



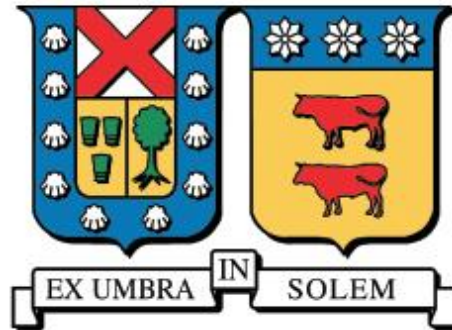
UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA  
DEPARTAMENTO DE OBRAS CIVILES

# **EVALUACIÓN DE CONDICIÓN DE PAVIMENTOS URBANOS UTILIZANDO MACHINE LEARNING**

**PAULINA BEATRIZ GÓMEZ CONTI**

**INGENIERO CIVIL**

**2023**



UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA  
DEPARTAMENTO DE OBRAS CIVILES

**EVALUACIÓN DE CONDICIÓN DE PAVIMENTOS URBANOS  
UTILIZANDO MACHINE LEARNING**

Memoria de titulación presentada por

**PAULINA BEATRIZ GÓMEZ CONTI**

Como requisito parcial para optar título de

**INGENIERO CIVIL  
MAGÍSTER EN CIENCIAS DE LA INGENIERÍA CIVIL**

Profesor Guía

**ALELÍ OSORIO LIRD**

AGOSTO DE 2023

## **Dedicatoria**

iii

Para mis papás y Mackicito.

Quiero agradecer a mis papás por todo el apoyo brindado durante toda mi vida y haberme inculcado el gusto por el conocimiento.

Quiero agradecer a mi profesora guía Alelí Osorio por haber confiado en mi y haberme propuesto y apoyado en la investigación de este tema tan interesante. También al profesor Héctor Allende, quien me ayudó a llevar a puerto esta investigación.

Quiero agradecer a aquellos profesores y docentes que aportaron en mi formación, en especial a Juan Pablo Chacana y Pablo Garnica. Y una mención especial a aquellos profesores que dejaron una huella en mi, pero lamentablemente no se encuentran presentes para leer esto, Justo Margalet y Belisario Gutierrez.

Quiero agradecer a Diego, por haberme apoyado y confiado en mi y mis capacidades. Gracias por haber sido una motivación cuando más lo necesitaba.

Quiero agradecer a mis compañeros de carrera que se convirtieron en mis amigos de la vida, que lograron hacer la universidad un espacio mucho más lindo, en especial a Karla, Marco, Daniel.

Quiero agradecer al departamento de física, quienes de mi segundo año de Universidad me permitieron desempeñarme como ayudante, lo que se volvió parte de mi vida. Doy las gracias porque gracias a esta oportunidad conocí muchos profesores que se convirtieron en amigos, Juan Pablo Ramos, José Miguel Pinto y Guillermo Gómez. Y no puedo olvidar a la Cristi, quien estuvo desde el primer día con una sonrisa apoyándome.

In pavement management, it is essential to have a good database with information on the condition of the roads that compose the corresponding network. With this information, it is possible to implement pavement condition models over time and thus, to have preventive maintenance instead of reactive maintenance, which implies a lower cost and thus, an optimization in the use of resources.

At present, standard manual surveys for asphalt pavements evaluate only 24% of the pavement, since a field operator must evaluate a 6 m long strip every 25 m of the total pavement. On the other hand, there is an automatic evaluation that has been applied in concessioned highways, which is carried out with the "Laser Crack Measurement System" instrument, that obtains images in three dimensions and then evaluates them. This technology, although it is very efficient, is also very expensive.

In Chile there is no official database on the condition of urban pavements. The most similar database is the SGPU, but it has not been implemented due to the lack of data. There is also no technology capable of evaluating the condition of urban pavements in an efficient way. On one hand, manual monitoring is a slow process, it depends on the evaluator and it is difficult to evaluate the entire network. On the other hand, automatic evaluation is unaffordable for local agencies. This research proposes a more efficient and low-cost technology capable of assessing the condition of urban asphalt pavements at the network level.

On this research, more than 50,000 images of 13.5 x 2.6 meters of asphalt pavement from different zones of Santiago, Chile were obtained. These images were processed, and the following distresses were manually labeled with two different levels of severities: patches, potholes and transversal, longitudinal and fatigue cracking.

The data was used to train and evaluate convolutional neural networks, specifically object detection models: YOLOv5 and YOLOv7. Five different trainings were carried out for each model with the images distributed in the training, validation and test sets differently for each training. The final results of each model represent the average of these five training sessions. Finally, when analyzing a new pavement section and with the inference of the network, a summary file is provided with the number of distresses, their severities, and dimensions. This information can be used in a pavement condition index.

En la gestión de pavimentos, es esencial contar con una base de datos de la condición de las calles que componen la red vial. Con esta información, se pueden implementar modelos de condición de pavimentos a través del tiempo y así, contar con mantenimiento preventivo en vez de reactivo, el cual implica un costo menor y así, una optimización en el uso de los recursos.

En la actualidad, las auscultaciones manuales normadas en el caso de los pavimentos asfálticos evalúan solo un 24% del pavimento, dado que un operador en terreno debe evaluar una franja de 6 m de largo cada 25 m del pavimento total. Por otro lado, existe una evaluación automática que se ha aplicado en autopistas concesionadas, que se lleva a cabo con el instrumento “Laser Crack Measurement System”, el cual obtiene imágenes en tres dimensiones que luego evalúa. Esta tecnología, si bien es muy eficiente, presenta costos muy elevados.

En la actualidad, en Chile no existe una base de datos oficial de la condición de los pavimentos urbanos, lo más cercano es el SGPU, pero que no ha sido implementado debido a la falta de datos y tampoco se ha implementado una tecnología capaz de evaluar la condición de pavimentos urbanos de forma eficiente. Por un lado, la auscultación manual es un proceso lento, depende del evaluador y difícilmente se puede evaluar la red en su totalidad. Y por otro lado, la evaluación automática es inasequible para agencias locales. Esta investigación propone una tecnología más eficiente y de bajo costo capaz de evaluar la condición de pavimentos urbanos asfálticos a nivel de red.

En esta memoria, más de 50.000 imágenes de 13,5 x 2,6 metros de pavimentos asfálticos de distintas zonas de Santiago, Chile fueron obtenidas. Estas imágenes fueron procesadas, y los siguientes deterioros fueron etiquetados de forma manual con dos niveles distintos de severidad: parches, baches y grietas longitudinal, transversal y de fatiga.

Estos datos fueron usados para entrenar y evaluar una red neuronal convolucional, particularmente usando modelos de detección de objetos: YOLOv5 y YOLOv7. Se llevaron a cabo cinco diferentes entrenamientos para cada modelo con las imágenes distribuidas de forma diferente en sets de entrenamiento, validación y prueba para cada entrenamiento. Los resultados finales de cada modelo representan las métricas promediadas obtenidas por estos cinco entrenamientos. Finalmente, al analizar una nueva sección de pavimentos con la inferencia de la red ya entrenada, se entrega un resumen con la cantidad de deterioros, sus severidades y dimensiones, información que se puede aplicar para un índice de condición de pavimentos urbanos.

1	Introducción.....	1
2	Estado Actual de la Auscultación de Pavimentos .....	2
2.1	Monitoreo y Evaluación de Pavimentos en Forma Manual .....	2
2.2	Importancia de la Auscultación de Pavimentos.....	3
2.3	Deterioros en Pavimentos de Asfalto en Zonas Urbanas .....	4
2.4	Tecnología Actual.....	6
2.5	Definición del Problema y Oportunidades de Investigación .....	16
3	Propuesta de Investigación .....	1
3.1	Hipótesis Científica .....	1
3.2	Objetivos.....	1
3.3	Metodología de Investigación .....	2
3.4	Estructura del Trabajo de Memoria de Título .....	3
4	Conformación de la Base de Datos y Diseño Experimental.....	4
4.1	Estudio Preliminar Cantidad de Deterioros .....	4
4.2	Metodología de Obtención de Datos .....	4
4.3	Procesamiento de los Videos.....	13
4.4	Etiquetado de Imágenes.....	21
4.5	Procesamiento Etiquetas para Entrenar en Formato YOLO.....	22
4.6	Split Datos para Entrenamiento, Validación y Prueba .....	23
4.7	Aumento de Datos .....	25
4.8	Herramientas Por Utilizar para la Red Neuronal.....	28
4.9	Implementación de Modelos YOLOv5 y YOLOv7 .....	29
4.10	Planificación Entrenamiento Red Neuronal .....	31
5	Resultados.....	1
5.1	Datos Obtenidos (lugar, cantidad de videos, cantidad de imágenes) .....	1
5.2	Etiquetas Totales (cantidad de deterioros totales) .....	2
5.3	Cantidad Final Imágenes y Etiquetas para Entrenar (post aumento de datos) .....	3
5.4	Resultado Red Neuronal (YOLO) .....	4
5.5	Inferencia Red Neuronal.....	16
6	Análisis de Resultados.....	18
6.1	Análisis Comparativo entre YOLOv5 y YOLOv7 .....	18
6.2	Análisis de Desempeño por Deterioros .....	22
6.3	Análisis de Resultados por Cada Split.....	23
6.4	Análisis de Resultados Comparación Grietas Longitudinales + Transversales vs Grietas Lineales .....	24
6.5	Análisis Cuantitativo .....	27
7	Conclusiones y Recomendaciones .....	30
7.1	Conclusiones.....	30
7.2	Recomendaciones .....	31
7.3	Investigación Futura .....	32
8	Lista de referencias .....	33
9	Anexos .....	41

9.1	Anexo 1. Tabla con Planificación de Recorrido para Auscultaciones .....	41viii
9.2	<i>Anexo 2. Deterioros y Singularidades del Pavimento Urbano Asfáltico Detectadas</i> 43	
9.3	Anexo 3. Técnicas de aumento de datos utilizadas por tipo de deterioro .....	53
9.4	Anexo 4. Resultados Entrenamiento .....	58

Tabla 2.1: Deterioros de pavimentos asfálticos a analizar ..... 5

Tabla 2.2: Representación del significado de los conceptos TP, TN, FP y FN. 0: nada, 1: detección..... 12

Tabla 2.3. Números de imágenes usadas para entrenar las redes neuronales con sus respectivas frecuencias. .... 14

Tabla 4.1: Cantidad de deterioros con sus respectivas severidades. B: Severidad baja, M: Severidad media y A: Severidad alta..... 4

Tabla 4.2. Uso de memoria de un video para cada calidad de grabación..... 8

Tabla 4.3. Duración máxima de un video según la calidad de grabación ..... 8

Tabla 4.4. Capacidad de grabación de una memoria de 64 GB según la calidad de grabación ..... 8

Tabla 4.5. Duración de las baterías según la calidad de grabación ..... 8

Tabla 4.6. Altura de montaje de la cámara en función del número de pasajeros ..... 12

Tabla 4.7: Tiempo que tardan en correr distintos códigos de la transformación de imágenes ..... 13

Tabla 4.8 Hiperparámetros asignados al momento de entrenamiento usando YOLOv5 ..... 29

Tabla 4.9 Hiperparámetros utilizados en el entrenamiento de YOLOv5 ..... 29

Tabla 4.10 Hiperparámetros asignados al momento de entrenamiento usando YOLOv5 ... 30

Tabla 4.11 Parámetros utilizados en el entrenamiento de YOLOv7 ..... 30

Tabla 4.12: Problemas computacionales y sus soluciones al entrenar YOLO ..... 30

Tabla 5.1: Distancias grabadas por comuna para entrenar la red ..... 1

Tabla 5.2: Cantidad de deterioros etiquetados en las imágenes obtenidas de las grabaciones en terreno ..... 2

Tabla 5.3: Cantidad de imágenes por set para cada split..... 4

Tabla 5.4: Cantidad de deterioros que no requirieron aumento de datos por cada split (grietas de fatiga y transversales). M: Severidad media, A: Severidad alta. .... 4

Tabla 5.5: Cantidad de deterioros que requirieron aumento de datos por cada split para grietas longitudinales severidad media ..... 4

Tabla 5.6: Cantidad de deterioros que requirieron aumento de datos por cada split para grietas longitudinales severidad alta ..... 5

Tabla 5.7: Cantidad de deterioros que requirieron aumento de datos por cada split para parches severidad media..... 5

Tabla 5.8: Cantidad de deterioros que requirieron aumento de datos por cada split para parches severidad alta..... 5

Tabla 5.9: Cantidad de deterioros que requirieron aumento de datos por cada split para baches ..... 6

Tabla 5.10: Cantidad de singularidades presentes en cada split..... 6

Tabla 5.11: Métricas de desempeño al evaluar el conjunto test para cada split entrenado con YOLOv5 ..... 7

Tabla 5.12: Métricas de desempeño al evaluar el conjunto test para el split 1 entrenado con YOLOv5 y grietas lineales unidas..... 10

Tabla 5.13: Métricas de desempeño al evaluar el conjunto test para cada split entrenado con YOLOv7 .....	11
Tabla 5.14: Métricas de desempeño al evaluar el conjunto test para el split 1 entrenado con YOLOv7 y grietas lineales unidas.....	14
Tabla 6.1: Resultados de prueba de Jarque-Bera para métricas obtenidas de YOLOv5 .....	19
Tabla 6.2: Resultados de prueba de Jarque-Bera para métricas obtenidas de YOLOv7 .....	19
Tabla 6.3: Resultados de la prueba F de Fisher para las métricas de precisión, recall y F1	20
Tabla 6.4: Resultados de la prueba F de Fisher para las métricas de mAP50 y mAP50-95	20
Tabla 6.5: Prueba t-Student para muestras asumiendo varianzas iguales para las métricas de precisión y recall.....	21
Tabla 6.6: Prueba t-Student para muestras asumiendo varianzas distintas para las métricas de F1, mAP.50 y mAP.50-95 .....	21
Tabla 6.7: Error cuadrático medio de las distintas métricas al comparar resultados de YOLOv5 y YOLOv7.....	22
Tabla 6.8: Resumen de resultados en base a F1 .....	22
Tabla 6.9: Resultados de medidas de dispersión para métricas obtenidas con YOLOv5 ....	23
Tabla 6.10: Resultados de medidas de dispersión para métricas obtenidas con YOLOv7 ..	24
Tabla 6.11: Comparación de grietas de severidad media detectadas usando el ponderado del entrenamiento con etiquetas “grietas longitudinales” y etiquetas “grietas transversales” vs el entrenamiento con “grietas lineales” con YOLOv5.....	25
Tabla 6.12: Comparación de grietas de severidad alta detectadas usando el ponderado del entrenamiento con etiquetas “grietas longitudinales” y etiquetas “grietas transversales” vs el entrenamiento con “grietas lineales” con YOLOv5.....	25
Tabla 6.13: Comparación de grietas de severidad media detectadas usando el ponderado del entrenamiento con etiquetas “grietas longitudinales” y etiquetas “grietas transversales” vs el entrenamiento con “grietas lineales” con YOLOv7.....	26
Tabla 6.14: Comparación de grietas de severidad alta detectadas usando el ponderado del entrenamiento con etiquetas “grietas longitudinales” y etiquetas “grietas transversales” vs el entrenamiento con “grietas lineales” con YOLOv7.....	26
Tabla 6.15: Valores de significancia al calcular el estado del pavimento asfáltico según el ICPU, a mayor valor, mayor afecta en el cálculo del estado. (N°) indica el orden ascendente del peso del deterioro en el índice, a mayor número más afecta la nota del deterioro.....	28
Tabla 9.1: Ejemplos de grietas de fatiga según su severidad .....	43
Tabla 9.2: Ejemplos de grietas transversales según su severidad.....	44
Tabla 9.3: Ejemplos de grietas longitudinales según su severidad .....	45
Tabla 9.4: Ejemplos de baches .....	46
Tabla 9.5: Ejemplos de parches según su severidad.....	47
Tabla 9.6: Ejemplos de tapas de alcantarillado .....	47
Tabla 9.7: Ejemplos de sumideros.....	48
Tabla 9.8: Ejemplos de zonas con extracción de testigos.....	48
Tabla 9.9: Ejemplos de pérdida de material, catalogado como deterioros misceláneos .....	49
Tabla 9.10: Ejemplos de imágenes descartadas por no presentar continuidad.....	50
Tabla 9.11: Ejemplos de imágenes descartadas por ser de pavimentos de hormigón.....	51

Tabla 9.12: Ejemplos de imágenes descartadas por tener vehículos que cubren el pavimento	xi
.....	52
Tabla 9.13: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 1 obtenido con YOLOv5	63
Tabla 9.14: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 1 obtenido con YOLOv7	63
Tabla 9.15: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 2 obtenido con YOLOv5	64
Tabla 9.16: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 2 obtenido con YOLOv7	64
Tabla 9.17: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 3 obtenido con YOLOv5	65
Tabla 9.18: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 3 obtenido con YOLOv7	65
Tabla 9.19: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 4 obtenido con YOLOv5	66
Tabla 9.20: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 4 obtenido con YOLOv7	66
Tabla 9.21: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 5 obtenido con YOLOv5	67
Tabla 9.22: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 5 obtenido con YOLOv7	67

**Figura 2.1:** Caso general para auscultar pavimentos asfálticos según MANVU. Fuente propia ..... 2

**Figura 2.2:** Progresión del estado de un pavimento en el tiempo ..... 3

**Figura 2.3:** Ciclo de vida típico de un pavimento ..... 4

**Figura 2.4:** Laser Crack Measurement System ..... 7

**Figura 2.5:** Foto referencial de la cámara a utilizar en el proyecto ..... 7

**Figura 2.6.** Explicación básica de inteligencia artificial, Machine Learning y Deep Learning ..... 9

Figura 2.7. Técnica de detección de objetos..... 10

Figura 2.8. Técnica de segmentación de objeto..... 10

Figura 2.9. Gráfico de barras de la tabla 2.1 a (investigaciones con imágenes en 2D y 3D)15

Figura 2.10: Gráfico de barras de la tabla 2.1 b (investigaciones que solo consideran imágenes en 2D) ..... 15

Figura 3.1: Diagrama que representa la metodología de investigación..... 2

Figura 4.1. Pruebas de detección del color obtenido en Python para distintos pliegos de goma eva. Las figuras de la izquierda son con imágenes obtenidas en resolución 1080 pixeles, y las figuras de la derecha fueron obtenidas con 4K. .... 5

Figura 4.2. Fotografía de la figura circular de goma eva azul..... 6

Figura 4.3. Fotos de la misma zona de pavimento agrietado. La imagen de la izquierda fue obtenida con grabación en 1920 pixeles, y la de la derecha en 4K. .... 7

Figura 4.4. Binarización de imágenes de pavimentos deteriorados. La imagen de la izquierda fue grabada con 1920 pixeles, y la de la derecha con 4K..... 7

Figura 4.5. Representación de zona a transformar para llevar a cabo la auscultación. .... 9

Figura 4.6. Dimensiones definidas del montaje para automatizar la transformación de videos a imágenes ..... 10

Figura 4.7. La imagen de la izquierda muestra la placa de PLA con un láser apuntando a la misma dirección que el centro del lente de la cámara. La imagen de la derecha muestra cómo se ve el tercer láser en el pavimento. .... 11

Figura 4.8. Modelo en Autodesk® Fusion 360 de la placa para el montaje ..... 11

Figura 4.9: Esquema del uso de la placa PLA con el láser para el montaje del sistema ..... 12

**Figura 4.10.** Esquema de coordenadas definidas para transformar las imágenes a una vista en planta ..... 14

Figura 4.11. Imagen que muestra el funcionamiento de la detección automática de pixeles azules ..... 15

Figura 4.12. Ejemplo de imágenes normalizadas, obtenidas por T. Tran et al (2020) [14]. Las imágenes en (a) son las imágenes originales, y las imágenes en (b) son las normalizadas ..... 16

Figura 4.13: Imágenes seleccionadas para llevar a cabo el análisis de histograma. Se seleccionaron 4 imágenes por cada video: una imagen con sombra, otra con sol, otra con sol y sombra y la última, con alta presencia de deterioros. .... 17

Figura 4.14. Histogramas de videos obtenidos con 15 horas de diferencia, para imágenes con sombra, sol, sombra y sol y pavimento agrietado..... 18

Figura 4.15. Desplazamiento del histograma para normalizar las imágenes. Los valores azules representan el histograma original, y los naranjos el normalizado.....	19
Figura 4.16. Imágenes de pavimentos a modo de ejemplo para demostrar la normalización. Las imágenes (a) corresponde a las originales, y las imágenes (b) a las normalizadas. ....	20
Figura 4.17: Diferencia de coordenadas VGG con YOLO en una etiqueta. El punto azul representa las coordenadas de VGG y el punto verde representa las coordenadas de YOLO, las cuales además deben estar normalizadas. ....	22
Figura 4.18: Diagrama para obtener las etiquetas de entrenamiento en formato YOLO a partir del etiquetado en programa VIA .....	23
Figura 4.19: Diagrama de flujo para llevar a cabo el split de imágenes en sets de entrenamiento, validación y prueba.....	24
Figura 4.20: Ejemplo de un crop llevado a cabo en un bache. El cuadrado rojo representa la etiqueta original y el resto de la imagen consiste en los 200 pixeles extra por lado. ...	26
Figura 4.21: Todas las opciones posibles de técnicas a nivel espacial de aumento de datos. ....	27
Figura 4.22: Representa la importancia del orden en las transformaciones. A la figura a) se le realiza una rotación y luego un flip; a la figura b) se le realiza un flip y luego una rotación. Se puede ver que los resultados no son iguales.....	28
Figura 5.1: Diagrama circular de cantidad de deterioros etiquetados en las imágenes obtenidas de las grabaciones en terreno .....	2
Figura 5.2: Promedio de las matrices de confusión obtenidas de los 5 splits para YOLOv5.	8
Figura 5.3: Resultados por deterioro al evaluar el conjunto test y promediar las métricas de los 5 splits en YOLOv5 .....	9
Figura 5.4: Matriz de confusión del split 1 entrenado con YOLOv5 usando grietas lineales unidas.....	10
Figura 5.5: Promedio de las matrices de confusión obtenidas de los 5 splits para YOLOv7. ....	12
Figura 5.6: Resultados por deterioro al evaluar el conjunto test y promediar las métricas de los 5 splits en YOLOv7 .....	13
Figura 5.7: Matriz de confusión del split 1 entrenado con YOLOv7 usando grietas lineales unidas.....	15
Figura 5.8: Ejemplo de output final de resultados.....	16
Figura 5.9: Ejemplo de tabla de output de la inferencia.....	17
Figura 6.1: Ejemplos de etiquetas solapadas. a) representa la etiqueta manual asignada y b) la predicción de la red.....	27
Figura 6.2: Imágenes que demuestran la confusión en la clasificación de deterioros. a) representa la etiqueta manual asignada y b) la predicción de la red.....	29
Figura 6.3: Imágenes que demuestran la confusión en la clasificación de deterioros. a) representa la etiqueta manual asignada y b) la predicción de la red.....	29
Figura 9.1: Matriz de confusión del conjunto test de Split 1 entrenado con YOLOv5 .....	58
Figura 9.2: Matriz de confusión del conjunto test de Split 2 entrenado con YOLOv5 .....	58
Figura 9.3: Matriz de confusión del conjunto test de Split 3 entrenado con YOLOv5 .....	59
Figura 9.4: Matriz de confusión del conjunto test de Split 4 entrenado con YOLOv5 .....	59

Figura 9.5: Matriz de confusión del conjunto test de Split 5 entrenado con YOLOv5 ...	60xiv
Figura 9.6: Matriz de confusión del conjunto test de Split 1 entrenado con YOLOv7 .....	60
Figura 9.7: Matriz de confusión del conjunto test de Split 2 entrenado con YOLOv7 .....	61
Figura 9.8: Matriz de confusión del conjunto test de Split 3 entrenado con YOLOv7 .....	61
Figura 9.9: Matriz de confusión del conjunto test de Split 4 entrenado con YOLOv7 .....	62
Figura 9.10: Matriz de confusión del conjunto test de Split 5 entrenado con YOLOv7 .....	62

## ***1 Introducción***

El principal propósito de la gestión de pavimentos es el desarrollo de programas de conservación, mantenimiento y rehabilitación a lo largo de todo el ciclo de vida de ellos para asegurar un buen nivel de servicio y optimizar su comportamiento (De Solminihac, 2018). Para lograr esto, es necesario emplear metodologías efectivas para evaluar y modelar su condición a través del tiempo.

En general, las entidades que gestionan los pavimentos urbanos, como por ejemplo los municipios, emplean evaluaciones manuales para un gran porcentaje de kilómetros de su red vial. Las evaluaciones de pavimento manuales presentan una serie de complicaciones ya que son lentas y difícilmente permite evaluar la red completa. Por otro lado, su confiabilidad depende netamente del evaluador y es inseguro para quien lo efectúa (De Solminihac, 2018; Osorio, 2015).

Es por estas razones que, a lo largo de los años, distintos autores han trabajado para encontrar la forma de automatizar la detección de deterioros en distintos tipos de pavimentos. Lo que estos autores han concluido, es que la inteligencia artificial, particularmente el Deep Learning, es una gran herramienta para llevar esto a cabo. Algunos estudios (Berrio & Ussa, 2012) trabajaron con una red neuronal propia, mientras otros (Jenkins et al., 2018), (Carr et al., 2018), (J. Yin et al., 2019), (Tran et al., 2022) trabajaron con distintas arquitecturas previamente definidas y entrenadas de redes neuronales para lograr el mismo objetivo. Cabe destacar que en lo que están de acuerdo estos autores es que, para implementar esta tecnología en una diferente zona geográfica, es necesario entrenar la red con imágenes locales. En Chile, se usan equipos de alto rendimiento para determinar de forma automática el estado de los pavimentos, sin embargo, poseen costos muy elevados inasequibles para agencias locales.

Para llevar esto a cabo en suelo nacional, se está desarrollando el proyecto FONDECYT 11201150 “*Evaluación y Modelación del Comportamiento de Pavimentos Urbanos Usando Machine Learning*” (Osorio, 2015). Una primera etapa del proyecto desarrolló una metodología de evaluación de pavimentos urbanos de forma semi-automática (Rozas, 2023) con cámaras de bajo costo (Venegas, 2022). La segunda etapa, en la cual se enmarca esta investigación, propone evaluar automáticamente los pavimentos urbanos a través de la utilización de algoritmos de Deep Learning para detectar automáticamente los deterioros desde imágenes obtenidas en terreno con tecnología de bajo costo.

Dado el alcance de esta investigación de Magíster y el tiempo disponible para el mismo, el enfoque será en pavimentos urbanos asfálticos.

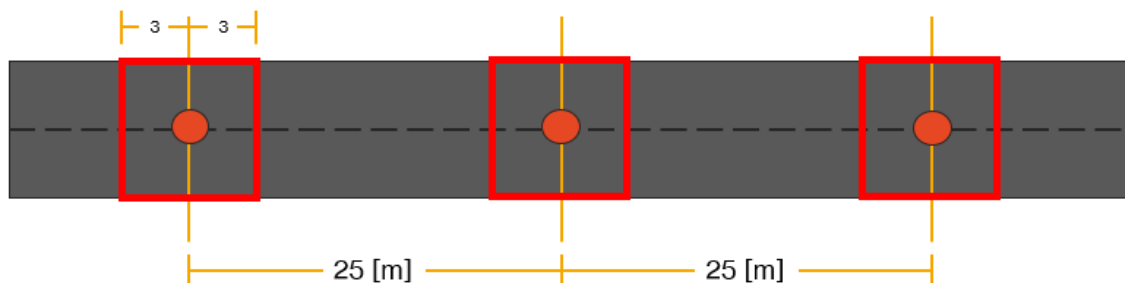
## 2 Estado Actual de la Auscultación de Pavimentos

Actualmente, las metodologías automatizadas de auscultación de pavimentos no son pagables por los municipios de forma regular, por lo que no existe una base de datos con la condición histórica de los pavimentos. Para subsanar esto, diversas no existe ningún estudio en Santiago para determinar el estado de los pavimentos. El tema de investigación al cual se están invirtiendo los esfuerzos son los sistemas de gestión, lo que conlleva a obtener una metodología de evaluación, muestreo y condición de los pavimentos, pero no directamente a llevar a cabo la evaluación como tal.

Por lo tanto, dado que no existen datos directos del estado actual del pavimento en Santiago, se podría decir que queda en la incertidumbre si se encuentran en buen o mal estado.

### 2.1 Monitoreo y Evaluación de Pavimentos en Forma Manual

Para auscultar pavimentos asfálticos urbanos, se selecciona una franja de 6 [m] de largo cada 25 [m] como muestra la *Figura 2.1* por lo que, en este caso, se estaría auscultando el 24% del total del pavimento.



*Figura 2.1: Caso general para auscultar pavimentos asfálticos según MANVU. Fuente propia*

Esta auscultación se realiza de forma manual por un operario en terreno que debe medir la cantidad de grietas y/o el área que ocupan, dependiendo del tipo de deterioro.

