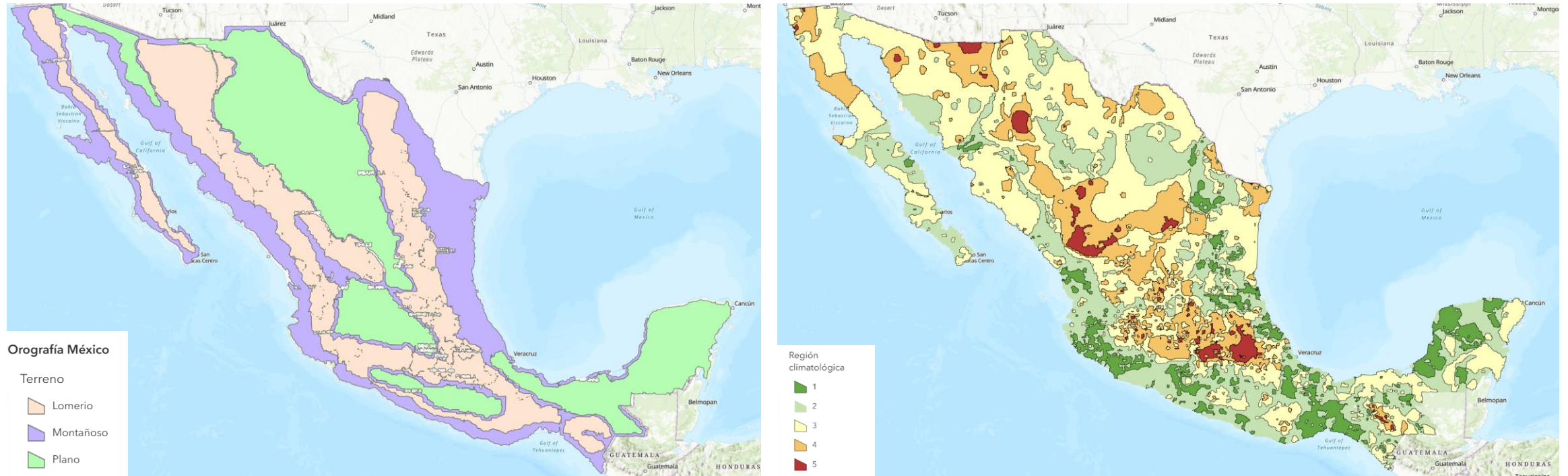
*cuerpo*: número de sentidos de la carretera

*región*: clasificación según la zona meteorológica del país

*orografía*: clasificación según su tipo de suelo

*año*: costos respecto al año (2020-2026)

# Orografía nacional y regiones climatológicas - SICT



Mediante técnicas de georreferenciación se obtiene la clasificación de suelo. Que lo dividen en plano, lomerío y montañoso. Asimismo, la clasificación de zonas meteorológicas. Categorizan al país por condiciones de clima y terreno para la construcción

# Tipo de carretera - SICT

La SICT muestra para los 15 corredores federales principales su tipo de carretera. Para el resto de carreteras en México no hay información disponible



Corredores principales federales

Carreteras admin federal


## TEMA 3


---

# Aproximación mediante la Auditoría Superior de la Federación



# Extracción de información – ASF

 Se extrae información a partir de la Auditoría Superior de la Federación a la SICT.

 Se une a la Red Nacional de Caminos para asignarle la valuación económica a los tramos carreteros

## Información disponible

Carretera	Nombre del proyecto	Inicio de construcción	Fin de construcción	Año fiscal	Entidad federativa	Muestra Auditada	Universo seleccionado	Tipo de proyecto	Longitud de la vía (km)	Ubicación
Mexicali-San Felipe	ASF.2012.PT.	20/02/2012	30/11/2012	2012	Baja California	187,695.00	229,887.40	Incluir más carriles	131.7	Entre el km 38+200 y 60+000

La base incluye detalles del proyecto, material de construcción e identificadores viales

# Proyectos auditados 2000 – 2025

Los proyectos auditados forman parte de los 15 principales corredores federales. Por lo que es posible hacer una **comparativa entre el costo estimado y la valuación económica**



