



UNIVERSIDAD
NACIONAL
DE COLOMBIA

Entendiendo comportamientos ciudadanos viales a partir de datos abiertos e inteligencia artificial

Juan Camilo Del Río Cuervo

Universidad Nacional de Colombia

Facultad de Minas

Departamento de Ciencias de la Computación y de la Decisión

Medellín, Colombia

2022

Entendiendo comportamientos ciudadanos viales a partir de datos abiertos e inteligencia artificial

Juan Camilo Del Río Cuervo

Trabajo final presentado como requisito parcial para optar al título de:
Magister en Ingeniería - Analítica

Director:

Ph.D. Juan David Ospina Arango

Línea de Investigación:

Inteligencia artificial

Universidad Nacional de Colombia

Facultad de Minas

Departamento de Ciencias de la Computación y de la Decisión

Medellín, Colombia

2022

Dedicatoria

A mi madre y sus innumerables esfuerzos por verme alcanzar mis sueños.

Resumen

A nivel mundial la accidentalidad vial sigue siendo la segunda causa de muerte en adultos, siendo los peatones y motociclistas las personas con mayor índice de mortalidad y lesiones en los incidentes de tránsito. Es por esto por lo que el presente trabajo tiene como objetivo generar una metodología basada en inteligencia artificial para la identificación y cuantificación automática del comportamiento vial, tanto para peatones como para motociclistas. Para esto, se hizo uso de datos abiertos con el fin de obtener información relevante que permitió el entrenamiento de un modelo de detección de objetos con el cual se puedan extraer características suficientes para medir si las acciones de las personas en la vía son las adecuadas o no. Una vez entrenado el modelo de detección, cuya sensibilidad y especificidad fue mayor del 90%, se logró establecer que el comportamiento de los peatones en el área de estudio tiende a ser inadecuado, ya que más del 70% de estos incumplen las normas de tránsito al encontrarse ubicados por donde deben de transitar los vehículos. Igualmente, se encontró que más de la mitad de los motociclistas realizan acciones que pueden aumentar la probabilidad de verse involucrados en algún incidente de tránsito, siendo transitar entre carriles la acción que mayor nivel de incurrancia tiene. Con los resultados obtenidos a partir de esta prueba de concepto, se logró establecer que los métodos de reconocimiento de patrones basados en inteligencia artificial pueden ser de utilidad al momento de reconocer y cuantificar los comportamientos de los agentes viales en tiempo real.

Palabras clave: Movilidad, Datos Abiertos, Reconocimiento de Objetos, Inteligencia Artificial.

Abstract

Understanding road citizen behaviors from open data and artificial intelligence

Worldwide, road accidents continue to be the second cause of death in adults, with pedestrians and motorcyclists being the people with the highest rate of mortality and injuries in traffic incidents. Therefore, the present work aims to generate a methodology based on artificial intelligence for the automatic identification and quantification of road behavior, both for pedestrians and motorcyclists. For this, open data was used to obtain relevant information that allowed the training of an object detection model with which sufficient characteristics can be extracted to measure whether the actions of the people on the road are appropriate or not. Once the detection model was trained, whose sensitivity and specificity were greater than 90%, it was possible to establish that the behavior of pedestrians in the study area tends to be inappropriate, since more than 70% of these do not comply with traffic regulations because they are located where vehicles must travel. Likewise, it was found that more than half of the motorcyclists carry out actions that can increase the probability of being involved in a traffic incident, with traveling between lanes being the action with the highest level of incurrence. With the results obtained from this proof of concept, it was possible to establish that pattern recognition methods based on artificial intelligence can be useful when recognizing and quantifying the behavior of road agents in real-time.

Keywords: Mobility, Open Data, Object Recognition, Artificial Intelligence.

Contenido

	Pág.
1. Introducción	1
1.1 Motivación.....	3
1.2 Objetivos del trabajo	5
1.2.1 Objetivo General	5
1.2.2 Objetivos Específicos	5
1.3 Estructura del documento	6
2. Antecedentes	7
3. Metodología	10
3.1 Recolección y preprocesamiento de la información	10
3.2 Extracción de la fecha y hora	11
3.3 Etiquetado de la información.....	12
3.4 Detección automática de actores viales	14
3.5 Clasificación de comportamientos viales.....	15
4. Resultados	18
4.1 Detección de objetos	18
4.2 Clasificación de comportamientos viales.....	21
5. Conclusiones y trabajos futuros	29
5.1 Conclusiones	29
5.2 Trabajos futuros	31
Anexo 1. YOLOv5	33
Bibliografía	38

Lista de imágenes

	Pág.
Imagen 3.1. Ejemplo de imágenes recolectadas.....	10
Imagen 3.2. Ejemplo de imágenes recortadas.....	11
Imagen 3.3. a) Valor medio para cada píxel por número. b) Desviación estándar para cada píxel por número.....	12
Imagen 3.4. Etiquetado de imágenes recolectadas.....	13
Imagen 3.5. Cantidad de actores viales por imagen.....	14
Imagen 3.6. Clasificación de píxeles.....	15
Imagen 3.7. a) Motociclista tipo I. b) Motociclista tipo II. c) Peatón tipo III.	17
Imagen 4.1. Curva ROC de entrenamiento.	18
Imagen 4.2. Curva ROC de validación.....	19
Imagen 4.3. Matriz de confusión.	20
Imagen 4.4. Muestra de detecciones en conjunto de validación.	21
Imagen 4.5. Distribución de motociclistas según su posición vial.....	22
Imagen 4.6. Distribución de peatones según su posición vial.	22
Imagen 4.7. Motociclistas detectados por hora.	23
Imagen 4.8. Distribución de motociclistas por hora.	23
Imagen 4.9. Peatones detectados por hora.	24
Imagen 4.10. Distribución de peatones por hora.....	25
Imagen 4.11. Motociclistas detectados por día.	26
Imagen 4.12. Distribución de motociclistas por día.	26
Imagen 4.13. Peatones detectados por día.....	27
Imagen 4.14. Distribución de peatones por día.	28
Imagen A1.1. Arquitecturas genéricas para detección de objetos.....	33
Imagen A1.2. Cross Stage Partial Network. Tomado de (Wang, y otros, 2019).	34
Imagen A1.3. PANet. Tomado de (Liu, Qi, Qin, Shi, & Jia, 2018).....	35
Imagen A1.4. YOLOv1. Tomado de (Redmon, Divvala, Farhadi, & Girshick, 2016).....	35
Imagen A1.5. Variación de la sensibilidad con las épocas.....	37
Imagen A1.6. Variación de la precisión con las épocas.	37

Lista de tablas

	Pág.
Tabla 3.1. Estadísticas de etiquetas por categoría.	13

1.Introducción

Según datos publicados por la Organización Panamericana de la Salud en cuanto a seguridad vial, la segunda causa de mortalidad en adultos en la región de las Américas corresponde a accidentes de tránsito (Organización Panamericana de la Salud, 2019). La anterior estadística es de especial importancia si se analiza en conjunto con los Objetivos de Desarrollo Sostenible (ODS) establecidos para el año 2030 por la Organización de Naciones Unidas, en donde la meta 6 del ODS 3 establecía inicialmente una reducción a la mitad del número de muertes y lesiones a raíz de accidentes de tránsito para el año 2020 (Organización de las Naciones Unidas, 2018), cifra que la OPS confirmó que no se alcanzó. Al analizar en detenimiento las razones del porqué de este incumplimiento se destaca la incidencia que tienen los motociclistas, peatones y ciclistas, ya que estos tres actores viales son los de mayor vulnerabilidad al estar expuestos por completo ante cualquier incidente vial, y los datos recopilados por la OPS estiman que en conjunto superan el 50% de las víctimas fatales en las vías del continente, siendo el actor más vulnerable de estos los motociclistas ya que aportan el 23% de las muertes totales y el de mayor aumento en los últimos años.

A raíz de la importancia que representa a nivel mundial la seguridad vial debido a la cantidad de muertes y lesiones anuales registradas en relación con los accidentes de tránsito surge la necesidad de establecer estrategias que logren disminuir estas cifras. Para esto se deben establecer políticas públicas claras por parte de los gobiernos, al igual que planes de acción en donde se logre concientizar a las personas de cuáles son aquellas situaciones cotidianas que pueden aumentar el riesgo de verse involucrado en un incidente vial. Para lo anterior resulta conveniente realizar una caracterización de los comportamientos que aumentan la probabilidad de ocurrencia de un accidente y que puedan desencadenar en una lesión o incluso la muerte, al igual que una metodología capaz de evaluar de manera constante en el tiempo la evolución de los comportamientos de los actores viales y así, establecer si las acciones emprendidas por parte de los gobiernos han sido de utilidad, o en caso de que no sea así, establecer una nueva hoja de ruta en cuanto a las medidas necesarias que deben de ser adoptadas.

Con el objetivo de entender estos comportamientos viales, en el presente trabajo de profundización se desarrolla una metodología con la cual caracterizar comportamientos de agentes viales y se implementa una prueba de concepto en el Valle de Aburrá. La metodología desarrollada podrá servir como insumo para la toma de decisiones al momento de la elaboración de planes de acción que busquen la reducción de incidentes viales en la ciudad. Dado que los dos actores viales que presentan una mayor vulnerabilidad debido a las lesiones y muertes registradas anualmente son los peatones y motociclistas, el presente trabajo se enfocará en evaluar las acciones llevadas a cabo por estos grupos.

Para prueba de concepto llevada a cabo en el Valle de Aburrá, se hizo uso de la iniciativa que ha venido desarrollando la ciudad de Medellín en los últimos años de brindar acceso a la ciudadanía a su plataforma de datos abiertos Medata. Igualmente, por la naturaleza misma del problema abordado en este trabajo, se utilizaron imágenes de lo sucedido en las calles de la ciudad, para lo cual la base de datos idónea es la llamada “*Cámaras CCTV Movilidad – SIMM*” (Alcaldía de Medellín, 2021), base de datos que alberga las imágenes en tiempo real de 73 cámaras de vigilancia vial distribuidas a lo largo de la ciudad. Mediante la información que se recolectó de este portal, se lograron extraer los descriptores apropiados para describir el comportamiento de los ciudadanos, con el fin último de caracterizar la forma en que los usuarios vulnerables interactúan en las vías de la ciudad, y así generar información de utilidad que sirva para la toma de decisiones por parte de las autoridades al momento de planear acciones encaminadas a mejorar la seguridad vial en el territorio.

Debido a la cantidad y naturaleza de la información disponible para el presente trabajo, se hizo uso de técnicas de visión por computador basadas en inteligencia artificial para la extracción automática de los descriptores de comportamiento de los ciudadanos que transitan en las vías. Con esto se obtuvo de manera automática el contexto de una escena en particular en cuanto al posicionamiento tanto de los motociclistas como de los peatones, cuantificando así el nivel de seguridad vial. Para lo anterior, se emplearon técnicas de reconocimiento de objetos en donde se logró identificar la posición relativa de los peatones como de los motociclistas, evaluando así si su comportamiento en presencia de los demás

usuarios de la vía es el adecuado o no, y de esta forma dar un indicio si las acciones llevadas a cabo por estos usuarios pueden aumentar el riesgo de ocasionar un incidente vial.

