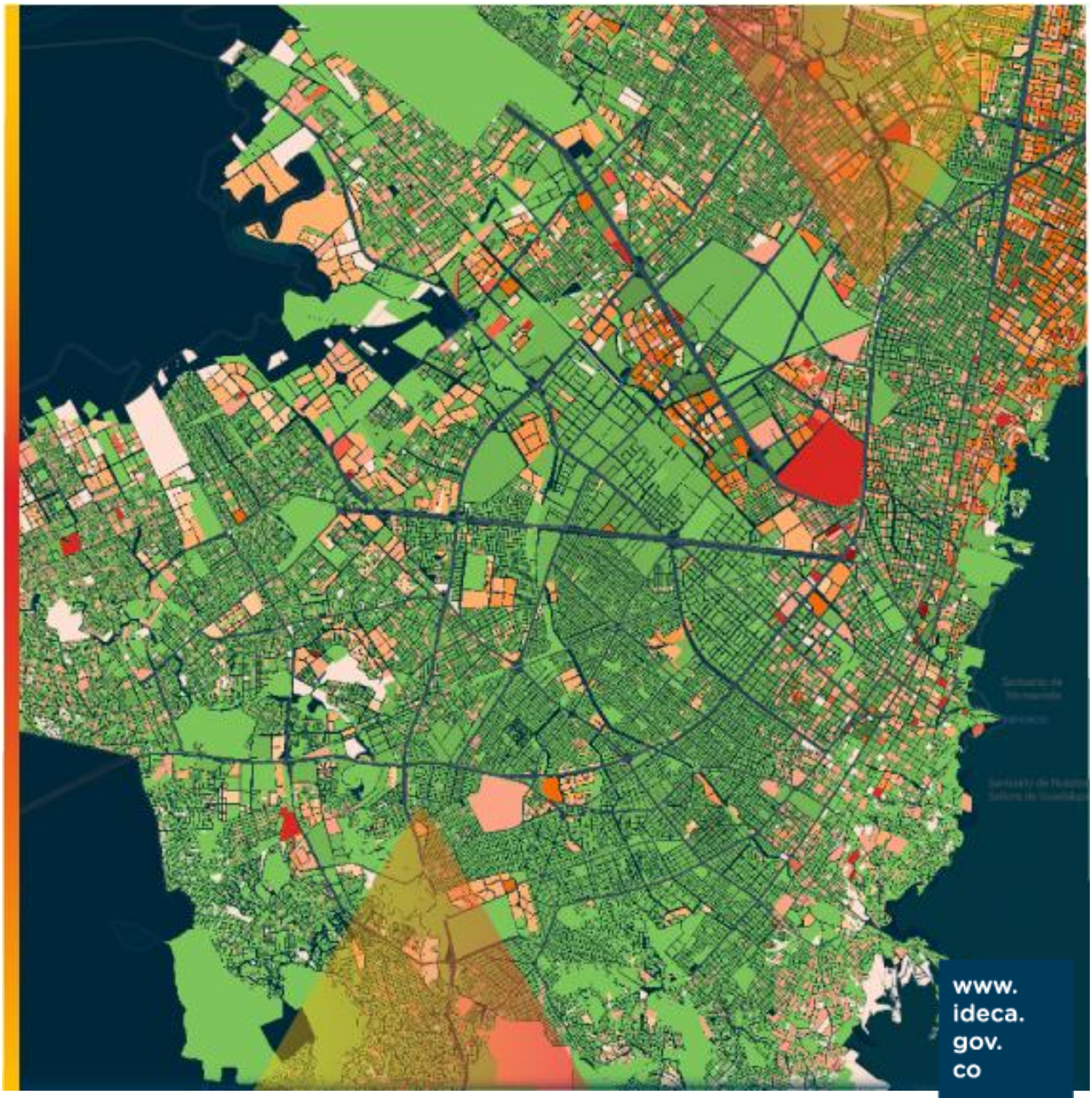


—

Clasificación del Uso Predial Usando Imágenes de GSV y Modelos de Lenguaje Visual

Bogotá, D.C., 30/06/2025/

Clasificación del Uso Predial Usando Imágenes de GSV y Modelos de Lenguaje Visual



Gerente IDECA
Salomé Naranjo Luján

Subgerente de (Operaciones o Analítica de Datos)
Diego Ricardo Ibarra Rodríguez

Equipo de Trabajo

Contratista
Víctor Alfonso Arias Vanegas

Fecha de creación o actualización: 30/06/2025

Página web: www.ideca.gov.co

Correo electrónico: ideca@catastrobogota.gov.co

Licencia: Attribution 4.0 International (CC BY 4.0)

Control de Versiones

CAMBIOS

Fecha	Autor	Versión	Cambio efectuado
30/06/2025	Víctor Alfonso Arias Vanegas	1.0	Primera versión del documento. No hay cambios para registrar.

REVISORES

Nombre	Dependencia
Diego Ricardo Ibarra Rodríguez	Subgerente de Analítica de Datos

Tabla de Contenido

Tabla de Contenido **4**

Resumen ejecutivo **5**

Introducción **6**

Antecedentes **7**

Alcance **9**

Objetivos **9**

1. Inicio **10**
1.1. Requerimientos 10
1.2. Estudio de Mercado 11

2. Idear **12**
2.1. Proceso de Ideación 12
2.2. Entendimiento de los Datos 13
2.3. Ingeniería de Características 20
2.4. Procesamiento de imágenes 21
2.5. Plataforma Tecnológica 23
2.6. Requisitos / Objetivos 24
2.7. Propuesta 24
2.8. Revisión con el cliente 25
2.9. Actividades 25

3. Prototipar **26**
3.1. Prototipo / Modelado 26
3.2. Versión 0.1 31
3.3. Testeo 0.1 32
3.4. Incidencias 0.1 32
3.5. Ajustes 33
3.6. Revisión Cliente 34

4. Operacionalizar – Industrializar **36**
4.1. Verificación Técnica y Validación de la Solución 36
4.2. Transferencia 36
4.3. Retrospectiva 37
4.4. Finalizacion 38

5. Resultados **39**

6. Conclusiones **41**

7. Bibliografía (ICONTEC) **42**

Resumen ejecutivo

El proyecto "Clasificación del Uso Predial Usando Imágenes de GSV y Modelos de Lenguaje Visual" tiene como objetivo desarrollar una API REST capaz de clasificar automáticamente el uso de predios en Bogotá a partir de múltiples imágenes de fachada, utilizando modelos de lenguaje visual multimodales (VLM).

El enfoque aborda retos como la alta desproporción en las clases, etiquetas poco confiables, múltiples entradas por predio y la necesidad de interpretabilidad, lo cual se soluciona mediante la reducción de clases a 20 grupos de uso definidos por expertos y el uso del modelo Kimi-VL-A3B-Thinking, capaz de razonar sobre descripciones textuales y entradas visuales.

La solución permite clasificar predios en contextos desbalanceados y sin etiquetas precisas, apoyándose en razonamiento visual y descripciones de clases, con una respuesta estandarizada en formato JSON.

Se propone desplegar el sistema como un servicio web, con etapas futuras que incluyen optimización de prompts, interfaz gráfica, evaluación en producción y posible fine-tuning con datos verificados, asegurando así una solución técnica robusta, escalable y útil para la gestión catastral de la ciudad.

Introducción

La correcta identificación del uso que se le da a un predio dentro del tejido urbano de una ciudad es un insumo fundamental para la planificación territorial, la gestión catastral y el desarrollo de políticas públicas. En ciudades como Bogotá, donde la dinámica urbana es diversa y cambiante, esta clasificación se ha venido realizando principalmente a través de métodos manuales, procesos administrativos o levantamientos de campo, los cuales presentan altos costos, demoras logísticas y, en muchos casos, inconsistencias en los datos registrados. La necesidad de una herramienta automatizada y confiable que permita clasificar los predios de manera eficiente fue el punto de partida de esta investigación.

Este proyecto surge del interés por aprovechar los avances recientes en inteligencia artificial, particularmente en modelos de lenguaje visual multimodal (VLM), para apoyar la clasificación del uso predial a partir de imágenes obtenidas desde Google Street View (GSV). A diferencia de aproximaciones tradicionales, esta propuesta integra elementos de visión por computador y procesamiento de lenguaje natural para interpretar tanto las características visuales de las fachadas como las descripciones textuales de los usos posibles.

La meta principal de esta investigación es construir una API REST capaz de recibir múltiples imágenes por predio y retornar una clasificación en formato estandarizado, con un enfoque multiclase e interpretable. Esto no solo tiene un valor técnico y académico, sino también un fuerte impacto práctico al ofrecer una solución escalable y reproducible para entidades públicas encargadas de la gestión del territorio urbano.