Proyectos auditados



Carreteras admin federal



## TEMA 4

---

**Identificación de plantas de  
tratamiento y subestaciones  
eléctricas a través de la  
clasificación de objetos en  
imágenes satelitales.**



# Inventario Nacional de Infraestructura

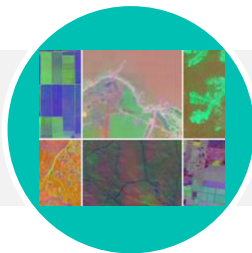
## Desafíos:

- No existe polígono
- Dato georreferenciado en número exterior
- Mala georreferenciación



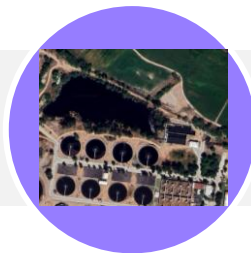
Objeto espacial	Valor
plantas_tratadoras_ags	
nom_objeto	CIUDAD
(Derivado)	
(coordenada X pinchada)	-11392185
(coordenada Y pinchada)	2491553
ID de objeto	8
X	-11392186
Y	2491553
(Acciones)	
fid	14912
id	201030200000008
id_sector	2
des_sector	Secundario
id_tema	01
des_tema	Agua, drenaje y alcantarillado
id_cat	03
des_cat	Tratamiento
id_objeto	02
des_objeto	Planta de tratamiento
nom_objeto	CIUDAD
id_inst	601
des_inst	Gobierno Municipal
id_ocupa	NULL
des_ocupa	NULL
id_confis	NULL
des_confis	NULL
id_ordgob	3
des_ordgob	Municipal
sup_terr	0
sup_cons	0
ano_cons	0
ano_inicio	0
vida_util	0
avaluo	0
ano_avaluo	0
valor_est	0
cve_ent	01
des_ent	Aguascalientes
cve_mun	001
des_mun	Aguascalientes
cve_loc	NULL
des_loc	NULL
ageb	1138
manzana	NULL
latitud	21.833556000000002
longitud	-102.337750000000000
clasescian	221312
id_fuente	104
des_fuente	Censo Nacional de Gobiernos Municipale...
fec_incorp	01/06/2023
idobfuente	01001026PT

# Clasificación de objetos en imágenes satelitales



## Embeddings de AlphaEarth

- Vectores matemáticos de alta dimensión que comprimen el contenido visual de imágenes satelitales en datos numéricos comparables.
- En lugar de procesar píxeles individuales, operan sobre el "significado" de la escena, permitiendo identificar y localizar infraestructuras similares.
- Ofrecen una resolución de 10 metros por 10 metros



## Segmentación en Imágenes alta resolución

- Se centran en la precisión geométrica y la interpretación visual humana y de máquinas.
- RGB de alta resolución que se limita al espectro del ojo humano.
- El tamaño de pixel varía entre 50 y 30 cm por pixel