Motivación

En la actualidad, el área metropolitana de la ciudad de Medellín se encuentra inmersa en diversas obras de infraestructura con el objetivo de mejorar la movilidad vial. Al consultar la cantidad de obras que en la actualidad se vienen desarrollando es posible encontrar que existen 23 adecuaciones, ampliaciones o construcciones en ejecución, 156 en etapa de diseño y 5 más en donde se están adquiriendo los lotes para su posterior desarrollo (Alcaldía de Medellín, 2021). Estos planes de adecuación de la infraestructura vial tienen como objetivo entre otros, estar preparados para evitar un posible colapso en la movilidad ante el constante aumento del parque automotor en el área metropolitana, dado que a corte de julio del 2020 se contaba con 321.902 de vehículos registrados solo en la ciudad de Medellín (Alcaldía de Medellín, 2020), adicional a la cifra revelada por el observatorio Medellín Cómo Vamos en su informe sobre movilidad y espacio público del año 2019, en donde indica que la cantidad estimada del parque automotor en el valle de Aburrá para ese año ascendía a 1.756.893 vehículos (Medellín cómo vamos, 2020), en donde resalta que cerca del 56% de estos son motocicletas. Ante el incremento constante del parque automotor que se ha desarrollado en la región durante los últimos años, la percepción ciudadana acerca de la movilidad es pesimista debido a que consideran que el aumento de la cantidad de vehículos circulando por las calles de la ciudad no va en concordancia con el aumento de las vías y obras de infraestructura vial requerida.

Al tener en cuenta la cantidad de obras de infraestructura vial que se están desarrollando actualmente en la ciudad, el aumento constante del parque automotor y la percepción ciudadana pesimista en cuanto a la movilidad al interior del territorio se observa que confluyen diversos factores de riesgo que pueden ocasionar una pérdida de la cultura por parte de los actores viales, lo que a su vez puede ser uno de los posibles causantes que originan incidentes de tránsito en la ciudad. Lo anterior ha tenido relevancia al interior de

los entes de control, siendo evidencia de esto la página web dispuesta por parte de la alcaldía de Medellín en donde se cuenta con material audiovisual dirigido a los actores viales con el objetivo de mejorar su consciencia vial (Alcaldía de Medellín, 2021), al igual que se venían realizando una gran cantidad de talleres sobre movilidad segura antes de comenzar las restricciones de movilidad y conglomeración de personas debido a la pandemia de la covid-19 (Alcaldía de Medellín, 2018).

Por otro lado, al analizar con detalle las estadísticas acerca de los incidentes viales es posible observar la importancia que representa los motociclistas y peatones, dado que son estos grupos de actores viales los que representan un mayor porcentaje de los lesionados y muertos. La distribución de víctimas fatales debido a incidentes viales durante los años 2009 y 2019 en Colombia presenta que en primer lugar se encuentran los motociclistas con 32.043 decesos y en segundo lugar los peatones con 19.342 (Agencia Nacional de Seguridad Vial, 2020). Igualmente, al analizar esta situación de manera local el panorama es idéntico, dado que el porcentaje tanto de lesionados como de muertos al momento de ocurrido un incidente vial en la ciudad supera el 70%, sumando peatones y motociclistas (Alcaldía de Medellín, 2021). Estas cifras más allá de ser un hecho aislado en nuestro país representan un común denominador a nivel mundial, ya que según estadísticas recopiladas por la Organización Mundial de la Salud los “usuarios vulnerables de las vías”, denominación que se le dan a los peatones y motociclistas, representan la mitad de las muertes por accidente de tránsito (Organización Mundial de la Salud, 2017).

Aunque si bien el tener las estadísticas claras acerca de las vidas que se ven afectadas año a año debido a incidentes viales es un buen inicio para poder mejorar la seguridad vial en las carreteras dado que brinda un panorama general de esta, no es suficiente. De acuerdo con (Hyder, 2019), la clave para el mejoramiento de la seguridad vial es la implementación de medidas, en donde se destaca que definir programas de educación para las personas puede no tener el efecto deseado por parte de los gobiernos ya que generalmente estos esfuerzos se realizan de manera descontextualizada y no se tiene en cuenta las características propias de cada zona, por lo que enfatiza que se debe de realizar un análisis contextual de los comportamientos de las personas en cada región con el fin de poder desarrollar medidas específicas que sean de alto impacto, donde la fuente

principal de la información requerida para guiar las políticas públicas debe de ser recolectada directamente en campo.

Por lo mencionado anteriormente, en la presente trabajo se tiene como objetivo proponer una metodología con la cual se pueda realizar seguimiento de manera automática de los comportamientos ciudadanos la pueda servir de soporte al momento de realizar tomas de decisiones por parte de los entes de control encargados de establecer las políticas públicas y los mecanismos de prevención y control requeridos para aumentar la seguridad vial, con la que se espera poder brindar un contexto realista de los comportamientos que tienen los usuarios de las vías que puedan ocasionar un incremento en la probabilidad de ocurrencia de un incidente vial, enfocándose únicamente en los motociclistas y peatones dado su alto porcentaje en las fatalidades y lesiones. Finalmente, se realizará una prueba de concepto con la información recopilada de una cámara perteneciente al CCTV que pone a disposición de los ciudadanos la alcaldía de Medellín, específicamente con la situada en la Carrera 50FF con Calle 8 Sur.

Objetivos del trabajo

Objetivo General

Desarrollar una metodología basada en datos abiertos e inteligencia artificial para hacer seguimiento a los actores viales.

Objetivos Específicos

- I. Identificar aquellos comportamientos de los actores viales que puedan ser usados para la caracterización de la seguridad vial de motociclistas y peatones.
- II. Analizar las fuentes de datos abiertos disponibles que sean propicias para la obtención de información relevante en la investigación.

- III. Generar una metodología basada en inteligencia artificial para el seguimiento a los comportamientos de motociclistas y peatones.
- IV. Desarrollar una prueba de concepto de la metodología con la información disponible del municipio de Medellín.

Estructura del documento

El presente documento se encuentra organizado de la siguiente manera: en el capítulo 1 se encuentra la introducción, el capítulo 2 presenta los antecedentes del problema a investigar, el capítulo 3 describe el método para la detección de actores viales y la cuantificación de sus comportamientos, el capítulo 4 muestra los resultados obtenidos de, y el capítulo 5 presenta las conclusiones y recomendaciones propuestas.

2. Antecedentes

Para poder llevar a cabo la medición de la seguridad vial y los comportamientos de los actores viales, es necesario contar con herramientas que permitan identificar de manera automática a estos con el objetivo de establecer posteriormente si su posición dentro de una escena vial es la correcta o no. Por esto, se hace necesario identificar aquellas investigaciones que se han llevado a cabo durante los últimos años con el objetivo de extraer automáticamente características que sean útiles para la descripción del comportamiento tanto de peatones como de motos.

Aunque han sido numerosas las investigaciones en donde el objetivo principal es identificar peatones y/o motociclistas en imágenes viales, los principios de estas son similares. Dentro de las primeras aproximaciones que se realizaron se encuentran la detección automática de peatones haciendo uso de filtros con los que se puedan extraer características que sirven de insumo para algoritmos de aprendizaje automático. Autores como (Cheng & Cheng, 2010), (Xu & Gao, 2010), (Kuo, Yang, & Yen, 2012), (Guo, Ge, Zhang, Li, & Zhao, 2012), (Binh & Kim, 2013) y (Zhao, He, Zhang, & Liang, 2015) hacen uso de filtros de Haar para detección de características tales como líneas, bordes y diagonales construyendo así un espacio de características propias de peatones para posteriormente usarlas como insumos de diversos clasificadores tales como máquinas de soporte vectorial (SVM), regresión logística, AdaBoost, clasificadores en cascada, entre otros, logrando así identificar de manera automática la posición relativa de peatones en una imagen. De manera similar, autores como (Chang & Hsu, 2010), (Silva, y otros, 2013) y (Anaya, Ponz, García, & Talavera, 2017) aplican los mismos conceptos, extrayendo de manera automática características y haciendo uso de estas en algoritmos de aprendizaje automático con motociclistas. Si bien esta aproximación para extraer de manera automática las coordenadas relativas a una imagen en donde se localizan peatones y motociclistas puede alcanzar precisiones por encima del 90%, presentan algunas desventajas que pueden hacerlos obsoletos, destacando la necesidad de sensores especiales en las cámaras que registran la información como el tiempo requerido para analizar una sola imagen, tardando en ocasiones varios segundos para dicha tarea.

Posteriormente, con el auge que tuvieron las redes neuronales convolucionales en la segunda mitad de la década pasada se desarrollaron diversos algoritmos para detección de objetos de manera automática, los cuales presentan velocidades de procesamiento superior, variando de unos cuantos cuadros por segundo hasta cientos, y pudiendo ser usados con imágenes tomadas con cámaras sin ningún sensor especial. Dentro de las metodologías desarrolladas que hacen uso de dichas redes convolucionales se encuentran Fast R-CNN (Girshick, 2015), Single Shot Multibox Detector (SSD) (Liu, y otros, 2016), You Only Look Once (YOLO) (Redmon, Divvala, Girshick, & Farhadi, 2016), Faster R-CNN (Ren, He, Girshick, & Sun, 2017), entre otras modificaciones que han tenido estos a lo largo de los años para mejorar su desempeño. Estos algoritmos comparten cierta parte de su idea estructural, la cual hace uso de redes neuronales convolucionales que son entrenadas con anterioridad en algún conjunto de datos abiertos como ImageNet (Russakovsky, y otros, 2015) o COCO (Lin, y otros, 2014), lo cual provee a los investigadores que deseen hacer uso de dichas arquitecturas un modelo sencillo de entrenar en algún problema específico posteriormente. Autores como (Espinosa, Valestín, & Branch, 2019), (Velasstín, Fernández, Espinosa, & Bay, 2020), (Jin, y otros, 2021) y (Dai, y otros, 2021) han desarrollado investigaciones donde se hace uso directa o indirectamente de las arquitecturas mencionadas anteriormente para realizar la detección automática de peatones y/o motociclistas y alcanzando en ocasiones precisiones por encima del 95%.

Si bien las investigaciones que se han desarrollado con el objetivo de realizar la detección automática de peatones y motociclistas han arrojado resultados sobresalientes en cuanto a la precisión que se logra alcanzar, dichas detecciones generalmente no se usan para contextualizar los comportamientos viales de las personas. Al indagar acerca de investigaciones que hayan sido desarrolladas con el objetivo de brindar información adicional acerca de los comportamientos ciudadanos, (Pramanik, Sarkar, & Maiti, 2021) hacen uso de técnicas de detección de objetos para poder identificar de manera automática diversas conductas que pueden servir para la medición de la seguridad vial tales como parqueo de automóviles en zonas prohibidas, cargue y descargue de pasajeros en zonas prohibidas, violación del límite de velocidad establecido, conducción en contravía y adelantamiento de un vehículo sobre otro. Para lograr esta detección de conductas peligrosas, los investigadores proponen una metodología de dos etapas, en donde la primera consta de una detección de actores viales (vehículos, camiones, peatones) con un

algoritmo de detección automática de objetos, y posteriormente se aplica un una discriminación basada en un umbral el cual puede variar de acuerdo al comportamiento que se está evaluando, con el objetivo de determinar si la detección realizada se puede clasificar como un comportamiento inadecuado o no. Por otro lado, autores como (Wilhelm & Lin, 2020), (Lin, Deng, Albers, & Wilhelm, 2020) y (Jaganathan, prabhu, Chinnasamy, & Suresh, 2020) hacen uso de técnicas de detección de objetos para determinar cuándo un motociclista lleva puesto el casco de seguridad o no, permitiendo así tener una medida de la cantidad de personas que cumplen con esta norma de seguridad vial en un entorno específico. De manera similar, (Zhang, 2020) y (de Goma, Bautista, Eviota, & Lopena, 2020) hacen uso de técnicas de aprendizaje profundo para detectar cuando hay infracción de cruzar un semáforo en rojo, en donde se detecta automáticamente al actor vial de interés y su posición relativa en la escena, teniendo en cuenta al tiempo los tiempos en que el semáforo se encuentra disponible para pasar.