En términos metodológicos, se construyó una base de datos representativa a partir de muestras de predios No Propiedad Horizontal (NPH), agrupadas en 20 categorías definidas por expertos. Se seleccionó un modelo de razonamiento visual open source, el cual fue probado y ajustado en una serie de experimentos controlados que permitieron validar su precisión, su capacidad para integrar múltiples imágenes y su nivel de interpretación.

Este documento describe los antecedentes que motivaron la investigación, el desarrollo técnico realizado, los retos encontrados durante su implementación y las posibles rutas de mejora futura, con el fin de consolidar una herramienta innovadora para el análisis urbano desde un enfoque basado en inteligencia artificial.

Antecedentes

La necesidad de desarrollar una herramienta automática para la clasificación del uso predial parte de la identificación de problemáticas en la gestión territorial de Bogotá: prevalecen procesos manuales costosos y tardíos para registrar los usos de los predios, lo que genera inexactitudes y baja interoperabilidad entre entidades. Este escenario motivó el diseño del proyecto actual, que surge con base en la imperiosa demanda de soluciones que permitan automatizar y mejorar la toma de decisiones en el ámbito catastral y de planificación urbana.

El interés por este tema se reforzó al explorar fuentes de innovación en el análisis geoespacial, particularmente tras revisar el estado del arte en visión por computador y modelos de lenguaje visual multimodal, que ofrecen una capacidad prometedora para integrar imágenes con descripciones textuales y generar clasificaciones inteligentes. Estos avances motivaron conceptualizar un flujo técnico que combinara imágenes de fachada extraídas de Google Street View con razonamiento textual sobre categorías de uso predial.

En cuanto a los antecedentes institucionales, este proyecto encaja perfectamente con la estrategia de IDECA – la Infraestructura de Datos Espaciales del Distrito Capital. IDECA promueve la integración, estandarización y aprovechamiento de datos geográficos producidos por entidades del Distrito, así como la creación de recursos geográficos de valor agregado mediante innovación tecnológica.

Su Plan Estratégico 2020-2024 destaca objetivos como fortalecer el gobierno de los recursos geográficos, fomentar la innovación a partir de datos, promover la interoperabilidad y descentralizar la gestión de datos Ideca. De esta manera, el presente proyecto apoya directamente esas líneas de acción al diseñar una API inteligente que clasifica automáticamente el uso del suelo, genera un recurso geográfico enriquecido y es interoperable mediante interfaces REST estándar.

Adicionalmente, IDECA ha impulsado la adopción de tecnologías avanzadas como la migración a ArcGIS Enterprise en Microsoft Azure (LINK), con el fin de aumentar disponibilidad, rendimiento e interoperabilidad de sus servicios geoespaciales Catastro Bogotá. Este entorno tecnológico proporciona un contexto adecuado para la integración futura de nuestra API, contribuyendo a ampliar servicios de clasificación automatizada y enriquecer la plataforma de Mapas Bogotá y servicios temáticos.

En síntesis, estos antecedentes permiten comprender el origen del problema, el camino que condujo a su formulación y los elementos esenciales de la solución planteada, al tiempo que se establece su compatibilidad clara con los objetivos estratégicos y operativos de IDECA para el fortalecimiento de la infraestructura de datos espaciales en Bogotá.

Alcance

Presentar los antecedentes que introduzcan al problema: cómo o de dónde surge, el camino que llevó a su formulación y los elementos más importantes para explicar en qué consiste.

Objetivos

Objetivo General

Desarrollar una solución automatizada basada en inteligencia artificial que permita clasificar el uso predial de los inmuebles de Bogotá a partir de imágenes de fachada, integrando modelos de lenguaje visual multimodal.

Objetivos Específicos

1. Diseñar una arquitectura técnica que permita integrar múltiples imágenes de fachada por predio como entrada para un sistema de clasificación multiclase.
2. Implementar un modelo de lenguaje visual (VLM) open source capaz de interpretar imágenes y descripciones textuales de usos prediales para emitir clasificaciones con razonamiento explícito.
3. Construir una base de datos de entrenamiento representativa y balanceada a partir de predios No Propiedad Horizontal (NPH), agrupados en 20 categorías definidas por expertos.
4. Desarrollar una API REST capaz de recibir entradas visuales, procesarlas mediante el modelo y retornar una clasificación estructurada en formato JSON schema.
5. Evaluar el desempeño del modelo en términos de precisión, razonamiento e interpretabilidad de resultados, considerando los retos del desbalance de clases y la ausencia de etiquetas verificadas.
6. Proponer mecanismos de despliegue escalable y reutilizable para la solución, facilitando su integración en plataformas como Azure ML.

1. Inicio

La iniciativa surge como respuesta a una necesidad prioritaria identificada en el marco de la gestión catastral del Distrito Capital. La clasificación del uso del suelo predial, particularmente en los predios No Propiedad Horizontal (NPH), ha dependido históricamente de recorridos de campo y procesos administrativos que demandan altos recursos humanos, logísticos y financieros. Esta situación limita la capacidad de las entidades públicas para mantener una caracterización actualizada, precisa y oportuna del territorio, lo que impacta negativamente la planeación urbana, el recaudo tributario y la gestión de servicios públicos.

El caso de uso se formuló con base en el siguiente requerimiento funcional:

¿Cómo automatizar la clasificación del uso del suelo predial en Bogotá, utilizando imágenes urbanas e inteligencia artificial, para mejorar la eficiencia operativa y la calidad del catastro?

La solución propuesta responde a esta necesidad mediante el desarrollo de un sistema basado en modelos de lenguaje visual multimodal (VLM), capaz de analizar imágenes de fachada y clasificar predios en categorías funcionales (residencial, comercial, mixto, etc.), de forma escalable, precisa y transparente.

1.1. Requerimientos

Durante la fase inicial, se llevaron a cabo reuniones técnicas entre el equipo de desarrollo, expertos catastrales y áreas funcionales clave. De este diálogo surgieron los siguientes requerimientos mínimos:

- Utilizar múltiples imágenes de fachada como entrada por cada predio, para mejorar la precisión del modelo.
- Clasificar automáticamente los predios en 20 grupos funcionales definidos por expertos, reduciendo la complejidad y desbalance del universo de clases.

- Generar salidas interpretables y estructuradas, bajo un formato estándar tipo JSON schema.
- Implementar el sistema como una API REST interoperable, adaptable a entornos locales, en la nube o integrable con la plataforma IDECA de ser posible.
- Validar técnicamente los resultados mediante revisión por expertos catastrales, contrastando con metodologías tradicionales.
- Asegurar el cumplimiento normativo en el manejo de datos personales e imágenes, garantizando anonimato y tratamiento ético de la información.

1.2. Estudio de Mercado

Durante el proceso de diseño de la solución, se llevó a cabo un análisis de literatura técnica, soluciones internacionales comparables y tecnologías emergentes:

Benchmark internacional:

En ciudades como Ámsterdam, Nueva York o Seúl, se han implementado soluciones basadas en imágenes satelitales y de calle para segmentación, identificación de infraestructura o monitoreo urbano. Sin embargo, la clasificación semántica del uso predial a partir de imágenes urbanas sigue siendo un campo emergente y poco explorado, especialmente en América Latina.

Estado del arte en IA multimodal:

Se evaluaron modelos como CLIP, Flamingo y Qwen-VL, los cuales mostraban avances en correlación texto-imagen, pero carecían de razonamiento explícito o no permitían múltiples imágenes de entrada. Finalmente, se seleccionó el modelo Kimi-VL-A3B-Thinking, por su arquitectura Mixture-of-Experts, su razonamiento tipo "chain-of-thought" y su capacidad de adaptarse a descripciones conceptuales, sin requerir entrenamiento supervisado exhaustivo.

Revisión de enfoques locales: A nivel distrital, se identificó que los procesos de caracterización predial siguen siendo en su mayoría manuales. El sistema propuesto busca precisamente complementar ese trabajo, reduciendo la carga operativa, mejorando la trazabilidad y anticipando zonas con probabilidad de cambio en el uso del suelo.

2. Idear

2.1. Proceso de Ideación

Aunque el proceso formal de ideación fue reducido, el proyecto surgió a partir de una necesidad concreta planteada por el equipo técnico de Catastro Distrital: automatizar la clasificación del uso predial en predios No Propiedad Horizontal (NPH) mediante imágenes y modelos de inteligencia artificial. En lugar de desarrollar una fase extendida de lluvia de ideas, el equipo pasó directamente a la formulación técnica de una hipótesis de trabajo:

"Es posible clasificar automáticamente el uso predial de un inmueble urbano en Bogotá, utilizando imágenes de fachada provenientes de Google Street View y un modelo de lenguaje visual multimodal (VLM) sin necesidad de entrenamiento supervisado extenso."

