Logotipo, Icono

Descripción generada automáticamente

Aprendizaje Profundo

Proyecto 2

"Aplicación de Redes Neuronales para el Análisis de Imágenes Satelitales en la Predicción de Incidencias Delictivas en México"

Juan Pablo Acevedo Robles

Carlos Agustín Manzo Rosas

Daniel Francisco Nuño Álvarez

**Descripción del problema social o ambiental. Resaltando su importancia y justificando porque es necesario utilizar Aprendizaje Profundo para desarrollar una propuesta de solución al problema.**

Según las últimas cifras oficiales del Instituto Nacional de Estadística y Geografía (INEGI), hasta marzo de 2023 se registraron 23,783 incidencias delictivas en México relacionadas con el narcomenudeo, rapto y secuestro. Este creciente problema se puso de manifiesto aún más en 2022 cuando, según el ranking del Consejo Ciudadano para la Seguridad Pública y la Justicia Penal, nueve de las diez ciudades más violentas del mundo se encontraban en México.

La seguridad en México es un tema de gran relevancia y delicadeza, evidenciado por el constante incremento de actividades delictivas, siendo el narcotráfico una de las más predominantes y notorias en nuestra sociedad. Según la última encuesta del INEGI sobre sensación de inseguridad de marzo de 2023, el 62% de la población considera que vivir en el país es inseguro. En ciertas regiones, como Fresnillo, Zacatecas, esta percepción de inseguridad es aún más pronunciada, con un 96% de los encuestados expresando inquietud.

En la era actual de avances tecnológicos, tenemos acceso a una amplia gama de fuentes de información, incluyendo imágenes satelitales que cubren casi todo el planeta. Este vasto corpus de datos puede ser analizado y aprovechado mediante el uso de Redes Neuronales, una de las principales ramas del Aprendizaje Profundo.

Estas tecnologías ofrecen la posibilidad de detectar patrones en las imágenes satelitales que pueden indicar cambios en las áreas residenciales, permitiendo anticipar la transformación de zonas tranquilas en focos de inseguridad y narcotráfico. Al identificar estos patrones a tiempo, las autoridades pueden tomar medidas preventivas y responder de manera más eficaz.

**Descripción breve de los enfoques de solución existentes en la literatura. Se espera incluya citas y las referencias de los trabajos que encontró en la literatura.**

De acuerdo con el documento “Using Deep Learning and Satellite Imagery to Quantify the Impact of the Built Environment on Neigborhood Crime Rates”, plantean que usando reportes de crímenes geo-referenciados e imágenes satelitales de 3 diferentes ciudades de Estados Unidos pudieron demostrar que usando redes neuronales convolucionales pueden explicar hasta un 82% en la variación de los índices de delincuencia en los vecindarios.

Su proceso se describe a continuación:

En primer lugar, se obtuvieron registros de delitos de 2016 georreferenciados y con marca de tiempo proporcionados por los departamentos de aplicación de la ley para cada una de las ciudades estudiadas. Los delitos se agregaron a los sectores censales según los límites establecidos por el Censo de Estados Unidos de 2010. Algunos delitos se separaron en categorías de delitos personales y contra la propiedad. También se obtuvieron estimaciones de características socioeconómicas y de población de la Encuesta de la Comunidad Americana (ACS). El número de delitos por sector censal se dividió por las estimaciones de población de la ACS para obtener el número de incidentes delictivos por cada 1,000 personas, conocido como "tasas de delincuencia".

Luego, se recopilaron cerca de 100,000 imágenes satelitales de Google Static Maps API que abarcan cada sector censal. Estas imágenes no estaban etiquetadas. Para superar el desafío de trabajar con datos sin etiquetar, se utilizó un marco de aprendizaje por transferencia, similar al utilizado por Jean et al. (2016) para predecir la pobreza. Se utilizó una red neuronal convolucional previamente entrenada en la base de datos ImageNet para diferenciar entre 1,000 categorías de objetos. La red se ajustó específicamente para el problema de la predicción de delitos al entrenar el modelo con imágenes de sectores censales con tasas de delincuencia altas y bajas. El modelo actualizado identificó características pertinentes para describir estructuras de vecindarios (como cobertura verde, edificios y carreteras), que podrían ser útiles para asociar las tasas de delincuencia con el medio ambiente.