3. Metodología

3.1 Recolección y preprocesamiento de la información

Para el caso de estudio, se recolectó información de acceso público que dispone la iniciativa de datos abierto *Medata* de la Alcaldía de Medellín (Alcaldía de Medellín, 2022), identificando que la cámara de video localizada en la Carrera 50FF # 8 Sur brinda la información necesaria para la prueba de concepto que busca brindar el presente trabajo. Una de las desventajas encontradas al momento de recolectar la información necesaria es el tiempo que existe entre la publicación de dos fotogramas diferentes, ya que *Medata* únicamente actualiza esta aproximadamente cada 6 minutos, por lo que se logró recolectar 2470 fotogramas desde el 6 de marzo al 27 de marzo de 2021. En la Imagen 3.1, se muestra un ejemplo de las imágenes que fueron recolectadas para el presente trabajo.

Imagen 3.1. Ejemplo de imágenes recolectadas.



La información que fue recolectada abarca dos sentidos de flujo vehicular, sin embargo, para el presente trabajo se tendrá en cuenta únicamente el sentido que se encuentra

ubicado a la derecha de las imágenes, donde se realizó un recorte a todas las imágenes en el píxel 500 del eje x. En la Imagen 3.2, se observa un ejemplo del resultado de dicho recorte.

Imagen 3.2. Ejemplo de imágenes recortadas.

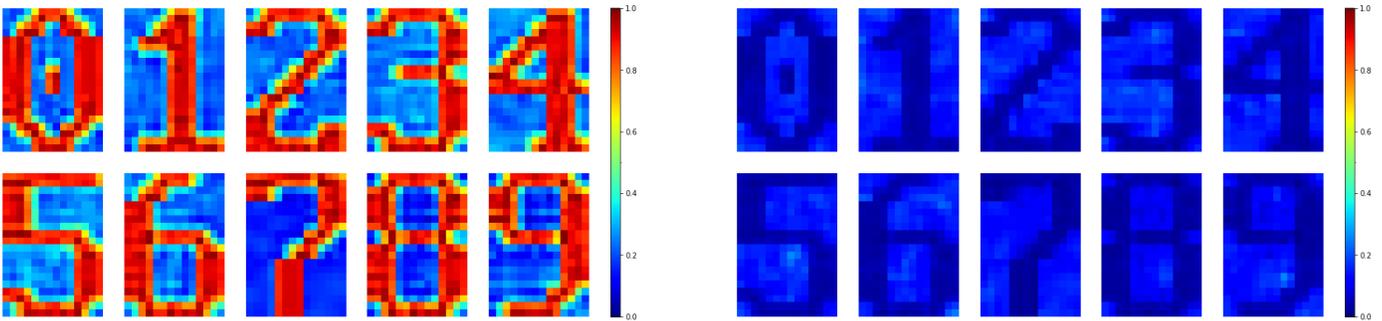


3.2 Extracción de la fecha y hora

Para la correcta caracterización y toma de decisiones en cuanto a los planes de acción a tomar por parte de las autoridades, se hace necesario la identificación de fecha y hora en la que cada imagen fue recolectada, esto con el fin de poder establecer si existe algún patrón en cuanto a las conductas ciudadanas que sean dependientes de alguna hora en específico, de algún día de la semana, entre otros. Para esto, se hará uso de la información que se encuentra en la parte superior derecha de cada imagen en donde se registra la fecha y hora en el formato DD/MM/AAAA HH:mm:ss. Así, en cada imagen es posible extraer 14 dígitos que varían entre 0 y 9, los cuales deberán de ser clasificados.

Para esto, se guardó independientemente 100 imágenes por cada dígito. En las Imágenes 3.3.a y 3.3.b, se observa cual es el valor promedio y la desviación que presenta cada dígito en cada píxel.

Imagen 3.3. a) Valor medio para cada píxel por número. **b)** Desviación estándar para cada píxel por número.

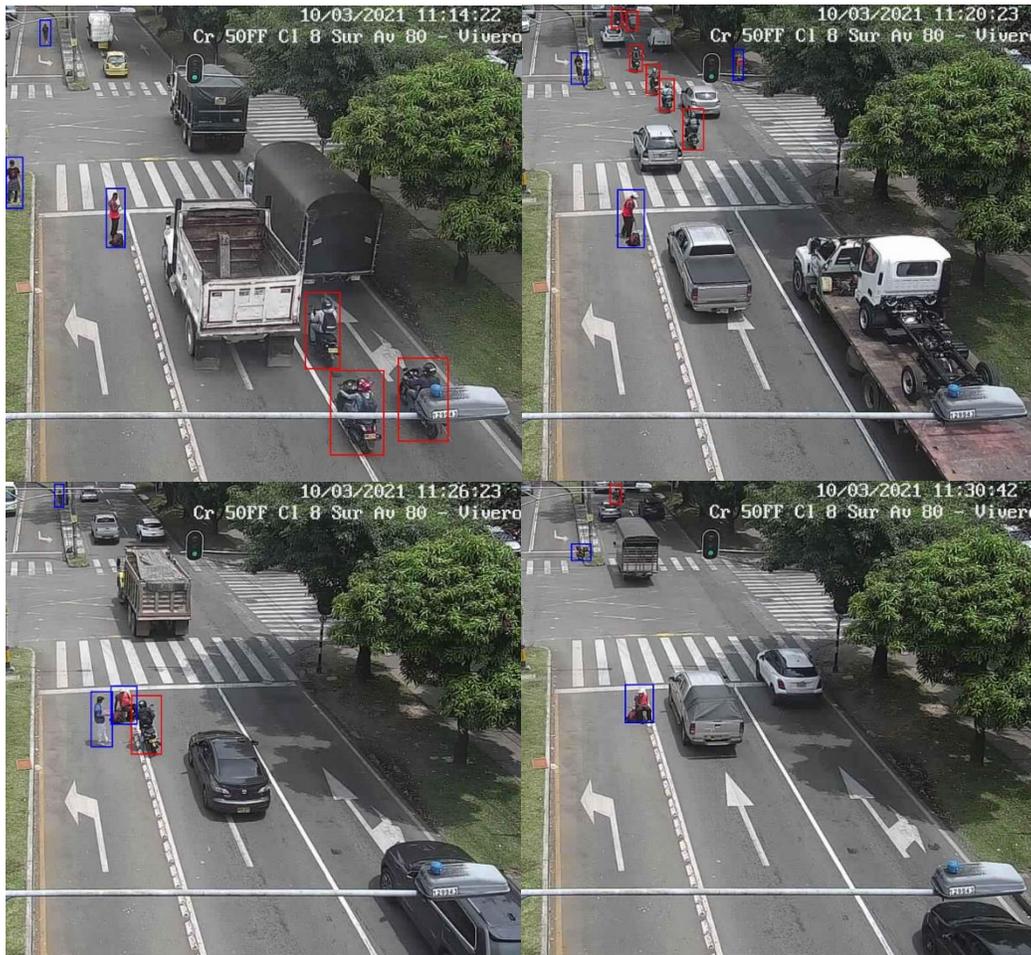


En las Imágenes 3.3.a y 3.3.b es posible observar que en promedio los valores presentan un valor de 1 en las zonas que forman cada dígito y la desviación en dicha zona es cercana a 0, por lo que dichos dígitos se mantienen constantes para las imágenes. Por lo anterior se optó por generar un modelo de vecinos cercanos (KNN) con parámetro K de 1 para la clasificación de cada dígito, y de esta manera extraer automáticamente la fecha y hora de cada imagen. Con esta metodología, se logró obtener una exactitud y precisión de 1, por lo que se tiene certeza de que la fecha extraída es la correcta, y por lo tanto los análisis posteriores que tomen en cuenta dicha información será confiables.

3.3 Etiquetado de la información

Una vez se recolectó y recortó las imágenes, se procedió a realizar el etiquetado de estas. Dado que el problema abordado requiere de la identificación de objetos al interior de una imagen, las etiquetas en este caso corresponden a la creación de cajas que encierren en su interior al objeto y objetos de interés, los cuales en el presente trabajo son dos: peatones y motociclistas. Para esto, se hizo uso del proyecto de código abierto *LabelImg* (tzutalin, 2015), en donde se logra obtener el centro de cada objeto (x, y) junto al ancho y la altura de este (w, h) . En la Imagen 3.4, se muestran algunas imágenes en donde se realizó dicho etiquetado.

Imagen 3.4. Etiquetado de imágenes recolectadas.



En la Ilustración 3.4 se observan 4 diferentes momentos en donde se realizó el etiquetado de los peatones y motociclistas que aparecen hacen parte de la vía, en color azul y rojo respectivamente. En la Tabla 3.1 se detalla las estadísticas de las etiquetas, tanto para peatones como para motociclistas por cada fotograma registrado, y en la Imagen 3.5 se muestra gráficamente la distribución de etiquetas por cada uno de estos.

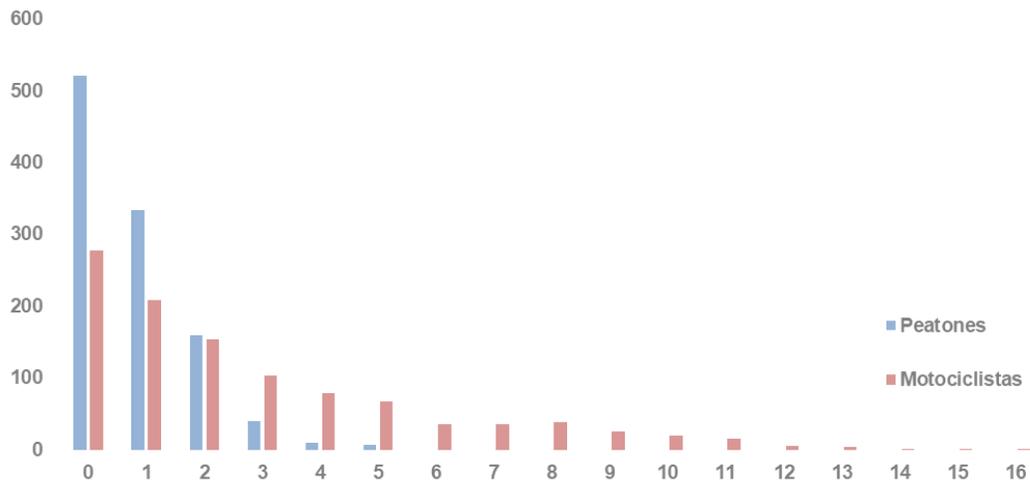
Tabla 3.1. Estadísticas de etiquetas por categoría.

	Peatones	Motociclistas
Mínimo	0	0
P25	0	0

Promedio	0.79	2.79
P50	1	2
P75	1	4
Máximo	5	16
Total Etiquetas	841	2980
Total Imágenes	1069	

Imagen 3.5. Cantidad de actores viales por imagen.

Actor vial por imagen



3.4 Detección automática de actores viales

Una vez las imágenes fueron recortadas y etiquetadas, se procedió a implementar el algoritmo YOLOv4 (Anexo 1), en donde se hizo uso de *transfer learning* para aumentar la precisión de las detecciones realizadas. Por defecto el framework empleado realiza una discriminación del objeto detectado cuando la confianza de la predicción está por encima de 0.5, sin embargo, este parámetro puede ajustarse con el objetivo de obtener mejores resultados, procedimiento que se explicará con más detalle en el capítulo 4.