La hipótesis se fundamentó en experiencias previas del equipo con modelos de visión por computador y en la revisión de tecnologías emergentes. La definición del problema, el diseño de la solución y el establecimiento de los objetivos específicos surgieron de reuniones técnicas con expertos en catastro, desarrolladores y analistas de datos.

El listado de tareas iniciales para construir el MVP fue:

1. Definir las clases objetivo (grupos de uso).
2. Crear un conjunto de datos representativo de predios NPH.
3. Seleccionar y configurar un modelo VLM razonador.
4. Diseñar prompts con descripciones semánticas de las clases.
5. Implementar pruebas con imágenes reales de predios.
6. Desarrollar una API REST funcional.

2.2.Entendimiento de los Datos

Antes de hablar de la implementación es necesario entender el problema en cuestión, para luego entender por qué se abordó la solución de la manera en como se abordó. Todo predio en la ciudad de Bogotá pertenece a dos categorías: PH (Propiedad Horizontal) y NPH (No Propiedad Horizontal), siendo esta última la de interés en el proyecto.

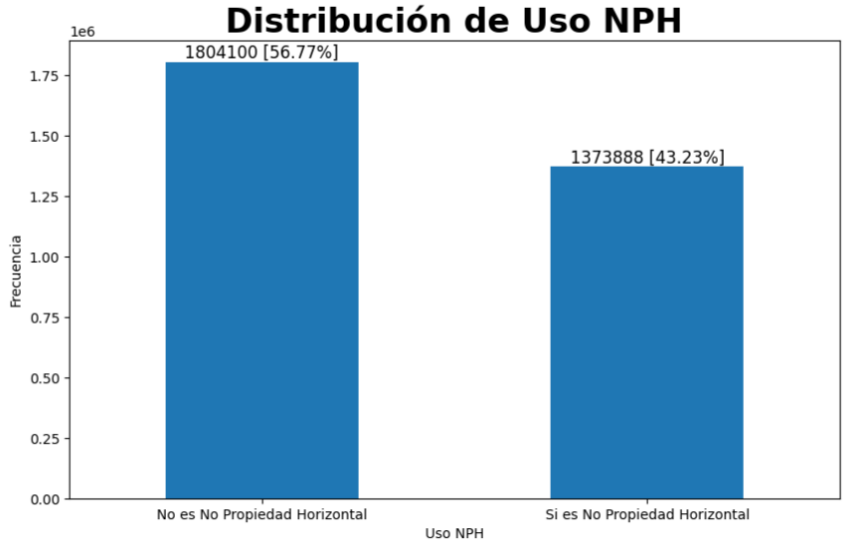


Figura 1. Distribución de las categorías del total de Predios en la ciudad de Bogotá, nuestro interés se centrará en el 43.23% que es NPH.

Estas categorías a su vez se dividen en el uso que se les da, es decir, la destinación que se le da a los elementos materiales de la estructura urbana en las distintas actividades ciudadanas. Siendo de gran interés todo lo que corresponde con actividades económicas que se desarrollen en cada una de las edificaciones de un predio (al momento de su reconocimiento).

DESCRIPCION DEL USO	CODIGO USO	
	NPH	PH
Habitacional menor o igual a 3 pisos	01	37
Habitacional mayor o igual a 4 pisos	02	38
Comercio Puntual	03	39
Comercio en Corredor Comercial	04	40
Oficinas operativas	05	82
Centro Comercial Mediano	06	41
Centro Comercial Grande	07	42
Grandes almacenes o Bodega Comercial	08	91
Actividad Artesanal	09	19
Industria Mediana	10	28
Industria Grande	11	34
Institucional Puntual	12	44
Colegios y Universidades de 1 a 3 pisos	13	35
Iglesias	14	53
Oficinas y Consultorios Oficiales	15	92
Colegios y Universidades de 4 pisos o más	16	
Clínicas Hospitales/Centros Médicos Grandes	17	43
Instalaciones Militares	18	
Oficinas y Consultorios particulares	20	45
Hoteles	21	46
Depósitos de Almacenamiento	22	98
Teatros y Cinemas	23	47
Edificio de Parqueo	24	50

DESCRIPCION DEL USO	CODIGO USO	
	NPH	PH
Bodega de Almacenamiento	25	93
Moteles Amoblados y Residencias	26	27
Parques de Diversión	29	36
Clubes de Mayor Extensión	30	
Piscinas	31	52
Coliseos	32	
Bodega Económica	33	97
Cementerios	55	
Restaurantes	56	60
Pista aeropuerto	62	
Aulas de Clases	64	
Clubes Pequeños	65	
Plazas de Mercado	66	
Museos	67	
Enramadas Cobertizos Caneyes	70	
Galpones Gallineros	71	
Establos Pesebreras	72	
Cocheras Marraneras Porquerizas	73	
Beneficiadores	74	
Secadores	75	
Kioscos	76	
Silos	77	
Oficinas en Bodegas y/o Industrias	80	81
Centro Comercial Pequeño	94	95
Parqueadero Cubierto		49
Parqueadero Cubierto	96	
Parqueo libre		48
Deposito Lockers		51
Predios sin construir en PH		90
Área de mezzanine en PH		57
Culto religioso	58	59

Figura 2. Además de ser un predio PH y NPH, las propiedades son clasificadas por su código de uso.

Tal y como lo muestra la Figura 2 los códigos de uso sirven para identificar cada uno de los usos que se le da a un predio. Una vez consultado los datos de los códigos de uso de cada uno de los predios de la ciudad de Bogotá, es necesario conocer la distribución de estos códigos de uso para responder a una primera pregunta de la investigación:

Pregunta de Investigación 1: ¿Cuántos predios por cada código de uso se encuentran disponibles en las bases de datos de la ciudad de Bogotá?

Y para responder esa pregunta fue necesario cargar todos los datos referentes a predios y asociarlos con las clasificaciones disponibles para cada uno de ellos, en este punto es importante entender algo: Un predio puede tener más de un código de uso (predios mixtos), por ejemplo, una familia que tiene su vivienda en un segundo piso y un pequeño local comercial en el primero. Una vez completada la tarea se grafica el top 20 de los tipos de uso en la Figura 3.

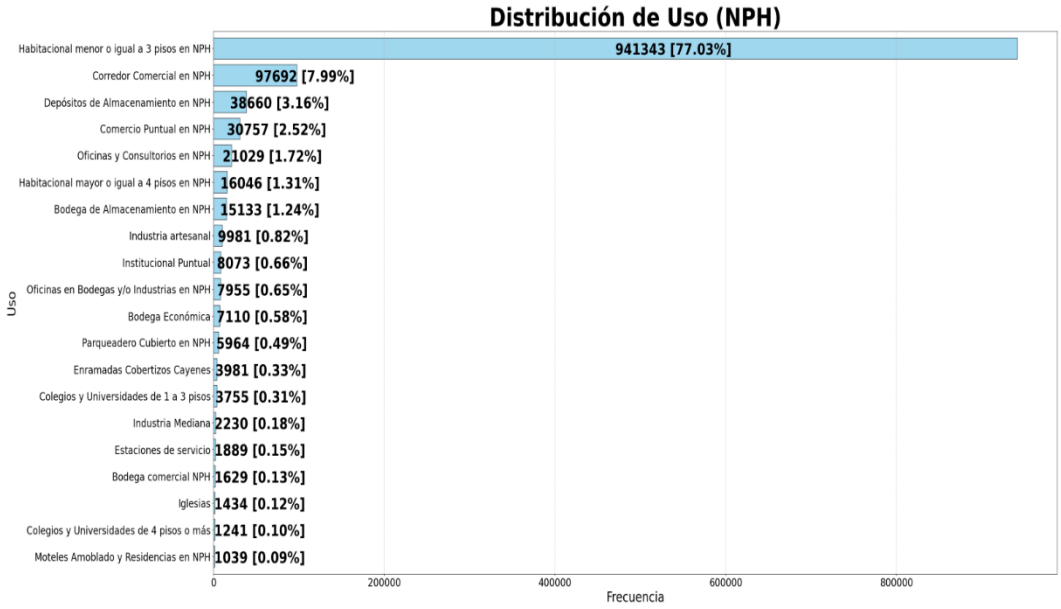


Figura 3. Distribución del tipo de Uso Predial (Top 20), en este gráfico solo se tuvieron en cuenta los NPH.

Una vez observada la distribución se llegó a la conclusión de que muy pocos tipos de uso agrupan la mayor cantidad de predios, es decir, tenemos una distribución de clases o tipos de uso altamente desbalanceada.