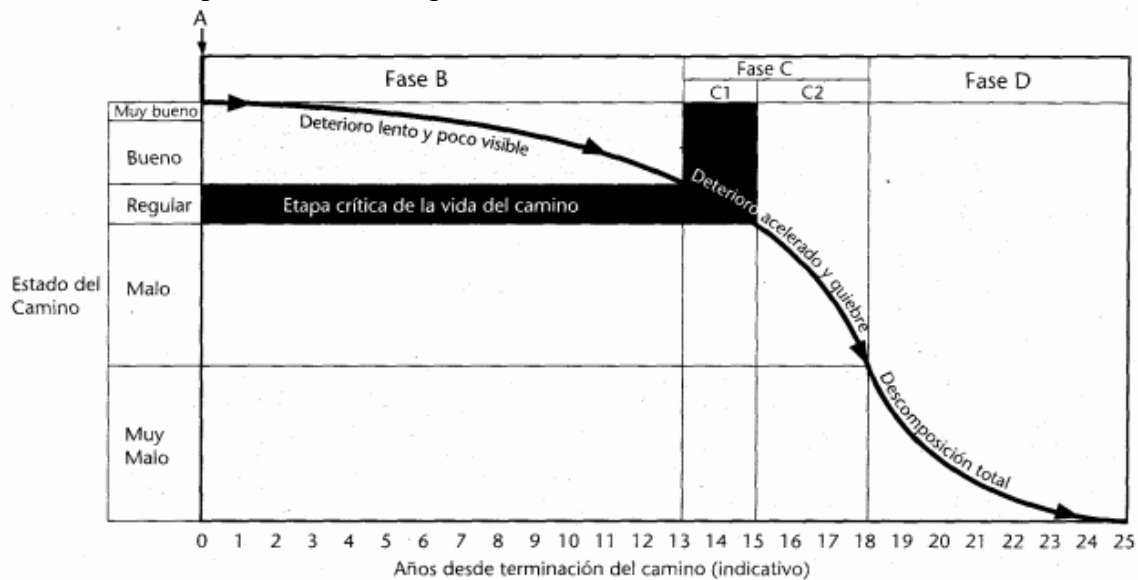
Sólo analizando este procedimiento de auscultación, se pueden encontrar varios problemas:

1. El pavimento auscultado representa un porcentaje muy bajo, tanto para pavimentos de hormigón como asfálticos.
2. Las mediciones de los deterioros dependen del operador que, si bien se intenta regular con el manual de metodología de auscultación de pavimentos, puede haber variaciones en las mediciones dependiendo de quién haga las auscultaciones.
3. Se debe detener el tránsito del área para poder medir las grietas y deterioros de la mejor forma posible.
4. Es peligroso para el operador a cargo de la medición, dado que debe estar de pie sobre los pavimentos realizando las mediciones correspondientes.

## 2.2 Importancia de la Auscultación de Pavimentos

Los pavimentos tienen como propósito servir al tránsito en forma segura, confortable y eficiente (De Solminihaç, 2018). Los caminos son esenciales para la comunidad y su condición afecta directamente el bienestar de las personas. Particularmente, la vialidad es el sustrato físico del sistema de transporte sobre la cual circula el tránsito.

En la **Figura 2.2**, se muestra en forma general el deterioro del pavimento asfáltico en el tiempo (Bull & Schliessler, 1994). En la imagen, se puede notar que en los primeros años (fase B) el deterioro es lento y poco visible. Luego (fase C), hay un periodo de tiempo en que el deterioro se acelera y se produce un quiebre para que finalmente, en la fase D, se genera la descomposición total del pavimento.



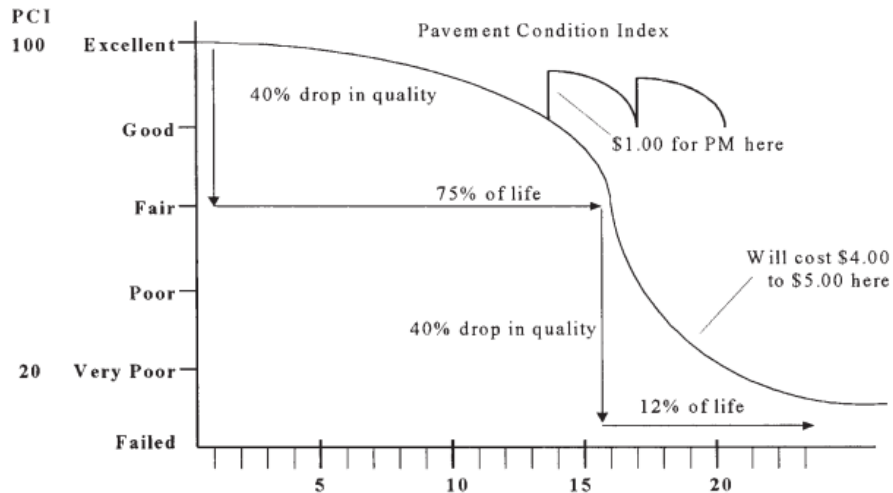
**Figura 2.2:** Progresión del estado de un pavimento en el tiempo

Con el fin de evitar esto, se cuenta con dos tipos de mantenimiento de pavimentos (Hicks et al., 1999), los cuales son:

1. Preventivo: Se realiza para arreglar deterioros menores, fallas que pueden progresar en el tiempo y reducir la necesidad de mantenimiento correctivo.
2. Correctivo o reactivo: Se lleva a cabo luego que el pavimento presente deterioros muy avanzados, como pérdida de roce, ahuellamiento moderado a severo o grietas muy extensivas.

A pesar de que se necesiten ambos para llevar a cabo un completo programa de mantenimiento de los pavimentos, los esfuerzos deberían estar enfocados en realizar un mantenimiento preventivo, es decir, antes de que el pavimento llegue a necesitar un mantenimiento correctivo debido a las grandes diferencias de costos de éstos, como se muestra en la **Figura 2.3**.

Se puede notar que, si se trabaja con mantenimiento correctivo, el costo de llevar esto a cabo puede ser el doble que si se trabaja con mantenimiento preventivo.



**Figura 2.3:** Ciclo de vida típico de un pavimento

De esta forma, se recalca la importancia de ser capaz de modelar el comportamiento de los pavimentos para poder predecir cuándo es necesario llevar a cabo el mantenimiento preventivo, y evitar el mantenimiento correctivo. Y para llevar esto a cabo, lo principal es poseer datos suficientes que sean capaces de respaldar los modelos de predicción, lo que actualmente es complicado debido a la falta de información por el método de auscultación vigente.

### **2.3 Deterioros en Pavimentos de Asfalto en Zonas Urbanas**

Con el fin de determinar los deterioros que se esperan detectar, se hizo un estudio en base al ICPU (Osorio, 2015), seleccionando así los deterioros presentados en la Tabla 2.1.

El cálculo del ICPU para pavimentos asfálticos viene dado según las siguientes ecuaciones, dependiendo si los datos son recolectados de forma manual o automática.

$$ICPU_{manual} = 10 - 0,038 FC - 0,049 TRC - 0,046 DP - 0,059 R - 0,237 P$$

$$ICPU_{auto} = 10 - 0,031 FC - 0,040 TRC - 0,028 DP - 0,082 R - 0,143 IRI$$

Donde:

FC: Grieta de fatiga (%)

TRC: Suma de grietas longitudinales o transversales (%)





DP: Parches deteriorados (%)


R: Ahuellamiento en mm, calculado como el promedio del ahuellamiento e segmentos en una unidad muestral

P: Baches (%)

IRI: Índice internacional de rugosidad en m/km, calculado como el promedio de la rugosidad de los segmentos de una unidad muestral

**Tabla 2.1: Deterioros de pavimentos asfálticos a analizar**

Nombre Deterioro	Ejemplo
Agrietamiento por Fatiga	
Grietas Transversales	
Grietas Longitudinales	
Baches	

Nombre Deterioro	Ejemplo
Parches	

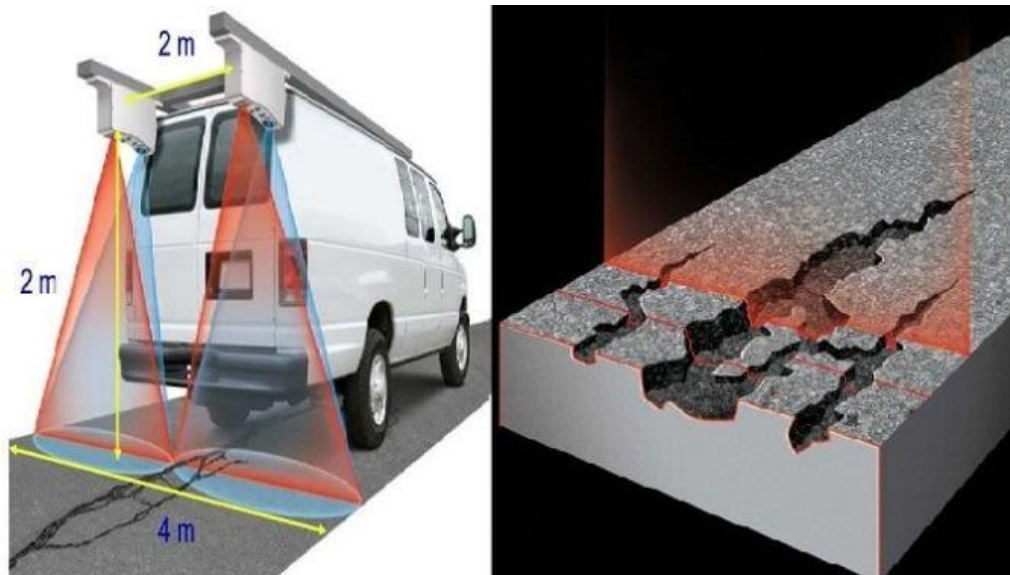
## 2.4 Tecnología Actual

A continuación se presentan distintas tecnologías que son evaluadas para ser aplicadas con el fin de llevar a cabo esta investigación, comenzando desde la tecnología de recolección de datos (laser crack measurement system y cámaras de bajo costo) y finalizando con un repaso de distintos tópicos de la inteligencia artificial.

### 2.4.1 LCMS (Laser Crack Measurement System)

Un ejemplo de tecnología vanguardista en la automatización de la auscultación de pavimentos que se usa actualmente, es el “LCMS”, cuyas siglas vienen de las palabras en inglés “Laser Crack Measurement System”. Este equipo es capaz de realizar un escaneo de la superficie en 3 dimensiones, como muestra la *Figura 2.4*.

Sin embargo, a pesar de ser un equipo con el que se pueden obtener muy buenos resultados, posee costos bastante elevados, lo que lo convierte inasequible para municipios y se descarta su uso para este proyecto.



*Figura 2.4: Laser Crack Measurement System*

#### **2.4.2 Cámaras de Bajo Costo**

Dado que el instrumento LCMS se escapa del presupuesto de los municipios, se propone utilizar una cámara de bajo costo que sea asequible para éstos. De esta forma, se utiliza una cámara GoPro Hero 8 (Venegas, 2022) como se presenta en la *Figura 2.5*, que posee un costo de US \$280 al momento de la adquisición.



*Figura 2.5: Foto referencial de la cámara a utilizar en el proyecto*

En la etapa anterior del proyecto, se definió el montaje del sistema de recopilación de imágenes, el cual consiste en la cámara mencionada sobre un selfie stick y un porta bicicletas, lo que permite que el sistema pueda ser montado sobre cualquier vehículo liviano.

La mitad del lente de la cámara, debe estar a una altura de 1.80 m medidos desde el pavimento con una inclinación de 30° respecto la vertical. (Venegas, 2022). El montaje también cuenta con dos láseres verdes que apuntan al pavimento, los cuales sirven como calibración dado que son una medida conocida en la realidad, y de esta forma se puede transformar la medida de pixeles a centímetros. (Rozas, 2023).

Luego, se lleva a cabo una grabación de video con el vehículo desplazándose a no más de 20 km/h para evitar imágenes borrosas. El horario de grabación es, para días soleados, desde el amanecer hasta una hora después del amanecer y una hora antes del atardecer, y para días nublados el horario es dentro del amanecer y el atardecer. Esto con el fin de evitar mucho contraste de luz y sombra dentro de las imágenes que se puedan confundir con deterioros propios del pavimento.

Luego, el video es procesado y con ayuda de la telemetría de la grabación, se extraen fotogramas en los milisegundos precisos con tal de obtener imágenes de la sección completa del pavimento a grabado.

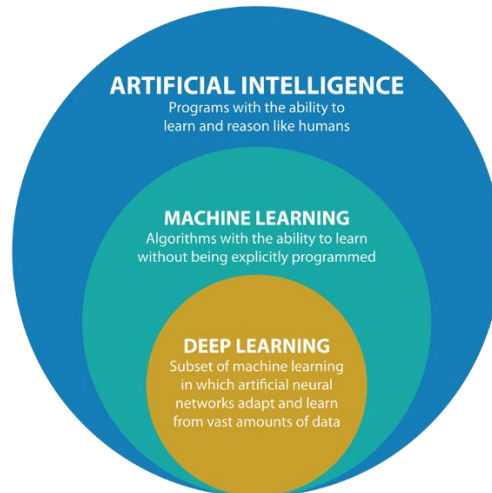
Finalmente, estas imágenes son procesadas a través de una transformación de perspectiva, para obtener la vista en planta para así poder medir los deterioros (Rozas, 2023).

### **2.4.3 Inteligencia Artificial**

Si las imágenes de pavimentos urbanos se siguen analizando manualmente por operadores, se habría automatizado solo una parte del proceso. Es por esta razón que, con el fin de resolver este problema, se propone el uso de inteligencia artificial, particularmente Deep Learning que es muy usado actualmente en el procesamiento de imágenes.

Para entender Deep Learning, primero es importante entender los siguientes conceptos, con ayuda de la **Figura 2.6**:

- **Inteligencia Artificial**: Combinación de algoritmos planteados con el propósito de crear máquinas que presenten las mismas capacidades que el ser humano. Actualmente está presente en variados ámbitos de nuestra vida, tales como: asistentes personales virtuales, logística y transporte, sanidad, comercial, finanzas, climáticas y agrícolas, entre otros.
- **Machine Learning**: Disciplina del ámbito de la Inteligencia Artificial que crea sistemas que aprenden automáticamente, es decir, identifican patrones complejos en millones de datos. Básicamente, la máquina aprende un algoritmo que revisa los datos y es capaz de predecir comportamientos futuros. A su vez, esto implica que estos sistemas mejoran de forma autónoma con el tiempo, ya que van obteniendo más información.
- **Deep Learning**: Lleva a cabo del proceso de Machine Learning, pero en varios niveles distintos, usando una red neuronal artificial que se compone de un número de niveles jerárquicos de la siguiente forma: En el nivel inicial, la red aprende algo sencillo y envía esta información al siguiente nivel, el cual toma esta información simple, la combina y compone una información un poco más compleja, y lo pasa al siguiente nivel, y así sucesivamente. De esta forma es capaz de procesar información compleja y mejorar su rendimiento de aprendizaje.



*Figura 2.6. Explicación básica de inteligencia artificial, Machine Learning y Deep Learning*

Ahora, para conocer cuándo un problema puede ser resuelto mediante Machine Learning, se debe entender su esencia, que establece 3 condiciones (Abu-Mostafa, 2012):

- Debe existir un patrón entre el problema y la solución.
- La solución del problema no puede ser calculado matemáticamente.
- Se debe poseer datos que representen el patrón.

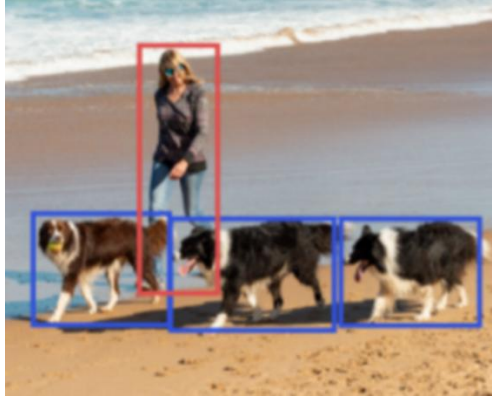
#### **2.4.4 Análisis de Imágenes de Pavimentos con Deep Learning**

En cuanto a autores que ya han llevado a cabo la metodología explicada previamente, existen muchos que la han utilizado con imágenes en tres dimensiones, como las que se obtienen con el LCMS. Sin embargo, como se especificó anteriormente, este tipo de imagen fue descartado debido a su alto costo, por lo que se trabajará con imágenes en dos dimensiones. Las distintas técnicas de análisis de imágenes se detallan a continuación.

#### 2.4.4.1 Análisis de imágenes

En cuanto al análisis propio de las imágenes a través de algoritmos de Deep Learning, se tienen a lo menos dos técnicas distintas (Qualcomm, s. f.):

1. Detección de Objetos: Esta técnica detecta mediante rectángulos los objetos en las imágenes, como muestra la Figura 2.7.



*Figura 2.7. Técnica de detección de objetos*

2. Segmentación de Objetos: Esta técnica, a diferencia de la anterior, tiene como fin marcar los píxeles del objetivo a detectar, como se puede ver en la Figura 2.8.



*Figura 2.8. Técnica de segmentación de objeto*

#### 2.4.4.2 Arquitecturas de Redes Convolucionales

Las arquitecturas de redes convolucionales clásicas se dividen en 2 grupos distintos:

1. De una etapa: Como lo dice su nombre, esta red trabaja haciendo un solo análisis por las imágenes (una etapa). La gran ventaja de estas arquitecturas es que suelen trabajar en detección de objetos muy rápido, por lo que sirven para hacer análisis en tiempo real. Algunos ejemplos de arquitecturas que trabajan de esta forma son:
  - YOLO (You Only Look Once)
  - SSD (Single Shot Detection)
  - RetinaNet
  
2. De dos etapas: Se hace un primer recorrido de análisis de la imagen, en la cual se determina la probabilidad de que el objeto a estudiar se encuentre en los distintos cuadros, para luego volver a analizar los cuadros con mayor probabilidad y determinar si efectivamente se encuentra ahí el objeto o no. Como es de esperarse, las arquitecturas que funcionan siguiendo estos pasos suelen ser menos rápidas, pero más asertivas. Ejemplos de redes que trabajan siguiendo estos pasos son:
  - VGG-16
  - R-CNN
  - U-Net
  - Mask RCNN

Con el fin de determinar las redes neuronales que se han utilizado con fines similares, se hizo una revisión de literatura de investigaciones similares encontrando las siguientes, que han usado diversas redes neuronales, en orden cronológico:

- Red propia (Berrio & Ussa, 2012), es decir, una arquitectura creada de cero y entrenada por ellos mismos.
- U-Net (Jenkins et al., 2018)
- RetinaNet (Carr et al., 2018)
- VGG16 (J. Yin et al., 2019)
- Mask RCNN (Tran et al., 2022)

En Chile, se han llevado a cabo dos investigaciones similares con imágenes de pavimentos interurbanos obtenidas del LCMS (Laser Crack Measurement System), uno con detección de objetos (Cáceres, 2021) y el otro con segmentación (Contreras, 2020). Los resultados de la investigación de segmentación obtuvieron mejores resultados que aquella con detección de objetos. Sin embargo, el etiquetado de imágenes con segmentación (pixel a pixel) requiere mucho más tiempo que los de detección de objetos (bounding boxes).

Para esta investigación se decide usar la arquitectura de YOLO, la cual trabaja como detección de objetos, dado que es fácil de entrenar con nuevas bases de datos y es amigable para nuevos usuarios. En particular, se usarán las versiones 5 y 7. La 5, dado que es la que

más ha sido usada experimentalmente y la 7 porque es una de las más nuevas y ha obtenido mejores resultados que sus versiones anteriores.

### Métricas de Desempeño de Redes Neuronales Artificiales para Detección de Objetos

A continuación, se explican algunas métricas que sirven para configurar o medir el desempeño de las redes neuronales artificiales usadas para detección de objetos.

1. Intersection Over Union; Intersección sobre unión, representa qué tanta exigencia se le pide a la red. Su forma de calcular es la razón entre la intersección y la unión de las etiquetas predichas y las reales/asignadas.

$$iou = \frac{\text{intersección}}{\text{unión}}$$

2. Verdadero positivo, verdadero negativo, falso positivo, falso negativo

Se trabajan con sus siglas en inglés, por lo tanto:

- Verdadero positivo → True positive (TP)
- Verdadero negativo → True negative (TN)
- Falso positivo → False positive (FP)
- Falso negativo → False negative (FN)

Estos conceptos se entienden mejor con la ayuda de la Tabla 2.2.

**Tabla 2.2: Representación del significado de los conceptos TP, TN, FP y FN. 0: nada, 1: detección.**

		Predicción	
		0	1
Realidad	0	TN	FP
	1	FN	TP

3. Precisión: Representa el porcentaje de predicciones correctas.

$$Precision = \frac{TP}{TP + FP} = \frac{TP}{\text{Total resultados positivos}}$$

4. Recall: Representa qué tan buenos son los positivos

$$Recall = \frac{TP}{TP + FN} = \frac{TP}{\text{Total casos reales}}$$

5. F1: Representa la exactitud del modelo

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

6. mAP: Representa la media del área bajo la curva del gráfico precisión vs recall.

#### **2.4.5 Cantidad de Datos para Trabajar con una Red Neuronal**

En cuanto a los datos, se realiza una extensa revisión bibliográfica para determinar cuántas imágenes serán las necesarias para llevar a cabo de forma exitosa el entrenamiento de la red, encontrando resultados de cuántas imágenes de entrenamiento usaron para cada investigación que se expresan en la Tabla 2.3, Figura 2.9 y Figura 2.10. En total se analizaron 28 distintas investigaciones (Bang et al., 2018; Carr et al., 2018; Escalona et al., 2019; Fan et al., 2018; Fei et al., 2019; Gopalakrishnan et al., 2017; Hsieh Yung-An & Tsai Yichang James, 2020; Jenkins et al., 2018; Li et al., 2018; J. Liu et al., 2020; Y. Liu et al., 2019; Maeda et al., 2018; Majidifard et al., 2020; Mandal et al., 2018; Marcelino et al., 2021; Nguyen et al., 2018; Roberts et al., 2020; Sarmiento, 2021; Song & Wang, 2021; Van Phuc et al., 2021; Wang & Hu, 2017; Ye et al., 2021; J.-L. Yin et al., 2019; N. Yusof et al., 2019; N. A. M. Yusof et al., 2018; A. Zhang et al., 2019; A. A. Zhang et al., 2017; Zhang Allen et al., 2018; L. Zhang et al., 2016).

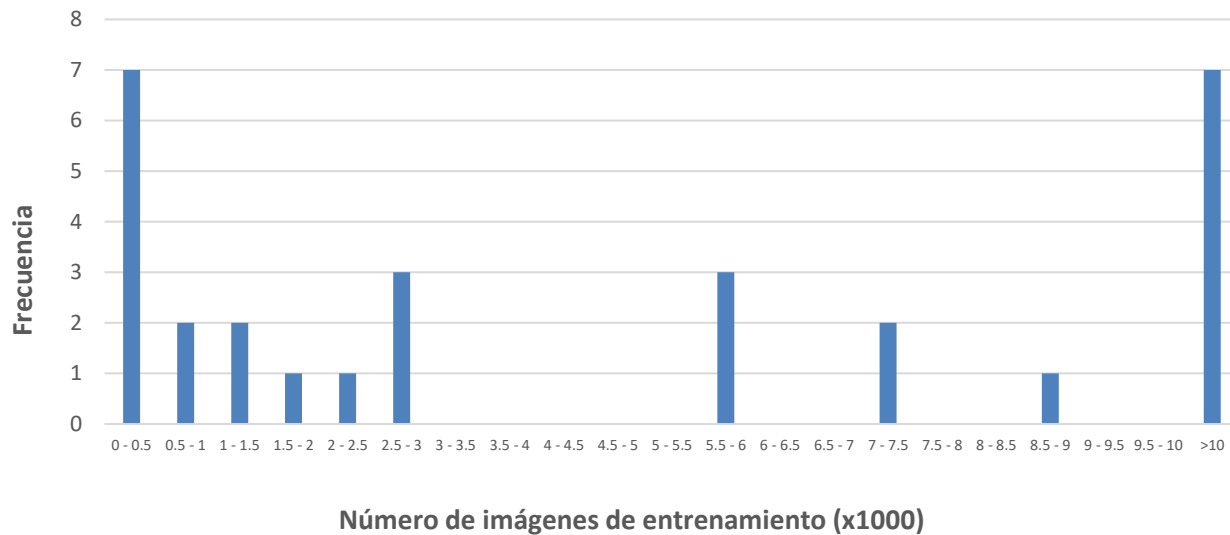
**Tabla 2.3. Números de imágenes usadas para entrenar las redes neuronales con sus respectivas frecuencias.**

Nº de imágenes de entrenamiento	Frecuencia	Nº de imágenes de entrenamiento	Frecuencia
0 - 500	7	0 - 500	7
500 - 1000	2	500 - 1000	2
1000 - 1500	1	1000 - 1500	2
1500 - 2000	0	1500 - 2000	1
2000 - 2500	0	2000 - 2500	1
2500 - 3000	1	2500 - 3000	3
3000 - 3500	0	3000 - 3500	0
3500 - 4000	0	3500 - 4000	0
4000 - 4500	0	4000 - 4500	0
4500 - 5000	0	4500 - 5000	0
5000 - 5500	0	5000 - 5500	0
5500 - 6000	3	5500 - 6000	3
6000 - 6500	0	6000 - 6500	0
6500 - 7000	0	6500 - 7000	0
7000 - 7500	2	7000 - 7500	2
7500 - 8000	0	7500 - 8000	0
8000 - 8500	0	8000 - 8500	0
8500 - 9000	1	8500 - 9000	1
9000 - 9500	0	9000 - 9500	0
9500 - 10000	0	9500 - 10000	0
>1000	7	>10000	7

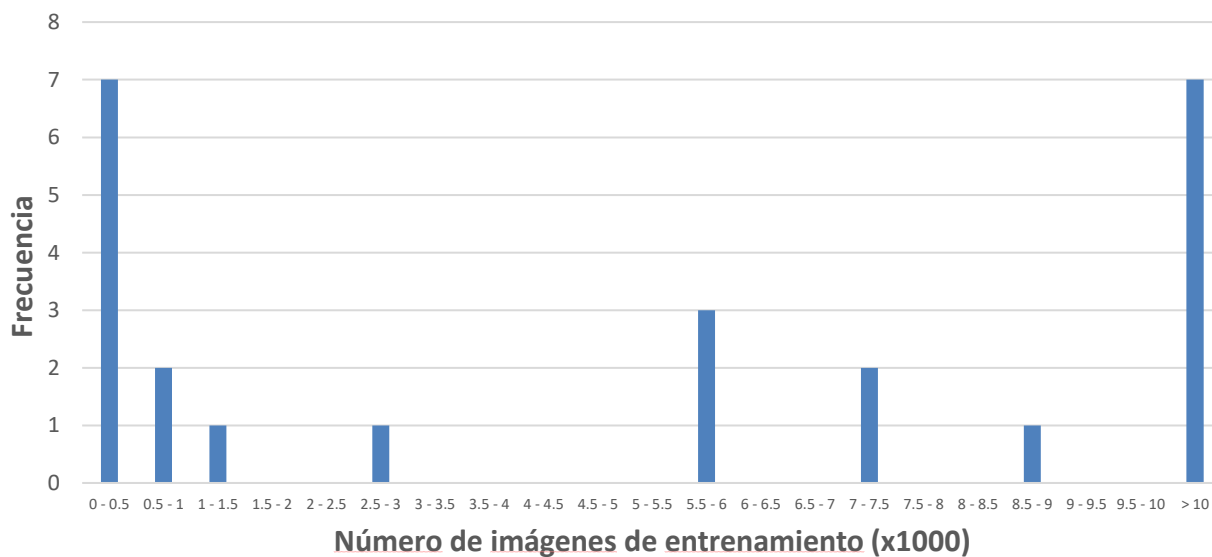
a) b)

En la Tabla 2.3, la sección a) representa investigaciones que han usado tanto imágenes en 2 dimensiones como en 3 dimensiones, mientras que la sección b) representa investigaciones que solo han usado imágenes en 2 dimensiones, lo que se acerca más a esta investigación.

Por otro lado, las Figura 2.9 y Figura 2.9 representan de forma más ilustrativa el contraste entre números de imágenes utilizadas para entrenar redes neuronales, dado que algunos usan pocas (menos de 500), y la misma cantidad de investigaciones usaron sobre 10.000 imágenes.



**Figura 2.9. Gráfico de barras de la tabla 2.1 a (investigaciones con imágenes en 2D y 3D)**



**Figura 2.10: Gráfico de barras de la tabla 2.1 b (investigaciones que solo consideran imágenes en 2D)**

Como se puede notar en los gráficos de barra, se presenta una gran variabilidad en el número de datos de entrenamiento, ya que algunas investigaciones usaron una cantidad menor a 500 datos de entrenamiento y otros un número mayor a 10.000 datos.

## **2.5 Definición del Problema y Oportunidades de Investigación**

En resumen, el problema actual es que la auscultación manual que se realiza actualmente evalúa un porcentaje pequeño de la totalidad de la red. Su medición depende del criterio del operador, es inseguro y se debe detener el tránsito. Por otro lado, la auscultación automática implica costos muy elevados e inasequibles para las agencias locales. Finalmente, la metodología desarrollada en el mismo proyecto evalúa los pavimentos de forma semi-automática, dado que la recolección de imágenes es automática, pero la evaluación de las imágenes requiere de un operador, es decir, falta un paso final para automatizar en su totalidad el proceso.

Se concluye que el tema de “Aplicación de Deep Learning para Automatizar la Auscultación de Pavimentos a través de Imágenes de Bajo Costo” es un tema de investigación que, de obtener buenos resultados, puede traer muy buenos beneficios a la comunidad.

Al tratarse de tecnología de bajo costo, ésta sería asequible para agencias locales encargadas de la gestión de pavimentos urbanos como lo son los municipios. Esto implicaría que sería más fácil auscultar, por lo que se podría aumentar su frecuencia y así, al poseer más información del estado de los pavimentos a través del tiempo, se podrían mejorar los modelos de comportamiento de éstos. Finalmente, en base a eso, los municipios y agencias serían capaces de determinar el momento oportuno para realizarle mantenciones preventivas a los pavimentos y con esto disminuir el costo de su mantenimiento.

Por otro lado, ya se encuentra definido el sistema de recolección de datos automático, encontrándose pendiente aún el automatizado de la detección de deterioros, dependiendo aún de un operador.

### ***3 Propuesta de Investigación***

A continuación se presentan tópicos generales de la investigación, dentro de ellos la hipótesis científica, objetivos generales y específicos, metodología de investigación y la estructura de trabajo.

#### **3.1 Hipótesis Científica**

Los algoritmos de Deep Learning detectan deterioros de pavimentos urbanos en imágenes recolectadas con una metodología de bajo costo y con esto, clasifican los respectivos deterioros y logran automatizar la evaluación.

#### **3.2 Objetivos**

A continuación se presentan los objetivos generales y específicos de esta investigación.