3.5 Clasificación de comportamientos viales

Una vez se logra identificar las coordenadas en las que se encuentran localizados tanto los peatones como los motociclistas al interior de cada imagen, se requiere clasificar si estos están inmersos en algún comportamiento que pueda aumentar el riesgo de sufrir un incidente de tránsito. Al consultar la normatividad colombiana, es posible encontrar aquellas acciones que están consideradas como infracciones de tránsito debido al riesgo que representan para la seguridad vial (Congreso de la República, 2002), de las cuales se seleccionarán las siguientes para ser detectadas automáticamente por ser consideradas las de mayor riesgo tanto para los peatones como para los motociclistas:

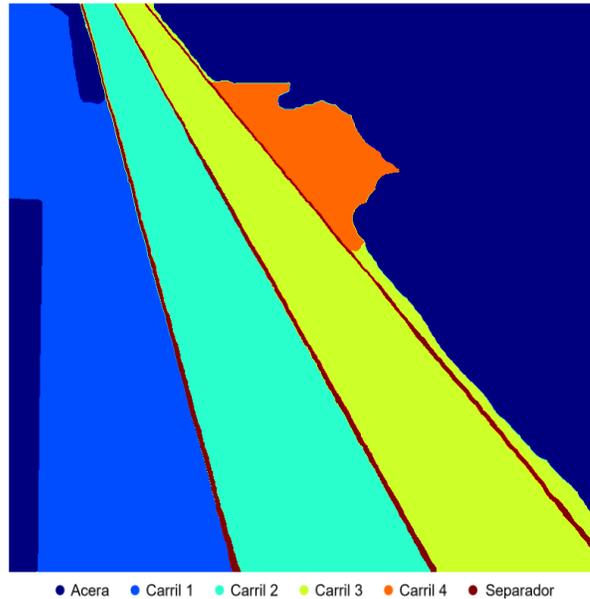
- I. **Infracción A9:** Transitar entre dos vehículos automotores que estén en sus respectivos carriles
- II. **Infracción D7:** Conducir realizando maniobras altamente peligrosas o irresponsables que pongan en peligro a las personas o las cosas.
- III. **Infracción F7:** Cruzar la vía atravesando el tráfico vehicular en lugares en donde existen pasos peatonales.

Con base en las anteriores infracciones, se identificaron los siguientes comportamientos a ser detectados automáticamente en las imágenes recolectadas:

- I. Aquellos motociclistas que transiten en medio de dos carriles.
- II. Aquellos motociclistas que transiten en un mismo carril junto a otro motociclista.
- III. Aquellos peatones que atraviesen la vía por lugares no permitidos o con el semáforo en rojo.

Para la detección de dichos comportamientos en la vía una vez se tiene las coordenadas de donde están ubicados los peatones y motociclistas, se hizo uso de reglas basada en la posición relativa al interior de cada imagen mediante la clasificación de cada píxel dentro

de la imagen, ubicando a estos en diversas categorías. En la Imagen 3.6 se observa una descripción gráfica de lo anterior.



Con base en esta segmentación mostrada en la Imagen 3.6, adicional a la detección de peatones y motociclistas se pueden generar las siguientes reglas para determinar si estos están inmersos en alguno de los comportamientos descritos con anterioridad:

1. Para un motociclista cuya base de detección en el eje x tenga pixeles de más de un carril diferente al carril 4, y el porcentaje máximo para ambos carriles sea menor a un umbral δ_{EC} . En este caso, el motociclista se encuentra entre la mitad de dos carriles y está inmerso la acción indebida I.
2. Para un par de motociclistas que transiten por un mismo carril, y compartan en el y una proporción mayor o igual a un umbral δ_{MC} . En este caso, ambos motociclistas se encuentran inmersos en la acción indebida II
3. Cuando la cantidad de pixeles clasificados como “Acera” que abarca un peatón sobre la base de su detección en el eje x tenga un porcentaje menor a un umbral δ_p . En este caso, este peatón se encuentra en la calle por donde transitan los vehículos y se encuentra inmerso en la acción indebida III.

En la Imagen 3.7, se muestra de manera gráfica la descripción de cada una de estas situaciones.

Imagen 3.7. a) Motociclista tipo I. b) Motociclista tipo II. c) Peatón tipo III.



Para la estimación de los umbrales δ_{EC} , δ_{MC} y δ_P no se cuenta con una métrica de evaluación objetiva que permita calcularlos de manera automática, por lo que obtención se realizó de manera subjetiva, mediante la evaluación visual de las clasificaciones realizadas.

4.Resultados

En esta sección se mostrarán los resultados obtenidos a partide la detección de objetos y clasificación de comportamientos viales, tanto para peatones como para motociclistas

4.1 Detección de objetos

En las Imágenes 4.1 y 4.2 se muestran las curvas ROC de entrenamiento y validación, respectivamente, en donde se observa el desempeño del framework empleado para la detección automática, tanto de los peatones como de los motociclistas.

Imagen 4.1. Curva ROC de entrenamiento.

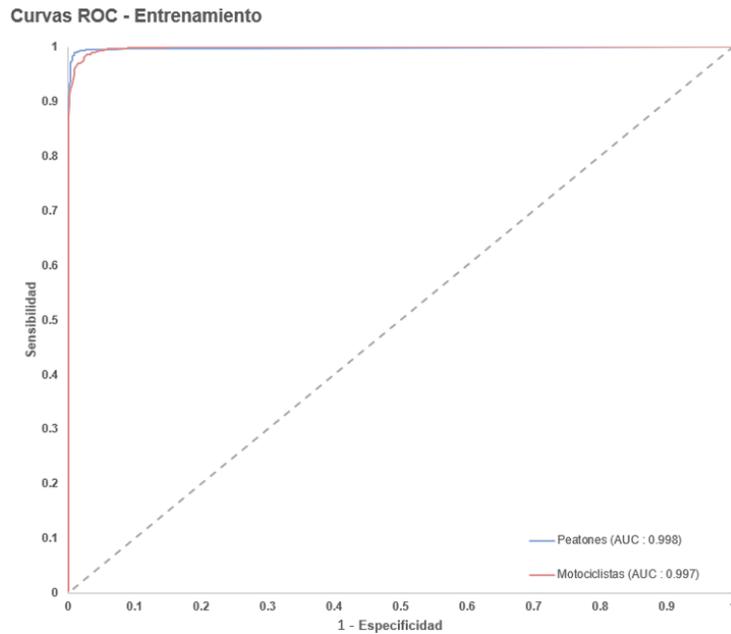
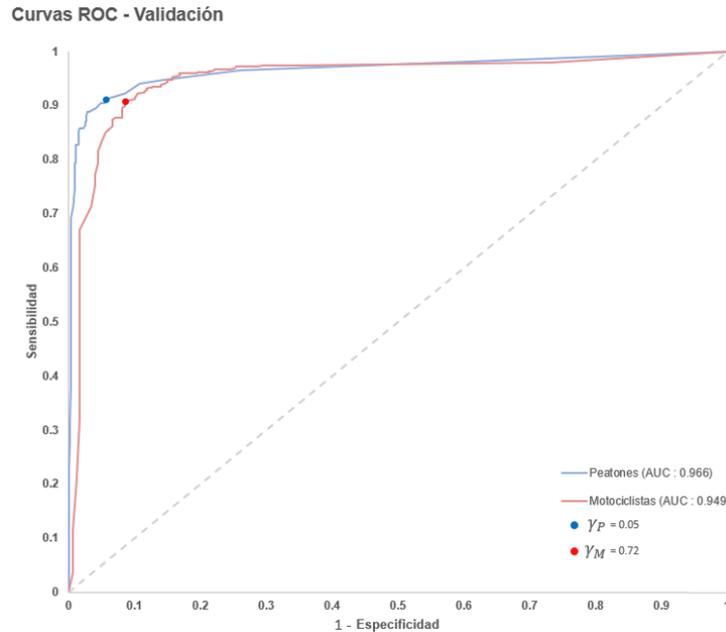


Imagen 4.2. Curva ROC de validación.



Al analizar los resultados obtenidos de validación, se identificaron los umbrales γ_P y γ_M , valores correspondientes al umbral de la confianza obtenida a partir de las predicciones realizadas por el modelo para la categoría Peatón y Motociclista, respectivamente, en el cual se obtiene la sensibilidad y especificidad más cercana al punto óptimo, es decir, una sensibilidad y especificidad de 1. Así, el valor óptimo de γ_P es 0.05, valor con el que se obtiene una sensibilidad y especificidad de 91.1% y 94.3% respectivamente para la categoría de Peatón; y de manera similar el valor óptimo de γ_M es 0.72 con una sensibilidad y especificidad de 90.4% y 91.4% respectivamente para la categoría Motociclista. Haciendo uso de los umbrales anteriores, se logra obtener la matriz de confusión para el conjunto de validación, la cual es mostrada en la Imagen 4.3.

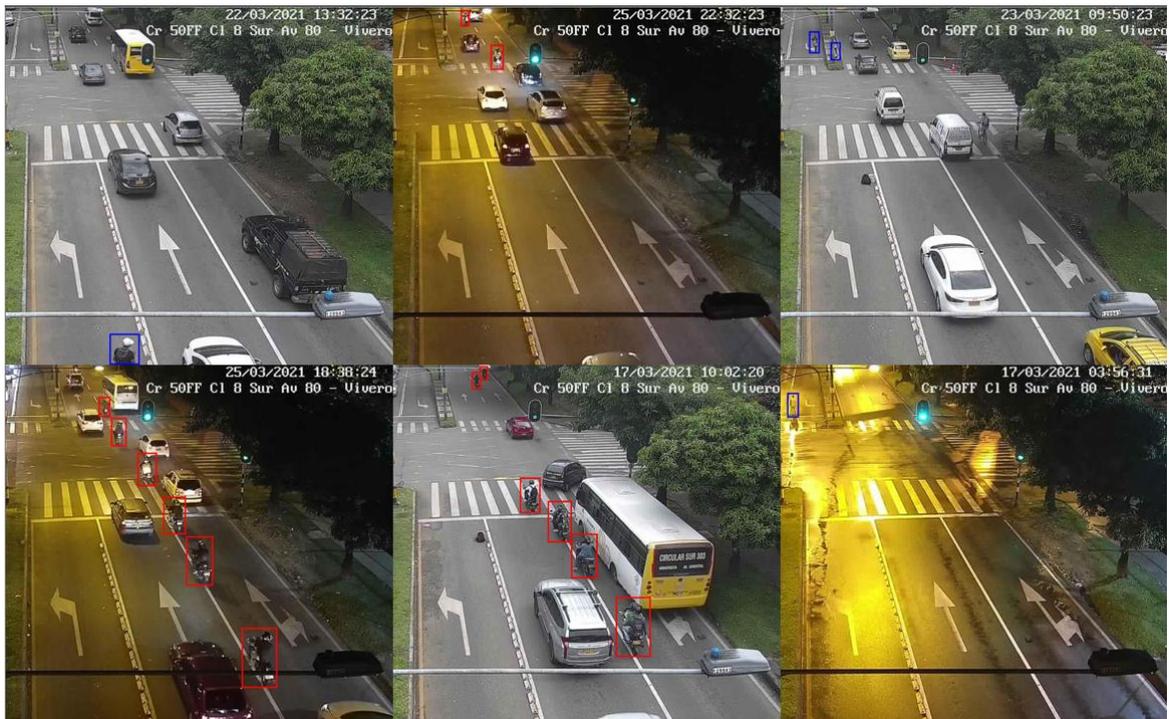
Imagen 4.3. Matriz de confusión.