Finalmente, se ajustó un modelo de regresión para evaluar cuánto explican las características extraídas las tasas de delincuencia. Para reducir la dimensionalidad de la matriz de características y identificar características predictivas relevantes, se utilizó la técnica de red elástica, que es tanto una técnica de regularización como de selección de variables. Se aplicó un enfoque riguroso de ajuste del modelo que implica un proceso de validación cruzada de cinco pliegues.

Para cada una de las tres ciudades, se ajustaron modelos individuales para predecir las tasas de delincuencia en cada ciudad utilizando únicamente las características extraídas de las imágenes satelitales. También evaluaron qué tan bien su modelo predice los delitos personales y contra la propiedad. Además, desarrollaron modelos de regresión adicionales para predecir las tasas de delincuencia utilizando variables demográficas y socioeconómicas que se han estudiado ampliamente en la literatura sobre delincuencia.

En el estudio titulado "Crime Mapping from Satellite Imagery via Deep Learning", se utilizó un enfoque diferente al mencionado anteriormente para predecir mapas de seguridad urbana a escala de ciudad. Este enfoque implicaba el uso de una plataforma de crowdsourcing online, donde se presentaban a los participantes pares aleatorios de 4019 imágenes de Google Street View. Las imágenes provenían de diversas ciudades, como Nueva York, Boston, Salzburgo y Linz.

Se pidió a los 7872 participantes que eligieran, de cada par, la imagen que percibían como más segura. Posteriormente, a cada imagen individual se le asignaba una puntuación de seguridad, basada en los vectores de preferencia acumulados. Cada una de estas imágenes se caracterizó con un conjunto de características visuales genéricas, las cuales fueron utilizadas para entrenar un modelo mediante la técnica de Regresión de Vectores de Soporte.

Este modelo se utilizó más adelante para generar mapas de seguridad a nivel de ciudad para 27 ciudades de los Estados Unidos. Estos mapas fueron predichos directamente a partir de las imágenes de Google Street View correspondientes a cada ciudad. De esta manera, se obtuvo una valiosa herramienta para evaluar y comparar la seguridad urbana a partir de la percepción ciudadana, reflejada en las imágenes de Street View.

**Descripción detallada del enfoque de solución seleccionado y su justificación.**

El enfoque seleccionado para nuestra solución se inspira en el método presentado en el primer documento, con algunas modificaciones clave que se detallan a continuación:

• Optaremos por utilizar imágenes nocturnas de áreas específicas de México, áreas que, basándonos en nuestra experiencia y en los informes noticiosos, sabemos que tienen una fuerte presencia de narcotráfico. Esto se basa en nuestra hipótesis de que podría existir una correlación entre la variación en la cantidad de luces en las coordenadas seleccionadas y la actividad del narcotráfico.

• Para la adquisición de imágenes, emplearemos Google Earth Engine, ya que nos permite segmentar las imágenes por año y coordenadas mediante un script relativamente sencillo. Esta herramienta facilitará la obtención de las imágenes necesarias para nuestro análisis.

• En lo que respecta a los datos delictivos, utilizaremos la base de datos de incidencia delictiva de México proporcionada por el INEGI. Esta base de datos contiene información relevante y precisa, que será de gran utilidad para nuestro estudio.

Este enfoque nos permitirá correlacionar los cambios observables en las imágenes satelitales con los datos de incidencia delictiva, proporcionando una visión más clara de la situación del narcotráfico en las zonas de interés seleccionadas.

A continuación, se describe la obtención de la información:

En el portal web del INEGI, existe una base de datos que proporciona información histórica sobre las incidencias delictivas, detallando la causa, el estado y el municipio, así como el número de incidencias, desde 2015 hasta marzo de 2023. La Figura 1.1 ofrece un ejemplo de cómo se presenta esta información.