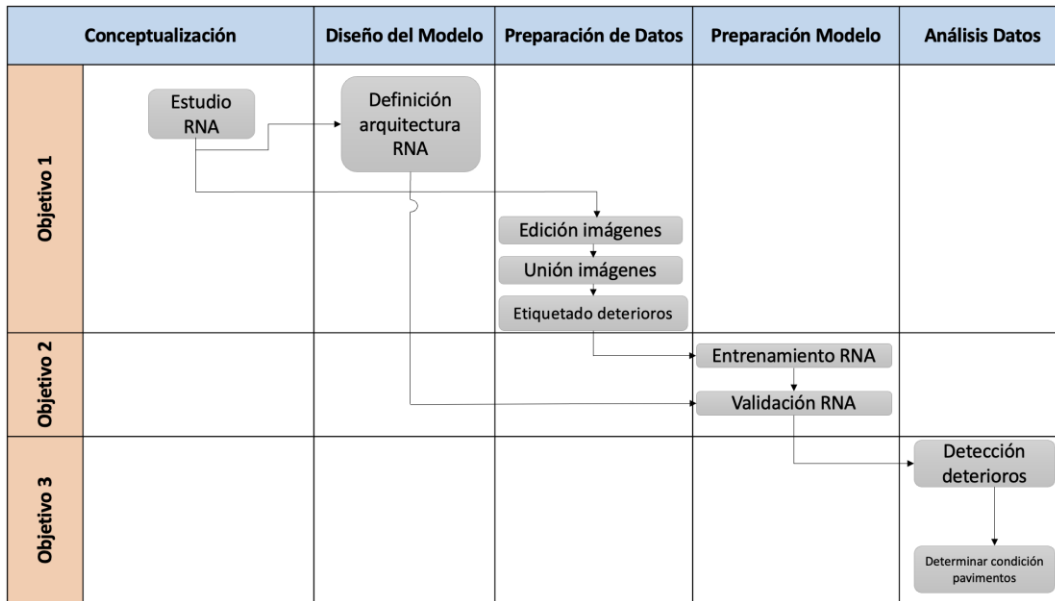
##### **3.2.1 Objetivo General**

Aplicar y validar algoritmos de Deep Learning para automatizar la detección de deterioros en pavimentos urbanos provenientes de imágenes de bajo costo obtenidas en terreno.

##### **3.2.2 Objetivos Específicos**

1. Selección de arquitectura óptima de Deep Learning que sea capaz de detectar y clasificar deterioros en imágenes obtenidas con instrumentos de bajo costo en base a la revisión de la literatura. Con esto, se debe determinar el procesamiento de las imágenes para que sean compatibles con la arquitectura seleccionada.
2. Entrenar, evaluar y validar los algoritmos de Deep Learning.
3. Determinar la condición de los pavimentos urbanos a partir de los valores de deterioros obtenidos en el procesamiento de las imágenes.

### 3.3 Metodología de Investigación



**Figura 3.1: Diagrama que representa la metodología de investigación**

La Figura 3.1 presenta el diagrama de la metodología de investigación.

El primer paso, que corresponde al objetivo 1 y conceptualización del problema, es llevar a cabo un estudio exhaustivo de las redes neuronales artificiales. El segundo paso, que incluye el diseño del modelo es la definición de arquitectura de la red neuronal artificial a utilizar. Finalmente, se deben llevar a cabo los pasos para que los datos sean compatibles con la red neuronal definida, los que consideran la edición y unión de imágenes y por último el etiquetado de deterioros.

Para completar el objetivo 2, que considera la preparación del modelo, se debe llevar a cabo el entrenamiento y validación de las redes neuronales artificiales.

Finalmente, en cuanto al objetivo 3 que conlleva el análisis de datos, considera la detección de deterioros y con esto determinar la condición de pavimentos.

El detalle de cada paso de la metodología se detalla a continuación:

1. Revisión exhaustiva del estado del arte y la práctica de la materia con el fin de seleccionar la arquitectura óptima de Deep Learning que sea capaz de detectar y clasificar deterioros en las imágenes de los pavimentos.
2. Capacitación para aprendizaje y uso de algoritmos de Deep Learning en Python.
3. Procesamiento de grabaciones obtenidas con una cámara de bajo costo montada a un vehículo que se desplaza mientras graba (Venegas, 2022). Los videos obtenidos se deberán procesar para obtener las imágenes que se etiquetarán y clasificarán según los deterioros que presenten.

4. Con estas imágenes etiquetadas se creará una base de datos con la que se entrenará a la red neuronal para que sea capaz de detectar y clasificar nuevas imágenes de pavimentos.
5. Luego, se deberá validar la red neuronal, lo que se realizará verificando que su precisión y exactitud estén dentro de un rango aceptable.
6. Finalmente, se medirán los deterioros en las imágenes detectados por las redes neuronales, para así poder calcular un índice de condición del pavimento y con esto, determinar su estado.

### **3.4 Estructura del Trabajo de Memoria de Título**

Este documento consta de 9 capítulos, los cuales se detallan a continuación:

1. Introducción: Se presenta una pequeña introducción del estudio que se llevó a cabo.
2. : Estado Actual de la Auscultación de Pavimentos: Se abarcan distintos tópicos que tienen como fin definir el problema y presentar la oportunidad de investigación del actual tema.
3. Propuesta de Investigación Se presentan puntos que definen el problema y cómo se llevó a cabo el trabajo, tal como hipótesis científica, objetivos y metodología de trabajo.
4. Conformación de la Base de Datos: Al trabajar con machine learning, los datos toman una gran importancia. En este capítulo se detalla la planificación para la obtención y procesamiento de éstos.
5. Resultados: Se presentan los resultados finales de esta investigación.
6. Análisis de Resultados: Se presentan análisis cualitativos y cuantitativos de los resultados obtenidos.
7. Conclusiones y Recomendaciones: Como detalla el título, se presentan las conclusiones y finales en base a los resultados obtenidos. También se presentan ideas de recomendaciones para trabajos futuros.
8. Lista de Referencias
9. Anexos

## 4 Conformación de la Base de Datos y Diseño Experimental

Dado que se trabaja con Deep Learning, es necesario recopilar una gran cantidad de información, que en este caso corresponde a imágenes de distintos tipos de deterioros. En este capítulo, primero se abarca la base de datos y luego el diseño experimental propio del algoritmo de inteligencia artificial.

### 4.1 Estudio Preliminar Cantidad de Deterioros

En primera instancia, para tener noción de los tipos de más repetidos en la red vial de Santiago, se realizó un inventario con los tipos de deterioros recolectados en auscultaciones manuales de terreno dentro del marco del proyecto en las comunas de Estación Central, Macul, Maipú y Ñuñoa, los que se presentan en la Tabla 4.1.

**Tabla 4.1: Cantidad de deterioros con sus respectivas severidades. B: Severidad baja, M: Severidad media y A: Severidad alta**

	Fatiga			Transversal			Longitudinal			Bache			Parche		
	B	M	A	B	M	A	B	M	A	B	M	A	B	M	A
Total	79	98	196	116	193	35	15	49	58	0	10	19	12	13	27
Total	373			344			122			29			52		

Como se puede analizar en la Tabla 4.1, de un total de 54 unidades muestrales auscultadas y un total de 920 deterioros presentes, aquellos más presentes son las grietas de fatiga y grietas transversales. Los deterioros con menor cantidad son los baches y parches, por lo tanto, se realiza una búsqueda de calles que presenten estos tipos de deterioros.

Para llevar a cabo la recolección de datos, se planificó un recorrido de grabación, que se presenta en Anexo 1. Tabla con Planificación de Recorrido para Auscultaciones, el que considera 52 calles distintos correspondientes a seis comunas distintas (Estación Central, La Reina, Macul, Maipú, Ñuñoa y Peñalolén) y un largo total de 246,5 km.

### 4.2 Metodología de Obtención de Datos

Dado que se requieren muchas imágenes para un buen desempeño de la red neuronal, se planifica un recorrido extenso a llevar a cabo con el montaje de sistema de grabación (Rozas, 2023; Venegas, 2022). Para llevar a cabo de la mejor forma posible estas grabaciones, se trabajó en optimizar el trabajo de terreno. Primero se presenta el paso a paso que se llevó a cabo para determinar el procedimiento óptimo, y finalmente el procedimiento final definido.

#### 4.2.1 Sistema de Referencia para Ajuste Ancho/Alto

En el montaje original, se definió el uso de una estructura metálica roja con forma cuadrada, para así poder hacer el ajuste en que la proporción ancho/alto de la imagen escale la realidad del terreno.

El contratiempo de esta estructura es que depende de que sus lados queden de forma alineada con los bordes de la pista. Por otro lado, al ser metálico, refleja la luz, por lo que su color se ve influido por la intensidad lumínica presente en el momento de la grabación. Para buscar una solución a esto, se realizan pruebas con distintos materiales, entre ellos, pintura en spray, cartulina y goma eva para determinar el material que tenga un mejor comportamiento, lo que significa en este caso, que mantenga su color independiente de la hora de grabación y el clima.

Finalmente, la goma eva es el material que mejor satisface estas condiciones, por lo que se realizan pruebas con pliegos de este material de distintos colores para determinar el que detecta mejor la cámara. Estas pruebas se muestran en la Figura 4.1.



*Figura 4.1. Pruebas de detección del color obtenido en Python para distintos pliegos de goma eva. Las figuras de la izquierda son con imágenes obtenidas en resolución 1080 pixeles, y las figuras de la derecha fueron obtenidas con 4K.*

Finalmente, se selecciona el color azul, ya que es un color difícil de encontrar en el pavimento (a diferencia del color amarillo que se puede encontrar como demarcación horizontal y el color verde como vegetación) y el color rojo también se descarta por no presentar bordes tan definidos.

Finalmente, la estructura para calibrar la imagen consiste en un círculo de goma eva azul de 400 mm de diámetro, el que cumple con poseer un color invariable a las distintas luminosidades del día y no requiere esfuerzo en su colocación en el terreno, ya que tiene las mismas dimensiones en cualquier orientación. La **Figura 4.2** muestra la estructura a utilizar para la calibración, la cual luego de tener definido el sistema, no se requerirá.



*Figura 4.2. Fotografía de la figura circular de goma eva azul*

#### **4.2.2 Ajuste en la Calidad de Grabación**

Con el fin de que la red sea capaz de detectar de mejor forma los deterioros, se busca obtener la mejor calidad de imagen posible. Dentro de los instrumentos ya definidos para el montaje, se hacen pruebas de grabaciones con calidad 4K (en reemplazo de los 1920 pixeles previamente definidos).

La grabación en 4K consiste en imágenes de 3840 x 2160 pixeles, a diferencia de la grabación de 1920 que consiste en 1920 x 1080 pixeles, los que abarcan una zona similar en ambos casos. De esta forma, tener más pixeles para una misma zona de grabación, se presume como una mejor calidad.

Un ejemplo de la diferencia de calidad entre ambos formatos de grabación es el mostrado en la Figura 4.1, en donde se puede notar que las imágenes en 4K definen de mejor forma los bordes de los objetos, en este caso, pliegos de goma eva, a diferencia de las imágenes en 1920, que presentan mayor reflejo por lo que sus bordes se ven difusos.

En la Figura 4.3 se muestra una comparación de imágenes obtenidas por las distintas calidades de la misma zona agrietada de un pavimento. Se puede notar que, si bien la imagen en 1920 pixeles muestra de forma más definidas las grietas, también se presenta más ruido en la imagen.



***Figura 4.3. Fotos de la misma zona de pavimento agrietado. La imagen de la izquierda fue obtenida con grabación en 1920 pixeles, y la de la derecha en 4K.***

Con el fin de complementar, se realizó una binarización a las imágenes de los pavimentos para determinar de forma más objetiva lo planteado en el párrafo anterior. El estudio de esto se presenta en la Figura 4.4, el que reafirma que en la imagen 4K se presenta menor ruido (cantidad de pixeles de pavimento en buen estado del mismo color de las grietas), y las grietas están más definidas (mayor número de pixeles rojos para cada grieta).



***Figura 4.4. Binarización de imágenes de pavimentos deteriorados. La imagen de la izquierda fue grabada con 1920 pixeles, y la de la derecha con 4K.***

Finalmente, es importante realizar un estudio de comparación del gasto de memoria/batería propios de cada tipo de grabación. Los resultados se presentan a continuación:

1. Uso de memoria

**Tabla 4.2. Uso de memoria de un video para cada calidad de grabación**

Resolución	MB/segundo de grabación
1920	5,6
4K	7,6

2. Capacidad duración de video

**Tabla 4.3. Duración máxima de un video según la calidad de grabación**

Resolución	Duración máxima de un video
1920	11 min 28 seg
4K	8 min 52 seg

Y considerando que la cámara posee una memoria SD de 64 GB:

**Tabla 4.4. Capacidad de grabación de una memoria de 64 GB según la calidad de grabación**

Resolución	Capacidad por memoria 64 GB
1920	2 hrs 56 min
4K	2 hrs 8 min

3. Duración baterías Go Pro

**Tabla 4.5. Duración de las baterías según la calidad de grabación**

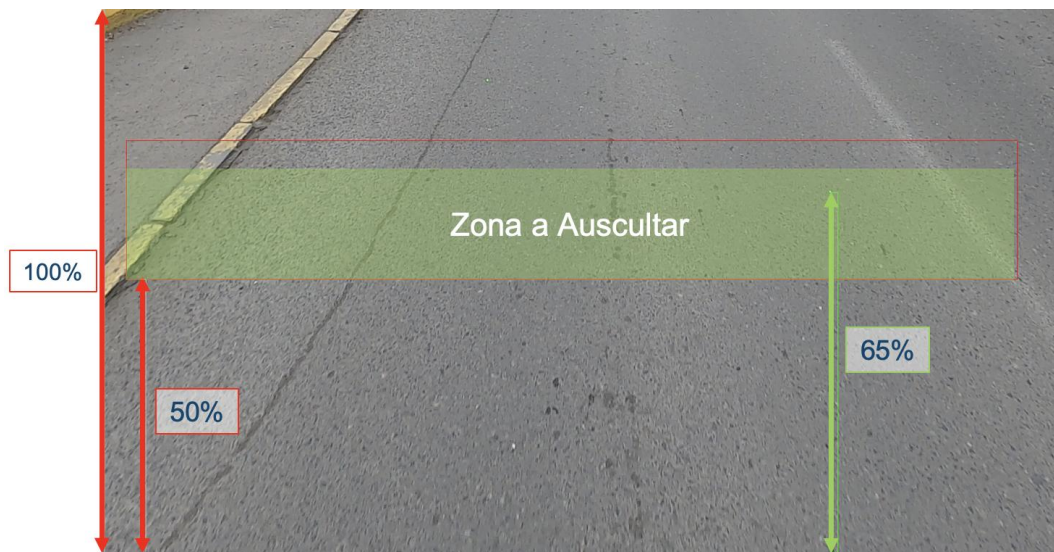
Resolución	Duración baterías, min
1920	88
4K	66

En base a los análisis expuestos, y si bien el gasto de batería y uso de memoria de una grabación en 4K es mayor que aquella realizada con 1920 pixeles, se concluye que no representa un impacto significativo en los recursos disponibles, pero sí una mayor ventaja en cuanto a la detección de los deterioros, por lo que se define que las grabaciones se realizarán con resolución 4K.

### 4.2.3 Posiciones de cada Láser

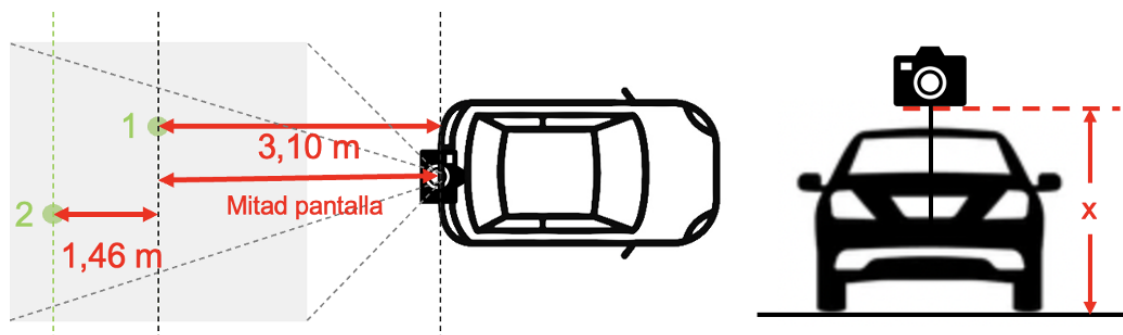
Las posiciones de cada láser no se encuentran definidas. Cada vez que se realiza el montaje, cada láser apunta a posiciones distintas. Sin embargo, se descubrió que dentro de un mismo día, las distintas grabaciones mantienen fijas las posiciones de cada láser, por lo que cada vez que se realiza el montaje, no se presenta una variación de posición de sus componentes.

Por otro lado, al estudiar las imágenes ya obtenidas, se determina que la zona media del lente de la cámara es la óptima para llevar a cabo la transformación, ya que zonas del pavimento más cercanas a la cámara (más abajo de la imagen) pueden no presentar el ancho total de la pista, y las zonas más lejanas (más arriba de la imagen) comprenden un menor número de píxeles para una mayor área, por lo que es de peor calidad. Este análisis se refleja en la **Figura 4.5**, donde el ancho de pista de esa sección es de 3,65 m y la flecha roja de 50% muestra la mitad del alto de la imagen.



**Figura 4.5. Representación de zona a transformar para llevar a cabo la auscultación.**

Por lo tanto, dado que la sección que muestra de mejor calidad el pavimento es la zona media del lente de la cámara, se busca que ambos láser estén siempre cercanos a esta zona, lo que se logra ajustando el porta bicicletas al realizar el montaje en terreno.



**Figura 4.6. Dimensiones definidas del montaje para automatizar la transformación de videos a imágenes**

Finalmente, el punto correspondiente al láser 1 (aquél más cercano al vehículo) se debe encontrar a 3,10 m de la cámara y el punto correspondiente al láser 2 (aquél más lejano al vehículo) se debe encontrar a 1,46 m del láser 1 (o 4,56 m de la cámara). Estas medidas son tomadas en el pavimento de forma longitudinal a la calle, como muestra el esquema de la Figura 4.6.

#### **4.2.4 Ajuste Inclinación de la Cámara**

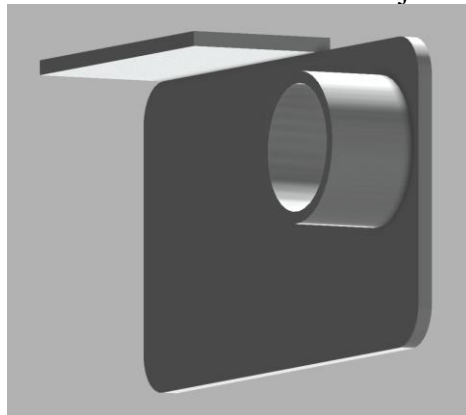
La posición de la cámara, a pesar de tener definida su inclinación, es muy sensible, por lo que ante una pequeña variación de su posición, cambia radicalmente la zona a grabar. Por lo tanto, cada vez que se retira la cámara para cambiar sus baterías o memoria, se debe llevar a cabo una nueva calibración en la transformación de la imagen, lo que implica determinar coordenadas y relación ancho/alto de la zona a extraer de las imágenes, lo que conlleva mayor tiempo de procesamiento humano y computacional.

Para solucionar esto, se diseñó una placa para poder determinar con mayor precisión la zona a la que está apuntando la cámara. La placa consta de una solapa en la zona superior, que sirve como límite para apoyar sobre la cámara, para que la posición entre la placa y la cámara sea siempre la misma. Además, cuenta con un cilindro hueco para introducir un tercer láser, cuyo centro se encuentra a la misma altura del centro del lente de la cámara, por lo que la zona del pavimento que apunta el láser corresponde a la misma zona del medio del lente de la cámara. La **Figura 4.7** muestra fotos de cómo se usa en terreno.



***Figura 4.7. La imagen de la izquierda muestra la placa de PLA con un láser apuntando a la misma dirección que el centro del lente de la cámara. La imagen de la derecha muestra cómo se ve el tercer láser en el pavimento.***

La placa fue modelada en Autodesk® Fusion 360, como muestra la Figura 4.8, y confeccionada en una impresora 3D con material PLA. El costo de la pieza, considerando materiales y energía requerida para su confección, es de \$342, manteniendo el alineamiento del proyecto que consiste en auscultación del terreno a bajo costo.



***Figura 4.8. Modelo en Autodesk® Fusion 360 de la placa para el montaje***

Finalmente, se llevaron a cabo pruebas con distintas grabaciones en que se retira la cámara del montaje, y se vuelve a colocar calibrando su posición con la placa PLA. El resultado fue que las mismas coordenadas y porcentaje de proporción ancho/alto fue válido para todas las grabaciones, logrando así la definición de los parámetros de transformación para todas las grabaciones.

#### 4.2.5 Procedimiento Final para Llevar a Cabo la Obtención de Datos en Terreno

Finalmente se define que las grabaciones se realizarán con resolución 4K, con las siguientes recomendaciones:

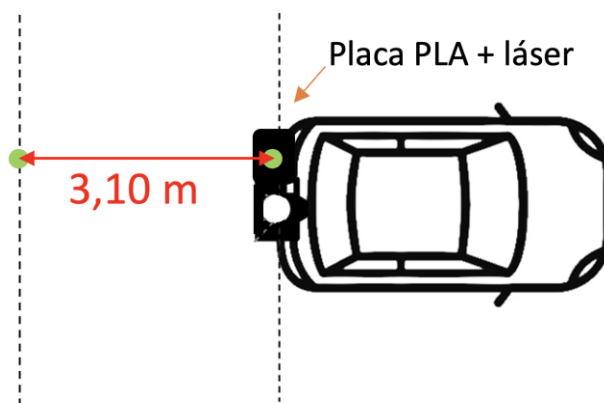
- Cambiar la batería de la cámara cada 60 minutos.
- Cambiar la memoria de la cámara cada 2 horas efectivas de grabación.

Por otro lado, la altura “x” de la cámara, mostrada en la Figura 4.6, depende del número de pasajeros que habrán en el vehículo al momento de obtener las grabaciones. La altura del montaje se debe fijar sin ningún pasajero en el vehículo, y debe ser según lo definido en la Tabla 4.6.

**Tabla 4.6. Altura de montaje de la cámara en función del número de pasajeros**

Personas en el Vehículo Cantidad	Altura base de Cámara m
1	1,73
2	1,74
3	1,75

La placa de PLA se debe usar para verificar la inclinación de la cámara. Para esto, se debe medir en el pavimento una distancia de 3,10 m desde donde está la cámara hacia atrás del vehículo, y el láser que se usa en la placa debe apuntar a ese punto, tal como muestra la Figura 4.9.



**Figura 4.9: Esquema del uso de la placa PLA con el láser para el montaje del sistema**

Gracias al uso de esta placa, ya no es necesario el uso de la goma eva azul ni los láseres verdes para calibrar las coordenadas para la transformación ni para conocer la relación

pixel/centímetros en el pavimento. De igual forma, se pueden utilizar según lo mencionado anteriormente para volver a validar los datos mencionados previamente.

### 4.3 Procesamiento de los Videos

La transformación de videos a imágenes se lleva a cabo en base a los pasos y códigos de Python (Rozas, 2023).

Con el fin de tener un aproximado de tiempo que tardará esta actividad, se calcula el tiempo que conlleva correr los códigos que conllevan mayor demanda computacional, los cuales se presentan en la Tabla 4.7.

**Tabla 4.7: Tiempo que tardan en correr distintos códigos de la transformación de imágenes**

Función del código	Segundos que toma por cada segundo de video
Obtención de fotogramas del video	3,5
Transformación de cada imagen para simular vista en planta	0,8
Normalización de las imágenes	33,0
Unión de 5 imágenes para obtener una sección	0,4

Por lo tanto, dado que es esperable que el procesamiento lleve mucho tiempo, se definen métodos para optimizarlo.

A continuación se enumera el paso a paso de la transformación de videos a imágenes (Rozas, 2023). Aquellos pasos a los que se les realizó un cambio con el fin de optimizar el proceso están marcados con asterisco (\*), y se detallará la modificación en el siguiente inciso.

1. Obtener la imagen para realizar la transformación
2. Ajuste de imagen para determinar la zona a transformar\*
3. Transformar la imagen basándose en la relación ancho/alto entregado por la figura en terreno\*
4. A partir de la telemetría del video, extraer las imágenes en el tiempo necesario para formar la sección
5. Transformar todas las imágenes
6. Normalizar las imágenes\*
7. Unir las imágenes para formar la sección

Las modificaciones en los pasos se detalla a continuación.

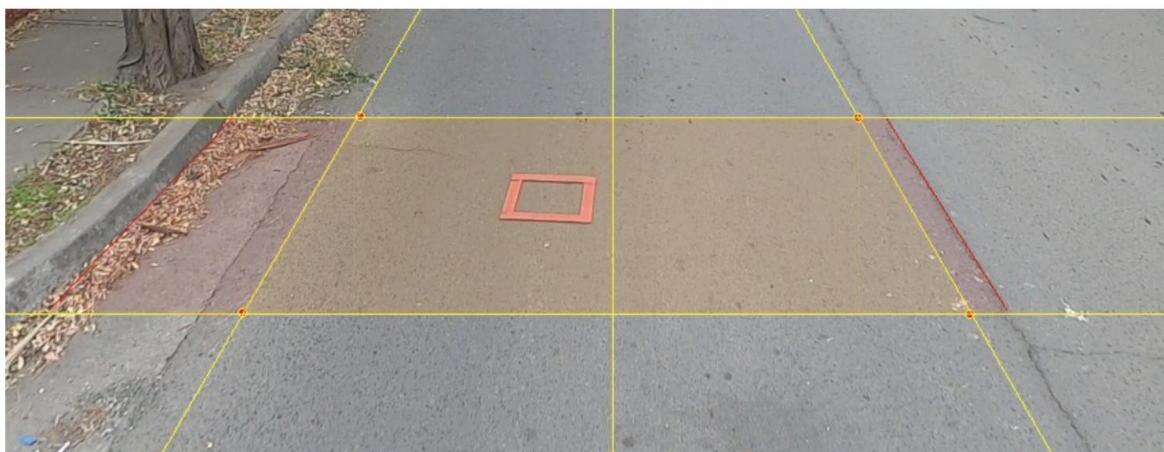
### 4.3.1.1 Ajuste de imagen para determinar la zona a transformar

En base a las modificaciones realizadas para la recopilación de datos en terreno, particularmente el uso de la placa PLA para fijar siempre la cámara a una misma posición y que la posición de cada láser esté en una posición establecida para todas las grabaciones, ahora es posible definir las coordenadas de la sección a auscultar.

Esta zona limita el ancho de sección auscultada, independiente del ancho de la calle. De esta forma, se evita el ruido que producen las aceras o bordes de las calles, que pueden perjudicar el desempeño de la red neuronal. Se define un ancho de auscultación de 2,66 m, el cual corresponde a un 88% de área de auscultación para calles con un ancho de pista de 3 m; y un 76% para calles con un ancho de pista de 3,5 m. Esta apreciación se muestra en la Figura 4.10, donde la zona amarilla sería la sección auscultada, y la zona roja es el ancho de pista que se omitiría.

Cabe destacar que el ancho auscultado no influye en el resultado final, dado que el output sirve de igual forma para hacer la auscultación a nivel de red, que es lo que se busca en este proyecto. Además, el porcentaje de superficie del pavimento que no se está evaluando es bajo en comparación a la cantidad de superficie que se puede auscultar de forma rápida y económica en las calles.

Por otro lado, para un análisis global, los videos y las imágenes son pueden ser útiles para otras investigaciones, por ejemplo, estado de los bordes de las calles, drenaje, entre otros.



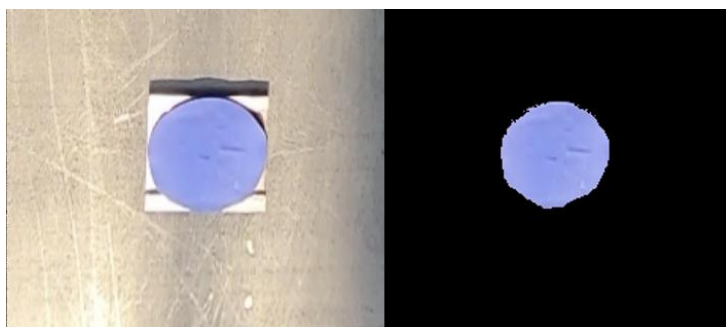
*Figura 4.10. Esquema de coordenadas definidas para transformar las imágenes a una vista en planta*

#### 4.3.1.2 Transformación de la imagen basándose en la relación ancho/alto entregado por la figura en terreno

Dado que el proceso actual de encontrar el ajuste de proporción ancho/alto de la imagen que escale la realidad del terreno es iterativo de forma manual, se define un código que sea capaz de detectar automáticamente la figura en el pavimento que ayuda a esta tarea.

Este código consiste en poder detectar aquellos pixeles que cumplen con un rango de color, brillo y saturación (HSV) definidos, por lo que la estructura de ajuste, como requisito, debe poseer un color invariable, que es algo que cumple el círculo de goma eva azul. Luego, el código itera automáticamente el ajuste de proporción de ancho y alto, hasta que encuentra valores en que la cantidad máxima de pixeles del color definido sean iguales tanto en el ancho como el alto de la imagen, por lo que el círculo en el terreno, también se ve como círculo en la imagen.

La Figura 4.11 muestra un ejemplo de cómo el código en Python detecta el círculo azul en una imagen obtenida a través de una grabación de la cámara Go Pro.