		Categoría		
		P	M	FP
Predicción	P	154	2	41
	M	3	580	13
	ND	11	57	-

P: Peatón M: Motociclista FP: Falsos positivos ND: No detectado

Debido a la naturaleza del problema abordado, en donde para cada imagen analizada puede existir más de un objeto de interés por cada categoría, específicamente peatones o motociclistas, se hizo necesario incluir dos “categorías” adicionales: *Falsos Positivos* y *No Detectados*. Así, los Falsos Positivos son aquellas predicciones que tienen una confianza mayor a 0.05 y cuya etiqueta es de Peatón, al igual de aquellas cuya confianza es mayor a 0.72 y su etiqueta es de Motociclista, pero que en realidad el objeto detectado no corresponde a un peatón o un motociclista. Igualmente, la categoría *No Detectados* son aquellos objetos que, estando presentes en las imágenes analizadas, no lograron ser detectados ya sea porque su confianza es menor a los umbrales establecidos o porque el modelo de detección no logró identificarlo. Así, se detectaron en total 2.006 peatones y 7.668 motociclistas en la totalidad de imágenes recolectadas.

En la Imagen 4.4, se muestra algunas predicciones de peatones y motociclistas realizadas en imágenes que no fueron usadas al momento del entrenamiento y validación del modelo de detección.

Imagen 4.4. Muestra de detecciones en conjunto de validación.

4.2 Clasificación de comportamientos viales

Como se mencionó en el capítulo anterior, para clasificar tanto los peatones como motociclistas de acuerdo con su posición dentro de la escena vial, se realizó la calibración de los umbrales δ_{EC} , δ_{MC} y δ_P de manera visual, en donde los valores “óptimos” fueron 0.91, 0.5 y 0.9, respectivamente. Con estos valores se lograron discriminar los peatones y motociclistas detectados en alguna de las categorías descritas en la sección 3.5.

Imagen 4.5. Distribución de motociclistas según su posición vial.

Distribución de motociclistas

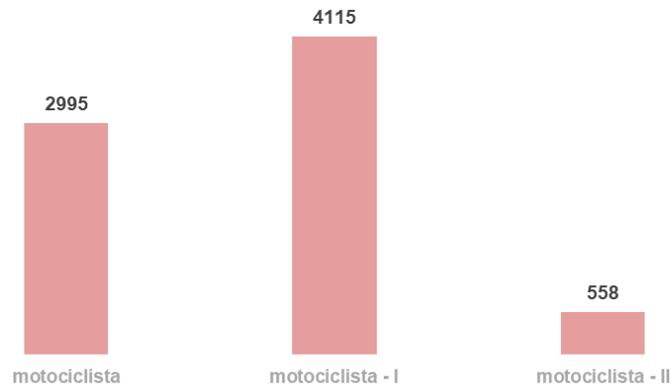
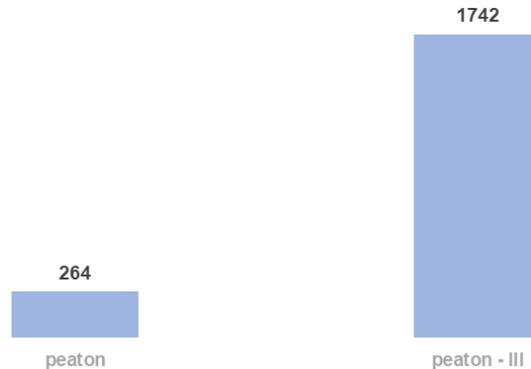


Imagen 4.6. Distribución de peatones según su posición vial.

Distribución de peatones



En las Imágenes 4.5 y 4.6 se logra observar que, tanto para los peatones como para los motociclistas, la mayor cantidad de detecciones se encuentran en categorías inmersas en alguna acción que aumenta la probabilidad de ocurrencia de un incidente vial, lo cual puede ser de interés para poder tomar acciones específicas por parte de la Alcaldía para disminuir la proporción de actores viales que incumplen las normas de tránsito especificadas con anterioridad.

Igualmente, se hizo uso de la información temporal extraída de manera automática para analizar el comportamiento de los actores viales según el día de la semana y la hora del día, con el objetivo de establecer si existe algún patrón que resulte relevante para la toma de decisiones a futuro. En las Imágenes siguientes se muestra la cantidad de actores viales detectados por hora y día, al igual de la proporción de cada categoría.

Imagen 4.7. Motociclistas detectados por hora.

Motociclistas

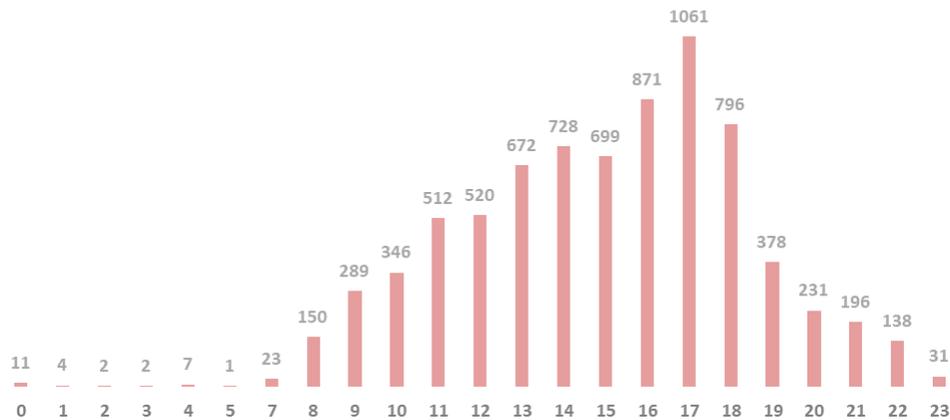
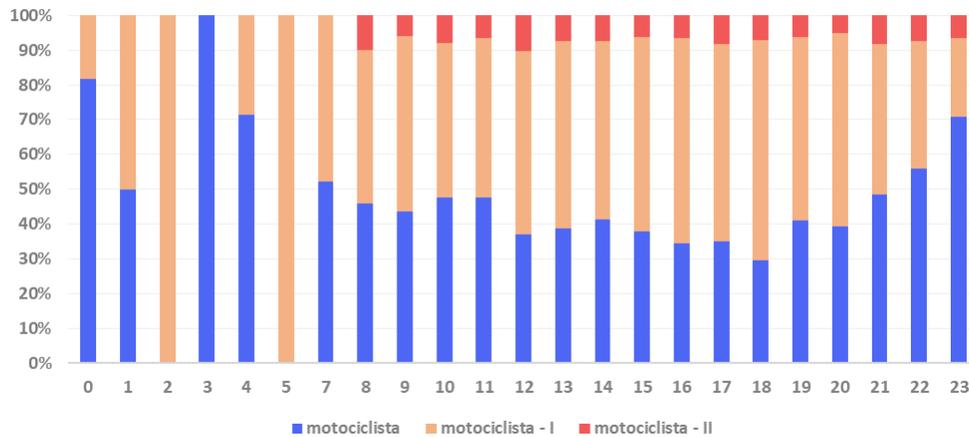


Imagen 4.8. Distribución de motociclistas por hora.

Distribución de motociclistas por hora

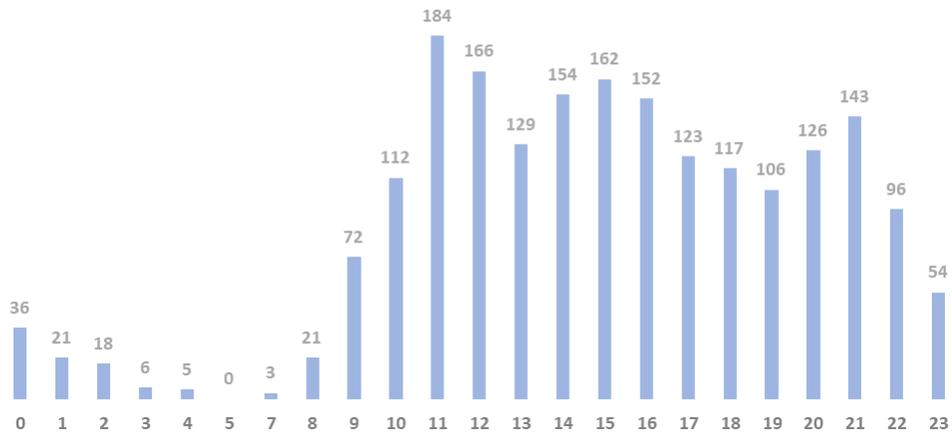


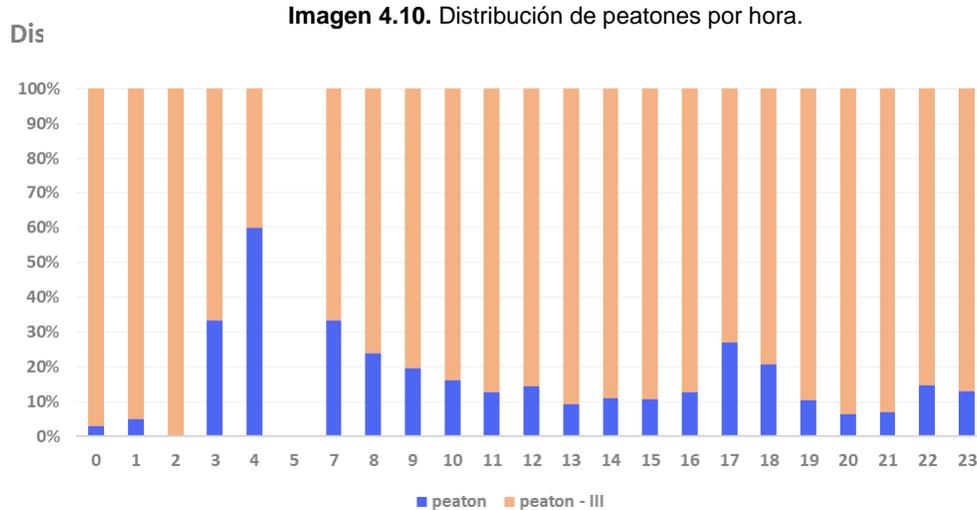
Con respecto a los resultados mostrados en las Imágenes 4.7 y 4.8 acerca de la cantidad de detecciones de motociclistas y su distribución a lo largo del día, se puede establecer:

- A. La cantidad de motociclistas que transitan por la calle monitoreada aumentan constantemente desde las 8 de la mañana teniendo un pico máximo a las 5 de la tarde, lo cual es congruente con los horarios laborales de la gran mayoría de personas en la ciudad. Adicionalmente, se observa que en horas de la madrugada el tránsito de motociclistas se reduce considerablemente y el cual es casi nulo.
- B. Al contrastar el porcentaje de motociclistas que no se encuentra inmerso en ninguno de los comportamientos riesgosos que fueron descritos anteriormente, se observa una tendencia decreciente a lo largo del día de manera similar al aumento de motociclistas que transitan por la calle.

Imagen 4.9. Peatones detectados por hora.