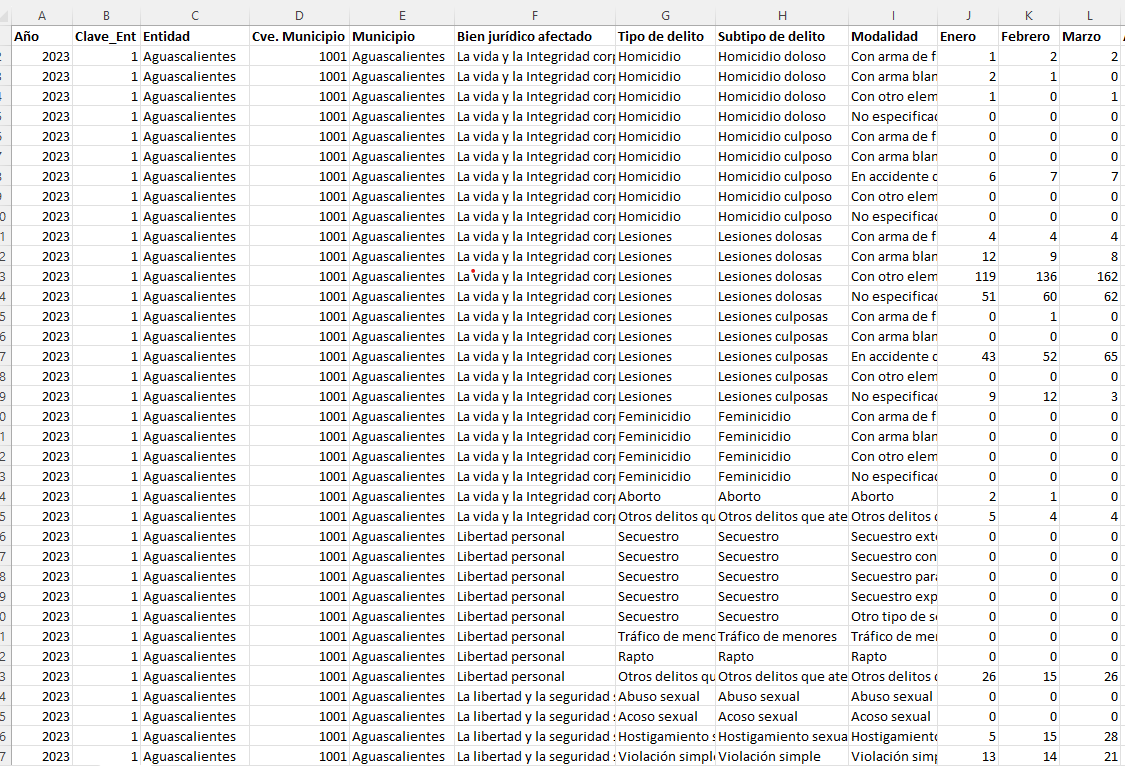


Figura 1.1

La base de datos consiste en un archivo zip, el cual contiene las carpetas separadas por año y un archivo que recopila toda la información todos en formato xlsx.

Para el objetivo que nos hemos propuesto, esta información resulta extremadamente valiosa. Nos permite filtrar los datos necesarios para las ubicaciones seleccionadas, gracias a que se proporcionan con georreferenciación. Sin embargo, hay que tener en cuenta una limitación: la granularidad de los datos geográficos se reduce únicamente hasta el nivel municipal.

En este proyecto, el archivo de Excel constituirá nuestra variable de salida, ya que se utilizarán para comparar las incidencias delictivas con los posibles cambios en los patrones de estas imágenes. Se seleccionaron para el estudio los siguientes lugares:

* Apaseo el Grande, Guanajuato
* Encarnación de Díaz, Jalisco
* Fresnillo, Zacatecas
* Uruapan, Michoacán
* Coyuca de Catalán, Guerrero
* Leonardo Bravo, Guerrero

Para optimizar el uso de nuestros recursos de datos, filtraremos y restringiremos la información a estos municipios seleccionados. Como se puede observar en la Figura 1.2, aunque el formato de los datos sigue siendo el mismo, hemos excluido la información de otros lugares para centrarnos únicamente en las áreas de interés para nuestro estudio.

Interfaz de usuario gráfica, Aplicación, Tabla, Excel

Descripción generada automáticamente

Figura 1.2

En este proyecto, se seleccionó la plataforma Google Earth Engine para la adquisición de imágenes satelitales, ya que proporciona una plataforma potente y flexible para el análisis de datos geoespaciales. Uno de los mayores atractivos de Google Earth Engine es su capacidad para manejar scripts, lo que nos permite automatizar y personalizar el proceso de obtención de imágenes.