# Embeddings de AlphaEarth (vectores de incorporación)

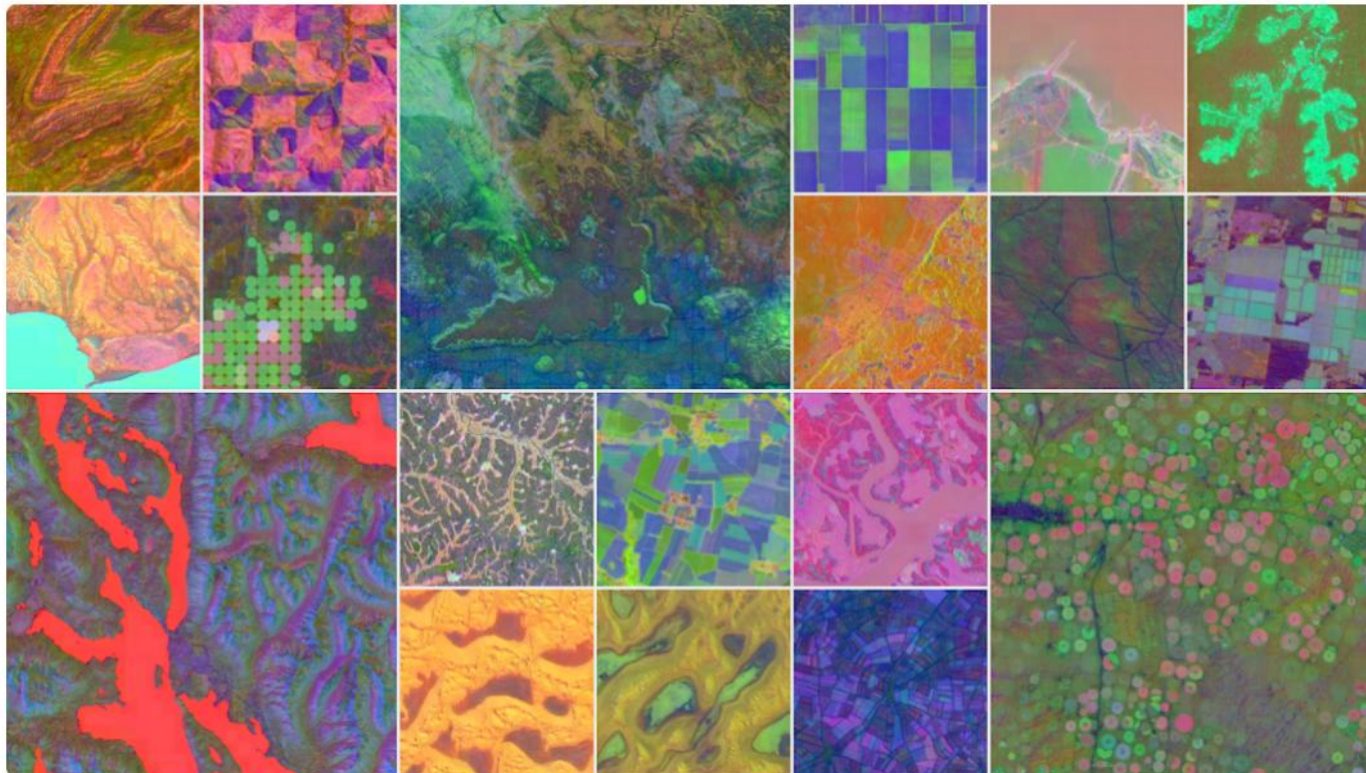
SCIENCE

AlphaEarth Foundations helps map our planet in unprecedented detail

30 JULY 2025

The AlphaEarth Foundations team

[Share](#)



- 64 dimensiones
- Distintos productos geoespaciales en uno sólo
- Disponibilidad anual para 2017-2024