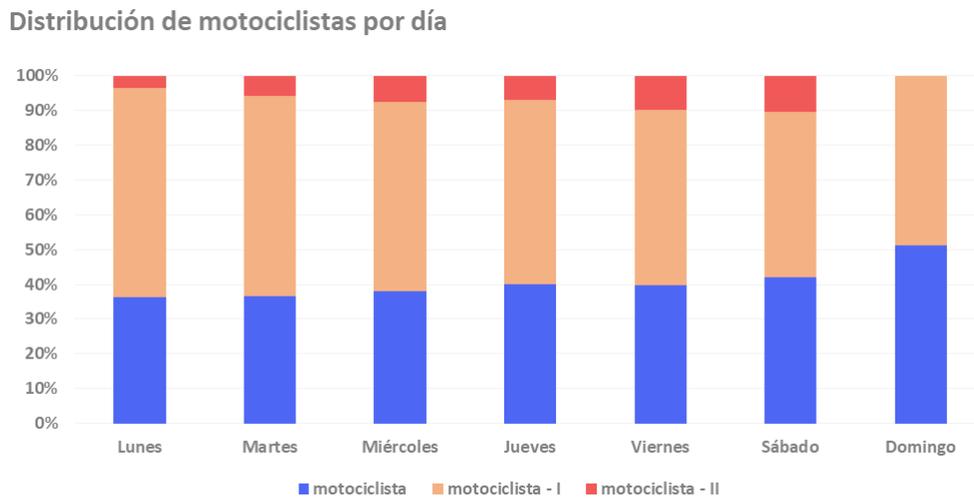
Peatones





Con respecto a los resultados mostrados en las Imágenes 4.9 y 4.10 acerca de la cantidad de detecciones de peatones y su distribución a lo largo del día, se puede indicar que:

- A. De manera similar a los motociclistas, la cantidad de peatones aumenta desde las 8 de la mañana, pero teniendo un pico máximo alrededor de las 11 de la mañana, y en horas de la tarde hasta las 9 de la noche permanece relativamente constante, para tener un descenso importante desde esa hora hasta la madrugada.
- B. Con respecto a la distribución de los comportamientos viales de los peatones, se evidencia que la gran mayoría de estos no se posicionan en el lugar que corresponde en la vía, es decir, transitan por los carriles que son específicamente para automóviles. Lo anterior se puede explicar al analizar la información recolectada, debido a que la gran mayoría de los peatones que fueron registrados en las imágenes son personas que trabajan de manera informal en los semáforos, ya sea vendiendo productos u ofreciendo servicios a los automóviles que pasan por dicha calle. Igualmente, se puede observar una relación inversamente proporcional entre la cantidad de peatones que se encuentran posicionados en el interior de la vía con la cantidad de motociclistas detectado, por lo que se puede establecer que a mayor cantidad de vehículos se tiende a respetar más las normas viales por parte de los peatones.

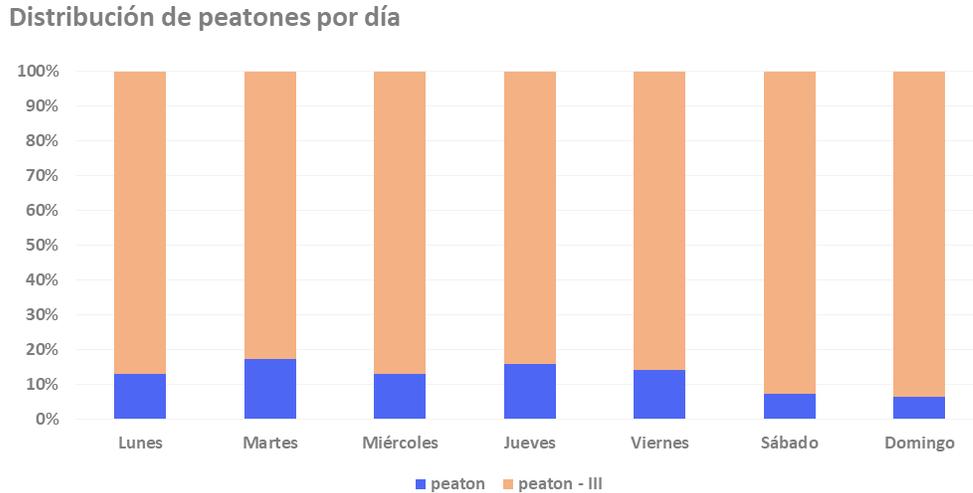
Imagen 4.11. Motociclistas detectados por día.**Imagen 4.12.** Distribución de motociclistas por día.

Con respecto a los resultados mostrados en las Imágenes 4.11 y 4.12 acerca de la cantidad de detecciones de motociclistas y su distribución de acuerdo con los días de la semana, es posible establecer que:

- A. Durante la semana, la cantidad de motociclistas aumentan constantemente de lunes a viernes, teniendo el pico máximo este último, pero decrece considerablemente los fines de semana.
- B. El patrón observado durante las horas del día se mantiene cuando se analiza por día, pero en menor proporción, corroborando que la cantidad de motociclistas que se encuentran inmersos en alguna acción indebida aumenta a medida que aumenta la cantidad de motociclistas transitando por la vía.

Imagen 4.13. Peatones detectados por día.



Imagen 4.14. Distribución de peatones por día.

Con respecto a los resultados mostrados en las Imágenes 4.13 y 4.14 acerca de la cantidad de detecciones de peatones y su distribución de acuerdo con los días de la semana, se concluye que:

- A. Si bien de lunes a viernes es cuando la gran mayoría de peatones son detectados, al contrario de lo sucedido con los motociclistas, el domingo se observa una gran cantidad de peatones en la vía analizada. Al analizar las imágenes, se encuentra que la gran mayoría de las personas detectadas estos días son trabajadores informales, los cuales generalmente se encuentran ubicados al interior de las vías por donde deben de transitar los automóviles.
- B. De acuerdo con la distribución de la categoría de los peatones, se observa que el domingo es el día en donde existe una menor cantidad de personas que cumplen con la norma de tránsito analizada en el presente trabajo, siendo congruente con lo analizado en el literal anterior.

5. Conclusiones y trabajos futuros

5.1 Conclusiones

A continuación, se explicará la manera en cómo se dio cumplimiento a cada uno de los objetivos específicos planteados en el presente trabajo:

- I. Identificar aquellos comportamientos de los actores viales que puedan ser usados para la caracterización de la seguridad vial de motociclistas y peatones.

Con base en la información recolectada al analizar la norma de tránsito vigente en el territorio nacional, se encontraron 3 acciones que pueden presentar la mayor incidencia en el aumento de la probabilidad de suceso de incidentes viales, los cuales están asociados a peatones y motociclistas. Estos se mencionarán a continuación:

- A. Aquellos peatones que transitan por los carriles en donde se movilizan los vehículos pueden ocasionar incidentes de tránsito al poner en riesgo su propia integridad al igual la de los conductores. Adicional a esto, en términos de movilidad generan una reducción del flujo vehicular debido a que los conductores disminuyen la velocidad para evitar una colisión con ellos.
- B. Aquellos motociclistas que conducen realizando prácticas inadecuadas, las cuales ponen en riesgo su integridad al aumentar la probabilidad de algún incidente. Cuando estos actores viales transitan en medio de dos carriles a través de la línea divisoria, representa un gran peligro para ellos debido a la posibilidad de colisión con algún automotor durante el trayecto, al igual de la inestabilidad que puede representar en el manejo debido al cambio de superficie de contacto que tienen las llantas al pasar por encima de la pintura, con la cual se demarca la separación de los carriles.

C. Finalmente, una práctica común llevada a cabo por motociclistas es compartir mismo carril junto a otro. Esto aumenta el riesgo de sufrir algún incidente debido al poco espacio de maniobrabilidad del que disponen los conductores, ya que al igual que en la acción descrita anteriormente se puede presentar con mayor facilidad una colisión entre dos vehículos.

II. Analizar las fuentes de datos abiertos disponibles que sean propicias para la obtención de información relevante en la investigación.

Se identificó como principal fuente de información para llevar a cabo el presente trabajo, las cámaras de seguridad a cargo de la Alcaldía de Medellín, en donde se registra en tiempo real el estado de las vías en zonas específicas de la ciudad. Sin embargo, no todas son de utilidad, debido a que algunas de estas presentan un bajo flujo de movilidad y otras tienen una baja frecuencia en la captura de la información. Así, la cámara de seguridad seleccionada fue la que se encuentra localizada en la Carrera 50FF con Calle 8 Sur, en donde se recolecta información cada 6 minutos aproximadamente de manera continua.

Con esto se logra demostrar que haciendo uso de la información disponible con la que cuenta la ciudad de Medellín, es posible generar análisis orientados a entender y caracterizar aquellos problemas cotidianos que de manera tradicional puede consumir gran cantidad de recursos, ya que exige la presencia de personal en las vías para que realicen mediciones de ciertas variables de interés.

III. Generar una metodología basada en inteligencia artificial para el seguimiento a los comportamientos de motociclistas y peatones.

Se demostró que haciendo uso técnicas de *web scrapping*, aprendizaje profundo y *transfer learning*, es posible generar una metodología para entender los comportamientos viales de peatones y motociclistas. Se logró así, generar una secuencia de tareas en la cual de manera automática se recolecta la información

generada en tiempo real. Posteriormente, es usada para la extracción automática de información temporal y espacial, lo cual permite analizar los comportamientos de los actores viales.

- IV.** Desarrollar una prueba de concepto de la metodología con la información disponible del municipio de Medellín.

Al validar la metodología desarrollada en el presente trabajo en el caso de aplicación de la ciudad de Medellín, se aumentó la comprensión de las conductas de los actores viales. Se identificó aquellos momentos en donde se pueden presentar incidentes viales con mayor proporción, lo cual puede servir como insumo para aquellas personas encargadas de generar planes de acción que busquen disminuir estas situaciones.

Al analizar la veracidad de la información generada, se puede concluir que es válida debido a las métricas de desempeño alcanzadas en la validación del modelo de detección, donde se obtuvo una sensibilidad y especificidad mayor al 90% tanto para peatones como para motociclistas.

5.2 Trabajos futuros

Si bien se logró detectar de manera automática tanto a los peatones y motociclistas al interior de las imágenes analizadas con una sensibilidad y especificidad mayor al 90% y clasificar a estos de acuerdo con sus acciones, queda abierta la puerta a trabajos futuros con el objetivo de poder replicar la metodología desarrollada en el presente trabajo en otras vías al interior de la ciudad. A continuación, se plantean algunas opciones de investigaciones para alcanzar dicho objetivo:

1. Para poder ser aplicado en diferentes partes de la ciudad, se requiere disponer de mayor variedad de situaciones viales diferentes a las recolectadas en para este trabajo.

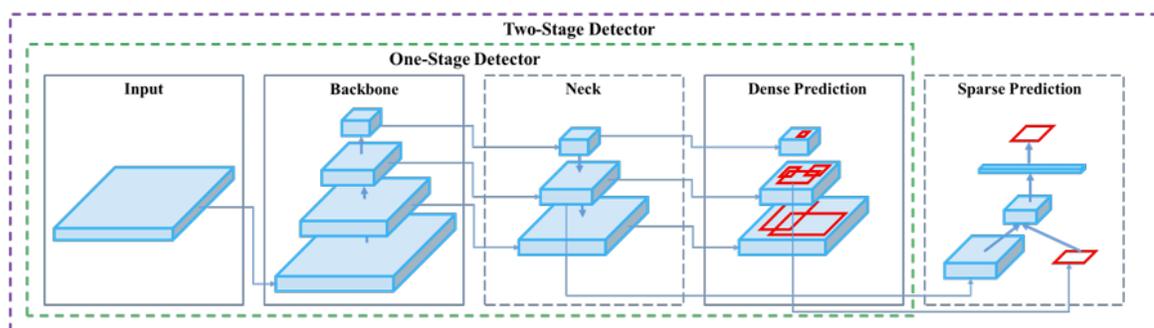
Por esto, se debe de localizar fuentes de datos adicionales, los cuales pueden hacer parte de la iniciativa Medata o no, que capturen en tiempo real la movilidad en diversos puntos de Medellín.

2. Dado que las condiciones del desplazamiento vial no son constantes y las interacciones sociales tampoco, se debe de poder generar una metodología que permita validar el desempeño del modelo de detección de manera constante y en caso de ser necesario, reentrenarlo para poder estar parametrizado de acuerdo con las condiciones actuales de la vía analizada.
3. Si bien la metodología presentada acá para la clasificación de los peatones y motociclistas según las normas de tránsito analizadas genera visualmente resultados adecuados, en el futuro podría ser de utilidad investigar un método de clasificación de acciones automático que sea robusto a la localización de la cámara, ya que puede darse el caso de que esta pueda ser desplazada de la posición original con la que capturó la información al momento de entrenar el modelo, generando inconvenientes al momento de realizar clasificaciones en un futuro.