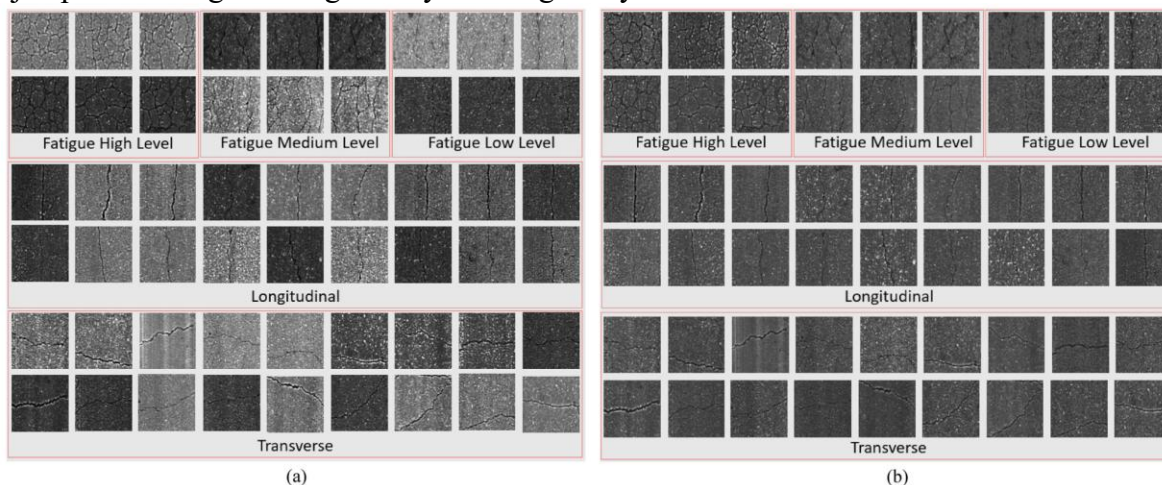


*Figura 4.11. Imagen que muestra el funcionamiento de la detección automática de pixeles azules*

En base a las nuevas coordenadas y la transformación definida, se tiene que el lado de 1 pixel de la imagen equivale a 1,44 mm. Por lo tanto, 1 pixel equivale a 2.07 mm<sup>2</sup>.

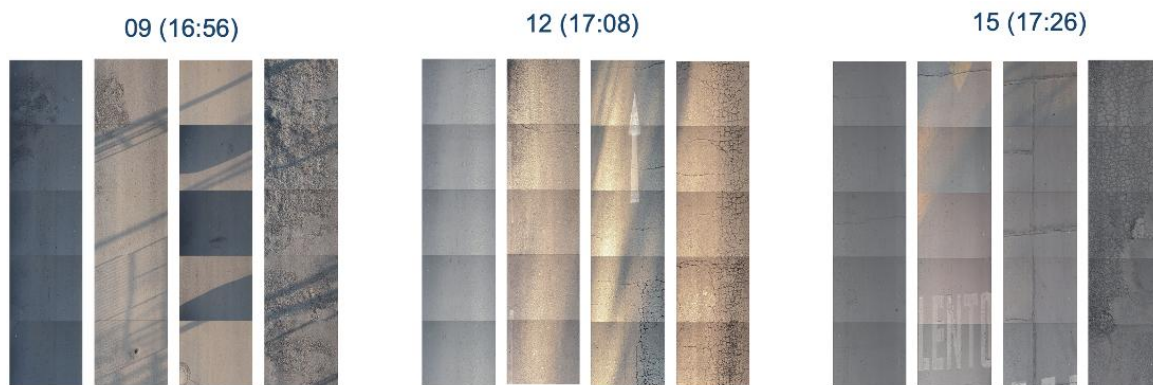
### 4.3.1.3 Normalización de las Imágenes

Este paso no se encuentra en la transformación de imágenes original, y se agrega debido a que el modelo Mask RCNN entrega mejores resultados cuando el input está lo más uniforme y estandarizado posible (Tran et al., 2022). En la Figura 4.12, se muestran ejemplos de imágenes originales y las imágenes ya normalizadas.



**Figura 4.12. Ejemplo de imágenes normalizadas, obtenidas por T. Tran et al (2020) [14]. Las imágenes en (a) son las imágenes originales, y las imágenes en (b) son las normalizadas**

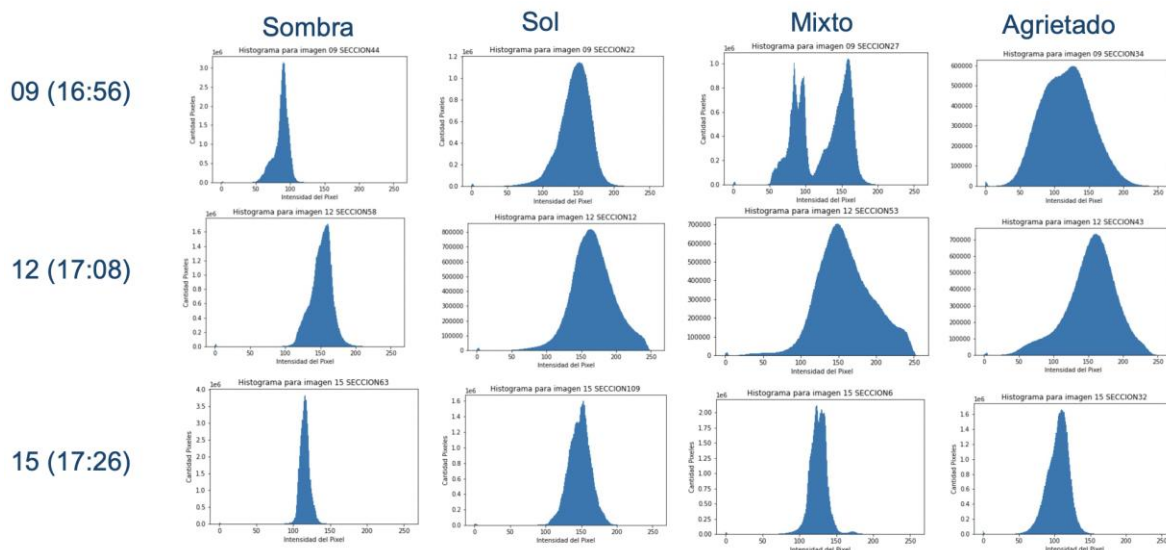
Debido a esto, se hace un análisis de qué tan homogéneas son las imágenes ya obtenidas, para lo cual, se obtienen histogramas (cantidad de píxeles por cada valor de intensidad) en 3 videos obtenidos en Maipú el 23/07/2022 con 15 minutos de diferencia aproximadamente cada uno (16:56, 17:08 y 17:26 hrs.), en una misma tarde con clima soleado. Para analizar el espectro total de posibilidades que se pueden encontrar en una misma grabación, se buscaron imágenes con sombra, imágenes con sol, imágenes mixtas (sol y sombra) e imágenes de pavimentos con un alto nivel de agrietamiento, como muestra la Figura 4.13.



***Figura 4.13: Imágenes seleccionadas para llevar a cabo el análisis de histograma. Se seleccionaron 4 imágenes por cada video: una imagen con sombra, otra con sol, otra con sol y sombra y la última, con alta presencia de deterioros.***

Los histogramas ya mencionados se encuentran en la Figura 4.14. En base a esto, se realizan las siguientes observaciones:

- Como era de esperarse, los píxeles en imágenes con sombra están más concentrados a la izquierda del histograma (valores que representan tonos más oscuros) que las imágenes con sol.
- Para el video grabado más temprano (16:56 hrs), particularmente la imagen mixta (sombra y sol), se presenta un mayor contraste entre las zonas de sombra y sol, pudiéndose notar dos peaks distintos dentro del histograma (uno para la zona con sombra y otro para la zona con sol). En las grabaciones de más tarde, esta diferencia disminuye y se concentran los píxeles en un menor número de intensidades.
- Para los pavimentos agrietados, era de esperarse que apareciera un peak en los valores de intensidad de pixel más bajos (que representan intensidades de pixel más oscuros), dado que representan los píxeles de las grietas, sin embargo, esto es casi imperceptible para los tres videos.

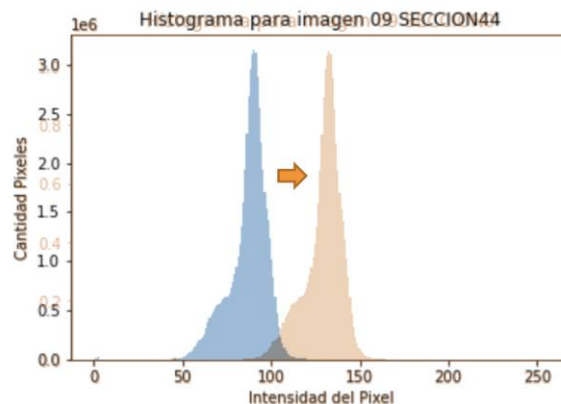


**Figura 4.14. Histogramas de videos obtenidos con 15 horas de diferencia, para imágenes con sombra, sol, sombra y sol y pavimento agrietado**

Como conclusión a este análisis, se determina que las imágenes actuales no presentan una homogeneidad, y varían dependiendo de la hora de grabación. También, hay diferencias de sombra y sol dentro de una misma imagen.

Por lo tanto, se procede a determinar una forma de poder homogeneizar las imágenes, para que se vean iguales independiente de la hora en que es tomada la grabación.

Para lograr a cabo lo anterior, se trabaja con los histogramas de las imágenes, y se decide que el promedio de las intensidades de los píxeles serán trasladados al valor 128, que corresponde a la mitad del histograma, como se muestra en la Figura 4.15. En base al histograma original, se puede afirmar que en un principio la imagen era de tonos más oscuros, y luego de ser normalizada, queda de tonos más claros que los originales, dado que para este caso el histograma se desplazó hacia la derecha (hacia tonos más claros).

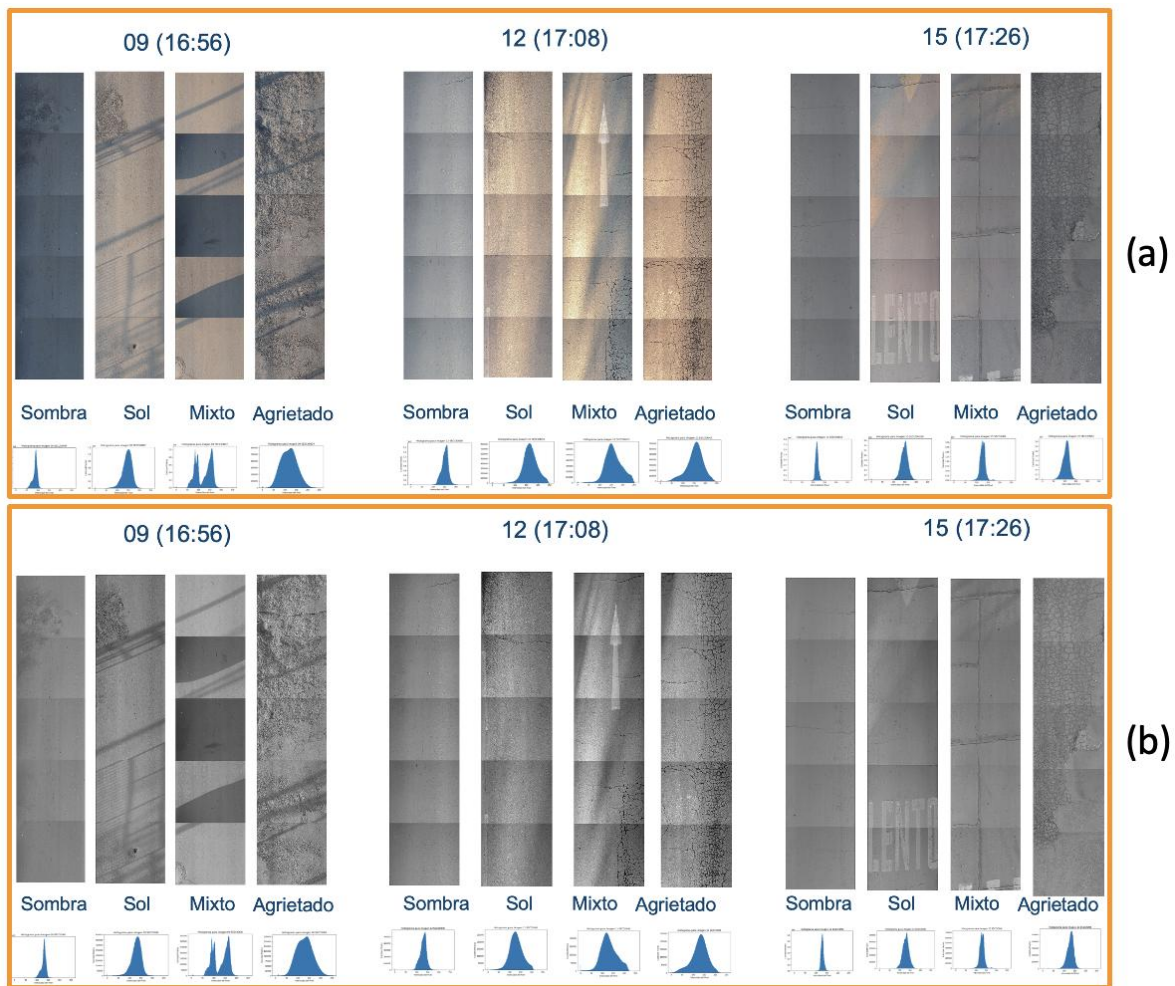


***Figura 4.15. Desplazamiento del histograma para normalizar las imágenes. Los valores azules representan el histograma original, y los naranjos el normalizado***

En la Figura 4.16, se muestra cómo las imágenes del pavimento se ven antes y después de la normalización. Luego de normalizar, no se nota una mayor diferencia entre imágenes que son completamente con sol e imágenes que son completamente con sombra, lo que es un buen resultado. Sin embargo, se sigue teniendo el gran contraste en imágenes que poseen zonas de sombra y sol a la vez.

Si bien la normalización de imágenes fue un paso que tomó tiempo y esfuerzo evaluar y definir, una vez programado no representa un alto costo computacional, por lo que se decide mantener, dado que además presenta beneficios significativos para el desempeño de la red neuronal.

Finalmente, se confeccionó un archivo de notebook de Python con el paso a paso de la transformación de videos a imágenes con el fin de que pueda ser replicable por cualquier usuario.



**Figura 4.16. Imágenes de pavimentos a modo de ejemplo para demostrar la normalización. Las imágenes (a) corresponde a las originales, y las imágenes (b) a las normalizadas.**

#### 4.4 Etiquetado de Imágenes

Para poder entrenar una red neuronal que sea capaz de reconocer distintos tipos de deterioros con sus respectivas severidades, es importante enseñar con variados ejemplos. Esta enseñanza se realiza mostrándole imágenes que deben estar etiquetadas con lo que se desea obtener.

Para llevar a cabo el etiquetado, se usa el programa VGG Image Annotator (VIA) (Dutta & Zisserman, 2019), el cual fue seleccionado porque presenta un entorno amigable con el usuario y es de código abierto basado en HTML, por lo que se ejecuta en cualquier navegador como Chrome, Safari y Firefox.

Por otro lado, se pueden usar distintas formas para llevar a cabo el etiquetado, para este caso se usará el “rectángulo”.

Los datos del output del etiquetado, que se descargan en formato CSV, son los siguientes:

- Nombre archivo
- Tamaño archivo
- Atributos del archivo
- Cantidad de zonas etiquetadas
- ID de la zona
- Forma de la zona
  - Formato: {"name": "rect", "x":\_, "y":\_, "width":\_, "height":\_}
  - Las coordenadas “x” e “y” corresponden a la posición de la esquina superior izquierda de la etiqueta.
- Atributos de la zona
  - Formato: {"Deterioro":\_}

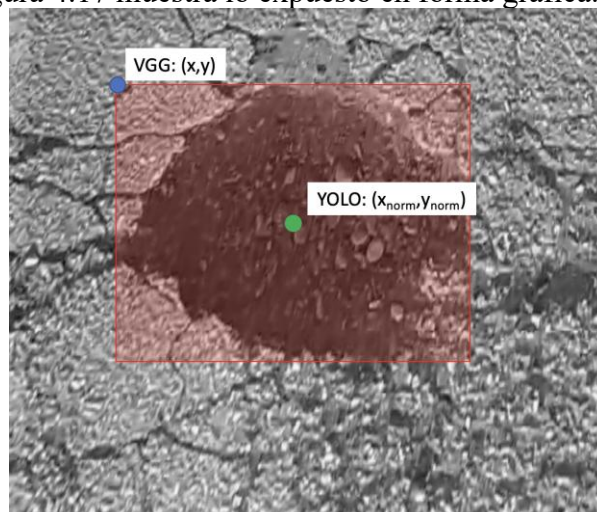
El etiquetado fue llevado a cabo por un equipo de alumnos de ingeniería civil, a los cuales se les realizó una capacitación y se les entregó un manual con el paso a paso.

Para lograr que la red posea un mejor desempeño, se etiquetarán de igual forma algunas singularidades del pavimento para evitar una confusión con los deterioros, las que se detallan más adelante.

Las severidades de los deterioros vienen dadas por distintos factores, según el tipo de deterioro. En base a la calidad de la imagen entregada por la cámara, los deterioros se clasifican en severidad media y alta, dado que aquellas grietas de severidad baja son, en su mayoría, imperceptibles. El detalle de cada uno se explica en la sección *Anexo 2. Deterioros y Singularidades del Pavimento Urbano Asfáltico Detectadas*.

#### 4.5 Procesamiento Etiquetas para Entrenar en Formato YOLO

Las etiquetas entregadas por el programa VIA son distintas a las necesarias para entrenar con YOLO. La diferencia radica es que las coordenadas “x” e “y” de VGG corresponden a la posición de la esquina superior izquierda de la etiqueta, mientras que las coordenadas “x” e “y” necesarias para entrenar YOLO deben corresponder al centro de la etiqueta, normalizadas. La Figura 4.17 muestra lo expuesto en forma gráfica.



**Figura 4.17: Diferencia de coordenadas VGG con YOLO en una etiqueta. El punto azul representa las coordenadas de VGG y el punto verde representa las coordenadas de YOLO, las cuales además deben estar normalizadas.**

En resumen:

- Formato entregado por programa etiquetado VIA:

filename	shape	clase	x	y	width	height
----------	-------	-------	---	---	-------	--------

- Formato etiquetas YOLO:

filename	clase	x_center_norm	y_center_norm	width_norm	height_norm
----------	-------	---------------	---------------	------------	-------------

Por lo tanto, para obtener las coordenadas en formato YOLO, a partir de las coordenadas entrenadas por el programa de etiquetado VIA, se usan las siguientes relaciones:

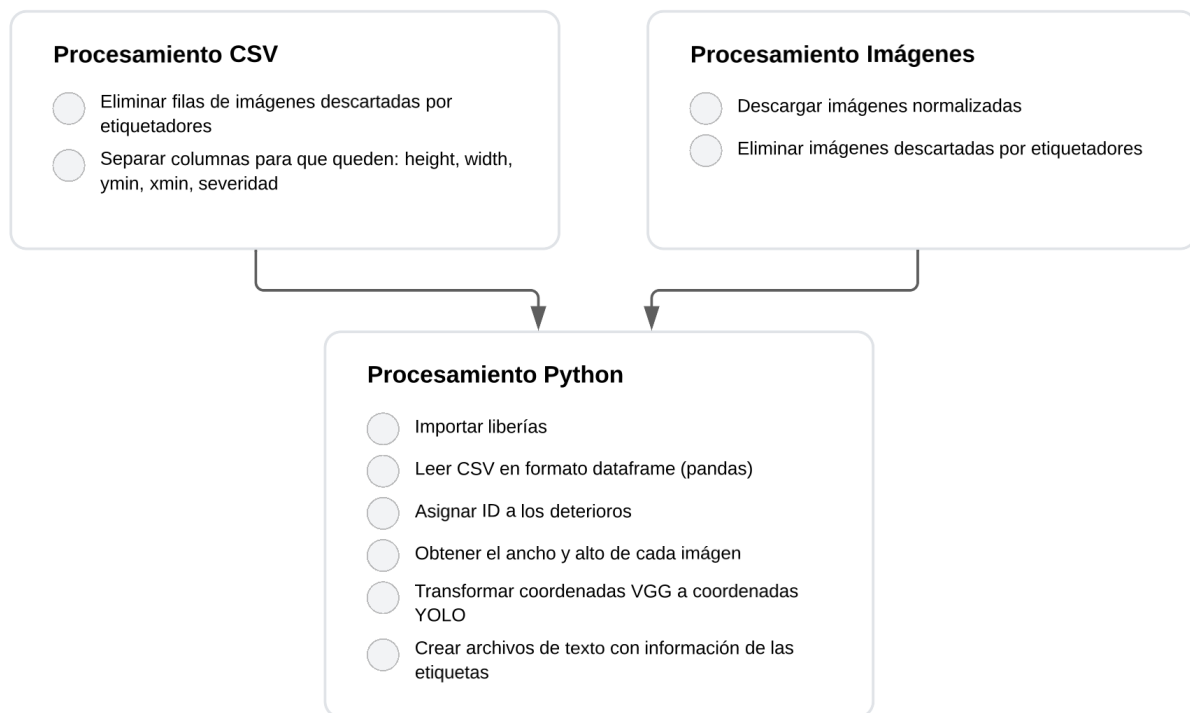
$$x_{center,norm} = \frac{x_{min} + \frac{width_{label}}{2}}{width_{img}}$$

$$y_{center,norm} = \frac{y_{min} + \frac{height_{label}}{2}}{height_{img}}$$

$$width_{norm} = \frac{width_{label}}{width_{img}}$$

$$height_{norm} = \frac{height_{label}}{height_{img}}$$

El procesamiento para obtener los archivos de texto de etiquetas requeridos por YOLO para llevar a cabo el entrenamiento se presenta en la Figura 4.18.



**Figura 4.18: Diagrama para obtener las etiquetas de entrenamiento en formato YOLO a partir del etiquetado en programa VIA**

#### 4.6 Split Datos para Entrenamiento, Validación y Prueba

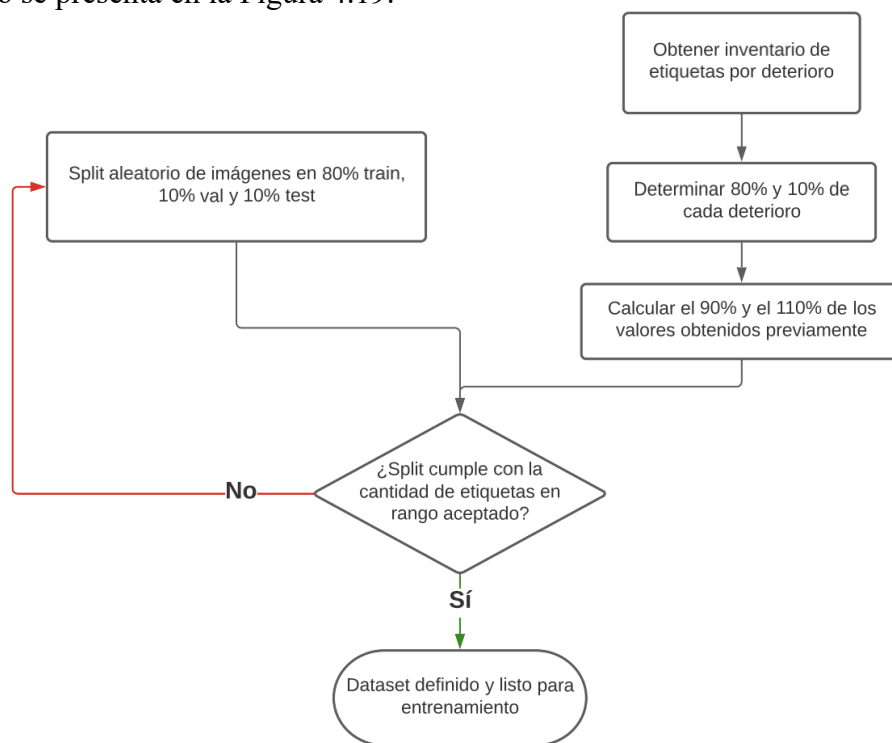
Como indica la literatura, se trabajará con los datos con una distribución de 80% entrenamiento, 10% validación y 10% prueba.

Dado que se está trabajando con detección de objetos en imágenes, y cada imagen puede tener más de una etiqueta de distintas clases, el split de las imágenes requiere de un pequeño procesamiento, con las siguientes restricciones:

- Cada imagen solo puede pertenecer a un set (entrenamiento, validación o prueba).
- Las imágenes se separarán en 80% entrenamiento, 10% validación y 10% prueba.
- Idealmente cada clase también se debe encontrar balanceada en 80% entrenamiento, 10% validación y 10% prueba.

El punto más complicado es, al separar las imágenes, que las clases se encuentren balanceadas entre los distintos set. Para facilitar eso, se define un rango de aceptación, el cual comprende un 10%. Por lo tanto, se tiene en cuenta la cantidad total de etiquetas por cada deterioro, a esa cantidad se le calcula su 80% y su 10% y finalmente a cada uno de esos porcentajes se les resta y suma un respectivo 10%. De esta forma, al separar las imágenes de forma aleatoria, si cada set posee ese rango permitido de etiquetas para cada deterioro, se acepta la distribución planteada.

El proceso se presenta en la Figura 4.19.



**Figura 4.19: Diagrama de flujo para llevar a cabo el split de imágenes en sets de entrenamiento, validación y prueba**

#### **4.7 Aumento de Datos**

Debido al desbalance de clase, y con el fin de mejorar el desempeño de la red, se decide llevar a cabo un aumento de datos con la librería Albumentations (Buslaev et al., 2020).

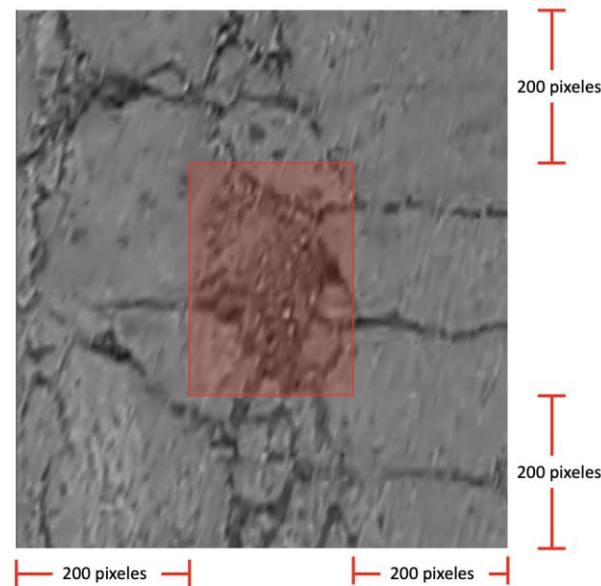
Existen 2 tipos de técnicas:

- Nivel Espacial: Son aquellas técnicas que cambian la posición de los píxeles, como por ejemplo: rotación, flip, cambio de tamaño, entre otros.
- Nivel Pixel: Son aquellas técnicas que mantienen la posición de los píxeles, pero varían su intensidad, como por ejemplo, variar el brillo, contraste, oscurecer, agregar sombras, efectos de lluvia, entre otros.

Debido a la naturaleza de estas técnicas, las etiquetas de las imágenes aumentadas a nivel espacial deben ser procesadas para mantener la etiqueta del deterioro a estudiar, no así con las técnicas a nivel de pixel. Afortunadamente, la librería a utilizar considera esto, por lo que se deben ingresar como input las coordenadas de la etiqueta original, y entrega como output la nueva etiqueta de la imagen editada.

En general, el aumento de datos se lleva a cabo a las mismas imágenes utilizadas para entrenar la red, sin embargo, en este caso, en una imagen puede haber varios deterioros distintos. Y dado que la idea es balancear las clases, este método no es el apropiado. Por ejemplo, algunos baches se crean en zonas donde hay grietas de fatiga de alta severidad, ya que al tener tanta interacción de grietas, zonas del pavimento terminan desprendiéndose. Es por esto que si se quiere realizar una técnica de aumento de datos a una imagen con un bache, lo más probable es que se aumente también la cantidad de etiquetas de las grietas de fatiga.

Dicho esto, se decide realizar “crop” a las imágenes en base a las coordenadas de las etiquetas, para tener nuevas imágenes que contengan solamente el deterioro. Para no perder el contexto del deterioro, se realiza el crop a las coordenadas de las etiquetas sumando, en caso de ser posible, 200 píxeles para cada lado de la imagen, teniendo el cuidado de mantener la etiqueta en la zona original. Todo esto se lleva a cabo con Python. Un ejemplo de lo mencionado se muestra en la Figura 4.20.



***Figura 4.20: Ejemplo de un crop llevado a cabo en un bache. El cuadrado rojo representa la etiqueta original y el resto de la imagen consiste en los 200 pixeles extra por lado.***

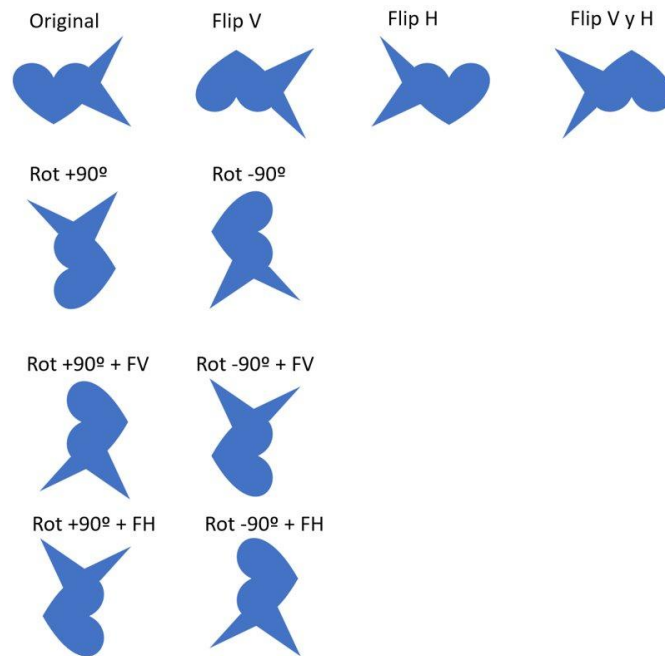
La técnica de aumento de datos a utilizar depende del tipo de deterioro a aumentar. Para los distintos elementos se tienen las siguientes restricciones:

- Las grietas longitudinales deben permanecer de forma vertical, por lo que no se pueden rotar en  $90^\circ$  o  $270^\circ$ .
- Las grietas transversales deben permanecer de forma horizontal, por lo que no se pueden rotar en  $90^\circ$  o  $270^\circ$ .
- Los sumideros, en su gran mayoría, están contruidos para ir de forma longitudinal a la calle, por lo que se evita su rotación en  $90^\circ$  o en  $270^\circ$ . Por otro lado, poseen siempre una característica forma de rejillas en ángulo recto, por lo que no se pueden distorsionar de ninguna forma.