Por lo tanto, cualquier proyecto relacionado con implementar una solución técnicamente viable que genere un clasificador automático de tipos de uso predial, tiene que tener en cuenta las restricciones en el abordaje de un conjunto muy grande de clases a predecir, con un alto desbalanceo, y con clases bastante limitadas con respecto a la cantidad de datos disponibles en el universo de predios.

Esto último es importante aclararlo, nuestra base de datos de entrenamiento no puede componerse del total de predios de la ciudad de Bogotá, sino solo una muestra, ya que si tomamos todos los datos disponibles, ¿cómo evaluamos la capacidad predictiva del modelo? ¿Y qué utilidad tiene modelar todo el universo de datos disponibles?

Todo esto nos lleva la segunda pregunta de investigación:

Pregunta de Investigación 2: Dado los desafíos mencionados anteriormente ¿Cómo podemos desarrollar un universo de datos viable que nos permita obtener una base de datos para entrenar un clasificador multiclase?

La respuesta a esta pregunta de investigación vino por parte de los expertos en el área, que desarrollaron los llamados Grupos de Usos de la Construcción, estos grupos, como su nombre indica, agrupan usos de predios con características en común como ser de tipo residenciales o comerciales.

Aunque esto pierde algo de especificidad en el uso del predio, simplifica enormemente la cantidad de clases a predecir (llegamos a solo 20 clases), además el agrupamiento permite que cada una de estas clases tenga una mayor representación en el número de predios disponibles para entrenar el modelo. Finalmente, la nueva distribución de los grupos de uso resultante se muestra en la Figura 4.

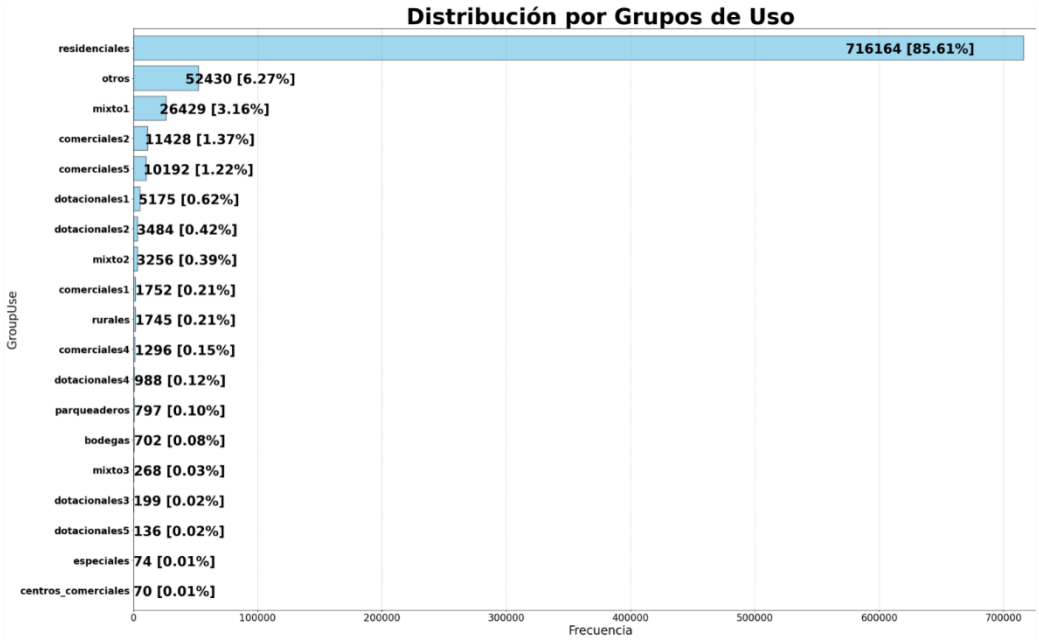


Figura 4. Los grupos de uso como manera de disminuir el número de clases a predecir al tiempo que incrementa el número de datos disponibles para entrenamiento por clase.

Una vez el universo de datos disponible está representado lo mejor posible, es momento de extraer una muestra del universo de predios, que será la base de datos de entrenamiento. Debido al desbalance inherente en los datos (4 órdenes de magnitud entre la clase con más predios y menos predios) se debió buscar una solución alternativa, esto nos lleva a la tercera pregunta de investigación:

Pregunta de Investigación 3: Dado los desafíos en el muestreo para obtener una base de datos de entrenamiento ¿Cómo podemos muestrear una representación lo más fidedigna posible del universo de datos para obtener una base de datos de entrenamiento?

Debido a que los resultados del muestreo estratificado no eran factibles para uso en el problema es cuestión, se buscó una alternativa que comprendía un muestreo por conveniencia, además, los predios de cada clase debían estar en lo posible uniformemente distribuida por la ciudad de Bogotá, y aquellas clases con menos de 150 predios, se tomarían todos los predios disponibles de la clase.

Finalmente, se pasó la muestra y se obtuvieron los datos con los que se va a entrenar. La Tabla 1 muestra el número de predios por cada uno de los grupos de uso.

Tabla 1. Distribución de la muestra por clases

Nombre del Grupo de Uso	Número de muestras
dotacionales3	193
residenciales	188
bodegas	174
mixto3	173
rurales	167
comerciales3	167
comerciales1	165
comerciales4	151
comerciales2	148
otros	140
mixto2	135
comerciales5	131
dotacionales5	131
parqueaderos	128
mixto1	123
dotacionales1	116
dotacionales2	114
dotacionales4	99
especiales	74
Centros_comerciales	70

La Figura 5 muestra la distribución de los predios por cada grupo de uso y como estos se distribuyen de manera homogénea en la ciudad de Bogotá.

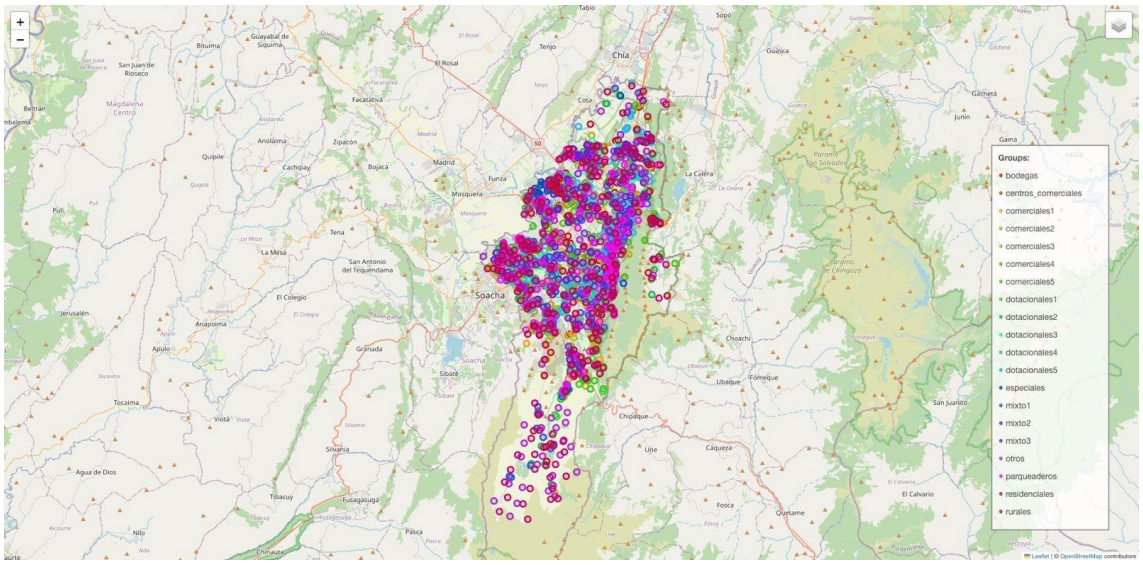


Figura 5. Distribución espacial de los grupos de uso de manera uniforme en la ciudad de Bogotá

Un desafío final debe ser tenido en cuenta, y es la poca confianza que se tiene en las etiquetas de los predios de la muestra de la base de datos, ya que dichas clases no fueron verificadas debido a la no disponibilidad de un experto en el área que revisara y etiquetara estos datos.

El problema entonces quedaría planteado de la siguiente manera:

- El modelo debe clasificar un predio como perteneciente a uno de los 20 grupos de uso (multi-clase).

- Cada predio puede tener una o más imágenes disponibles, el modelo debe poder consumirlas todas en la entrada (multi-input).

- Hay grupos de uso con 100's de miles de predios y otros con decenas de predios (desbalance de clases).

- El clasificador no solo debe tener en cuenta características generales para la clasificación, sino que además debe tener la capacidad de fijar su atención en detalles (Mecanismos de Atención).