Anexo 1. YOLOv5

Para poder explicar la arquitectura implementada en el presente trabajo para la detección de objetos al interior de las imágenes recolectadas, es necesario entender como está compuesta la arquitectura de cualquier modelo de detección. En la Imagen A1.1, se muestra dicha arquitectura.

Imagen A1.1. Arquitecturas genéricas para detección de objetos.

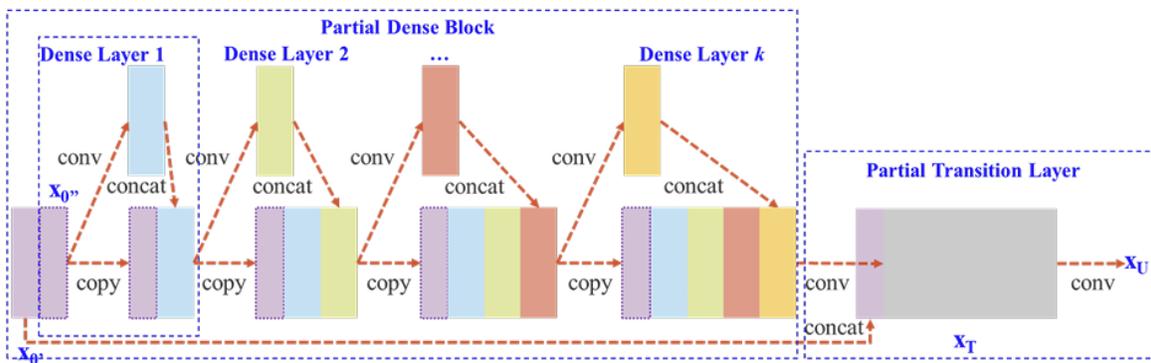


En la Imagen A1.1 se observa la existencia de dos tipos de arquitectura que pueden ser implementadas en la detección de objetos. La primera de estas es un detector de una sola etapa, en donde la detección y generación de las *Bounding Boxes* y la clasificación de estas en alguna de las categorías de interés es realizada de manera simultánea, mientras que en un detector de dos etapas se genera de manera independiente las posibles zonas de interés de los *Bounding Boxes*, las cuales son procesadas posteriormente de manera independiente para su clasificación y el refinamiento de la región de interés. Dado que el alcance del presente trabajo no es explicar en detalle las diferencias entre los posibles modelos existentes para la detección de objetos, se recomienda la lectura de la encuesta realizada por (Jiao, y otros, 2019) en donde se explica de manera detallada cada uno de estos.

En el framework YOLOv5, el cual tiene como base YOLOv1 y YOLOv3, se implementó un modelo de una etapa. A continuación, se explica el significado de cada etapa (*Backbone*, *Neck*, y *Dense Prediction*) al igual de la configuración específica para el modelo usado.

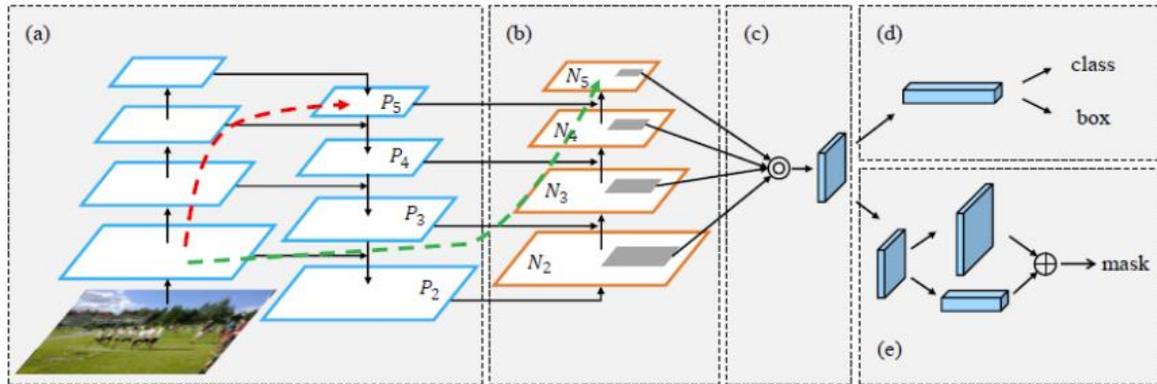
- I. **Backbone:** Hace referencia a la red neuronal que sirve como principal extractor de características de las imágenes ingresadas, la cual generalmente está compuesta por capas convolucionales. La selección de este componente es de suma importancia para la precisión que se obtendrá y generalmente, se hace uso de modelos pre-entrenados. En YOLOv5, se hace uso de la red *Cross Stage Partial Networks* (Wang, y otros, 2019), cuya representación gráfica se muestra en la Imagen A1.2.

Imagen A1.2. Cross Stage Partial Network. Tomado de (Wang, y otros, 2019).



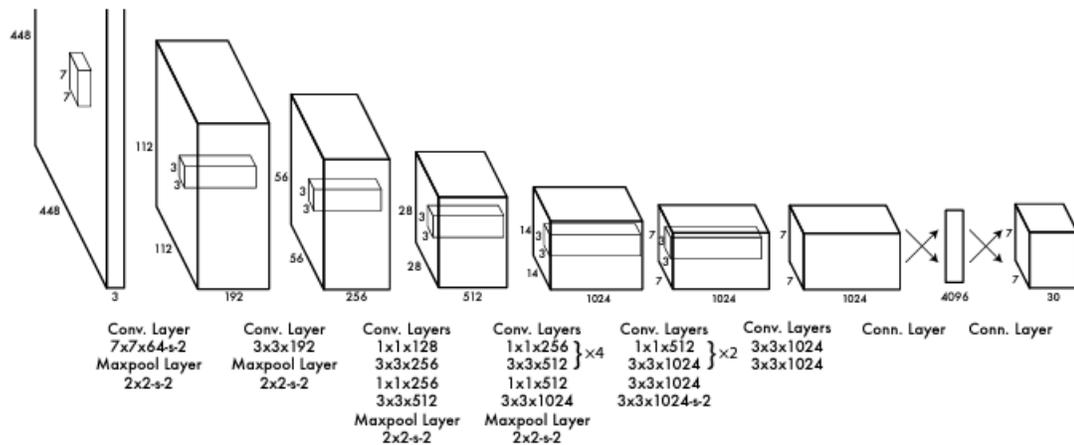
- II. **Neck:** Hace referencia al modelo empleado para la generación de pirámides de características en donde posiblemente se encuentren los objetos a ser detectados incluso a diferentes escalas. En YOLOv5, se hizo uso de *PANet* (Liu, Qi, Qin, Shi, & Jia, 2018), cuya representación gráfica se muestra en la Imagen A1.3.

Imagen A1.3. PANet. Tomado de (Liu, Qi, Qin, Shi, & Jia, 2018).



III. **Dense Prediction:** En esta última capa, se realiza la predicción densa del modelo, la cual está compuesta por un vector de coordenadas del *Bounding Box* del objeto detectado, la confianza de la predicción y la categoría asignada. La arquitectura usada para YOLOv5 consiste en la arquitectura completa de YOLOv1, la cual se muestra en la Ilustración A1.4

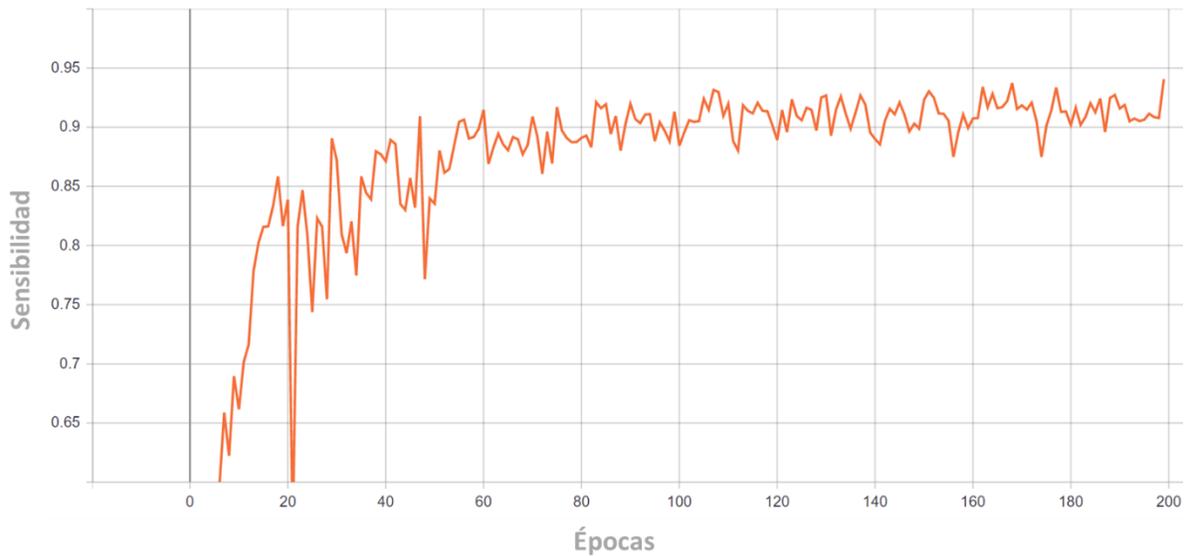
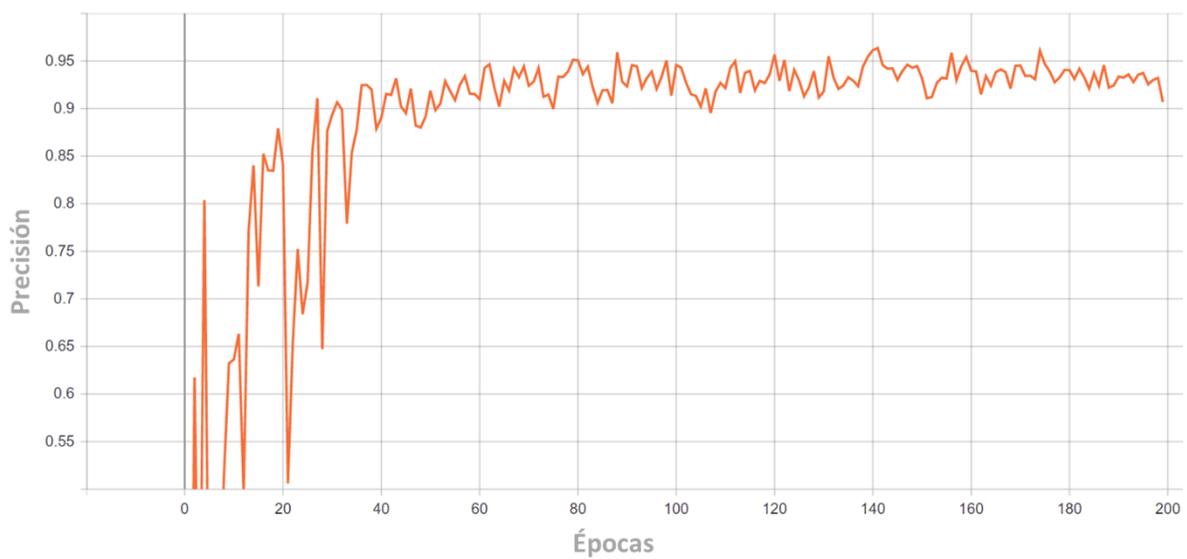
Imagen A1.4. YOLOv1. Tomado de (Redmon, Divvala, Farhadi, & Girshick, 2016).