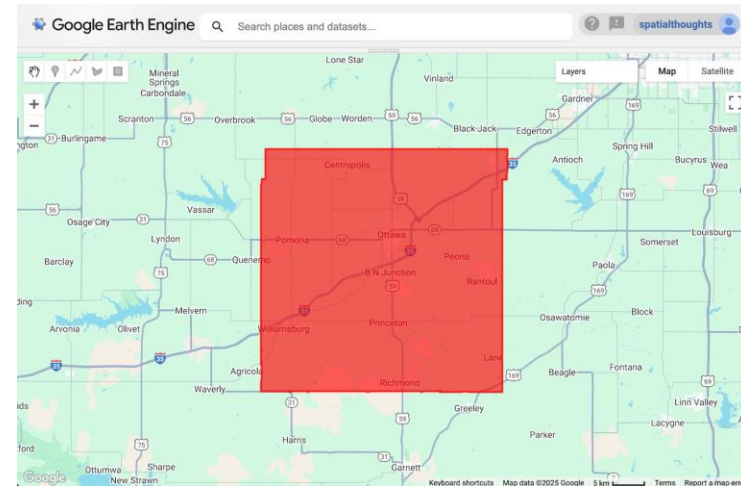
inegi.org.mx



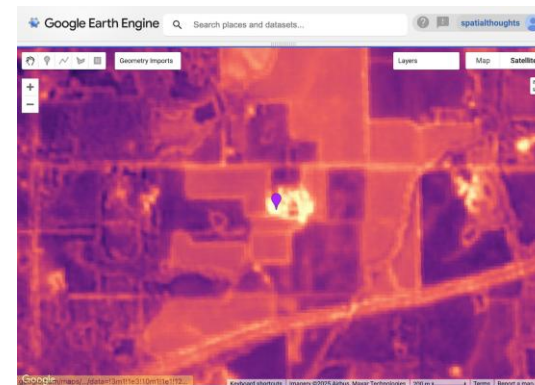
# Aplicaciones Embeddings



Objeto a clasificar

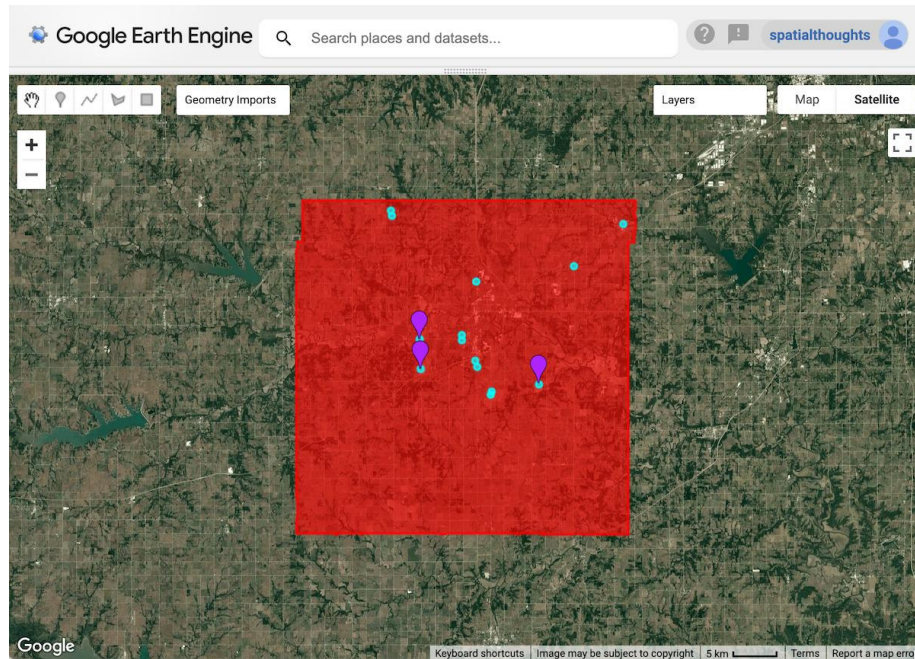


Ejemplos conocidos



Búsqueda de similitud

# Aplicaciones Embeddings (cont.)



Área de búsqueda

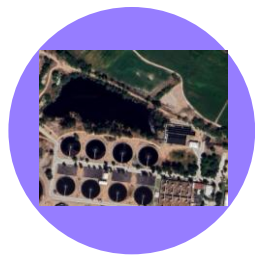


Verdaderos positivos



Falsos positivos





# Segmentación en Imágenes alta resolución

Usamos el "Segment Anything Model" (SAM) de Meta AI basado en una CNN.

Es un modelo masivo pre-entrenado.

Puede segmentar cualquier objeto en una imagen.

El problema: No sabe qué es el objeto.