El script específico que se utilizó en este proyecto está diseñado para recopilar y exportar una serie de imágenes del satélite VIIRS de NOAA. Este conjunto de datos es particularmente valioso porque proporciona imágenes de luz nocturna, que pueden revelar patrones de actividad humana, como la expansión urbana y las áreas de actividad económica (el script completo se muestra en el apéndice A).

En el script, primero se define un punto y una región de interés (ROI) basada en las coordenadas específicas que queremos investigar. Un ejemplo, sería usar las coordenadas para Uruapan, Michoacán, pero este valor podría cambiarse para cualquier otro lugar de interés.

A continuación, se carga el conjunto de datos del satélite VIIRS y aplicamos dos filtros: uno para seleccionar un rango de fechas específico y otro para limitar las imágenes a nuestra región de interés. Estos filtros permiten personalizar el conjunto de datos que recopilamos, asegurándonos de que las imágenes son relevantes para nuestro estudio y de que cubren el período de tiempo que queremos investigar.

Una vez que se tiene la colección de imágenes, se crea una función para exportar cada imagen a Google Drive. Esta función toma cada imagen, la recorta a la región de interés, y luego la exporta a Google Drive en formato GeoTIFF. Este formato es ideal para este tipo de datos porque puede manejar múltiples bandas de datos y también admite diferentes tipos de datos, como los datos de punto flotante utilizados en este script.

Por último, aplicamos esta función de exportación a cada imagen en nuestra colección. Esto resulta en una serie de imágenes de satélite de nuestra región de interés, cada una de las cuales corresponde a un mes diferente dentro de nuestro rango de fechas seleccionado.

A continuación, se muestra un ejemplo de una visualización de una de las imágenes obtenidas, es poco común encontrar que equipos con visualizadores de imágenes satelitales o en formato TIF, por lo que se utiliza una librería rasterio de Python para poder observarla, se muestra en la Imagen 1.1

Gráfico

Descripción generada automáticamente

Imagen 1.1

Algunos de los detalles técnicos de las imágenes son los siguientes:

* Formato de imagen: Las imágenes se exportan en formato GeoTIFF.
* Bandas: El conjunto de datos VIIRS DNB que se está utilizando tiene una sola banda. Sin embargo, los datos VIIRS en general pueden tener hasta 22 bandas.
* Región de interés: Las imágenes se generan para una región de interés definida en torno a las coordenadas proporcionadas, con un radio de 20 km.
* Resolución espacial: La escala de las imágenes exportadas es de 10 metros, lo que significa que cada píxel representa una superficie de 10x10 metros en el suelo.
* Fecha: El rango de fechas de las imágenes es del 1 de diciembre de 2006 al 31 de enero de 2023.

**Documentación de la arquitectura del modelo base propuesto.**

Para abordar este proyecto, se propone una arquitectura basada en Redes Neuronales Convolucionales (CNN, por sus siglas en inglés) para el análisis de las imágenes satelitales, junto con un sistema de correlación que vincule los cambios observados en las imágenes con los datos de incidencia delictiva.

La arquitectura propuesta consta de las siguientes partes:

1. Datos de entrada: Como se mencionó antes, los datos de entrada serán las imágenes satelitales nocturnas y los datos de incidencia delictiva. Las imágenes satelitales nocturnas serán procesadas y convertidas en un formato que puede ser alimentado en una red neuronal convolucional (CNN). Los datos de incidencia delictiva se utilizarán como etiquetas para el entrenamiento supervisado de la CNN.
2. Preprocesamiento de datos: Para las imágenes satelitales, se aplicará una serie de transformaciones para prepararlas para la CNN. Estas transformaciones pueden incluir la normalización de los píxeles, la eliminación de ruido y la segmentación de las imágenes en regiones más pequeñas para facilitar el entrenamiento de la red. Los datos de incidencia delictiva se preprocesarán para coincidir con las imágenes satelitales en términos de ubicación y tiempo.
3. Red Neuronal Convolucional (CNN): Las CNN son eficientes para el análisis de imágenes, ya que pueden detectar características locales en las imágenes y son invariantes a la translación. En este proyecto, se utilizará una CNN pre-entrenada en ImageNet como base, y se ajustará a nuestras imágenes satelitales mediante transfer learning. La última capa de la CNN se reemplazará con una capa de clasificación binaria, que se entrenará para predecir la presencia o ausencia de actividad delictiva basándose en las características extraídas por la CNN.
4. Entrenamiento y validación: El modelo se entrenará utilizando un conjunto de datos de entrenamiento, y se validará utilizando un conjunto de datos de validación separado. Se empleará una función de pérdida de entropía cruzada binaria para evaluar el rendimiento del modelo durante el entrenamiento, y se utilizará el descenso de gradiente estocástico para optimizar los parámetros del modelo.
5. Evaluación del modelo: Una vez entrenado el modelo, se evaluará su rendimiento en un conjunto de datos de prueba no visto. Se utilizarán métricas como la precisión, la sensibilidad, la especificidad y el área bajo la curva ROC para evaluar el rendimiento del modelo. Además, se realizará una matriz de confusión para visualizar la eficacia del modelo en la clasificación de los datos.
6. Aplicación del modelo: Una vez validado y evaluado el modelo, se aplicará a nuevas imágenes satelitales para predecir la presencia de actividad delictiva en diferentes regiones y en diferentes momentos.

Finalmente los dataset tanto de las imágenes satelitales como los documentos de Excel necesarios se encuentran en la liga de Drive en el apéndice B.

Apéndice A

// Crear un punto con las coordenadas proporcionadas

var punto = ee.Geometry.Rectangle([-100.87677614231562, 20.41943784870882, -100.56675569553828, 20.610116800016385]);

// Crear una región de interés (ROI) alrededor del punto con un radio de 10 km (puedes ajustar el radio según tus necesidades).

var roi = punto.buffer(20000); // 20000 metros = 20 km

// Cargar el conjunto de datos del satélite VIIRS y filtrar por fecha y región.

var collection = ee.ImageCollection('NOAA/VIIRS/DNB/MONTHLY\_V1/VCMCFG')

  .filterDate('2006-12-01', '2023-01-31') // Ajusta las fechas según tus necesidades.

  .filterBounds(roi);

// Función para exportar cada imagen en la colección a Google Drive.

var exportToDrive = function(image) {

  var dateString = new Date(image.get('system:time\_start').getInfo()).toISOString().slice(0, 10);

  var imageName = 'Nighttime\_Coordinates\_' + dateString + '\_todaysbatch';

  Export.image.toDrive({

    image: image.clip(roi).toFloat(),

    description: imageName,

    scale: 10,

    maxPixels: 1e13

  });

};

// Obtener información sobre las imágenes en la colección y aplicar la función de exportación a cada imagen.

var images = collection.getInfo().features;

for (var i = 0; i < images.length; i++) {

  var image = ee.Image(images[i].id);

  exportToDrive(image);

}

Apéndice B

<https://iteso01-my.sharepoint.com/personal/cmanzo_iteso_mx/_layouts/15/onedrive.aspx?id=%2Fpersonal%2Fcmanzo%5Fiteso%5Fmx%2FDocuments%2FMCD%2F04%20Primavera%202023%2F02%20AP%2FProyecto%2002&ga=1>

**Bibliografía**

Secretariado Ejecutivo del Sistema Nacional de Seguridad Pública. (s.f.). Datos abiertos de incidencia delictiva. Gobierno de México. Recuperado de <https://www.gob.mx/sesnsp/acciones-y-programas/datos-abiertos-de-incidencia-delictiva>

A. Maharana, Q. C. Nguyen, E.O. Nsoesie. (2017). Using Deep Learning and Satellite Imagery to Quantify the Impact of the Built Environment on Neigborhood Crime Rates. ArXiv.org. Recuperado de <https://arxiv.org/ftp/arxiv/papers/1710/1710.05483.pdf>

A. Najjar, S. Kaneko, Y. Miyanaga (2018).Crime Mapping from Satellite Imagery via Deep Learning . ArXiv.org. Recuperado de <https://arxiv.org/pdf/1812.06764.pdf>

Instituto Nacional de Estadística y Geografía. (2023). Encuesta Nacional de Seguridad Urbana (ENSU). Recuperado de <https://www.inegi.org.mx/contenidos/programas/ensu/doc/ensu2023_marzo_presentacion_ejecutiva.pdf>