Dicho esto, se estudiaron todas las posibilidades de técnicas de aumento de datos a nivel espacial que no involucren variaciones en el tamaño. Para esto, se hicieron pruebas con una figura asimétrica en el eje horizontal y vertical, obteniendo 7 distintas combinaciones posibles, tal como se muestra en la Figura 4.21, y que se detallan a continuación:

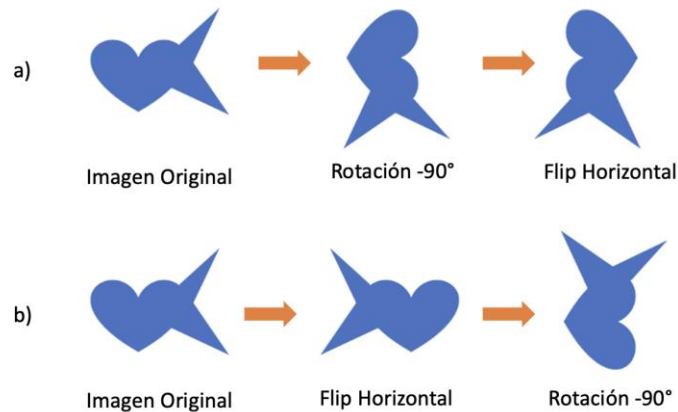
1. Flip vertical
2. Flip horizontal
3. Flip vertical + flip horizontal
4. Rotación  $+90^\circ$
5. Rotación  $-90^\circ$
6. Rotación  $+90^\circ$  + flip vertical = rotación  $-90^\circ$  + flip horizontal

7. Rotación  $+90^\circ$  + flip horizontal = rotación  $-90^\circ$  + flip vertical  
 Entiéndase por rotación  $+90^\circ$  como una rotación en sentido anti-horario, y rotación  $-90^\circ$  en sentido horario.



**Figura 4.21: Todas las opciones posibles de técnicas a nivel espacial de aumento de datos.**

Algo importante a recalcar, es la importancia del orden de las transformaciones, dado que con distinto orden se pueden obtener figuras distintas. Por ejemplo, al realizar rotación en  $-90^\circ$  y luego un flip horizontal, se obtiene una figura distinta que al realizar a la misma figura un flip horizontal y luego una rotación de  $-90^\circ$ , como se muestra en la Figura 4.22.



**Figura 4.22:** Representa la importancia del orden en las transformaciones. A la figura a) se le realiza una rotación y luego un flip; a la figura b) se le realiza un flip y luego una rotación. Se puede ver que los resultados no son iguales.

Estas nuevas imágenes serán incluidas solamente en el conjunto de entrenamiento, siempre y cuando la imagen original se encuentre en ese conjunto. Esto se realiza con el fin de evitar que la red vea el mismo deterioro al entrenar y al validar o entrenar, y de esta forma tener el desempeño verídico.

#### 4.8 Herramientas Por Utilizar para la Red Neuronal

Los modelos se trabajan en el prompt de Anaconda con las librerías propias de YOLO desarrollada por Ultralytics. Esta arquitectura de red neuronal destaca por tener detección rápida, buenos resultados y ser amigable para ser implementada en nuevos datasets.

En cuanto al hardware, se trabaja en un computador Lenovo Legion T5i Tower 6ta Gen, el cual posee un procesador de 11va generación Intel® Core™ i7-11700F, con una memoria de 16 GB UDIMM ARMOR DDR4 3200MHz y una tarjeta gráfica NVIDIA GeForce® RTX™ 3060 12GB GDDR6.

## 4.9 Implementación de Modelos YOLOv5 y YOLOv7

A continuación se presentan los detalles de los modelos YOLOv5 y YOLOv7.

### 4.9.1 YOLOv5

Los hiperparámetros que se le entregan a la red al momento del entrenamiento se presentan en la Tabla 4.8.

**Tabla 4.8 Hiperparámetros asignados al momento de entrenamiento usando YOLOv5**

Hiperparámetro	Valor
Epochs	300
Paciencia	30
Tamaño imagen	640
Pesos iniciales	Small
Tamaño Batch	-1

Cabe destacar que esta opción de YOLO permite asignar al entrenamiento un tamaño de batch de “-1”, el cual significa que YOLO determine automáticamente el máximo tamaño de batch permitido según las capacidades del PC. Lo ideal para tener un entrenamiento en un menor tiempo, es tener un batch mayor, pero esto implica también mayor uso de memoria.

Por otro lado, los hiperparámetros que vienen asignados por defecto en la red, se presentan en la Tabla 4.9.

**Tabla 4.9 Hiperparámetros utilizados en el entrenamiento de YOLOv5**

Hiperparámetro	Valor
Tamaño Batch usado	47
Momentum	0,937
Weight Decay	0,0005
Anchor_t	4.0
Optimizador	SGD

### 4.9.2 YOLOv7

Los hiperparámetros que se le entregan a la red al momento del entrenamiento se presentan en la Tabla 4.10.

**Tabla 4.10 Hiperparámetros asignados al momento de entrenamiento usando YOLOv5**

Hiperparámetro	Valor
Epochs	300
Tamaño imagen	640
Pesos iniciales	Yolov7
Tamaño Batch	16

Por otro lado, los hiperparámetros que vienen asignados por defecto en la red, se presentan en la Tabla 4.11.

**Tabla 4.11 Parámetros utilizados en el entrenamiento de YOLOv7**

Parámetro	Dato
Momentum	0,937
Weight Decay	0,0005
Anchor_t	4.0
Optimizador	SGD

### 4.9.3 Dificultades Computacionales Durante el Entrenamiento

Cabe destacar que hubo dificultades computacionales en el proceso de entrenamiento en el computador que usa la tarjeta gráfica NVIDIA GeForce® RTX™ 3060. Los problemas fueron identificados y resueltos según lo plasmado en la Tabla 4.12.

**Tabla 4.12: Problemas computacionales y sus soluciones al entrenar YOLO**

Problema	Solución
Red no aprendía (llevaba muchas epochs y precisión y recall 4%)	Instalando nuevamente entornos de trabajo. Entornos y carpetas independientes (clonados) para YOLOv5 y YOLOv7
Red no usaba GPU para entrenar (mayor tiempo de entrenamiento)	Volver a instalar programas CUDA y pytorch (versiones más compatibles según los recursos usados)
Cantidad de workers limitados con YOLOv5 (1/8)	Problema de Windows/Python, tiene problemas al trabajar con multiprocesos. Se soluciona modificando workers en archivo “dataloaders.py”

#### **4.10 Planificación Entrenamiento Red Neuronal**

Como se menciona en el punto 4.6 Split Datos para Entrenamiento, Validación y Prueba, se realiza un split en 80%, 10% y 10% para entrenamiento, validación y prueba.

Para validar el funcionamiento de la red y verificar que su desempeño no dependa de la selección de imágenes para cada conjunto, es que se realizan 5 splits aleatorios distintos.

## 5 Resultados

A continuación se presentan los resultados directos obtenidos en esta investigación. Primero, los resultados en la creación de la base de datos y luego los resultados propios de la red neuronal artificial.

### 5.1 Datos Obtenidos (lugar, cantidad de videos, cantidad de imágenes)

Se obtienen 103 grabaciones en las comunas de Colina, Estación Central, Macul, Maipú, Ñuñoa y Providencia. El detalle se presenta en la Tabla 5.1.

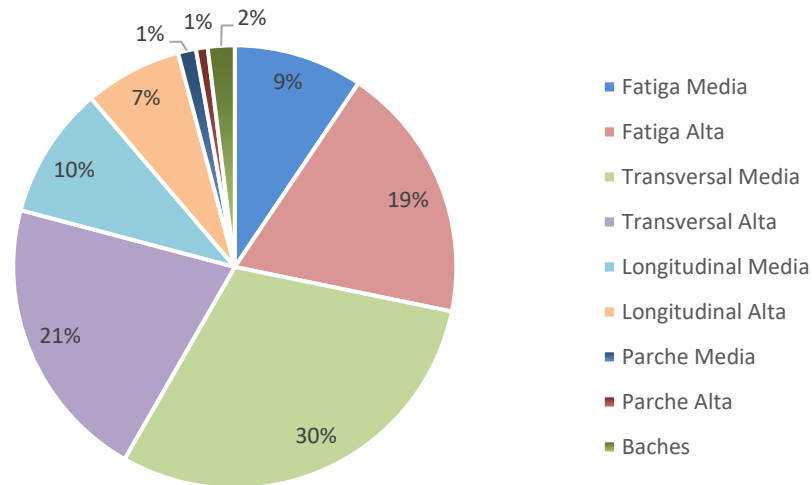
**Tabla 5.1: Distancias grabadas por comuna para entrenar la red**

Comuna	Distancia grabada m
Colina	51.826
Estación Central	5.318
Macul	18.532
Maipú	24.181
Ñuñoa	13.252
Providencia	2.420
Total	115.529

A pesar de haber recolectado un largo total de 115.529 m de secciones de pavimento, finalmente se trabaja con 104.037 m. Esta disminución se debe al descarte de imágenes por distintos motivos, los que se explican en el punto 9.2.3 Imágenes Descartadas.

## 5.2 Etiquetas Totales (cantidad de deterioros totales)

En cuanto a la cantidad de deterioros totales etiquetados en las imágenes de las grabaciones de terreno, se tiene un total de 22.562 etiquetas, las que se distribuyen en las distintas clasificaciones como muestra la Figura 5.1.



**Figura 5.1: Diagrama circular de cantidad de deterioros etiquetados en las imágenes obtenidas de las grabaciones en terreno**

El detalle de la cantidad por deterioro detectada en las grabaciones, se encuentra en la Tabla 5.2.

**Tabla 5.2: Cantidad de deterioros etiquetados en las imágenes obtenidas de las grabaciones en terreno**

Deterioro	Cantidad
Fatiga Media	2.128
Fatiga Alta	4.240
Transversal Media	6.783
Transversal Alta	4.701
Longitudinal Media	2.176
Longitudinal Alta	1.603
Parche Media	296
Parche Alta	193
Baches	442
<b>Total</b>	<b>22.562</b>

### **5.3 Cantidad Final Imágenes y Etiquetas para Entrenar (post aumento de datos)**

Debido al alto nivel de desbalance de clases, el cual se puede notar claramente en la Figura 5.1, se debe aplicar un aumento de datos a los baches, parches de severidad media y alta y grietas longitudinales de severidad media y alta.

Dado que el valor que se repite en cuanto a cantidad de deterioros es 4.000, se define como la cantidad objetivo para los deterioros previamente mencionados.

Considerando lo expuesto en el punto 4.7 Aumento de Datos, se aplicaron las siguientes transformaciones por tipo de deterioro:

1. Bache: flip horizontal, flip vertical, rotación  $+90^\circ$ , rotación  $-90^\circ$ , desenfoque, contraste aleatorio, compresión de imagen y CLAHE.
2. Parches: flip horizontal, flip vertical, rotación  $+90^\circ$ , rotación  $-90^\circ$ , desenfoque, contraste aleatorio, compresión de imagen, CLAHE, ruido Gaussiano, reducción de escala, gofrado, desenfoque en movimiento, ruido multiplicativo, niebla aleatoria, gamma aleatorio, lluvia aleatoria, sombras aleatorias, exceso de timbre, curva de tono aleatoria, agudización, solarización, superpixel y desagudizar.
3. Grietas longitudinales: flip horizontal, flip vertical, compresión de imagen, lluvia aleatoria, reducción de escala y CLAHE.

Cabe destacar que para llegar a 4.000 etiquetas por tipo de deterioro, fue necesario hacer combinaciones de técnicas de aumento de dato, las cuales se presentan en 9.3 Anexo 3. Técnicas de aumento de datos utilizadas por tipo de deterioro.

En estas tablas, las técnicas de distorsión óptica y de grilla están en color gris, lo que se debe a que, en un principio se establecía que estas técnicas eran compatibles con los bounding boxes de YOLO, es decir, al aplicar la técnica se podría obtener como output el bounding box del deterioro modificado. Sin embargo, al aplicarse no fue factible, por lo que finalmente se descartó como técnica.

#### 5.4 Resultado Red Neuronal (YOLO)

Se llevaron a cabo 5 splits distintos para ser entrenados con la red YOLOv5 y YOLOv7. La cantidad de imágenes por entrenamiento se presentan en la Tabla 5.3.

**Tabla 5.3: Cantidad de imágenes por set para cada split**

Split	Entrenamiento			Validación	Prueba
	Real	Artificial	Total	Real	Real
1	6648	17010	23658	833	831
2	6648	17356	24004	833	832
3	6649	16682	23331	833	831
4	6647	16338	22985	832	833
5	6649	17123	23772	832	833

Por otro lado, la cantidad de deterioros por cada split, se presenta en la Tabla 5.4 para aquellos que no fue necesario realizar aumento de datos, y en la Tabla 5.5 para aquellos deterioros que sí.

**Tabla 5.4: Cantidad de deterioros que no requirieron aumento de datos por cada split (grietas de fatiga y transversales). M: Severidad media, A: Severidad alta.**

Split	Fat-M	Fat-A	Transv-M	Transv-A
1	1713	3399	5568	3791
2	1692	3407	5477	3738
3	1732	3371	5417	3825
4	1687	3361	5321	3789
5	1703	3417	5420	3745

**Tabla 5.5: Cantidad de deterioros que requirieron aumento de datos por cada split para grietas longitudinales severidad media**

Split	Long-M		
	Real	Artificial	Total
1	1762	3948	5710
2	1741	4038	5779
3	1749	3864	5613
4	1742	3876	5618
5	1733	3930	5663

**Tabla 5.6: Cantidad de deterioros que requirieron aumento de datos por cada split para grietas longitudinales severidad alta**

Split	Long-A		Total
	Real	Artificial	
1	1297	3096	4393
2	1267	3232	4499
3	1278	3112	4390
4	1276	3216	4492
5	1282	3288	4570

**Tabla 5.7: Cantidad de deterioros que requirieron aumento de datos por cada split para parches severidad media**

Split	Parche-M		Total
	Real	Artificial	
1	241	3492	3733
2	240	3241	3481
3	242	3074	3316
4	234	2947	3181
5	233	3060	3293

**Tabla 5.8: Cantidad de deterioros que requirieron aumento de datos por cada split para parches severidad alta**

Split	Parche-A		Total
	Real	Artificial	
1	143	2750	2893
2	161	3135	3296
3	155	2964	3119
4	161	2907	3068
5	155	3135	3290

**Tabla 5.9: Cantidad de deterioros que requirieron aumento de datos por cada split para baches**

Split	Bache		Total
	Real	Artificial	
1	360	3724	4084
2	361	3710	4071
3	365	3668	4033
4	338	3318	3656
5	371	3180	3551

En cuanto a las singularidades en los pavimentos, la cantidad de etiquetas por split se presentan en la Tabla 5.10.

**Tabla 5.10: Cantidad de singularidades presentes en cada split**

Split	Tapa Alcantarillado	Sumidero	Testigo	Misceláneo
1	332	133	296	332
2	338	115	288	306
3	333	125	271	316
4	328	124	276	298
5	347	126	274	309

Cabe destacar que cantidad de deterioros o singularidades por cada split, no equivale a cantidad de imágenes de cada deterioro por cada split, dado que una imagen real puede tener más de un deterioro o singularidad, de la misma naturaleza o distintas.

Distinto es el caso a las imágenes artificiales, que son aquellas que provienen de técnicas de aumento de datos, en que sí cada deterioro equivale a una imagen artificial, dado que se trabajaron a partir de crop de los deterioros.

## 5.4.1 YOLOv5

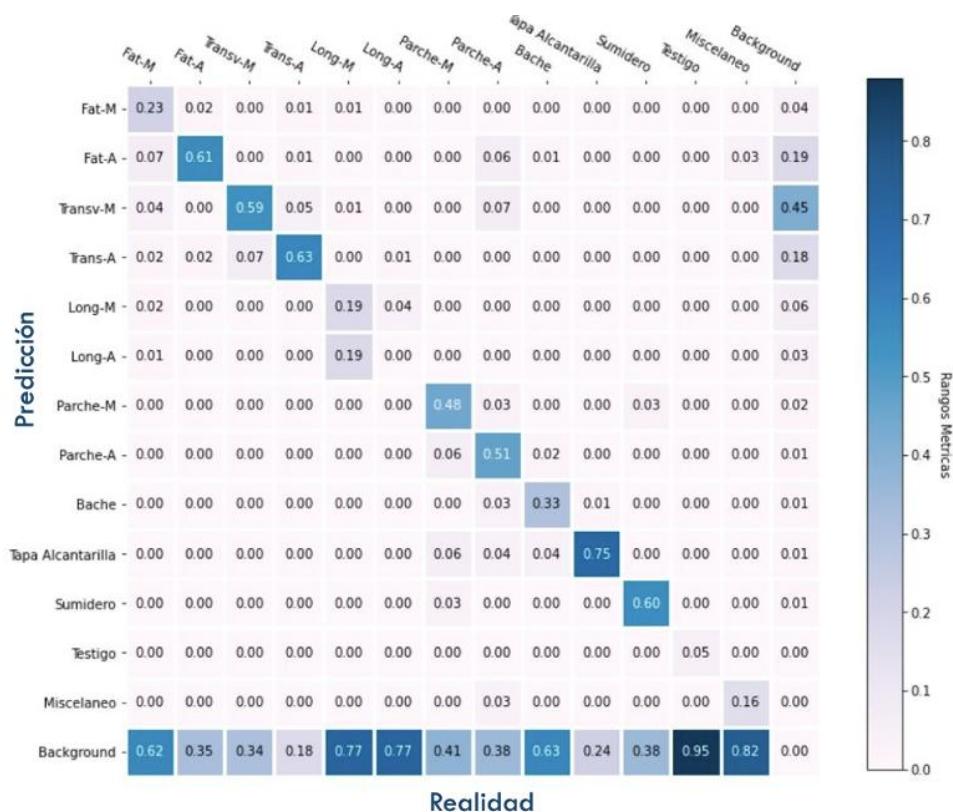
### 5.4.1.1 Entrenamiento con Etiquetas Grietas Longitudinales y Transversales

Los resultados de desempeño obtenidos por cada entrenamiento para el set de prueba se presentan en la Tabla 5.11 con un iou (intersection over unión) de 0,6.

**Tabla 5.11: Métricas de desempeño al evaluar el conjunto test para cada split  
entrenado con YOLOv5**

Parámetro	Desempeño por Split (%)				
	1	2	3	4	5
Precisión	41,4	47,5	42,6	61,3	66,5
Recall	43,3	40,3	44,2	52,6	57,8
F1	42,3	43,6	43,4	56,6	61,8
mAP50	37,4	36,8	37,0	55,3	62,1
mAP50-95	18,9	18,4	19,1	30,7	35,7

En la Figura 5.2 se muestra el promedio de resultados de las matrices de confusión correspondiente a los 5 splits. En el eje horizontal se muestran los deterioros realmente etiquetados y en el eje vertical los deterioros de la predicción de la red. Cada cuadrado representa el porcentaje de predicciones del deterioro presente en el eje vertical y que realmente era el deterioro del eje horizontal. Lo ideal es que el deterioro de la predicción coincida con el deterioro etiquetado, por lo que se busca que la diagonal sea 1.

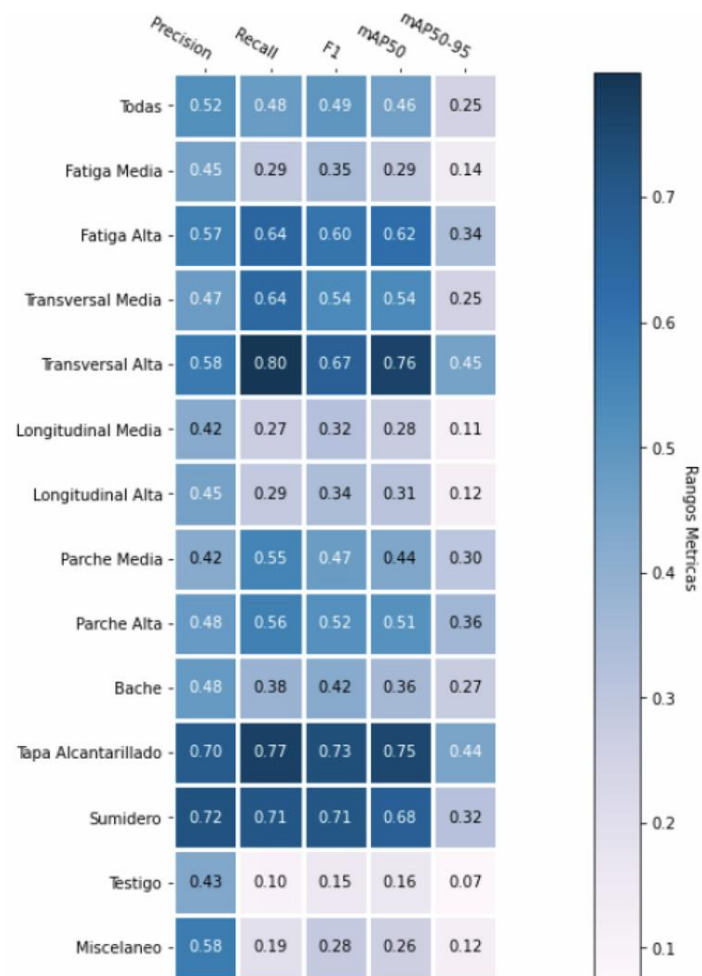


**Figura 5.2: Promedio de las matrices de confusión obtenidas de los 5 splits para YOLOv5.**

Las matrices de confusión de cada split obtenidas con YOLOv5 se encuentran en el 9.4 Anexo 4. Resultados Entrenamiento.

Por otro lado, para obtener más información respecto al desempeño de la red por cada tipo de etiqueta, se registraron los resultados de precisión, recall, F1, mAP50 y mAP50-95 al evaluar el conjunto test para los 5 splits y para cada tipo de YOLO. Los resultados se presentan en el Anexo 9.4.2 Resultados por Tipo de Deterioro.

La Figura 5.3 representa las métricas promediadas de los 5 splits al evaluar el conjunto test para YOLOv5.



**Figura 5.3: Resultados por deterioro al evaluar el conjunto test y promediar las métricas de los 5 splits en YOLOv5**

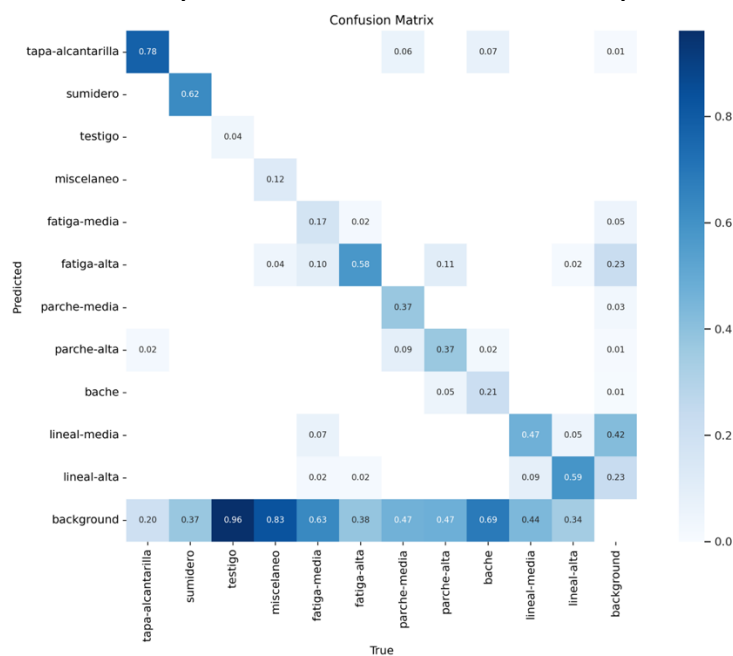
### 5.4.1.2 Entrenamiento con Etiquetas Grietas Lineales

En el caso del entrenamiento que se llevó a cabo usando “grietas lineales” en vez de “grietas longitudinales” y “grietas transversales” por separado, para el split 1, se presentan los resultados en la Tabla 5.12.

**Tabla 5.12: Métricas de desempeño al evaluar el conjunto test para el split 1 entrenado con YOLOv5 y grietas lineales unidas**

Parámetro	Desempeño por Split (%)
	1
Precisión	40,0
Recall	43,1
F1	41,4
mAP50	39,3
mAP50-95	21,2

La matriz de confusión correspondiente a este entrenamiento se presenta en la Figura 5.4.



**Figura 5.4: Matriz de confusión del split 1 entrenado con YOLOv5 usando grietas lineales unidas.**

## 5.4.2 YOLOv7

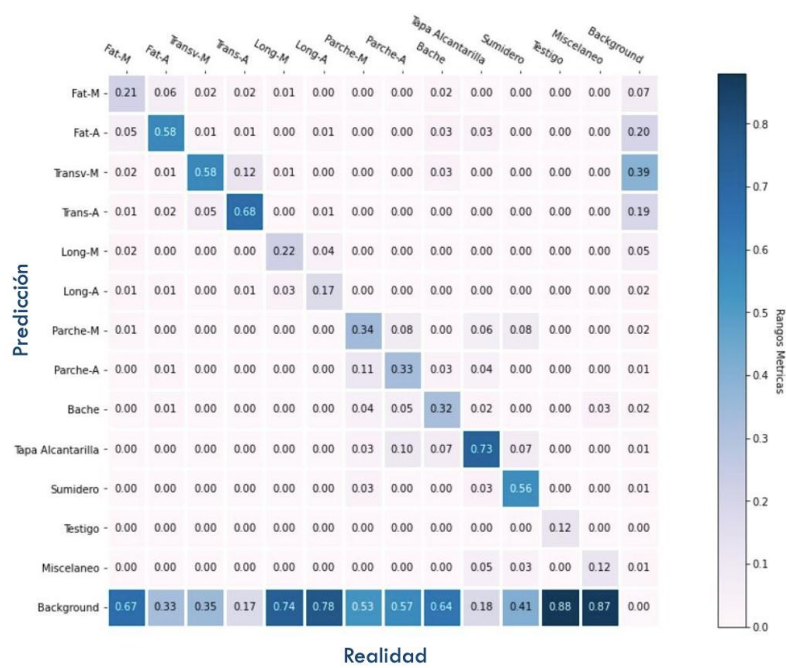
### 5.4.2.1 Entrenamiento con Etiquetas Grietas Longitudinales y Transversales

Los resultados de desempeño obtenidos por cada entrenamiento para el set de prueba se presentan en la Tabla 5.13.

**Tabla 5.13: Métricas de desempeño al evaluar el conjunto test para cada split entrenado con YOLOv7**

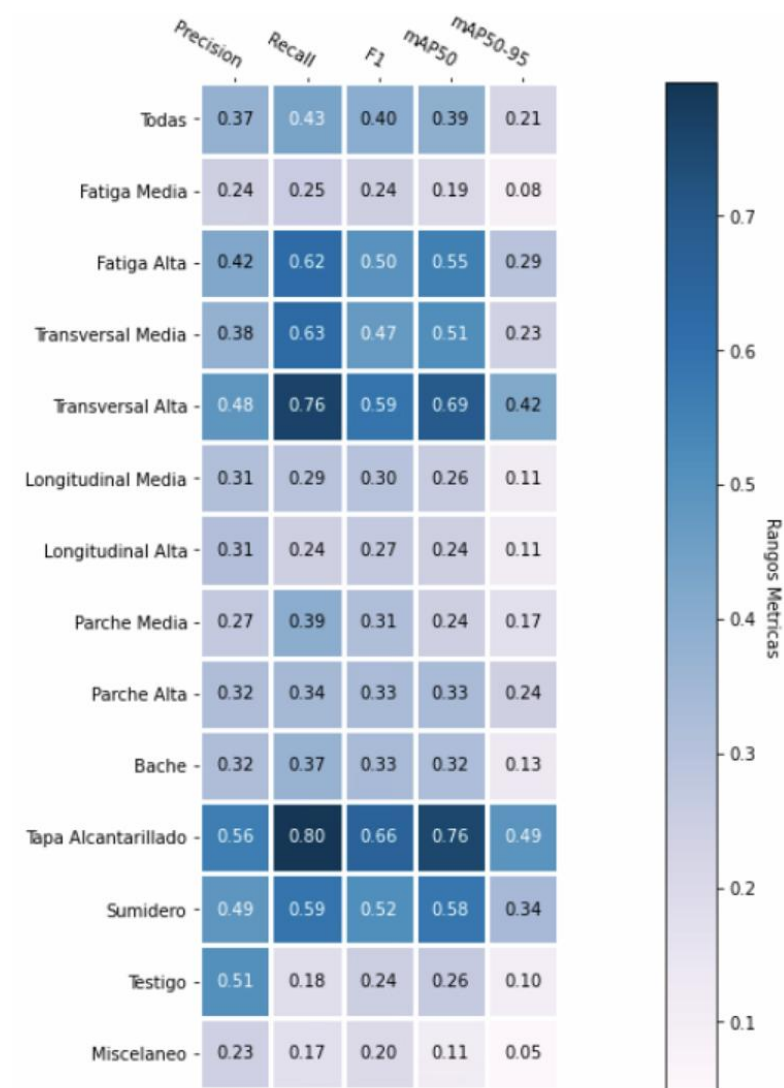
Parámetro	Desempeño por Split				
	1	2	3	4	5
Precisión	52,0	38,5	42,4	44,4	42,4
Recall	36,4	42,6	44,2	48,7	44,3
F1	42,8	40,4	43,3	46,5	43,3
mAP50	37,4	36,8	37,8	42,8	39,2
mAP50-95	21,2	20,1	19,6	23,8	25,0

En la Figura 5.5 se muestra el promedio de resultados de las matrices de confusión correspondiente a los 5 splits. Lo ideal en este tipo de matrices, es que la diagonal sea 1.