- Finalmente, la respuesta del modelo debería ser interpretable, ya que un entendimiento de las decisiones que tomo el modelo nos permiten iterar el prompt y lo que el modelo entiende como una respuesta razonable (MoE o Modelos Razonadores).

- El entrenamiento puede verse comprometido debido a la no disponibilidad de expertos etiquetadores y la no fiabilidad en el etiquetado actual (Zero-Shot Learning).

2.3. Ingeniería de Características

Durante la exploración inicial, se detectaron varias inconsistencias:

- Predios con múltiples códigos de uso sin jerarquía clara.
- Casos donde la fachada no reflejaba el uso registrado en la base.
- Variación en la calidad de las imágenes GSV según la zona.

Estas observaciones fueron discutidas con expertos catastrales, quienes aclararon que muchos predios son mixtos y que los códigos reflejan registros administrativos, no necesariamente observables desde la vía pública.

Para abordar esto:

- Se optó por usar los Grupos de Uso de la Construcción, que agrupan códigos similares en 20 categorías.
- Se eliminaron predios con fachada no visible o imágenes borrosas.
- Se construyó una matriz de entrada con múltiples imágenes por predio, asociadas a un prompt que contenía la descripción de cada clase.

2.4. Procesamiento de imágenes

La etapa de ingeniería de características en este proyecto fue además diseñada para asegurar que las imágenes de entrada se encuentren en un formato homogéneo, eficiente y compatible con los requerimientos del modelo multimodal Kimi-VL-A3B-Thinking, facilitando así un razonamiento preciso en el contexto de clasificación del uso predial.

1. Estandarización del Formato y Tamaño

Cada imagen recibida (ya sea cargada por archivo o en base64) pasa por una rutina de preprocesamiento implementada en el archivo classifier.py. El método resize_image() garantiza que las imágenes:

1. Se redimensionen manteniendo su relación de aspecto,
2. No excedan un tamaño máximo de 224x224 píxeles (ideal para cargas ligeras en modelos de lenguaje visual),
3. Se conviertan en imágenes PIL RGB limpias y estandarizadas.
4. Esto previene inconsistencias visuales y reduce el ruido que puede confundir al modelo durante la inferencia.

2. Decodificación y Almacenamiento Temporal

En el archivo app.py, las imágenes son decodificadas desde base64 (decode_base64_image) y guardadas localmente en una carpeta temporal (save_images_to_disk). Esta etapa permite:

1. Persistir las imágenes para su trazabilidad,
2. Crear un vínculo explícito entre imagen original y clasificación realizada,
3. Utilizar rutas de archivo como referencia visual dentro del prompt que se le entrega al modelo.

3. Preprocesamiento en Lote y Multientrada

El clasificador (Classifier.get_response()) permite trabajar con múltiples imágenes por predio, lo cual es clave para capturar diferentes fachadas o ángulos del mismo inmueble. Este enfoque:

1. Incrementa la robustez del análisis,
2. Reduce errores por imágenes parciales u ocultas,
3. Permite que el modelo razone sobre una representación más completa del entorno.
4. Las imágenes preprocesadas se agrupan y se entregan juntas al modelo como parte de una conversación multimodal.

4. Preparación del Prompt Multimodal

En la función prepare_messages(), se estructura un mensaje para el modelo que incluye:

1. Las imágenes como entradas tipo "image",
2. Un bloque de texto con el prompt de clasificación, que describe las 20 categorías de uso predial en inglés.
3. Este prompt es parte del contexto conversacional del modelo, utilizando el sistema de chat templates del procesador de HuggingFace para alinear texto e imágenes.

5. Vectorización y Razonamiento

Dentro de generate_model_response():

1. Las imágenes y el texto son codificados conjuntamente usando el processor del modelo Kimi,
2. Se genera una predicción utilizando .generate() sobre los input_ids del prompt + embeddings visuales,

3. El modelo devuelve un texto que incluye:
    3.1. Las clases predichas (mínimo 3),
    3.2. Sus confianzas,
    3.3. Y una explicación tipo "think" que justifica la decisión, p. ej.:

    "*The facade includes several glass displays, suggesting commercial use on the ground level, with residential floors above.*"

6. Salida Estructurada

La salida final es procesada como un objeto ImageData, que incluye:

classification: lista de clases con su confidence (entre 0 y 1),

think: explicación textual resumida.

Esto garantiza interpretabilidad y estructura JSON válida para interoperabilidad con otras plataformas como IDECA o dashboards internos.

2.5. Plataforma Tecnológica

El desarrollo se ejecutó sobre los siguientes componentes:

1. **Lenguaje y entorno**: Python 3.10 con notebooks de desarrollo.
2. **Modelos utilizado**s: Kimi-VL-A3B-Thinking (open source).
3. **Procesamiento en la nube**: GPU server de RunPod (versión serverless para pruebas), Spaces de Hugginface (con GPU L40S), Azure ML (Sin implementar por falta de acceso).
4. **Almacenamiento**: Repositorio Git privado y datasets en almacenamiento S3.
5. **Despliegue**: API REST Dockerizada, preparada para integrar con Azure ML, Hugginface Spaces, o RunPod.

2.6. Requisitos / Objetivos

Con base en los hallazgos y aclaraciones del cliente, se acordaron los siguientes requisitos mínimos viables:

- El sistema debe clasificar imágenes en una de las 20 clases definidas.
- Debe aceptar múltiples imágenes por predio.
- Debe entregar una respuesta en formato JSON estructurado e interpretable.
- Debe ser escalable y preparado para ejecución en la nube.
- Debe facilitar la validación técnica por parte de expertos.
- Los objetivos específicos definidos ya se mencionaron en la sección anterior, y están directamente alineados con estos requerimientos.

2.7. Propuesta

La propuesta consiste en construir un sistema de clasificación automática de uso predial que utiliza imágenes de fachada y un modelo de lenguaje visual multimodal capaz de razonar sobre descripciones textuales. El procesamiento se da en dos fases:

Entrada: Múltiples imágenes por predio + prompt con descripciones de clases.

Salida: JSON estructurado con las 3 a 5 clases más probables y razonamiento textual.

Se consideraron como alternativas:

- Usar modelos como CLIP o MiniGPT4, pero fueron descartados por su baja capacidad de razonamiento multiclase.
- Implementar un modelo entrenado localmente, pero se priorizó zero-shot reasoning por falta de etiquetas verificadas.

2.8. Revisión con el cliente

La propuesta fue bien recibida por los expertos catastrales, quienes destacaron su utilidad para focalizar esfuerzos de campo en zonas con alta probabilidad de cambio. También se discutieron dudas sobre la validez de las etiquetas actuales y la necesidad de un proceso futuro de validación manual.

Se acordó:

- Realizar pruebas piloto con muestras limitadas.
- Incorporar razonamiento explícito para facilitar auditoría.
- Desarrollar una interfaz que permita validar imágenes y resultados en una misma vista.

Sin embargo, quedo pendiente una etapa de retroalimentación basado en los hallazgos realizados una vez el producto sea utilizado por el cliente objetivo.

2.9. Actividades

Las principales actividades programadas fueron:

1. **Recolección y depuración de la base de datos de predios NPH.**
2. **Extracción automatizada de imágenes GSV.**
3. **Formulación de prompts descriptivos por clase.**
4. **Selección e implementación del modelo Kimi-VL-A3B-Thinking.**
5. **Pruebas con predios de distintas zonas y clases.**
6. **Desarrollo de la API REST y formato de respuesta estándar.**
7. **Validación interna y revisión por expertos catastrales.**
8. **Preparación para escalamiento en la nube (Spaces / RunPod / Azure ML).**

3. Prototipar

3.1. Prototipo / Modelado

La solución a nivel de ciencia se refiere a la selección del modelo utilizado para automatizar la correcta clasificación de grupos de uso a partir de un número N de imágenes de entrada. La respuesta no está sujeta al feedback del experto, y la información de cada clase no va más allá de unas cortas descripciones dadas por el experto encargado de la creación de los grupos de usos.

Aun así, estas cortas descripciones brindan una mayor seguridad a la hora de crear nuestro modelo que el etiquetado actual de los datos. Es por ello que un camino intermedio que se vislumbró fue usar una herramienta como CHATGPT Vision. Esta herramienta puede recibir las descripciones de cada una de las clases a predecir (Tabla 2) y las imágenes pertenecientes a cada predio y responder a qué clase pertenecen dichos predios. Finalmente, se compararían estos resultados con la etiqueta actual de los predios para saber si realmente son correctas.