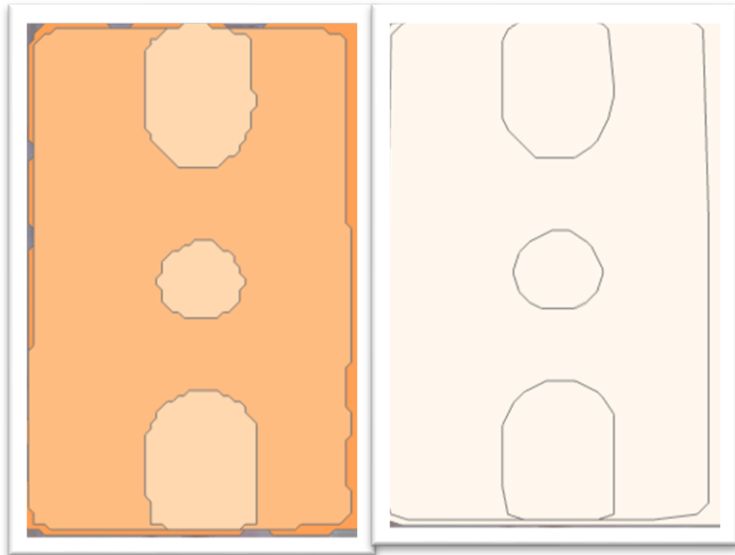




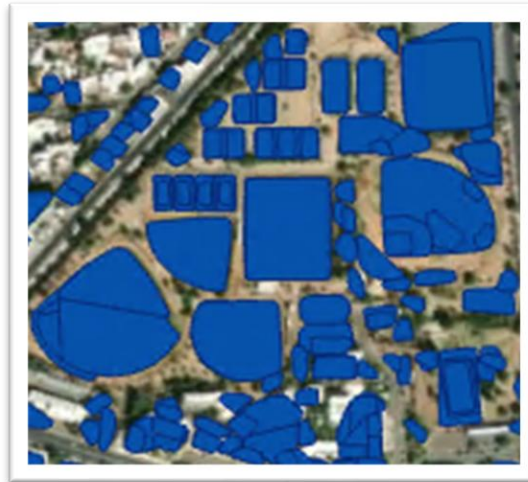
## Modelo no supervisado

El modelo analiza datos que **no tienen etiquetas previas**. El algoritmo debe explorar, organizar y descubrir por sí mismo los patrones, estructuras o similitudes ocultas dentro de la información

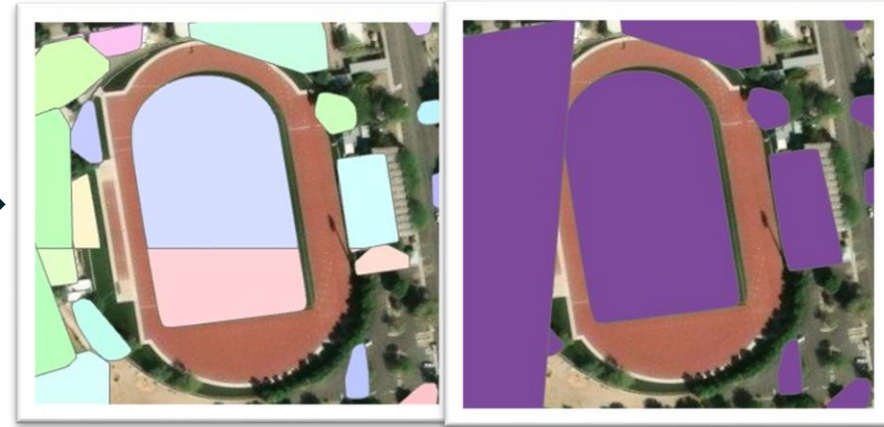
# Generalización topográfica



Mínima geometría envolvente



Eliminar polígonos de tamaño inusual



Disolver

# MOMENTOS HU DIMENSION FRACTAL

Los momentos Hu son descriptores invariantes que proporcionan información sobre sus características geométricas independientemente de su escala, posición y orientación.

La dimensión fractal es una medida utilizada para describir la complejidad y la irregularidad de formas geométricas.



HuMoment1N	HuMoment2N	HuMoment3N	HuMoment4N	HuMoment5N	HuMoment6N	HuMoment7N
0.002644686399...	2.650408454e-06	5.70498e-09	3.6311153e-08	1.75991374e-07	1.9684767092e-05	0.01317712433345
0.002970444723...	3.388391521e-06	4.021627e-09	2.0254657e-08	1.75991373e-07	1.9684712849e-05	0.013177124333...
0.002887118303...	3.873992804e-06	1.2681435e-08	4.4707987e-08	1.75991374e-07	1.9684839235e-05	0.013177124333...
0.000520458830...	3.98427099e-07	3.181e-12	3.018e-12	1.75991373e-07	1.9684619908e-05	0.013177124333...
0.000544017346...	4.29394385e-07	4.613e-12	5.707e-12	1.75991373e-07	1.9684619913e-05	0.013177124333...
0.004630965949...	1.147962602e-05	1.90689953e-07	1.50110388e-07	1.75991397e-07	1.9685866223e-05	0.0131771243321
0.001726320731...	8.4773603e-08	1.22250263e-07	4.7644595e-07	1.75991461e-07	1.9684964818e-05	0.013177124320...
0.001171125562...	3.18122545e-07	2.554304e-09	3.071665e-09	1.75991373e-07	1.9684623607e-05	0.013177124333...