**Figura 5.5: Promedio de las matrices de confusión obtenidas de los 5 splits para YOLOv7.**

La Figura 5.6 representa las métricas promediadas de los 5 splits al evaluar el conjunto test para YOLOv7.



*Figura 5.6: Resultados por deterioro al evaluar el conjunto test y promediar las métricas de los 5 splits en YOLOv7*

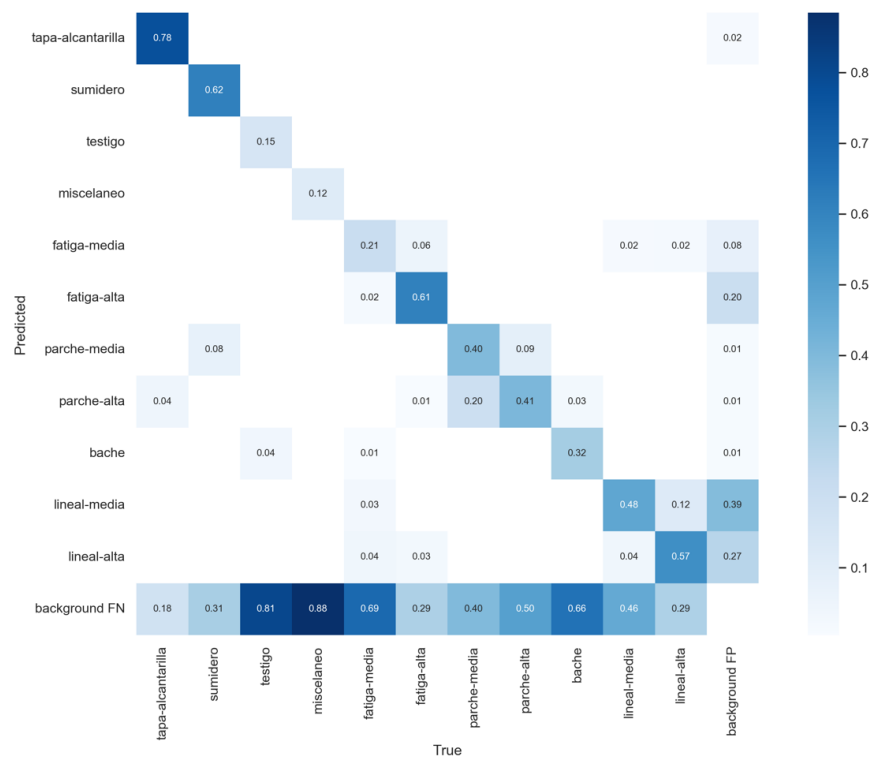
### 5.4.2.2 Entrenamiento con Etiquetas Grietas Lineales

En el caso del entrenamiento que se llevó a cabo usando “grietas lineales” en vez de “grietas longitudinales” y “grietas transversales” por separado, para el split 1, se presentan los resultados en la Tabla 5.14.

**Tabla 5.14: Métricas de desempeño al evaluar el conjunto test para el split 1 entrenado con YOLOv7 y grietas lineales unidas**

Parámetro	Desempeño por Split (%)
	1
Precisión	45,6
Recall	45,4
F1	45,5
mAP50	42,1
mAP50-95	23,8

La matriz de confusión correspondiente a este entrenamiento se presenta en la Figura 5.7.



**Figura 5.7: Matriz de confusión del split 1 entrenado con YOLOv7 usando grietas lineales unidas.**

## 5.5 Inferencia Red Neuronal

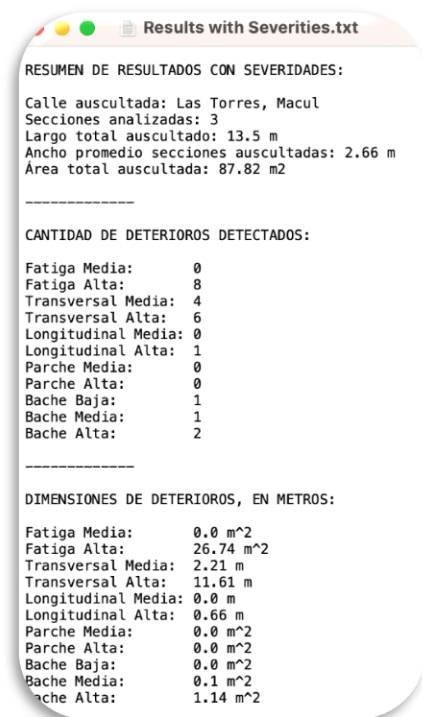
Se conoce como inferencia, cuando se le entrega una imagen o video nueva a la red neuronal, sin etiquetas, para que éste detecte lo que fue entrenado para detectar.

Dado que el objetivo es determinar la condición del pavimento, se trabajó un código capaz de procesar el output de la inferencia de la red y crear un archivo “.txt” con la información necesaria.

Los datos que entrega este documento son los siguientes:

- Calle analizada
- Secciones auscultadas (considerando cada sección de 13,55 m)
- Ancho promedio secciones auscultadas
- Área total auscultada
- Cantidad de deterioros detectados
- Dimensiones de los deterioros detectados, en metros

Un ejemplo de la visualización del documento se presenta en la Figura 5.8.



```

Results with Severities.txt
RESUMEN DE RESULTADOS CON SEVERIDADES:
Calle auscultada: Las Torres, Macul
Secciones analizadas: 3
Largo total auscultado: 13.5 m
Ancho promedio secciones auscultadas: 2.66 m
Area total auscultada: 87.82 m2

-----

CANTIDAD DE DETERIOROS DETECTADOS:
Fatiga Media:      0
Fatiga Alta:      8
Transversal Media: 4
Transversal Alta: 6
Longitudinal Media: 0
Longitudinal Alta: 1
Parche Media:     0
Parche Alta:     0
Bache Baja:      1
Bache Media:     1
Bache Alta:      2

-----

DIMENSIONES DE DETERIOROS, EN METROS:
Fatiga Media:     0.0 m^2
Fatiga Alta:     26.74 m^2
Transversal Media: 2.21 m
Transversal Alta: 11.61 m
Longitudinal Media: 0.0 m
Longitudinal Alta: 0.66 m
Parche Media:    0.0 m^2
Parche Alta:    0.0 m^2
Bache Baja:     0.0 m^2
Bache Media:    0.1 m^2
Bache Alta:     1.14 m^2

```

**Figura 5.8: Ejemplo de output final de resultados**

Por otro lado, se entrega una tabla, como se muestra en la Figura 5.9, exportable a Excel con los siguientes datos:

- xmin, ymin, xmax, ymax: Representan las coordenadas x mínimo y máximo e y mínimo y máximo de los bounding box detectados, en pixeles.

- Confidence: Representa la confianza de que el bounding box y el etiquetado sea correcto.
- Name: Nombre del deterioro detectado.
- Filename: Nombre del archivo donde se encontró ese bounding box.
- Width\_cm, height\_cm, area\_cm2: Representa el ancho, alto y área, respectivamente del bounding box detectado, en centímetros.

Unnamed: 0	xmin	ymin	xmax	ymax	confidence	class	name	filename	width_cm	height_cm	area_cm2
0	4.056485	1910.371216	1080.636719	4683.182617	0.891701	1	fatiga-alta	127 SECCION165	155.027554	399.284842	61900.152255
1	16.027870	6232.455078	974.187439	9133.340820	0.863439	1	fatiga-alta	127 SECCION165	137.974978	417.727547	57635.949051
2	1035.684692	8074.317871	1353.493164	8212.361328	0.684973	2	transversal-media	127 SECCION165	45.764420	19.878258	909.716938
3	1057.879272	6429.261230	2013.848755	6647.793945	0.628941	3	transversal-alta	127 SECCION165	137.659605	31.468711	4331.970332
4	999.670715	7083.458984	2058.482666	7391.333984	0.608733	3	transversal-alta	127 SECCION165	152.468921	44.334000	6759.557139

***Figura 5.9: Ejemplo de tabla de output de la inferencia***

## 6 *Análisis de Resultados*

En este capítulo se presenta el análisis y evaluación de los resultados a partir de los resultados directos presentados en el capítulo anterior.

### 6.1 Análisis Comparativo entre YOLOv5 y YOLOv7

Para comparar la diferencia entre los resultados obtenidos por YOLOv5 y YOLOv7, se realizan pruebas de comparación de medias. Para esto, se decide utilizar la prueba de t-Student, que posee las siguientes restricciones:

- Datos deben presentar normalidad
- Se realizan pruebas distintas en el caso que los datos presenten o no homocedasticidad (diferencia entre sus varianzas)

Primero, para verificar la normalidad de los datos, se lleva a cabo la prueba de Jarque-Bera. Esta es una prueba de bondad de ajuste que determina si los datos de la muestra tienen asimetría y curtosis que coinciden con una distribución normal (Statologos, 2021).

Las hipótesis de esta prueba son:

- $H_0$ : Los datos se distribuyen normalmente.
- $H_A$ : Los datos no se distribuyen normalmente.

La estadística de prueba JB se define como:

$$JB = \left(\frac{n}{6}\right) \cdot \left(S^2 + \left(\frac{C^2}{4}\right)\right)$$

Dónde:

- n: número de observaciones de la muestra
- S: asimetría de la muestra
- C: curtosis de muestra

Bajo la hipótesis nula de normalidad,  $JB \sim X^2(2)$ .

Si el valor p, que corresponde al estadístico de prueba es menor que algún nivel de significancia (en este caso,  $\alpha=0,05$ ), entonces se rechaza la hipótesis nula y se concluye que los datos no se distribuyen normalmente.

Bajo la hipótesis nula de normalidad, el estadístico de prueba JB sigue una distribución Chi-Cuadrado con 2 grados de libertad.

Dado que se desea hacer la prueba de comparación de medias entre todas las métricas obtenidas, se le realiza la prueba de normalidad a todos los datos. Los resultados obtenidos al realizar la prueba se muestran en las tablas Tabla 6.1 y Tabla 6.2 para YOLOv5 y YOLOv7, respectivamente.

**Tabla 6.1: Resultados de prueba de Jarque-Bera para métricas obtenidas de YOLOv5**

Dato	Precisión	Recall	F1	mAP50	mAP50-95
Observaciones	5	5	5	5	5
Sample Skewness	0.554	0.703	0.782	0.784	0.816
Sample Kurtosis	-2.509	-1.500	-2.240	-2.311	-2.103
JB test statistic	1.567	0.881	1.554	1.624	1.477
Valor-P	0.457	0.644	0.460	0.444	0.478

**Tabla 6.2: Resultados de prueba de Jarque-Bera para métricas obtenidas de YOLOv7**

Dato	Precisión	Recall	F1	mAP50	mAP50-95
Observaciones	5	5	5	5	5
Sample Skewness	1.193	-0.737	0.417	1.569	0.497
Sample Kurtosis	2.315	1.857	1.928	2.359	-2.279
JB test statistic	2.304	1.171	0.919	3.210	1.288
Valor-P	0.316	0.557	0.632	0.201	0.525

Dado que los valores-P no son menores que 0,05, no se rechaza la hipótesis nula, por lo que no hay evidencia suficiente para decir que el conjunto de datos no se distribuye normalmente.

Ahora, para evaluar la homocedasticidad de los datos, se usa la Prueba F de Fisher para comparar dos varianzas. Las hipótesis de esta prueba son:

- $H_0$ : Homocedasticidad en ambas muestras.
- $H_A$ : Heterocedasticidad en ambas muestras.

Para esta prueba también se asumió un valor de confianza del 95%. Y para que no se rechace la homocedasticidad, el valor de probabilidad asociado al contraste de homogeneidad de varianzas debe ser mayor a 0,05.

Los resultados se presentan en la Tabla 6.3 para las métricas de precisión, recall y F1 y en la Tabla 6.4 para mAP50 y mAP50-95.

**Tabla 6.3: Resultados de la prueba  $F$  de Fisher para las métricas de precisión, recall y  $F1$**

Valor	Precisión		Recall		F1	
	YOLOv5	YOLOv7	YOLOv5	YOLOv7	YOLOv5	YOLOv7
Promedio	0.519	0.439	0.476	0.432	0.496	0.433
Varianza	0.013	0.002	0.005	0.002	0.008	0.000
Observaciones	5	5	5	5	5	5
Df	4	4	4	4	4	4
F	5.202		2.686		17.856	
P( $F \leq f$ ) una cola	0.070		0.181		0.008	
F Crítico una cola	6.388		6.388		6.388	

**Tabla 6.4: Resultados de la prueba  $F$  de Fisher para las métricas de  $mAP50$  y  $mAP50-95$**

Valor	mAP50		mAP50-95	
	YOLOv5	YOLOv7	YOLOv5	YOLOv7
Promedio	0.457	0.388	0.246	0.219
Varianza	0.015	0.001	0.006	0.001
Observaciones	5	5	5	5
Df	4	4	4	4
F	25.299		11.766	
P( $F \leq f$ ) una cola	0.004		0.017	
F Crítico una cola	6.388		6.388	

Al analizar los resultados, se puede notar que el valor de  $P$  es mayor solamente para las métricas de precisión y recall, no así para  $F1$ ,  $mAP50$  y  $mAP50-95$ . Es por esta razón, que para las métricas de precisión y recall se les hará la prueba t-Student asumiendo varianzas iguales, y para las métricas de  $F1$ ,  $mAP50$  y  $mAP50-95$ , se les hará la prueba t-Student asumiendo varianzas distintas.

Para la prueba t-Student se tienen las siguientes hipótesis de prueba:

- $H_0$ : La métrica evaluada entre YOLOv5 y YOLOv7 son iguales.
- $H_A$ : La métrica evaluada entre YOLOv5 y YOLOv7 son distintas.

Los resultados se presentan en la Tabla 6.5 para las métricas de precisión, recall y F1 y en la Tabla 6.6 para mAP50 y mAP50-95.

**Tabla 6.5: Prueba t-Student para muestras asumiendo varianzas iguales para las métricas de precisión y recall**

Valor	Precisión		Recall	
	YOLOv5	YOLOv7	YOLOv5	YOLOv7
Promedio	0.5186	0.4394	0.4764	0.4324
Varianza	0.013	0.002	0.005	0.002
Observaciones	5	5	5	5
Varianza agrupada	0.008		0.004	
Diferencia media hipotética	0		0	
Df	8		8	
t Estadístico	1.426		1.153	
P(T<=t) dos colas	0.192		0.282	
t Crítico dos colas	2.306		2.306	

**Tabla 6.6: Prueba t-Student para muestras asumiendo varianzas distintas para las métricas de F1, mAP.50 y mAP.50-95**

Valor	F1		mAP.50		mAP.50-95	
	YOLOv5	YOLOv7	YOLOv5	YOLOv7	YOLOv5	YOLOv7
Promedio	0.496	0.433	0.457	0.388	0.246	0.219
Varianza	0.008	0.000	0.015	0.001	0.007	0.001
Observaciones	5	5	5	5	5	5
Diferencia media hipotética	0		0		0	
Df	4		4		5	
t Estadístico	1.515		1.255		0.695	
P(T<=t) dos colas	0.204		0.278		0.518	
t Crítico dos colas	2.777		2.776		2.57	

Dado que en los 5 casos el valor de t estadístico se encuentra dentro del rango [-t crítico, t crítico], se acepta la hipótesis nula que establece que no existe una diferencia significativa entre las métricas de YOLOv5 y YOLOv7.

Por otro lado, se realiza el cálculo del error cuadrático medio (RMSE) para determinar la diferencia entre las métricas obtenidas entre YOLOv5 y YOLOv7, obteniendo lo expuesto en la Tabla 6.7.

**Tabla 6.7: Error cuadrático medio de las distintas métricas al comparar resultados de YOLOv5 y YOLOv7**

Dato	RMSE
Precisión	0.146
Recall	0.071
F1	0.096
mAP50	0.117
mAP50-95	0.058

## 6.2 Análisis de Desempeño por Deterioros

En base a los resultados obtenidos por cada deterioro para YOLOv5 y YOLOv7, de las Figura 5.3 y Figura 5.6, respectivamente, se generó la Tabla 6.8, que presenta el resumen con las singularidades con mejor y peor desempeño, en base al parámetro F1, seleccionado porque considera precisión y recall.

**Tabla 6.8: Resumen de resultados en base a F1**

Parámetro (F1)	YOLOv5	YOLOv7
Etiqueta con mayor F1	Tapa alcantarillado (0,73)	Tapa alcantarillado (0,66)
Deterioro con mayor F1	Transversal alta (0,67)	Transversal alta (0,59)
Etiqueta con menor F1	Testigo (0,15)	Misceláneo (0,20)
Deterioro con menor F1	Longitudinal media (0,32)	Fatiga media (0,24)

Al analizar el desempeño por deterioros obtenido con YOLOv5, con la ayuda de la Figura 5.3, se detecta que la tapa de alcantarillado es la etiqueta que posee mejores resultados. En cuanto al deterioro con mejor resultado, corresponde a la grieta transversal de severidad alta. Por otro lado, la detección y clasificación de extracción de testigos fue la que obtuvo peores resultados, y en cuanto a deterioro, fue la grieta longitudinal de severidad baja. El promedio de precisión del desempeño de todos los deterioros fue de 48% con una desviación de solo 6%.

Por otro lado, para YOLOv7, se repiten los mismos resultados, salvo que la singularidad peor evaluada fue el “misceláneo” y el deterioro con peor detección y clasificación fueron las grietas de fatiga de severidad baja.

El promedio de precisión fue de 34% con una desviación de solo un 8%.

El deterioro de fatiga severidad media obtuvo un muy mal recall, es decir, la red reconoce pocos deterioros existentes, con un 35% de diferencia respecto al desempeño de detección de la fatiga severidad alta.

Las grietas longitudinales obtuvieron un mal recall para ambas severidades.

En general, los deterioros de severidad alta obtuvieron un mejor desempeño, lo que es esperable porque presentan características con mayor distinción y reconocimiento.

Por otro lado, al analizar las matrices de confusión, en particular las Figura 5.2 y Figura 5.5, se puede notar que no hay mucha diferencia entre ambas matrices, por lo que los desempeños de YOLOv5 y YOLOv7 fueron similares. En cuanto a los deterioros, se vuelve a cumplir que las grietas de severidad alta tienen mejor desempeño que las grietas de severidad media, a excepción de las grietas longitudinales.

Además, llama la atención que la red encontró grietas transversales de severidad media donde originalmente no había ninguna etiqueta.

En cuanto a las singularidades, las etiquetas de testigo y misceláneo poseen los peores resultados, lo que no afecta en la investigación dado que el objetivo es ser capaz de detectar deterioros.

### 6.3 Análisis de Resultados por Cada Split

Con el fin de analizar el nivel de influencia de selección de imágenes para cada conjunto de datos, se obtienen las medidas de dispersión para cada métrica por cada red neuronal, presentadas en las Tabla 6.9 y Tabla 6.10.

**Tabla 6.9: Resultados de medidas de dispersión para métricas obtenidas con YOLOv5**

Dato	Precisión	Recall	F1	mAP50	mAP50-95
Promedio	0.519	0.476	0.496	0.457	0.246
Valor mínimo	0.414	0.403	0.423	0.368	0.184
Valor máximo	0.665	0.578	0.618	0.621	0.357
Rango	0.251	0.175	0.195	0.253	0.173
Varianza	0.013	0.005	0.008	0.015	0.007
Desviación Estándar	0.114	0.073	0.090	0.121	0.081

**Tabla 6.10: Resultados de medidas de dispersión para métricas obtenidas con YOLOv7**

Dato	Precisión	Recall	F1	mAP50	mAP50-95
Promedio	0.439	0.432	0.433	0.388	0.219
Valor mínimo	0.385	0.364	0.404	0.368	0.196
Valor máximo	0.520	0.487	0.465	0.428	0.250
Rango	0.135	0.123	0.060	0.060	0.054
Varianza	0.002	0.002	0.000	0.001	0.001
Desviación Estándar	0.050	0.044	0.021	0.024	0.024

En las tablas, se puede analizar que la desviación estándar de los resultados para cada split va entre 0.021 y 0.121. Por otro lado, YOLOv7 presenta el valor mínimo y YOLOv5 el valor máximo. De igual forma, la variabilidad es más alta para todas las métricas de YOLOv5 que aquellas obtenidas en YOLOv7.

En base a esto, se recomienda el uso de YOLOv7, dado que es menos variable a la distribución de imágenes en entrenamiento, validación y prueba, lo que lo hace más confiable. Por otro lado, el resultado de YOLOv5 se ve más influenciado en esta distribución.

#### **6.4 Análisis de Resultados Comparación Grietas Longitudinales + Transversales vs Grietas Lineales**

Para llevar a cabo la comparación entre el entrenamiento que consideró solo grietas lineales con aquel que consideró grietas longitudinales + transversales, se calculó el desempeño ponderado del primero, que considera la cantidad de etiquetas por cada tipo de grieta (longitudinal y transversal) y su ponderación. Esto con el fin de tener claridad si el desempeño del entrenamiento llevado a cabo con “grietas lineales” realmente es mejor, o si solo es influenciado por la cantidad de etiquetas.

Para llevar a cabo esto, se usó la siguiente ecuación:

$$R_{lineal} = \frac{N_T \cdot R_T + N_L \cdot R_L}{N_T + N_L}$$

Donde el significado de cada elemento es:

- $R_T$ : Porcentaje de la cantidad de grietas lineales ponderadas detectadas
- $N_T$ : Cantidad de etiquetas de grietas transversales
- $R_L$ : Porcentaje de la cantidad de grietas transversales detectadas en el entrenamiento considerando “grietas transversales” y “grietas longitudinales” por separado
- $N_L$ : Cantidad de etiquetas de grietas longitudinales

- $R_L$ : Porcentaje de la cantidad de grietas longitudinales detectadas en el entrenamiento considerando “grietas transversales” y “grietas longitudinales” por separado

Se analizan los resultados por cada severidad:

#### 6.4.1 YOLOv5

Primero se presentan los resultados obtenidos por la versión 5 de YOLO.

##### 6.4.1.1 Grietas con Severidad Media

***Tabla 6.11: Comparación de grietas de severidad media detectadas usando el ponderado del entrenamiento con etiquetas “grietas longitudinales” y etiquetas “grietas transversales” vs el entrenamiento con “grietas lineales” con YOLOv5***

Valor	Severidad media	
	Transversal	Longitudinal
Cantidad Etiquetas	5,568	1,762
Porcentaje Cantidad Grietas Detectadas	58	20
Porcentaje Cantidad Grietas Detectadas Ponderadas		49
Porcentaje Cantidad Grietas Detectadas Obtenidas		47

##### 6.4.1.2 Grietas con Severidad Alta

***Tabla 6.12: Comparación de grietas de severidad alta detectadas usando el ponderado del entrenamiento con etiquetas “grietas longitudinales” y etiquetas “grietas transversales” vs el entrenamiento con “grietas lineales” con YOLOv5***

Valor	Severidad media	
	Transversal	Longitudinal
Cantidad Etiquetas	3,791	1,297
Porcentaje Cantidad Grietas Detectadas	70	22
Porcentaje Cantidad Grietas Detectadas Ponderadas		58
Porcentaje Cantidad Grietas Detectadas Obtenidas		59

## 6.4.2 YOLOv7

A continuación se presentan los resultados obtenidos por la versión 7 de YOLO.

### 6.4.2.1 Grietas con Severidad Media

**Tabla 6.13: Comparación de grietas de severidad media detectadas usando el ponderado del entrenamiento con etiquetas “grietas longitudinales” y etiquetas “grietas transversales” vs el entrenamiento con “grietas lineales” con YOLOv7**

Valor	Severidad media	
	Transversal	Longitudinal
Cantidad Etiquetas	5,568	1,762
Porcentaje Cantidad Grietas Detectadas	56	22
Porcentaje Cantidad Grietas Detectadas Ponderadas		48
Porcentaje Cantidad Grietas Detectadas Obtenidas		48

### 6.4.2.2 Grietas con Severidad Alta

**Tabla 6.14: Comparación de grietas de severidad alta detectadas usando el ponderado del entrenamiento con etiquetas “grietas longitudinales” y etiquetas “grietas transversales” vs el entrenamiento con “grietas lineales” con YOLOv7**

Valor	Severidad media	
	Transversal	Longitudinal
Cantidad Etiquetas	3,791	1,297
Porcentaje Cantidad Grietas Detectadas	67	20
Porcentaje Cantidad Grietas Detectadas Ponderadas		55
Porcentaje Cantidad Grietas Detectadas Obtenidas		57

Como se puede observar, en ningún caso los resultados del entrenamiento que consideró “grietas lineales” en vez de “grietas longitudinales” y “grietas transversales” obtuvo un mucho mejor desempeño, sino que se detectaron la misma cantidad de grietas. Por lo tanto, considerando que cada entrenamiento toma unos 5 días en llevarse a cabo, se descarta continuar indagando esta opción.

## 6.5 Análisis Cuantitativo

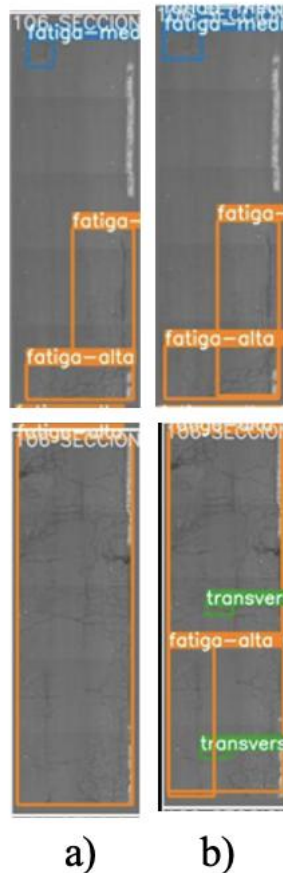
Al estudiar las imágenes etiquetadas y de predicción al evaluar los conjuntos test, se obtienen 2 temas a analizar: etiquetas solapadas y confusión entre deterioros. Estos puntos se comentan a continuación:

### 6.5.1 Solape entre Etiquetas

Las arquitecturas de redes neuronales artificiales YOLO, son entrenadas con imágenes en que se considera que objetos pueden estar a distinta distancia de la cámara, por lo que considera que un objeto puede estar detrás del otro.

Distinto ocurre con las imágenes de esta investigación, en que toda la zona analizada se encuentra plana a una misma profundidad, y cada zona del pavimento agrietada, se puede clasificar solamente como un tipo de deterioro, y no distintos.

Un par de ejemplos de la situación planteada se pueden ver en la Figura 6.1.



**Figura 6.1: Ejemplos de etiquetas solapadas. a) representa la etiqueta manual asignada y b) la predicción de la red**

Una solución, planteada a eso, proviene del creador de la arquitectura, quien plantea que con el comando “-agnostic-nms” al usar la inferencia de la red, se evita el solape entre etiquetas. Dado que es un comando que solo se puede aplicar en la inferencia, no se puede usar para mejorar las métricas de rendimiento de la red neuronal al evaluar el conjunto test.

Por otro lado, sería interesante poder decidir qué deterioro predomina sobre el otro, lo que se propone es decidir en base a los valores de significancia provenientes del cálculo del ICPU (Osorio, 2015) para pavimentos asfálticos. Estos valores se presentan en la Tabla 6.15, que indica los valores de significancia por cada deterioro.

**Tabla 6.15: Valores de significancia al calcular el estado del pavimento asfáltico según el ICPU, a mayor valor, mayor afecta en el cálculo del estado. (Nº) indica el orden ascendente del peso del deterioro en el índice, a mayor número más afecta la nota del deterioro**

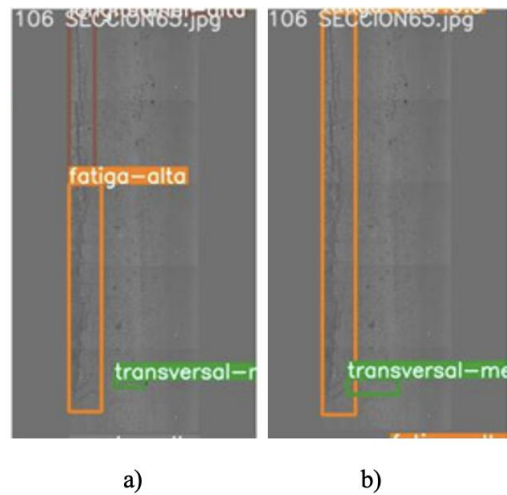
Deterioro	Severidad Baja/media	Severidad Alta
Fatiga	0,038 (1)	0,076 (5)
Transversal + Reflexión	0,049 (3)	0,098 (7)
Baches	0,237 (4)	0,474 (8)
Parches	0,046 (2)	0,092 (6)

### 6.5.2 Confusión entre Deterioros

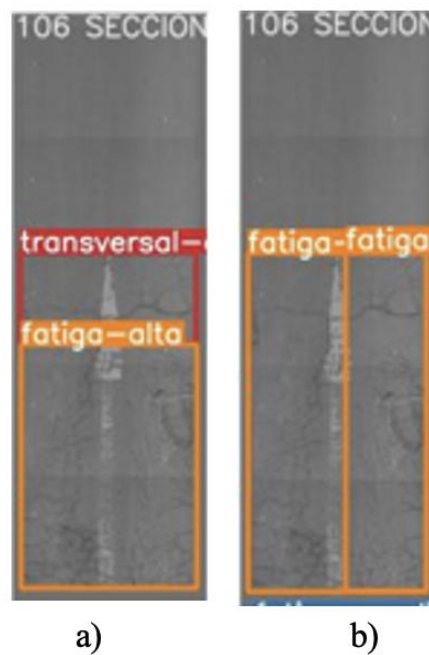
La red neuronal YOLO fue entrenada con un dataset denominado COCO (objetos comunes en contexto), los cuales se encuentran claramente definidos para el ojo humano, por ejemplo, todos los humanos estamos de acuerdo en cómo diferenciar un auto de un pájaro o de un avión. Esta situación no ocurre de igual forma para los deterioros en pavimentos, y existe un grado de subjetividad, aun cuando se busca evitarlo.