Tabla 2. Descripciones de los grupos de uso que comprenden las clases de la base de datos de entrenamiento

Grupos de Uso	Subcategoria	Descripciones
Residenciales	-	Buildings intended for housing - Houses, PH Buildings, Condominiums.
Comerciales1	-	Refers to the storage, distribution, or exchange of products, goods, or services with a commercial interest.
Comerciales2	-	Buildings where activities aimed at providing services are carried out.
Comerciales3	-	Buildings used for artisanal activities where raw materials are transformed on a local scale.
Comerciales4	-	Hotels, Motels, and Restaurants.
Comerciales5	-	Operational offices and warehouses.

Centros_Comerciales	-	Commercial premises located on properties of one or several buildings.
Bodegas	-	Buildings in warehouse-type constructions dedicated to commercial, industrial, or storage activities.
Parqueaderos	-	Buildings designed for vehicle parking.
Dotacionales1	-	Buildings where activities aimed at the welfare or service of a community are carried out.
Dotacionales2	-	Buildings designed to carry out educational or training activities.
Dotacionales3	-	Buildings with the necessary infrastructure to provide surgical and/or hospitalization services.
Dotacionales4	-	Buildings for religious worship owned by communities or religious congregations.
Dotacionales5	-	Theaters, cinemas, swimming pools, museums, sports, events, or shows.
Especiales	-	Military administrative areas, cemeteries, airport runways.
Moles	-	Large buildings in height (>4 floors) or area (>10,000 m²), usually under construction.
Rurales	-	Sheds, kiosks, shelters, barns, stables, silos, etc.
Mixto1	(Residencial + Comercial1)	Housing and commercial premises.
Mixto2	(Residencial + Comercial2)	Housing and offices.
Mixto3	(Comercial1 + Comercial2)	Commercial premises and offices.

Dado el limitado acceso a recursos de este tipo, se inició una búsqueda de modelos VLM (Visual Language Models) open source que pudieran ser configurados con las definiciones de la Tabla 2, recibir múltiples imágenes de

entrada y de ser posible tener la capacidad de retornar el razonamiento que lo llevo a la respuesta correcta como una alternativa al ChatGPT o1.

Los modelos de razonamiento son modelos capaces de resolver problemas muy complejos y con mucha atención a los detalles como el presente. Primero se vieron soluciones de este con los grandes modelos de Lenguaje (LLMs) y ahora con los modelos de visión y lenguaje (VLMs). Hasta 2025, solo existía un modelo multimodal de razonamiento open source: QVQ-72B-preview de Qwen. Era un modelo experimental desarrollado por el equipo Alibaba Qwen y venía con muchas advertencias.

Pero este mismo año fue la presentación de Kimi-VL-A3B-Thinking (actualizado en el mes de junio a su versión más actual: Kimi-VL-A3B-Thinking-2506), del equipo Moonshot AI. Este modelo utiliza MoonViT (SigLIP-so-400M) como codificador de imágenes y un decodificador Mixture-of-Experts (MoE) con 16 Billones de parámetros totales (pero solo 2.8B activos). Es una versión bien afinada con cadena de pensamiento extensa (long chain-of-thought) y posteriormente alineada mediante aprendizaje por refuerzo (RL).

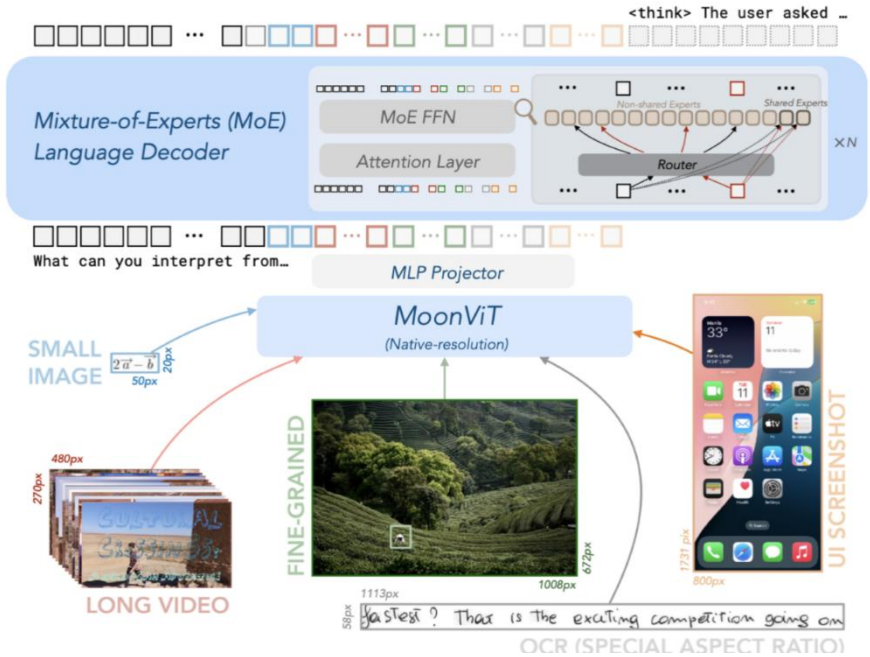


Figura 6. Kimi-VL Thinking permite la realización de múltiples tareas como detección, clasificación y segmentación sobre imágenes y videos en múltiples resoluciones usando un decodificador de tipo Mixture-of-Experts.

Una vez se seleccionó un modelo de razonamiento cuyo rendimiento podía ajustarse a las necesidades del problema en cuestión, era necesario crear un MVP, un notebook donde se pudiera ejecutar y evaluar las capacidades del modelo, para ello se usaron de ejemplo las imágenes en la Figura 7 y 8, donde se evaluaron las respuestas con una y múltiples imágenes.



Figura 7. Ejemplo de una imagen de un predio que se supone es de tipo **mixto 1**



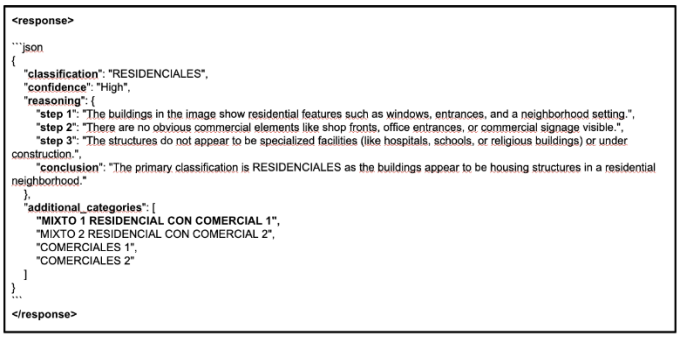


Figura 8. Entrada con múltiples imágenes al mismo tiempo para mejorar el resultado de la clasificación. El modelo empieza a darse cuenta de que quizás habría una posibilidad más allá de *RESIDENCIALES* con *MIXTO1*

Una vez probada la utilidad práctica del modelo, el siguiente paso lógico es crear un desarrollo para desplegar el modelo como un servicio que se pueda consumir. Antes de todo esto, hay varios desafíos por abordar en esta solución, algunos deben ser atacados inmediatamente y otros como trabajos futuros:

Debe haber un middleware que fuerce al modelo a dar la respuesta en un formato estandarizado, en este caso en un JSON schema tal y como se mostró en los objetivos, ya que la solución originalmente entrega un string largo tal y como se puede ver en la Figura 7.

Una vez ejecutado el endpoint debe poderse ejecutar en el background una interfaz gráfica capaz de recibir múltiples imágenes y consumir el servicio para entregar un resultado. Debe iterarse aún más el prompt en conjunto con el cliente para verificar que los resultados son consistentes con los requerimientos y el conocimiento acumulado de los expertos.

Es necesaria una ronda de optimización de la ejecución con el fin de evaluar el impacto en los resultados al limitar ciertas características o capacidades del modelo. Una etapa de fine-tuning que use los datos previamente verificados por un etiquetador manual para reforzar la capacidad predictora del modelo en esta base de datos.

Bien, una vez entendida la deuda técnica dejada atrás en el desarrollo, nos concentramos en las diferentes formas de abordar el despliegue, las cuales he ordenado de menor a mayor grado de complejidad:

- API REST dockerizada lista para ser desplegada en la nube o en local.

- **Modificación 1**: Adaptarla al modelo serverless de runpod (https://www.runpod.io/) para poder ser consumida dentro de la plataforma, siendo este el acceso más fácil a una GPU de alto rendimiento.

- **Modificación 2**: Adaptarla a un endpoint en Azure ML, siendo la solución de mayor complejidad debido a las limitaciones a la hora de acceder a máquinas de alto rendimiento dentro de la nube de Microsoft.

- **Modificación 3**: API REST dockerizada ejecutada según el formato de Hugginface Spaces.