Adicional a la estructura mostrada anteriormente, se implementó la siguiente configuración al momento de entrenar el modelo de predicción para la detección de motociclistas y peatones:

- I. Se inicializaron los pesos de las redes neuronales que hacen parte de la arquitectura con un entrenamiento previo realizado por parte de Ultralytics sobre la base de datos COCO, haciendo así uso de *transfer learning* para obtener mejores métricas de detección en comparación con iniciar el proceso de entrenamiento desde cero.
- II. YOLOv5 cuenta con algunas variantes desde la arquitectura *nano* hasta la arquitectura *XLarge*, cuya diferencia es el coste computacional para el entrenamiento del modelo de detección. Debido a limitaciones tecnológicas, dado que el entrenamiento del modelo de predicción del presente trabajo se realizó en la versión gratuita de Google Colab, únicamente se pudo obtener resultados con YOLOv5s, el cual cuenta con 7.5 millones de parámetros a ser entrenados.
- III. Como algoritmo para el entrenamiento de la arquitectura se usó *Stochastic Gradient Descent*
- IV. YOLOv5 únicamente recibe como entrada imágenes cuyo tamaño, tanto en ancho como en largo sea múltiplo de 32, debido al tamaño de los kernels utilizados en las convoluciones. Por esto, a todas las imágenes se les cambió el tamaño a 704 x 704 píxeles, siendo la medida que menor distorsionaba la información con respecto a la entrada original.
- V. El tamaño del batch de entrenamiento fue de 32 imágenes, los cuales fueron entrenados durante 200 épocas tomando 51 minutos.

En las Imágenes A1.5 y A1.6 se muestra la evolución de la sensibilidad y la precisión contra la cantidad de épocas, respectivamente.

Imagen A1.5. Variación de la sensibilidad con las épocas.**Imagen A1.6.** Variación de la precisión con las épocas.

Al analizar el comportamiento de la sensibilidad y la precisión en las Ilustraciones A1.5 y A1.6, se observa que dichas métricas tienden a estabilizar sus valores entre la época 80 y 100, indicando que el haber aumentado el número de las épocas más allá de 200 no hubiera sido de ayuda para la mejora en la capacidad de detección del modelo.

Bibliografía

- Agencia Nacional de Seguridad Vial. (2020). *Histórico de víctimas*. Obtenido de <https://ansv.gov.co/es/observatorio/estad%C3%ADsticas/historico-victimas>
- Alcaldía de Medellín. (2018). *Prácticas ciudadanas para una movilidad segura y sostenible*. Medellín.
- Alcaldía de Medellín. (2020). *Transporte matriculado en la secretaría de movilidad*. Medellín.
- Alcaldía de Medellín. (Abril de 2021). *Aprendizajes y experiencias sobre movilidad en la ciudad*. Obtenido de <https://medellin.gov.co/educacionvial/>
- Alcaldía de Medellín. (1 de Febrero de 2021). *Cámaras CCTV Movilidad*. Obtenido de <https://geomedellin-m-medellin.opendata.arcgis.com/datasets/camaras-cctv-movilidad-simm>
- Alcaldía de Medellín. (Febrero de 2021). *Incidentes viales en Medellín 2014 - 2021*. Obtenido de <https://www.medellin.gov.co/movilidad/m-en-linea/indicadores>
- Alcaldía de Medellín. (Marzo de 2021). *Mapa de proyectos*. Obtenido de <http://www.edu.gov.co/component/k2/item/59-mapa-de-proyectos>.
- Alcaldía de Medellín. (2022). *El portal de datos de Medellín*. Obtenido de [Medata: medata.gov.co/](https://medata.gov.co/)
- Alex, H. (2020). *Data augmentation and image understanding*. Innenstadt.
- Anaya, J., Ponz, A., García, F., & Talavera, E. (2017). Motorcycle detection for ADAS through camera and V2V Communication, a comparative analysis of two modern technologies. *Expert Systems With Applications*, 77, 148-159. doi:<http://dx.doi.org/10.1016/j.eswa.2017.01.032>
- Binh, T., & Kim, H. (2013). Novel and efficient pedestrian detection using bidirectional PCA. *Pattern Recognition*, 46(8), 2220-2227. doi:<https://doi.org/10.1016/j.patcog.2013.01.007>
- Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv:2004.10934*.

- Chang, W., & Hsu, K. (2010). Vision-Based Side Vehicle Detection from a Moving Vehicle. *2010 International Conference on System Science and Engineering*, 553-558. doi:10.1109/ICSSE.2010.5551779.
- Cheng, W., & Cheng, Y. (2010). Pedestrian Detection in Video Using Shape Features and Mixture of SMVs. *2010 3rd International Congress on Image and Signal Processing*, 442-446. doi:10.1109/CISP.2010.5646971
- Congreso de la República. (2002). *Ley 769 de 2002*. Bogotá, D.C.
- Dai, X., Hu, J., Zhang, H., Shitu, A., Luo, C., Osman, A., . . . Duan, Y. (2021). Multi-task faster R-CNN for nighttime pedestrian detection and distance estimation. *Infrared Physics and Technology*, 115, 103694. doi:https://doi.org/10.1016/j.infrared.2021.103694
- de Goma, J., Bautista, R., Eviota, M., & Lopena, V. (2020). Detecting Red-Light Runners (RLR) and Speeding Violation through Video Capture. *2020 IEEE 7th International Conference on Industrial Engineering and Applications(774-778)*. doi:10.1109/ICIEA49774.2020.9102059.
- Enyinna, C., Ijomah, W., Gachagan, A., & Marshall, S. (s.f.). Activation Functions: Comparison of Trends in Practice and Research for Deep Learning. *arXiv:1811.03378*.
- Espinosa, J., Valestín, S., & Branch, J. (2019). EspiNet V2: a region based deep learning model for detecting motorcycles in urban scenarios. *Dyna*, 86(211), 317-326. doi:http://doi.org/10.15446/dyna.v86n211.81639
- Girshick, R. (2015). Fast R-CNN. *2015 IEEE International Conference on Computer Vision*, 1440-1448. doi:10.1109/ICCV.2015.169
- Guo, L., Ge, P., Zhang, M., Li, L., & Zhao, Y. (2012). Pedestrian detection for intelligent transportation systems combining AdaBoost algorithm and support vector machine. *Expert Systems with Applications*, 39, 4274-4286. doi:10.1016/j.eswa.2011.09.106
- Hyder, A. (2019). Measurement is not enough for global road safety: implementation is key. *The Lancet Public Health*, 4(1), E12-E13.
- Jaganathan, D., prabhu, V., Chinnasamy, A., & Suresh, A. (2020). Detection of Two wheeler Driver Safety Using. *European Journal of Molecular & Clinical Medicine*, 7(6), 2292-2298.
- Jiao, L., Zhang, F., Liu, F., Yang, S., Li, L., Feng, Z., & Qu, R. (2019). A Survey of Deep Learning-based Object Detection. *arXiv:1907.09408*.
- Jin, Y., Zhang, Y., Cen, Y., Li, Y., Mladenovic, V., & Voronin, V. (2021). Pedestrian detection with super-resolution reconstruction for low-quality image. *Pattern Recognition*, 115, 107846. doi:https://doi.org/10.1016/j.patcog.2021.107846

- Kuo, Y., Yang, Z., & Yen, C. (2012). Fast pedestrian detection system with a two layer cascade of classifiers. *Computers and Mathematics with Applications*, 64, 1311-1323. doi:10.1016/j.camwa.2012.03.075
- Lin, H., Deng, J., Albers, D., & Wilhelm, F. (2020). Helmet Use Detection of Tracked Motorcycles Using CNN-Based Multi-Task Learning. *IEEE Access*, 8, 162073-162084. doi:10.1109/ACCESS.2020.3021357
- Lin, T., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., . . . Dollár, P. (2014). Microsoft COCO: Common Objects in Context. *2014 European Conference on Computer Vision*, 740-755. doi:https://doi.org/10.1007/978-3-319-10602-1_48
- Liu, S., Qi, L., Qin, H., Shi, J., & Jia, J. (2018). Path Aggregation Network for Instance Segmentation. *arXiv:1803.01534*.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C., & Berg, A. (2016). SSD: Single Shot MultiBox Detector. *2016 European Conference on Computer Vision*, 21-37. doi:https://doi.org/10.1007/978-3-319-46448-0_2
- Medellín cómo vamos. (2020). *Informe de calidad de vida de Medellín 2016-2019*. Medellín.
- Organización de las Naciones Unidas. (2018). *La Agenda 2030 y los Objetivos de Desarrollo Sostenible: una oportunidad para América Latina y el Caribe*. Santiago de Chile.
- Organización Mundial de la Salud. (Julio de 2017). *10 datos sobre la seguridad vial en el mundo*. Obtenido de <https://www.who.int/features/factfiles/roadsafety/es/>
- Organización Panamericana de la Salud. (2019). *Estado de la seguridad Vial en la Región de las Américas*. Washington, D.C.
- Pramanik, A., Sarkar, S., & Maiti, J. (2021). A real-time video surveillance system for traffic pre-events detection. *Accident Analysis and Prevention*, 154, 106019. doi:https://doi.org/10.1016/j.aap.2021.106019
- Redmon, J., Divvala, S., Farhadi, A., & Girshick, R. (2016). You Only Look Once: Unified, Real-Time Object Detection. *arXiv:1506.02640*.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *IEEE Conference on Computer Vision and Pattern Recognition*, 779-788. doi:10.1109/CVPR.2016.91
- Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(6), 1137-1149. doi:10.1109/TPAMI.2016.2577031

- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., . . . Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, 115, 211-252. doi:<https://doi.org/10.1007/s11263-015-0816-y>
- Silva, R., Aires, K., Veras, R., Santos, T., Kalyf, L., & Soares, A. (2013). Automatic Motorcycle Detection on Public Roads. *Clei Electronic Journal*, 16(3).
- tzutalin, D. (2015). *labellmg*. Obtenido de Github: <https://github.com/tzutalin/labellmg>
- Ultralytics. (18 de May de 2020). YOLOv5. Obtenido de GitHub: <https://github.com/ultralytics/yolov5>
- Velastín, S., Fernández, R., Espinosa, J., & Bay, A. (2020). Detecting, Tracking and Counting People Getting On/Off a Metropolitan Train Using a Standard Video Camera. *Sensors*, 20(21), 6251. doi:10.3390/s20216251
- Wang, C. Y., Liao, H. Y., Yeh, I. H., Wu, Y. H., Chen, P. Y., & Hsieh, J. W. (2019). CSPNET: A new backbone that can enhance learning capability of CNN. *arXiv:1911.11929*.
- Wilhelm, F., & Lin, H. (2020). Detecting motorcycle helmet use with deep learning. *Accident Analysis & Prevention*, 134, 105319. doi:<https://doi.org/10.1016/j.aap.2019.105319>
- Xu, F., & Gao, M. (2010). Human detection and tracking based on HOG and particle filter. *2010 3rd International Congress on Image and Signal Processing*, 1503-1507. doi:10.1109/CISP.2010.5646273
- Zhang, S. (2020). *Prediction of Pedestrians' Red Light Violations Using Deep Learning*. Orlando.
- Zhao, X., He, Z., Zhang, S., & Liang, D. (2015). Robust pedestrian detection in thermal infrared imagery using a shape distribution histogram feature and modified sparse representation classification. *Pattern Recognition*, 48, 1947-1960. doi:[dx.doi.org/10.1016/j.patcog.2014.12.013](https://doi.org/10.1016/j.patcog.2014.12.013)