Errores comunes de confusión de deterioros se pueden dar, por ejemplo, al diferenciar una grieta longitudinal de una grieta de fatiga, dado que sigue la huella del vehículo, tal como se puede ver en la Figura 6.2, en que el etiquetado manual detalla una mitad como grieta fatiga alta y la otra mitad como grieta longitudinal; a diferencia de lo predicho por la red que dice que todo ese deterioro corresponde a grieta de fatiga.

Otro ejemplo se puede ver en la Figura 6.3, en donde la etiqueta manual indica una grieta de fatiga y una grieta transversal, y la etiqueta de la red neuronal indica dos zonas grandes con grieta de fatiga, lo que también cobra sentido ya que ambas etiquetas siguen la huella del vehículo, lo que calza con la definición de dicho deterioro.



**Figura 6.2:** *Imágenes que demuestran la confusión en la clasificación de deterioros. a) representa la etiqueta manual asignada y b) la predicción de la red*



**Figura 6.3:** *Imágenes que demuestran la confusión en la clasificación de deterioros. a) representa la etiqueta manual asignada y b) la predicción de la red*

Los expuestos acá son pequeñas muestras de lo que ocurre con la predicción de la red, sin embargo, son cosas que se repiten a lo largo de las imágenes.

## 7 Conclusiones y Recomendaciones

A continuación se presentan las conclusiones, recomendaciones e investigación futura recomendados para continuar este trabajo.

### 7.1 Conclusiones

Como conclusión, se logró optimizar el tiempo y procesamiento de la obtención de datos para su posterior automatización de detección de deterioros para pavimentos urbanos.

En base a esto, se confeccionaron:

1. Instructivo con el paso a paso para poder recopilar videos en terreno
2. Notebook Python para llevar a cabo la transformación de videos a imágenes
3. Manual para realizar el etiquetado de deterioros
4. Videos explicativos de cómo entrenar y evaluar las redes YOLOv5 y YOLOv7

En cuanto a los resultados del entrenamiento con las distintas versiones de YOLO (v5 y v7), se concluye que en gran medida se logró un buen desempeño, dado que, considerando que se trabaja con 13 clases distintas, se supera con creces la clasificación aleatoria (7,7%). Claramente, las matrices de confusión muestran que la red tiene algunas dificultades para detectar los deterioros, ya que a menudo los confunde con el background o pavimento en buen estado. Por otro lado, la red neuronal es capaz de detectar con mayor precisión ciertos elementos como las tapas de alcantarillado, sumideros y grietas transversales de severidad alta. La razón detrás de esto puede ser que YOLO, el algoritmo utilizado para la detección de objetos, fue entrenado para identificar objetos que se presentan en dos dimensiones con características distintivas. Esto es diferente a las grietas longitudinales, que tienen una dimensión mucho mayor que la otra.

En cuanto al RMSE, se obtienen valores pequeños, confirmando que no se presenta una diferencia significativa en el desempeño de las versiones 5 y 7 de YOLO.

Al analizar las imágenes etiquetadas del conjunto test, se concluye que a pesar de que la red neuronal detecta las zonas deterioradas de manera adecuada, su clasificación puede resultar confusa tanto para la red como para los humanos.

Por otro lado, la red neuronal presenta métricas para cada split lo suficientemente similares, por lo que se concluye es lo suficientemente estable como para que su desempeño no se vea afectado por la distribución de las imágenes en el conjunto de entrenamiento, validación y prueba.

Es importante destacar que la red neuronal tiene un mejor desempeño en la detección de deterioros con severidad alta en comparación con aquellos de severidad baja.

El modelo YOLOv5, para los distintos splits, entregó mejores resultados que para YOLOv7, por lo que la influencia de la selección de imágenes en los conjuntos de entrenamiento, validación y prueba es mayor que para YOLOv7, que entregó resultados con menor variabilidad para sus distintos splits.

Por otro lado, llevar a cabo el entrenamiento usando “grietas lineales” en vez de “grietas longitudinales” y “grietas transversales” por separado, no significó una mejora en los resultados, por lo que se descartó continuar entrenamientos con esta metodología.

En resumen, aunque la red neuronal muestre buenos resultados en la detección de zonas deterioradas en la imagen, existen limitaciones y desafíos en su desempeño.

Finalmente, se concluye que se aplicaron y validaron algoritmos de Deep Learning, particularmente las versiones 5 y 7 de YOLO para automatizar la detección de deterioros en pavimentos urbanos provenientes de imágenes de bajo costo obtenidas en terreno como parte de la investigación. Con esta información, se logra automatizar un paso a paso para determinar la condición de los pavimentos urbanos a partir de la detección y medición de los deterioros.

## **7.2 Recomendaciones**

Como primera recomendación, se sugiere obtener más imágenes de distintos pavimentos asfálticos urbanos, especialmente aquellos que presenten los deterioros de menor cantidad (como grietas longitudinales, baches y parches), las que se deben usar para continuar el entrenamiento de los modelos.

De la mano con la recomendación anterior, se sugiere determinar la forma para poder aumentar el tiempo de grabación, ya que al ser 1 hora antes de la puesta de sol, es bastante limitado para cada día. Para esto, se sugiere el uso de filtros que se compran como accesorios a la cámara GoPro y también estudiar las distintas configuraciones propias que provee la cámara (balance de blancos, brillo, etc).

Por otro lado, se sugiere también buscar una forma para aumentar la velocidad de grabación, ya que al ser 20 km/h, es muy bajo en comparación a la velocidad normal de circulación de vehículos, por lo que se vuelve peligroso y en algunas calles por lo mismo no se puede llevar a cabo el proceso de grabación. Para esto, se sugiere evaluar las grabaciones con un mayor número de FPS, ya que así se obtienen más marcos de imágenes por segundo, disminuyendo las escenas borrosas.

También se recomienda entrenar más splits de imágenes divididas aleatoriamente en entrenamiento, validación y prueba con el fin de robustecer los resultados obtenidos.

Al momento de llevar a cabo la inferencia con los pesos de las redes, se recomienda usar el comando “-agnostic-nms” para evitar el solapamiento entre etiquetas.

Por otro lado, se sugiere probar otras redes neuronales artificiales sugeridas en la literatura, como, por ejemplo, Faster R-CNN, U-NET y Mask R-CNN.

Además, se sugiere llevar a cabo el entrenamiento con las mismas arquitecturas de redes neuronales artificiales usadas en esta investigación, pero disminuyendo la cantidad de clases etiquetadas, en particular las singularidades de los pavimentos que obtuvieron peores resultados (testigo, deterioros misceláneos).

Finalmente, para complementar el trabajo, se recomienda continuar este trabajo en paralelo, pero con imágenes de pavimentos de hormigón, los cuales también se presentan en un gran porcentaje en zonas urbanas. Esto implica la recolección de datos, etiquetado de los deterioros propios de este tipo de pavimento y entrenamiento de una red neuronal artificial. Una vez realizado esto, sería útil una red neuronal artificial de clasificación de imágenes capaz de discernir si la imagen analizada corresponde a un pavimento de asfalto u hormigón, para luego ocupar la red correspondiente. De esta forma, se termina de automatizar el proceso de detección, clasificación y medición de deterioros de pavimentos urbanos.

### **7.3 Investigación Futura**

Dentro de la investigación futura, se sugieren los siguientes pasos:

- Reducir las limitaciones al momento de grabación de los videos de pavimentos (ampliar horario de grabación y aumentar velocidad mínima de grabación).
- Llevar a cabo la tarea de detección de deterioros de pavimentos asfálticos con otras arquitecturas de Deep Learning.
- Realizar el símil de este trabajo con pavimentos de hormigón.

## 8 *Lista de referencias*

- Abu-Mostafa, Y. (Director). (2012). *Machine Learning Course—CS156, CalTech*.
- Bang, S., Park, S., Kim, H., Yoon, Y., & Kim, H. (2018, julio). *A Deep Residual Network with Transfer Learning for Pixel-level Road Crack Detection*.  
<https://doi.org/10.22260/ISARC2018/0103>
- Berrio, L., & Ussa, A. (2012). *Auscultación de pavimentos mediante interpretación de imágenes*. Universidad Militar Nueva Granada.
- Bull, A., & Schliessler, A. (1994). *Caminos: Un nuevo enfoque para la gestión y conservación de redes viales*. <https://repositorio.cepal.org/handle/11362/30314>
- Buslaev, A., Iglovikov, V. I., Khvedchenya, E., Parinov, A., Druzhinin, M., & Kalinin, A. A. (2020). Alumentations: Fast and Flexible Image Augmentations. *Information, 11*(2). <https://doi.org/10.3390/info11020125>
- Cáceres, X. (2021). *Detección y clasificación automática de grietas en pavimentos asfálticos por medio de un modelo de detección de objetos de Deep Learning*. Universidad Técnica Federico Santa María.
- Carr, T. A., Jenkins, M. D., Iglesias, M. I., Buggy, T., & Morison, G. (2018). Road crack detection using a single stage detector based deep neural network. *2018 IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS)*, 1-5. <https://doi.org/10.1109/EESMS.2018.8405819>

- Contreras, F. (2020). *Detección y clasificación automática de grietas en pavimentos asfálticos usando modelos de segmentación de deep learning*. Universidad Técnica Federico Santa María.
- De Solminihac, H. (2018). *Gestión de infraestructura vial: Tercera edición*. Ediciones UC.
- Dutta, A., & Zisserman, A. (2019). The VIA Annotation Software for Images, Audio and Video. *Proceedings of the 27th ACM International Conference on Multimedia*, 2276-2279. <https://doi.org/10.1145/3343031.3350535>
- Escalona, U., Arce, F., Zamora, E., & Sossa, H. (2019). Fully convolutional networks for automatic pavement crack segmentation. *Computacion y Sistemas*, 23(2), 451-460. <https://doi.org/10.13053/CyS-23-2-3047>
- Fan, Z., Wu, Y., Lu, J., & Li, W. (2018). *Automatic Pavement Crack Detection Based on Structured Prediction with the Convolutional Neural Network*.
- Fei, Y., Wang, K., Zhang, A., Chen, C., Li, J., Liu, Y., Yang, G., & Li, B. (2019). Pixel-Level Cracking Detection on 3D Asphalt Pavement Images Through Deep-Learning-Based CrackNet-V. *IEEE Transactions on Intelligent Transportation Systems*, PP, 1-12. <https://doi.org/10.1109/TITS.2019.2891167>
- Gopalakrishnan, K., Khaitan, S. K., Choudhary, A., & Agrawal, A. (2017). Deep Convolutional Neural Networks with transfer learning for computer vision-based data-driven pavement distress detection. *Construction and Building Materials*, 157, 322-330. <https://doi.org/10.1016/j.conbuildmat.2017.09.110>

- Hicks, R. G., Moulthrop, J. S., & Daleiden, J. (1999). Selecting a Preventive Maintenance Treatment for Flexible Pavements. *Transportation Research Record*, 1680(1), 1-12. <https://doi.org/10.3141/1680-01>
- Hsieh Yung-An & Tsai Yichang James. (2020). Machine Learning for Crack Detection: Review and Model Performance Comparison. *Journal of Computing in Civil Engineering*, 34(5), 04020038. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000918](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000918)
- Jenkins, M. D., Carr, T. A., Iglesias, M. I., Buggy, T. W., & Morison, G. (2018). A Deep Convolutional Neural Network for Semantic Pixel-Wise Segmentation of Road and Pavement Surface Cracks. *2018 26th European Signal Processing Conference (EUSIPCO)*, 2120-2124.
- Li, B., Wang, K. C. P., Zhang, A. A., Yang, E., & Wang, G. (2018). Automatic classification of pavement crack using deep convolutional neural network. *International Journal of Pavement Engineering*, 21, 457-463.
- Liu, J., Yang, X., Lau, S., Wang, X., Luo, S., Lee, V. C.-S., & Ding, L. (2020). Automated pavement crack detection and segmentation based on two-step convolutional neural network. *Computer-Aided Civil and Infrastructure Engineering*, 35(11), 1291-1305. <https://doi.org/10.1111/mice.12622>
- Liu, Y., Yao, J., Lu, X., Xie, R., & Li, L. (2019). DeepCrack: A deep hierarchical feature learning architecture for crack segmentation. *Neurocomputing*, 338, 139-153. <https://doi.org/10.1016/j.neucom.2019.01.036>

- Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., & Omata, H. (2018). Road Damage Detection and Classification Using Deep Neural Networks with Smartphone Images: Road damage detection and classification. *Computer-Aided Civil and Infrastructure Engineering*, 33. <https://doi.org/10.1111/mice.12387>
- Majidifard, H., Jin, P., Adu-Gyamfi, Y., & Buttlar, W. G. (2020). Pavement image datasets: A new benchmark dataset to classify and densify pavement distresses. *Transportation Research Record*, 2674(2), 328-339.
- Mandal, V., Uong, L., & Adu-Gyamfi, Y. (2018). Automated Road Crack Detection Using Deep Convolutional Neural Networks. 5212-5215. <https://doi.org/10.1109/BigData.2018.8622327>
- Marcelino, P., de Lurdes Antunes, M., Fortunato, E., & Gomes, M. C. (2021). Machine learning approach for pavement performance prediction. *International Journal of Pavement Engineering*, 22(3), 341-354.
- Nguyen, N. T. H., Le, T. H., Perry, S., & Nguyen, T. T. (2018). Pavement Crack Detection Using Convolutional Neural Network. *Proceedings of the Ninth International Symposium on Information and Communication Technology*, 251-256. <https://doi.org/10.1145/3287921.3287949>
- Osorio, A. (2015). *Development of performance models and maintenance standards of urban pavements for network management*. Pontificia Universidad Católica de Chile. Escuela de Ingeniería.

- Qualcomm. (s. f.). *Classification, Object Detection and Image Segmentation*. Qualcomm Developer Network. Recuperado 14 de octubre de 2022, de <https://developer.qualcomm.com/software/qualcomm-neural-processing-sdk/learning-resources/image-segmentation-deeplab-neural-processing-sdk/classification-object-detection-segmentation>
- Roberts, R., Giancontieri, G., Inzerillo, & Mino, G. (2020). Towards Low-Cost Pavement Condition Health Monitoring and Analysis Using Deep Learning. *Applied Sciences*, 10, 319. <https://doi.org/10.3390/app10010319>
- Rozas, D. (2023). *Validación de Metodología Experimental para la Medición de Deterioros Superficiales en Pavimentos Urbanos a partir de Imágenes Recopiladas por Instrumentos de Bajo Costo*.
- Sarmiento, J.-A. (2021). *Pavement Distress Detection and Segmentation using YOLOv4 and DeepLabv3 on Pavements in the Philippines*.
- Song, L., & Wang, X. (2021). Faster region convolutional neural network for automated pavement distress detection. *Road Materials and Pavement Design*, 22(1), 23-41. <https://doi.org/10.1080/14680629.2019.1614969>
- Statologos. (2021, mayo 19). *Cómo realizar una prueba de normalidad en Excel (paso a paso)*. Statologos: El sitio web para que aprendas estadística en Stata, R y Phyton. <https://statologos.com/test-de-normalidad-excel/>
- Tran, T. S., Tran, V. P., Lee, H. J., Flores, J. M., & Le, V. P. (2022). A two-step sequential automated crack detection and severity classification process for asphalt

- pavements. *International Journal of Pavement Engineering*, 23(6), 2019-2033.  
<https://doi.org/10.1080/10298436.2020.1836561>
- Van Phuc, T., Son, T., Lee, H., Kim, K., Baek, J., & Nguyen, T. (2021). One stage detector (RetinaNet)-based crack detection for asphalt pavements considering pavement distresses and surface objects. *Journal of Civil Structural Health Monitoring*, 11, 1-18. <https://doi.org/10.1007/s13349-020-00447-8>
- Venegas, J. (2022). *Desarrollo de una metodología de utilización de cámaras de bajo costo para la evaluación de pavimentos urbanos.*
- Wang, X., & Hu, Z. (2017). *Grid-based pavement crack analysis using deep learning.* 917-924. <https://doi.org/10.1109/ICTIS.2017.8047878>
- Ye, W., Jiang, W., Tong, Z., Yuan, D., & Xiao, J. (2021). Convolutional neural network for pothole detection in asphalt pavement. *Road Materials and Pavement Design*, 22(1), 42-58. <https://doi.org/10.1080/14680629.2019.1615533>
- Yin, J., Peng, T.-H., Kuan, J.-L., & Chen, B.-H. (2019). *Towards Perspective-Free Pavement Distress Detection via Deep Learning.* 661-662. <https://doi.org/10.1109/GCCE46687.2019.9015591>
- Yin, J.-L., Peng, T.-H., Kuan, J.-L., & Chen, B.-H. (2019). Towards Perspective-Free Pavement Distress Detection via Deep Learning. *2019 IEEE 8th Global Conference on Consumer Electronics (GCCE)*, 661-662.
- Yusof, N. A. M., Osman, M. K., Noor, M. H. M., Ibrahim, A., Tahir, N. M., & Yusof, N. M. (2018). Crack Detection and Classification in Asphalt Pavement Images using

- Deep Convolution Neural Network. *2018 8th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, 227-232. <https://doi.org/10.1109/ICCSCE.2018.8685007>
- Yusof, N., Ibrahim, A., Noor, M., Tahir, N., Abidin, N., & Osman, M. (2019). Deep convolution neural network for crack detection on asphalt pavement. *Journal of Physics: Conference Series*, 1349, 012020. <https://doi.org/10.1088/1742-6596/1349/1/012020>
- Zhang, A. A., Wang, K. C. P., Li, B., Yang, E., Dai, X., Peng, Y., Fei, Y., Liu, Y., Li, J. Q., & Chen, C. (2017). Automated Pixel-Level Pavement Crack Detection on 3D Asphalt Surfaces Using a Deep-Learning Network. *Computer-Aided Civil and Infrastructure Engineering*, 32.
- Zhang, A., Wang, K. C. P., Fei, Y., Liu, Y., Chen, C., Yang, G., Li, J. Q., Yang, E., & Qiu, S. (2019). Automated Pixel-Level Pavement Crack Detection on 3D Asphalt Surfaces with a Recurrent Neural Network. *Comput.-Aided Civ. Infrastruct. Eng.*, 34(3), 213-229. <https://doi.org/10.1111/mice.12409>
- Zhang Allen, Wang Kelvin C. P., Fei Yue, Liu Yang, Tao Siyu, Chen Cheng, Li Joshua Q., & Li Baoxian. (2018). Deep Learning–Based Fully Automated Pavement Crack Detection on 3D Asphalt Surfaces with an Improved CrackNet. *Journal of Computing in Civil Engineering*, 32(5), 04018041. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000775](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000775)

Zhang, L., Yang, F., Zhang, Y. D., & Zhu, Y. J. (2016). Road crack detection using deep convolutional neural network. *2016 IEEE International Conference on Image Processing (ICIP)*, 3708-3712.

## 9 Anexos

### 9.1 Anexo 1. Tabla con Planificación de Recorrido para Auscultaciones

Id	Comuna	Calle	Entre	Entre	N° Pistas	Unidireccional/ Bidireccional	m
1	Macul	Escuela Agrícola	Av. Marathon	Pedro de Valdivia	4	Bidireccional	450
2	Macul	Av Padre Luis Querber	Escuela Agrícola	Froilán Roa	3	Unidireccional	1600
3	Macul	Sergio Vieira de Mello	Escuela Agrícola	Froilán Roa	3	Unidireccional	1600
4	Macul	Froilán Roa	Sergio Vieira de Mello	Departamental	2	Unidireccional	850
5	Macul	Marathon	Departamental	Benito Rebolledo	4	Bidireccional	850
6	Macul	Exequiel Fernández	Departamental	Monseñor Carlos Casanueva	4	Bidireccional	1600
7	Macul	Av Quilin	Av Macul	Ramón Cruz Montt	2	Bidireccional	1200
8	Macul	Ramón Cruz	Av Quilin	Capitán Ignacio Carrera Pinto	2	Bidireccional	1000
9	Macul	Capitán Ignacio Carrera Pinto	Ramón Cruz	Rodrigo de Araya (Rotonda)	4	Bidireccional	600
10	Macul	Los Plátanos	Alcalde Jorge Monckeberg	Macul	2	Bidireccional	850
11	Macul	Los Plátanos	Macul	Pedro de Valdivia	2	Unidireccional	600
12	Macul	Pedro de Valdivia	Los Plátanos	Rotonda Pedro de Valdivia	2	Bidireccional	1400
13	Ñuñoa	Av Salvador	Av Providencia	Av Grecia	2	Bidireccional	3000
14	Ñuñoa	Campo de Deportes	Grecia	José Domingo Cañas	6	Bidireccional	550
15	Ñuñoa	José Domingo Cañas	Av José Pedro Alessandri	Av Grecia	2	Bidireccional	2600
16	Ñuñoa	Av Pedro de Valdivia	Grecia	Simón Bolívar	4	Bidireccional	1300
17	Ñuñoa	Simón Bolívar	Av Pedro de Valdivia	Av Presidente Jose Batlle y Ordoñez	2	Bidireccional	1300
18	Ñuñoa	Av Presidente Jose Batlle y Ordoñez	Simón Bolívar	Irrazával	2	Bidireccional	1200
19	Estación Central	Av Gladys Marin / Av. Pajaritos	Ruta 68	Av Teniente Cruz	4	Bidireccional	2700
20	Estación Central	Santa Corina	Av. Pajaritos	Av 5 de Abril	4	Bidireccional	1100
21	Estación Central	Av 5 de Abril	Santa Corina	Yelcho	4	Bidireccional	650
22	Estación Central	Presidente Balmaceda	Av 5 de Abril	Av Las Parcelas	2	Bidireccional	400
23	Estación Central	Av Las Parcelas	Curacaví	Aeropuerto	2	Bidireccional	1300
24	Estación Central	Las Violetas	Av. Las Rejas Sur	Aeropuerto	2	Unidireccional	550
25	Maipú	Marta Ossa Ruiz	Av La Estrella	Pasaje Alfonso X	1	Unidireccional	1000
26	Maipú	Alfonso X	Marta Ossa Ruiz	Av El Descanso	2	Bidireccional	290
27	Maipú	Av El Descanso	Av Longitudinal	Los Pajaritos	4	Bidireccional	1000
28	Maipú	Llancahue	Av El Descanso	Butachauques	2	Bidireccional	180
29	Maipú	Butachauques	Mar de Drake	Av Tte Cruz	2	Bidireccional	950

Id	Comuna	Calle	Entre	Entre	N° Pistas	Unidireccional/ Bidireccional	m
30	Maipú	Av Tte Cruz	Los Pajaritos	Laguna Sur	6	Bidireccional	900
31	Maipú	Av La Laguna Sur	Av Tte Cruz	Av Las Torres	4	Bidireccional	650
32	Maipú	Av Las Torres	Av Travesía	Av Simon Bolivar	4	Bidireccional	2300
33	Maipú	Av Las Parcelas	Av Santa Corina	Los Pajaritos	4	Bidireccional	1500
34	Maipú	Los Pajaritos	Vespucio Norte	Av Tte Cruz	4	Bidireccional	3100
35	Maipú	Av Hugo Bravo	Los Pajaritos	Vespucio Norte	2	Bidireccional	1600
36	Maipú	Anita Lizana	Av Hugo Bravo	Av El Descanso	1	Bidireccional	300
37	Maipú	Av El Rosal	Vespicio Norte	Isabel Riquelme	4	Bidireccional	3000
38	Maipú	Teatro Municipal	Av El Rosal	Av Jorge Alessandri	2	Bidireccional	1100
39	Maipú	Isabel Riquelme	Av Jorge Alessandri	Av La Sinfonía	2	Bidireccional	650
40	Maipú	Av La Sinfonía	Isabel Riquelme	Teatro Municipal	4	Bidireccional	1400
41	La Reina	Carlos Ossandón	Echeñique	Av Fernando Castillo Velasco	2	Bidireccional	1000
42	La Reina	Jorge Alessandri	Av Fernando Castillo Velasco	Av José Arrieta	4	Bidireccional	1400
43	Peñalolén	Av José Arrieta	Av Américo Vespucio	Diputada Laura Rodríguez	4	Bidireccional	2900
44	La Reina	Vicente Pérez Rosales	José Zapiola	Av Echeñique	2	Bidireccional	800
45	La Reina	Av Echeñique	Vicente Pérez Rosales	Monseñor Edwards	2	Unidireccional	750
46	La Reina	Monseñor Edwards	Av Echeñique	Av Principe de Gales	2	Bidireccional	650
47	La Reina	Av Principe de Gales	Monseñor Edwards	Av Ossa	3	Unidireccional	1100
48	Ñuñoa	Tobalaba	Av Ossa	Av Apoquindo/Av Providencia	4	Bidireccional	3300
49	Peñalolén	Av Las Torres/Av Dr Amador Neghme Rodríguez	Av Macul	Tobalaba	4	Bidireccional	3400
50	Macul	Av Macul/Av Jose Pedro Alessandri	El Líbano	Av Irrrazaval	4	Bidireccional	5400
51	Peñalolén	Caletera Américo Vespucio	Av Grecia	Av Las Torres	4	Bidireccional	3700
52	Macul	Monseñor Carlos Casanueva	Av Vicuña Mackenna	Froilán Roa	4	Bidireccional	1500
Km totales							246

## 9.2 Anexo 2. Deterioros y Singularidades del Pavimento Urbano Asfáltico Detectadas

A continuación, se presentan los deterioros y singularidades del pavimento detectadas. Y por último, ejemplos de imágenes descartadas.

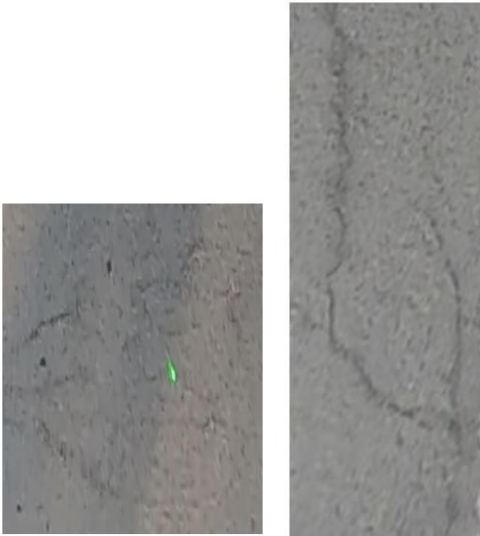

### 9.2.1 Deterioros

#### 9.2.1.1 Agrietamiento de Fatiga

El agrietamiento de fatiga o grieta de cocodrilo, son aquellas grietas en el pavimento que en general van por la huella de los vehículos, y su severidad viene dada por el nivel de interconexión entre ellas.

Ejemplos:

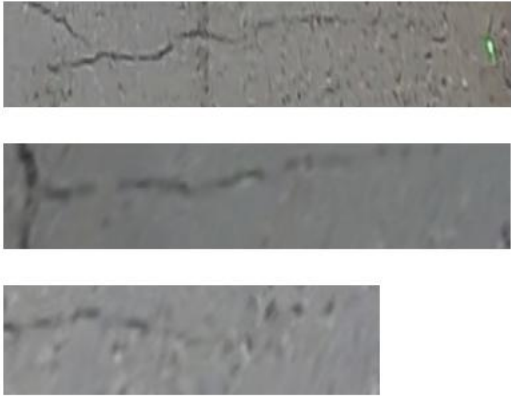

*Tabla 9.1: Ejemplos de grietas de fatiga según su severidad*

Grietas de Fatiga	
Severidad Media	Severidad Alta
	

### 9.2.1.2 Agrietamiento Transversal

Son aquellas grietas que, como su nombre lo dice, van de forma transversal a la calle. Su severidad viene dada por el ancho de la grieta, lo que en imágenes, se puede ver por su tonalidad, es decir, grietas más oscuras son de un ancho mayor.

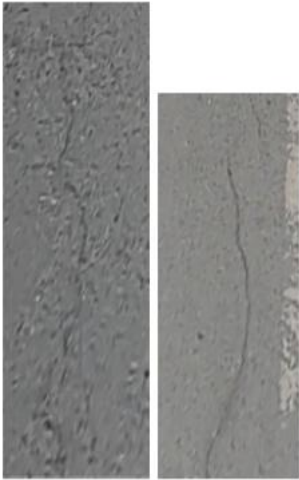
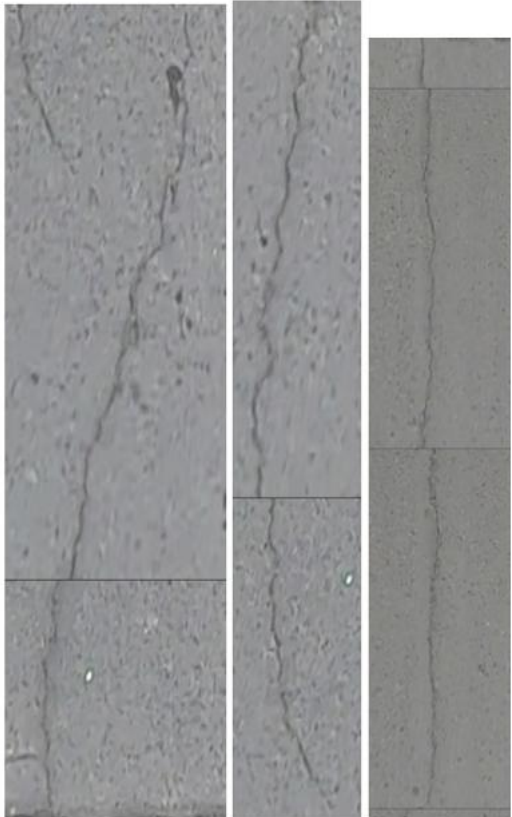
*Tabla 9.2: Ejemplos de grietas transversales según su severidad*

Grietas Transversales	
Severidad Media	Severidad Alta
	

### 9.2.1.3 Agrietamiento Longitudinal

En contraste con las grietas mencionadas anteriormente, las grietas longitudinales son grietas que se producen longitudinal a las calles. Su severidad también viene dada por el ancho de la grieta, que en imágenes se representa por su tonalidad, por lo que grietas más oscuras representan grietas más anchas.