3.2.Versión 0.1

La versión 0.1 del sistema fue desplegada como una aplicación funcional sobre una instancia de servidor basada en GPU (Hugginface Spaces), usando las siguientes tecnologías:

- **Modelo**: Kimi-VL-A3B-Thinking-2506 (Moonshot AI)
- **Lenguaje**: Python 3.10
- **Framework API**: FastAPI
- **Procesador de lenguaje multimodal**: HuggingFace Transformers
- **Entrada**: imágenes base64 o rutas de archivo
- **Salida**: JSON estructurado con clasificación y razonamiento

Esta primera versión demostró la integración exitosa entre el modelo VLM, las imágenes urbanas, el prompt semántico de clasificación y el servidor backend.

3.3. Testeo 0.1

Las pruebas de validación de esta versión se centraron en:

- Clasificar predios con imágenes multivista (entre 2 y 5 imágenes por inmueble),
- Verificar la integridad de la respuesta JSON, incluyendo etiquetas y confianza,
- Confirmar la coherencia del razonamiento textual generado por el modelo,
- Medir el rendimiento en tiempos de inferencia (inferencias promedio de ~1min por lote),
- Evaluar la interoperabilidad entre la API REST, la infraestructura de imágenes y el modelo.

Los resultados mostraron un comportamiento estable, procesamiento fluido y salidas consistentes con los usos prediales conocidos, especialmente en categorías frecuentes como "Residenciales", "Mixto1" o "Comerciales2".

3.4. Incidencias 0.1

Durante la etapa de pruebas se identificaron las siguientes incidencias relevantes:

- **Imágenes corruptas o con baja calidad**: algunas imágenes extraídas desde GSV presentaban fachadas poco visibles o desenfocadas, lo que afectaba la calidad del razonamiento.

- **Respuestas en lenguaje libre**: el modelo a veces generaba razonamientos con sintaxis libre o estructuras semánticas ambiguas, dificultando el análisis automatizado.
- **Desbalance de confianza**: en predios mixtos o con fachadas neutras, las predicciones tendían a sobreasignar categorías "Residenciales" por defecto.
- **Consumo de recursos elevado en pruebas locales**, por lo que se priorizó ejecución en GPUs bajo entornos en la nube (Spaces).

Estas incidencias fueron documentadas y usadas como insumo para ajustes y mejoras posteriores.

3.5. Ajustes

Tras el testeo inicial, se realizaron los siguientes ajustes:

- **Restricción del formato de salida** mediante el uso de pydantic y esquemas JSONSchema, asegurando que todas las respuestas cumplieran un formato estándar.

- **Validación de las imágenes antes de su procesamiento**, descartando entradas con dimensiones inválidas o calidad baja.

- **Reducción del tamaño máximo de imagen a 224x224 px**, lo cual equilibró la carga de procesamiento con el nivel de detalle necesario.

- **Ajuste del prompt** para reforzar el número mínimo de categorías predichas y asegurar que siempre se incluya un razonamiento estructurado.

La validación del modelo mostró niveles aceptables de rendimiento, especialmente en las categorías más representadas del set de prueba.

Durante la fase de pruebas y validación técnica del prototipo, se realizaron análisis exhaustivos para asegurar que el modelo **Kimi-VL-A3B-Thinking-2506** pudiera operar correctamente en un entorno de producción o pruebas intensivas. Este modelo multimodal, al contar con componentes de atención

visual, generación de lenguaje y razonamiento estructurado, presenta un **uso significativo de memoria GPU**.

Pruebas de consumo de memoria

Con apoyo de expertos técnicos, se empleó la herramienta **Model Memory Calculator** de Hugging Face para estimar el consumo de VRAM durante la inferencia.

- **Uso total estimado en inferencia (float16):** ~31.41 GB de VRAM.
- **Bloque más grande (residual layer):** 1.12 GB.
- **Entrenamiento completo (no aplicable en esta fase):** 125.62 GB.

- ***Elección de infraestructura: L40S de 40GB***

Inicialmente, se consideró como infraestructura de referencia una **NVIDIA A100 con 80GB de VRAM**, sin embargo, dadas las restricciones presupuestales y operativas, se optó por la implementación y pruebas en una **NVIDIA L40S**, que cuenta con **40GB de VRAM**.

Esta decisión fue viable porque:

- El modelo no requiere entrenamiento (solo inferencia).
- Las pruebas mostraron que el consumo máximo del modelo en inferencia está por debajo del umbral de la L40S.
- Se aplicaron optimizaciones como un límite de generación a 1024 tokens y procesamiento por lotes controlado.

3.6. Revisión Cliente

En la socialización con el subdireccion se destacaron los siguientes puntos:

- El sistema **cumple con el objetivo de entregar clasificaciones razonadas y trazables**, lo que representa un avance respecto a los procesos manuales.

- La visualización de múltiples imágenes por predio y el razonamiento textual facilitan la **auditoría y la toma de decisiones expertas**.

- Se valoró la posibilidad de **utilizar una interfaz gráfica para pruebas rapidas**, detectando áreas con alta probabilidad de cambio en el uso del suelo.

- Se sugirió avanzar en una segunda fase con integración de interfaz gráfica, uso de la nube de Azure ML y mejoras en el prompting.

La presentación final fue bien recibida, y se reconoció el potencial del modelo como herramienta innovadora para modernizar la gestión del territorio urbano.

4. Operacionalizar – Industrializar

4.1. Verificación Técnica y Validación de la Solución

Durante la etapa de operacionalización, se llevó a cabo un proceso integral de verificación del código fuente, validación de fórmulas, integridad de la API y funcionalidad general del sistema. La solución, construida sobre FastAPI, fue probada en entornos locales y remotos utilizando herramientas como uvicorn, nvidia-smi y validaciones en producción.

Se garantizó:

- La correcta ejecución del endpoint principal (/classify) para clasificación del uso predial.
- La consistencia de la respuesta JSON generada por el modelo visual-lingüístico.
- La interoperabilidad entre módulos: classifier.py, model.py, types_io.py y app.py.
- El correcto preprocesamiento de imágenes, decodificación de entradas base64 y procesamiento batch en GPU.
- La verificación incluyó pruebas con imágenes reales, comparación contra casos de uso catastral y análisis del rendimiento mediante monitoreo con nvidia-smi y revisión de logs.

4.2. Transferencia

La solución fue desplegada con éxito en un entorno Hugging Face Spaces, utilizando aceleración por GPU NVIDIA L40S, y siguiendo buenas prácticas de contenedorización con Docker (para mayor info DEPLOYMENT.pdf). Este despliegue permitió un acceso remoto confiable, de fácil mantenimiento y actualización.

Componentes transferidos al ambiente de producción:

1. app.py: Aplicación FastAPI.
2. classifier.py: Lógica de inferencia.
3. model.py: Cargue y definición del modelo multimodal.

4. types_io.py: Tipado estricto de entradas y salidas.
5. requirements.txt: Dependencias del entorno.
6. Dockerfile: Configuración para contenedores optimizada para GPU.
7. README.md: Guía de uso del Space.
8. Deployment Guide: Documento con instrucciones completas de despliegue.

Entregables adicionales:

- Manual técnico de uso de API (/classify, /health, /docs).
- Propuesta de mantenimiento mediante auto-escalado en Hugging Face y despliegue alternativo en Docker Compose o Azure ML.
- Variables de entorno documentadas para control de recursos (MAX_BATCH_SIZE, CUDA_VISIBLE_DEVICES, etc.).

4.3. Retrospectiva

Lecciones aprendidas:

- La modularidad del sistema permitió una rápida adaptación del modelo a múltiples entornos.
- Hugging Face Spaces resultó ser un excelente entorno de pruebas, pero requiere atención especial al consumo de memoria.
- La elección de una GPU L40S de 40GB fue adecuada, ya que equilibró costos y rendimiento.
- El uso de contenedores facilitó la transferencia técnica, el escalado futuro y la reproducibilidad de la solución.

Aspectos a mejorar:

- Definir tempranamente la arquitectura objetivo permite evitar reprocesos.

- Incluir herramientas de observabilidad más robustas para ambientes productivos (como Prometheus o Loki para FastAPI).

- Pensar a futuro la implementación en Azure ML.

- Desde el principio tener todas las herramientas listas para el despliegue de una versión experimental (los despliegues deben ser mensual).

4.4.Finalizacion

El proyecto **"Clasificación del Uso Predial de Bogotá"** ha finalizado su desarrollo y ha sido **completamente documentado, probado y desplegado** en un entorno funcional. La solución se encuentra actualmente disponible en:

https://huggingface.co/spaces/viarias/bogota_land_sp **Acceso:** Público (demo) – Privado (producción institucional futura)

La solución se considera **industrializada**, dado que cuenta con:

- Manual de despliegue y uso.
- Código contenedorizado y portable.
- Soporte GPU verificado.
- Documentación de variables, rutas de API y estructura de inferencia.