# AGRUPACIÓN Y SELECCIÓN

---

La agrupación de K-Means agrupa polígonos en función de sus características inherentes y revelando patrones de variación y similitud.

Realizamos una inspección visual de los clusters generados y seleccionamos aquellos clusters que, de forma clara y distintiva, nos permitieron identificar y separar los campos deportivos del resto de elementos presentes en las imágenes.





## Modelo supervisado

El modelo aprende a partir de datos históricos que ya tienen la respuesta correcta (etiquetas). Funciona como un estudiante con un profesor que le muestra el problema y la solución.

# El Modelo "Experto" (Mask R-CNN)

Dado que SAM es demasiado grande y general, necesitamos un modelo ligero y especializado.

Usamos Mask R-CNN, una arquitectura de "segmentación de instancias".

Lo entrenamos para que sea un experto que solo identifique las clases etiquetadas, los estanques circulares y los estanques rectangulares.

Etiquetamos 32 círculos y 38 rectángulos de un total de 29,609 polígonos segmentados en diferentes áreas



# Resultados

El modelo final ignora objetos (edificios, techos, vegetación, sombras) y detecta estanques de tratamiento de agua.

Las máscaras de instancia nos dan la geometría casi precisa de cada estanque, permitiendo el cálculo aproximado de área y perímetro.



**Average Precision Score:**

**Circulos: 0.93**

**Rectangulos: 0.64**

# Resultados Aguascalientes



Planta de tratamiento Los Sauces



Planta de tratamiento Arboledas de Paso Blanco

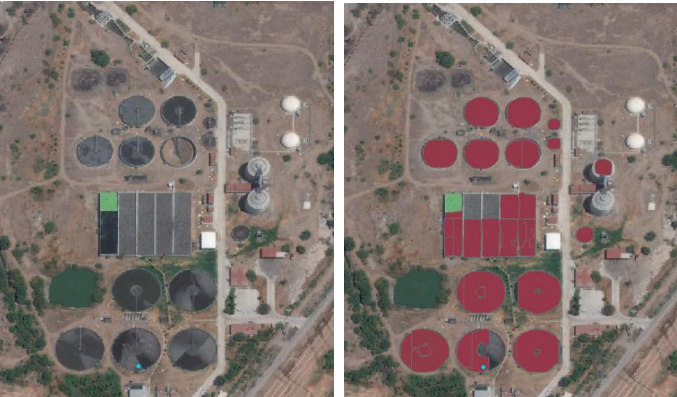


Planta de tratamiento Pocitos



Planta de tratamiento Tres centurias

# Resultados Hermosillo



Planta de tratamiento de Aguas residuales Hermosillo



Planta de tratamiento Los Arroyos



Planta de tratamiento Los Saucedo

# Resultados CDMX



# Trabajo en progreso

Edificios en 3D:

1. Extracción de información de registros administrativos.
2. Medición de altura con imágenes de satélite.
3. Modelo para valoración económica.

# Conclusiones

---

Fortalecimiento de la Plataforma de Infraestructura Pública Nacional:

1. Apoyo a la sistematización de información.
2. Geocodificación para evaluar e identificar direcciones.
3. Modelo para valuación económica de autopistas.
4. Aprovechamiento de imágenes satelitales para la identificación de objetos.



# ¡GRACIAS!

CONOCIENDO  
**MÉ  
XI  
CO**

800 111 46 34

[www.inegi.org.mx](http://www.inegi.org.mx)

[atencion.usuarios@inegi.org.mx](mailto:atencion.usuarios@inegi.org.mx)



**INEGI**INFORMA



[inegi.org.mx](http://inegi.org.mx)