Con esto, **se cierra la fase de operacionalización y se completa el proceso de industrialización** del caso de uso.

5. Resultados

Al finalizar el proyecto "Clasificación del Uso Predial en Bogotá mediante Modelos Multimodales", se obtuvieron los siguientes resultados técnicos, operativos y estratégicos:

1. Resultados cuantitativos y verificables

- Prototipo funcional completamente operativo para clasificación de uso del suelo con imágenes de fachada, disponible en Hugging Face Spaces.
- Procesamiento automatizado de imágenes multivista con predicciones razonadas y estructuradas en JSON.
- Evaluación del rendimiento técnico con pruebas reales en GPU NVIDIA L40S, validando estabilidad y escalabilidad.
- Reducción estimada de tiempos y costos operativos en tareas de campo para la entidad catastral.

2. Consistencia con los objetivos propuestos

- Se cumplió el objetivo general de desarrollar una solución que automatiza la clasificación del uso del suelo con IA, promoviendo decisiones basadas en datos.
- Los objetivos específicos también se cumplieron, al desarrollarse, validarse y testearse un sistema escalable y replicable, centrado en la mejora de procesos catastrales.

3. Aportes al campo disciplinar y multidisciplinar

- Se introdujo el uso de modelos de lenguaje visual multimodal en contextos urbanos latinoamericanos, abriendo una línea de innovación para el análisis territorial automatizado.

- El enfoque puede ser adoptado por disciplinas como geografía urbana, planificación territorial, ingeniería catastral, informática urbana y políticas públicas.
- Se establece un precedente para desarrollar sistemas de autovalúo, monitoreo urbano inteligente y alertas de cambio predial, con posibles beneficios para otros gobiernos locales en América Latina.

6. Conclusiones

La investigación y desarrollo del sistema de clasificación automatizada del uso predial con inteligencia artificial permite concluir que:

1. **La tecnología de modelos multimodales aplicada a imágenes urbanas tiene alto potencial** para modernizar procesos catastrales, permitiendo análisis más ágiles, precisos y menos dependientes de la inspección en campo.

2. El modelo desarrollado **es capaz de interpretar imágenes y emitir juicios razonados** sobre el uso del suelo, entregando información que no solo predice, sino que también explica, facilitando la validación por parte de expertos humanos.

3. **La solución demuestra ser replicable, modular y escalable**, y puede ser adaptada a otros territorios o incluso a otros dominios de análisis urbano (como uso del espacio público, identificación de tipologías constructivas, etc.).

4. Se logró cerrar el ciclo de desarrollo completo, desde el entrenamiento base y diseño del sistema hasta su despliegue en un entorno accesible, con documentación técnica, propuesta de mantenimiento y prueba de sostenibilidad técnica.

5. El trabajo colaborativo entre expertos técnicos, instituciones públicas y actores académicos demostró ser clave para el éxito del proyecto, y sienta las bases para nuevas iniciativas similares en el marco de **IDECA**, la red de observatorios y otras estrategias distritales de transformación digital.
6. Finalmente, se destaca como **hallazgo inesperado** la capacidad del sistema para detectar patrones en fachadas que podrían incluso sugerir subdivisiones no reportadas, cambios no autorizados de uso o construcción informal, abriendo nuevas oportunidades para su uso como herramienta de fiscalización preventiva.

7. Bibliografía (ICONTEC)

1. *CHEN, Shuyan; LI, Xiaodong; LIANG, Junhao. Kimi: An Open Multimodal LLM for Thinking. 1ª ed. Arxiv.org. Cornell University, 2024. Disponible en: https://arxiv.org/abs/2403.09752 [Consulta: 12 de junio de 2025].*

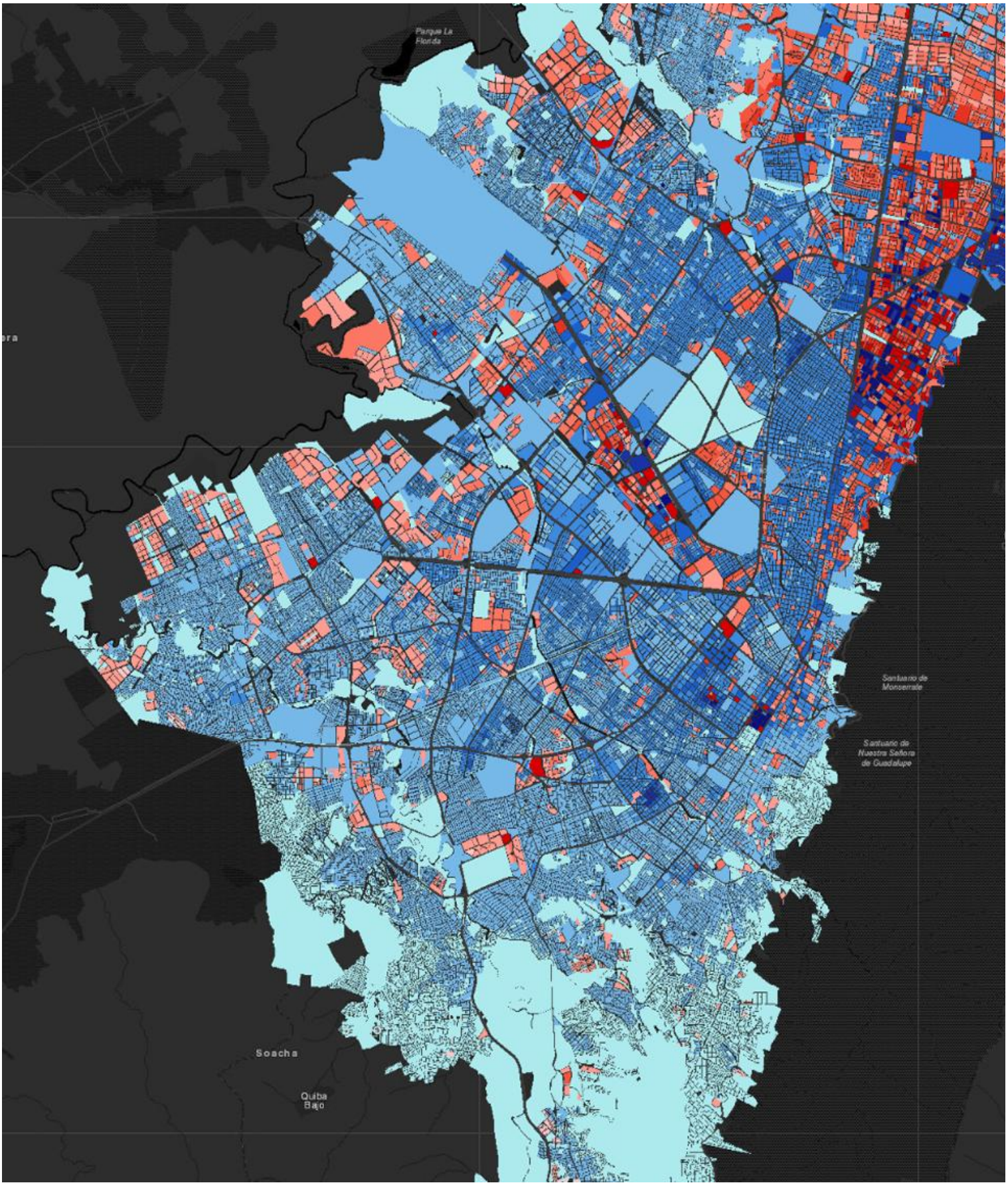
2. *HUGGING FACE INC. Hugging Face Spaces - bogota_land_space. 1ª ed. París: Hugging Face, 2025. Disponible en: https://huggingface.co/spaces/viarias/bogota_land_space [Consulta: 15 de julio de 2025].*

3. *INSTITUTO GEOGRÁFICO AGUSTÍN CODAZZI – IGAC. Manual Técnico de Catastro Multipropósito. 2ª ed. Bogotá: IGAC, 2022. 214 p.*

4. *IDECA – Infraestructura de Datos Espaciales de Bogotá. Portal Oficial. 1ª ed. Bogotá: Alcaldía Mayor de Bogotá, 2025. Disponible en: https://www.ideca.gov.co/ [Consulta: 10 de julio de 2025].*

5. *OPENAI. API Documentation – GPT-4o. 1ª ed. San Francisco: OpenAI, 2024. Disponible en: https://platform.openai.com/docs [Consulta: 14 de julio de 2025].*

6. *GONZÁLEZ, María; RAMÍREZ, Luis. Análisis y uso de imágenes urbanas para sistemas catastrales. 1ª ed. Medellín: Editorial EAFIT, 2023. 168 p.*



ideca

ALCALDÍA MAYOR DE BOGOTÁ D.C.

ideca