**Tabla 9.3: Ejemplos de grietas longitudinales según su severidad**

Grietas Longitudinales	
Severidad Media	Severidad Alta
	

Dado que tanto las grietas longitudinales como las grietas transversales son consideradas grietas lineales, se evalúa en paralelo llevar a cabo el entrenamiento con ambas etiquetas como una sola, “grietas lineales”.

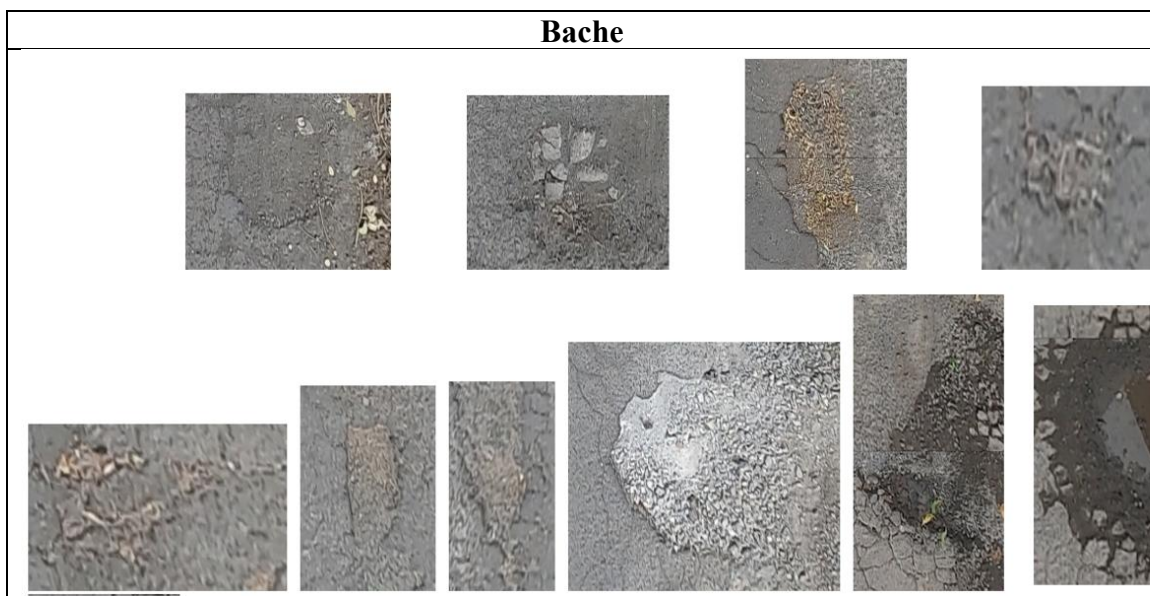
### 9.2.1.4 Baches

Son aquellas zonas del pavimento donde se carece de la capa superficial. Su severidad viene dada por el área propia del bache, por lo que no es necesario clasificar en el etiquetado su severidad, y se determina después en base al área de la zona etiquetada.

Las severidades de los baches, como se establece en el ICPU (Osorio, 2015), vienen dadas por las siguientes relaciones:

$$\begin{aligned} \text{Severidad baja:} & \quad r < 15 \text{ cm} = A < 707 \text{ cm}^2 \\ \text{Severidad media:} & \quad 15 < r < 30 \text{ cm} = 707 < A < 2827 \text{ cm}^2 \\ \text{Severidad alta:} & \quad 30 < r \text{ cm} = 2827 < A \text{ cm}^2 \end{aligned}$$



*Tabla 9.4: Ejemplos de baches*



### 9.2.1.5 Parches

Son aquellas regiones del pavimento en que se realizó una reparación para tapar un bache o intervención. Su severidad viene dada por el estado del parche, por lo tanto, si se encuentra agrietado, es de severidad alta. De lo contrario, es de severidad media.

*Tabla 9.5: Ejemplos de parches según su severidad*

Parches	
Severidad Media	Severidad Alta
	

## 9.2.2 Singularidades del Pavimento

### 9.2.2.1 Tapas de Alcantarillado

*Tabla 9.6: Ejemplos de tapas de alcantarillado*

Tapas de Alcantarillado


### 9.2.2.2 Sumideros

*Tabla 9.7: Ejemplos de sumideros*



### 9.2.2.3 Extracción de Testigos

*Tabla 9.8: Ejemplos de zonas con extracción de testigos*



#### 9.2.2.4 Deterioros Misceláneos

Se etiquetará como “deterioro misceláneo” aquellos deterioros que no entran en la clasificación detallada anteriormente. Ejemplo, pérdida de material

*Tabla 9.9: Ejemplos de pérdida de material, catalogado como deterioros misceláneos*



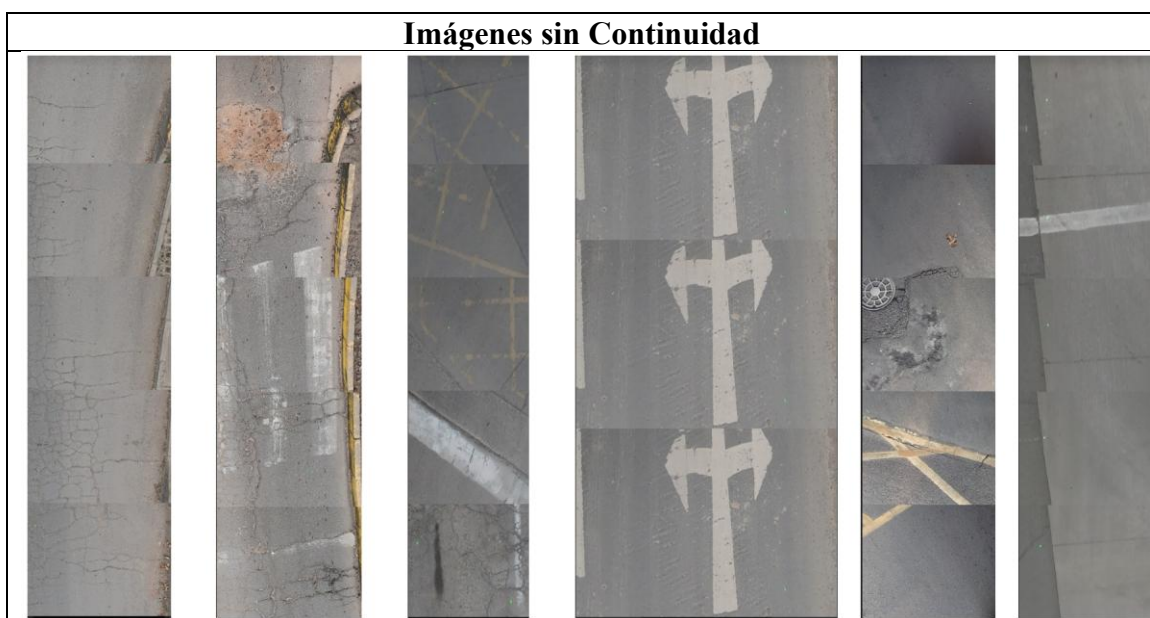
### 9.2.3 Imágenes Descartadas

Finalmente, al etiquetar las imágenes, también se encontraron imágenes que podrían perjudicar el desempeño de la red neuronal. Las imágenes descartadas se clasifican en 3 categorías, detalladas a continuación.

#### 9.2.3.1 Imágenes sin Continuidad

Se descartan las imágenes que presentan nula continuidad al unir 5 de ellas y formar una sección.

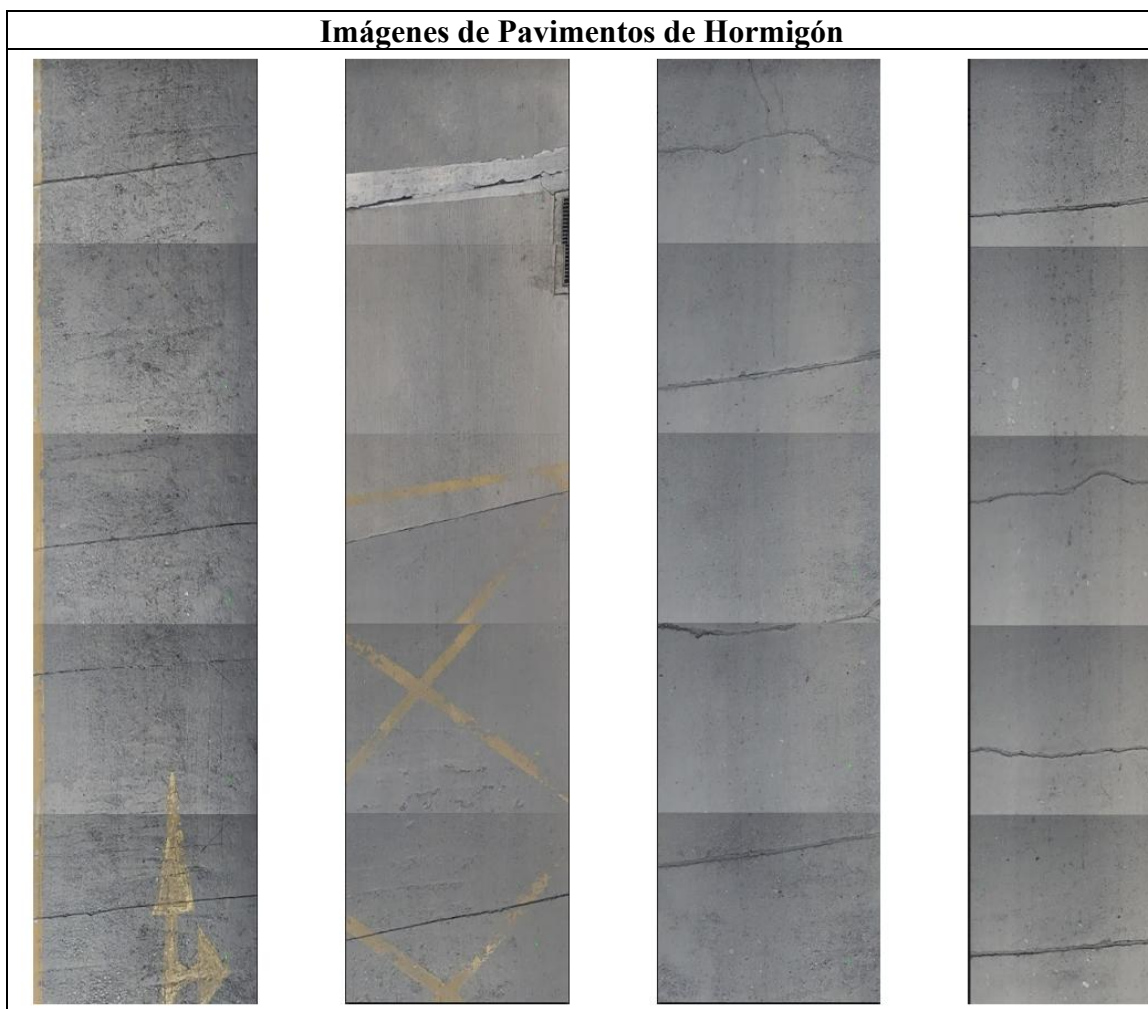
*Tabla 9.10: Ejemplos de imágenes descartadas por no presentar continuidad*



### 9.2.3.2 Imágenes de Pavimentos de Hormigón

A efecto de esta parte del estudio, que se está trabajando netamente con pavimentos asfálticos, por lo que se descartan las imágenes de pavimentos de hormigón, guardándolas para el futuro del proyecto.

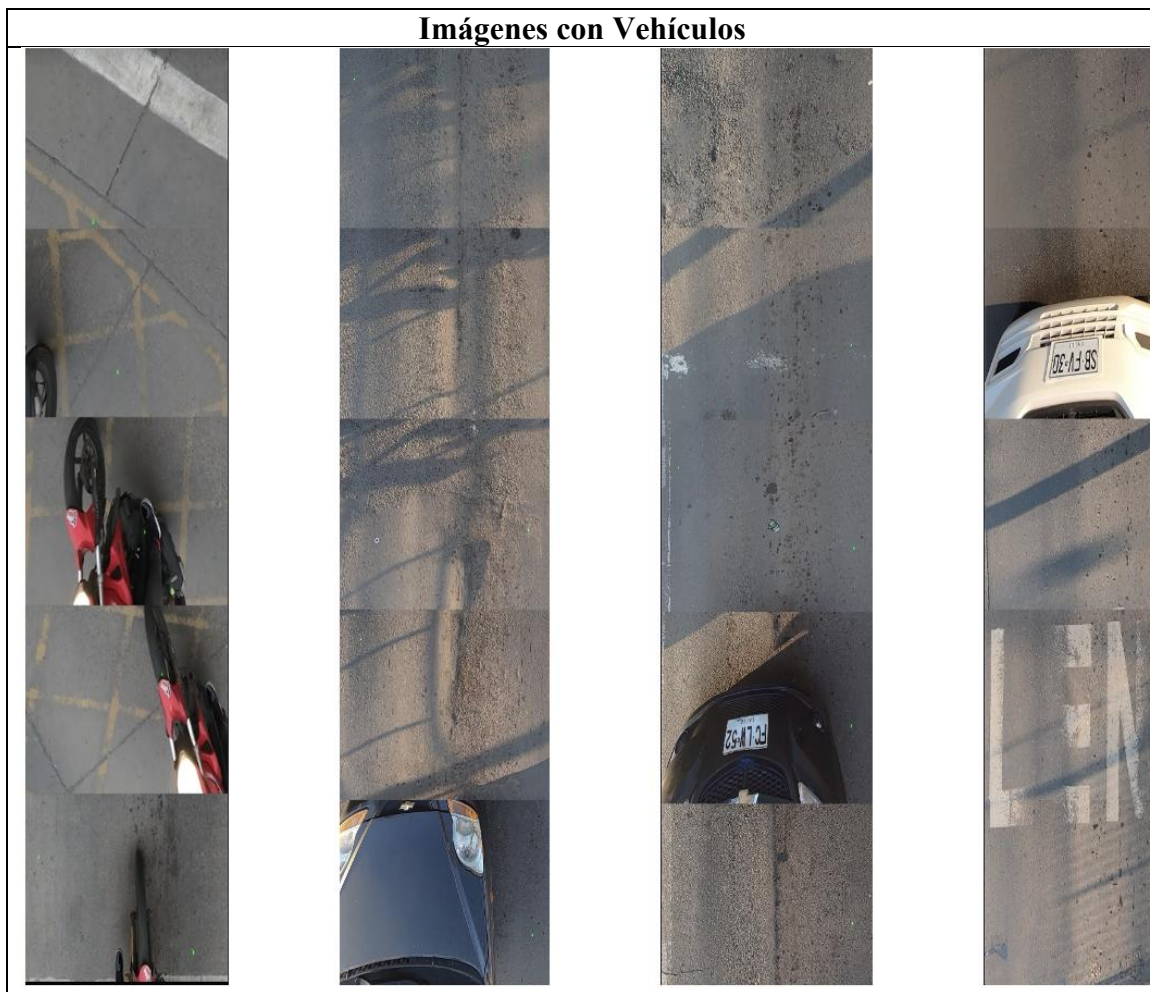
*Tabla 9.11: Ejemplos de imágenes descartadas por ser de pavimentos de hormigón*



### 9.2.3.3 Imágenes con Vehículos

Por último, se decide descartar imágenes donde una región del pavimento no se puede ver por estar tapada por algún vehículo.

*Tabla 9.12: Ejemplos de imágenes descartadas por tener vehículos que cubren el pavimento*



### 9.3 Anexo 3. Técnicas de aumento de datos utilizadas por tipo de deterioro

#### 9.3.1 Bache

		Bache									
		Spatial Level									
		Original	Flip H	Flip V	Flip ambos	Rotar +90°	Rotar -90°	Rot +90° + FV	Rot +90° + FH	Optical Distortion	Grid Distortion
Original		1									
Spatial Level	Flip H		1								
	Flip V			1							
	Flip ambos				1						
	Rotar +90°					1					
	Rotar -90°						1				
	Rot +90° + FV										
	Rot +90° + FH										
	Optical Distortion										
Grid Distortion											
Pixel Level	Blur	1									1
	Motion Blur										0
	Random Contrast			1							1
	JPEG compression		1								1
	CLAHE				1						1
	Gauss Noise										0
	Downscale										0
	Emboss										0
	Multiplicative Noise										0
	Random Fog						1				1
	Random Gamma										0
	Random Rain					1					1
	Random Shadow										0
	Ringing Overshoot										0
	Random Tone Curve										0
	Sharpen										0
Solarize										0	
Superpixel										0	
Unsharpen										0	
		2	2	2	2	2	2	0	0	0	0
		12									

Augmentations	Imágenes Originales	Total
12	315	3780

### 9.3.2 Parche Severidad Media

		Parche Severidad Media									
		Spatial Level									
		Original	Flip H	Flip V	Flip ambos	Rotar +90°	Rotar -90°	Rot +90° + FV	Rot +90° + FH	Optical Distortion	Grid Distortion
	Original	1									
Spatial Level	Flip H		1								
	Flip V			1							
	Flip ambos				1						
	Rotar +90°					1					
	Rotar -90°						1				
	Rot +90° + FV							1			
	Rot +90° + FH								1		
	Optical Distortion										
Grid Distortion											
Pixel Level	Blur	1		1		1					3
	Motion Blur		1		1				1		3
	Random Contrast		1					1			2
	JPEG compression				1				1		2
	CLAHE	1						1			2
	Gauss Noise					1			1		2
	Downscale						1	1			2
	Emboss					1			1		2
	Multiplicative Noise	1		1							2
	Random Fog		1		1	1					3
	Random Gamma	1		1							2
	Random Rain	1				1	1				3
	Random Shadow			1			1	1			3
	Ringing Overshoot		1						1		2
	Random Tone Curve	1		1							2
	Sharpen		1		1						2
	Solarize			1	1						2
Superpixel					1	1				2	
Unsharpen						1	1			2	
		7	6	7	6	7	6	6	6	0	0

Augmentations	Imágenes Originales	Total
51	73	3723

### 9.3.3 Parche Severidad Alta

		Parche Severidad Media									
		Spatial Level									
		Original	Flip H	Flip V	Flip ambos	Rotar +90°	Rotar -90°	Rot +90° + FV	Rot +90° + FH	Optical Distortion	Grid Distortion
	Original	1									
Spatial Level	Flip H		1								
	Flip V			1							
	Flip ambos				1						
	Rotar +90°					1					
	Rotar -90°						1				
	Rot +90° + FV							1			
	Rot +90° + FH								1		
	Optical Distortion										
	Grid Distortion										
Pixel Level	Blur	1			1		1				3
	Motion Blur		1				1				2
	Random Contrast			1			1	1			3
	JPEG compression				1	1		1			3
	CLAHE			1		1			1		3
	Gauss Noise		1				1		1		3
	Downscale	1			1			1			3
	Emboss							1	1		2
	Multiplicative Noise						1		1		2
	Random Fog		1			1		1			3
	Random Gamma				1		1				2
	Random Rain	1		1		1					3
	Random Shadow		1	1	1						3
	Ringing Overshoot	1		1							2
	Random Tone Curve		1						1		2
	Sharpen	1						1			2
	Solarize	1					1				2
Superpixel		1			1					2	
Unsharpen			1	1						2	
		7	7	7	7	7	7	7	6	0	0

Augmentations	Imágenes Originales	Total
55	68	3740

### 9.3.4 Grieta Longitudinal Severidad Media

		Longitudinal Media									
		Spatial Level									
		Original	Flip H	Flip V	Flip ambos	Rotar +90°	Rotar -90°	Rot +90° + FV	Rot +90° + FH	Optical Distortion	Grid Distortion
Spatial Level	Original	1									
	Flip H		1								
	Flip V			1							
	Flip ambos				1						
	Rotar +90°										
	Rotar -90°										
	Rot +90° + FV										
	Rot +90° + FH										
	Optical Distortion										
	Grid Distortion										
Pixel Level	Blur										0
	Motion Blur										0
	Random Contrast										0
	JPEG compression	1									1
	CLAHE										0
	Gauss Noise										0
	Downscale										0
	Emboss										0
	Multiplicative Noise										0
	Random Fog										0
	Random Gamma										0
	Random Rain	1									1
	Random Shadow										0
	Ringing Overshoot										0
	Random Tone Curve										0
	Sharpen										0
	Solarize										0
	Superpixel										0
Unsharpen										0	
		3	1	1	1	0	0	0	0	0	0

Augmentations	Imágenes Originales	Total
6	818	4908

### 9.3.5 Grieta Longitudinal Severidad Alta

		Longitudinal Alta									
		Spatial Level									
		Original	Flip H	Flip V	Flip ambos	Rotar +90°	Rotar -90°	Rot +90° + FV	Rot +90° + FH	Optical Distortion	Grid Distortion
	Original	1									
Spatial Level	Flip H		1								
	Flip V			1							
	Flip ambos				1						
	Rotar +90°										
	Rotar -90°										
	Rot +90° + FV										
	Rot +90° + FH										
	Optical Distortion										
	Grid Distortion										
	Pixel Level	Blur									
Motion Blur											0
Random Contrast											0
JPEG compression		1									1
CLAHE					1						1
Gauss Noise											0
Downscale				1							1
Emboss											0
Multiplicative Noise											0
Random Fog											0
Random Gamma											0
Random Rain			1								1
Random Shadow											0
Ringing Overshoot											0
Random Tone Curve											0
Sharpen											0
Solarize											0
Superpixel											0
Unsharpen										0	
		2	2	2	2	0	0	0	0	0	0

Augmentations	Imágenes Originales	Total
8	505	4040

## 9.4 Anexo 4. Resultados Entrenamiento

### 9.4.1 Matriz de Confusión por Cada Split

#### 9.4.1.1 YOLOv5

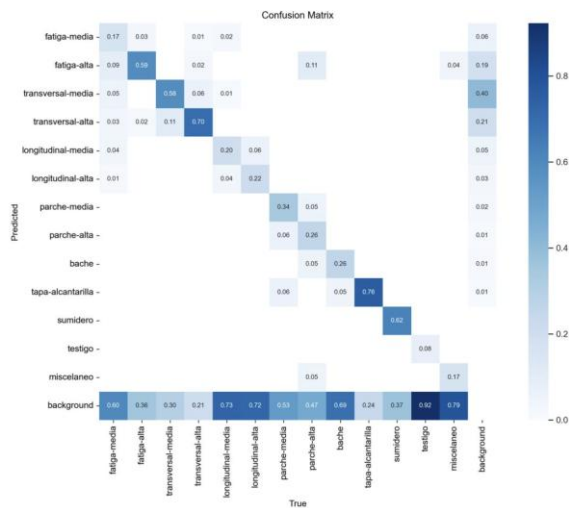


Figura 9.1: Matriz de confusión del conjunto test de Split 1 entrenado con YOLOv5

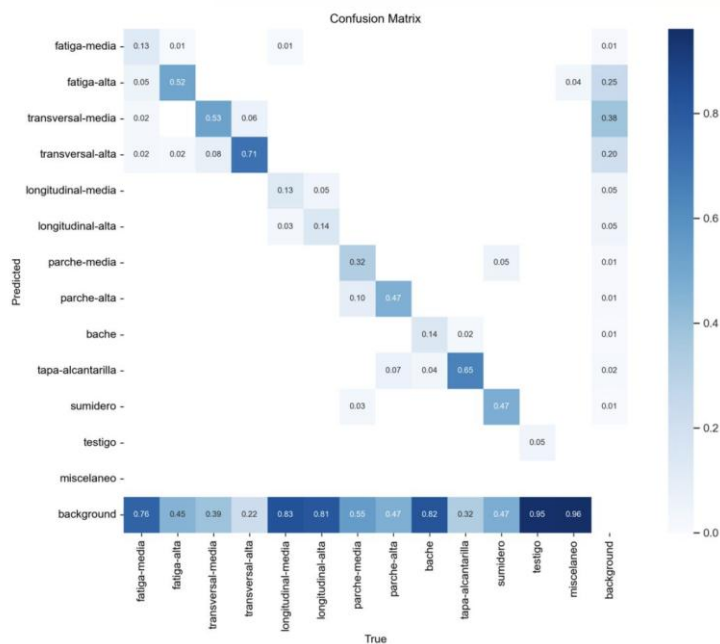


Figura 9.2: Matriz de confusión del conjunto test de Split 2 entrenado con YOLOv5

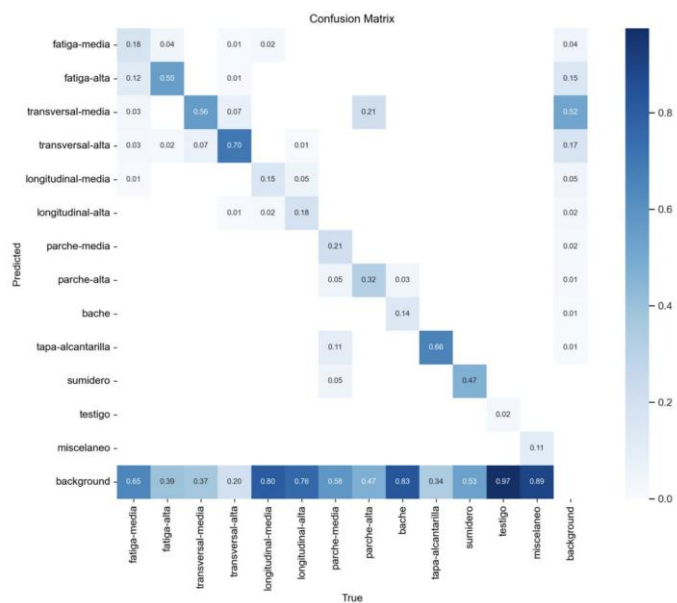


Figura 9.3: Matriz de confusión del conjunto test de Split 3 entrenado con YOLOv5

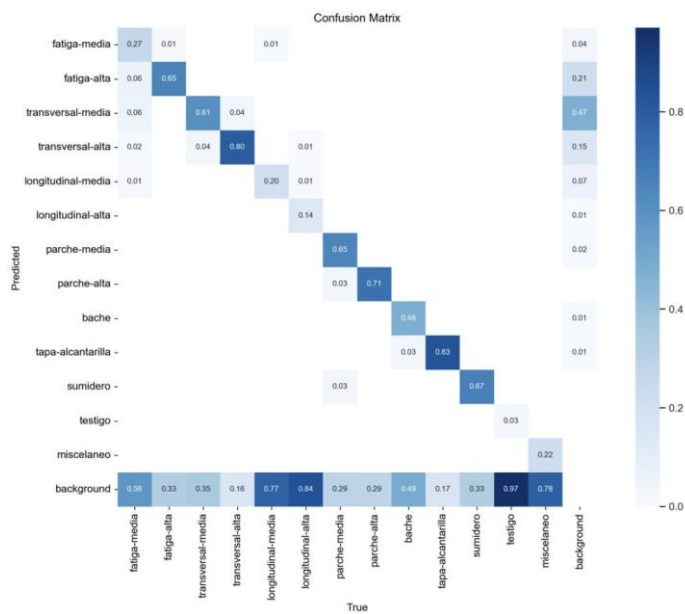


Figura 9.4: Matriz de confusión del conjunto test de Split 4 entrenado con YOLOv5

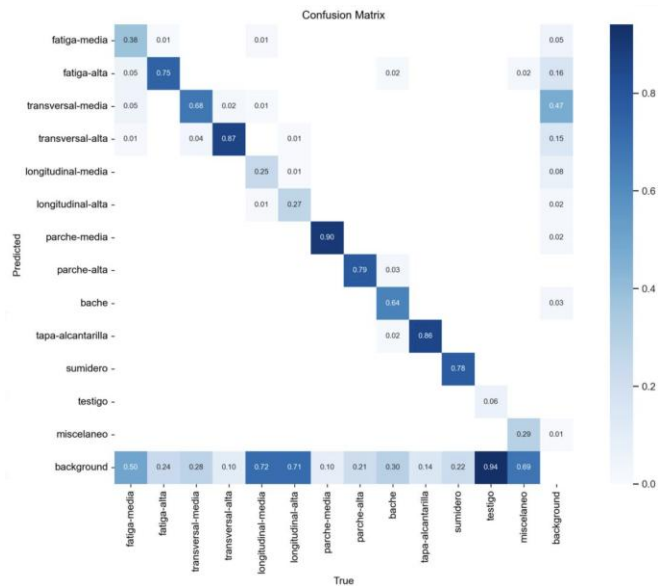


Figura 9.5: Matriz de confusión del conjunto test de Split 5 entrenado con YOLOv5

#### 9.4.1.2 YOLOv7

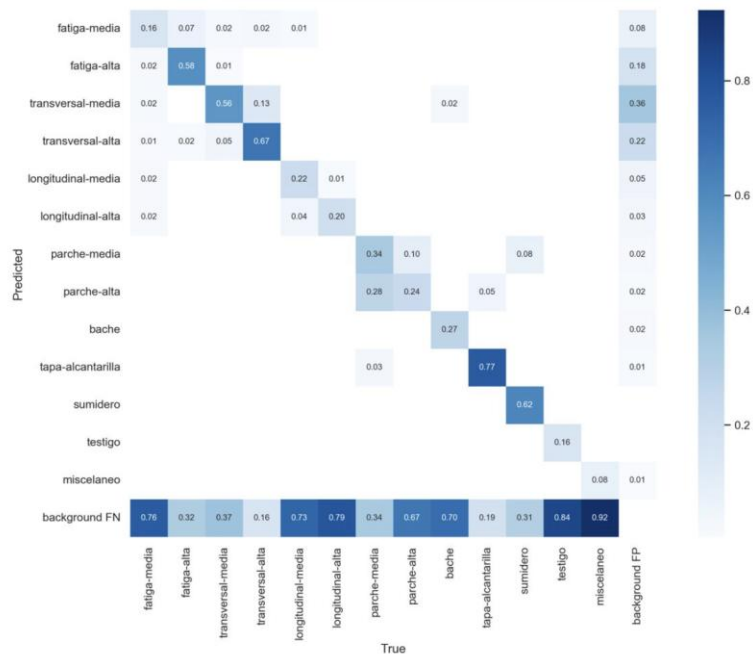


Figura 9.6: Matriz de confusión del conjunto test de Split 1 entrenado con YOLOv7

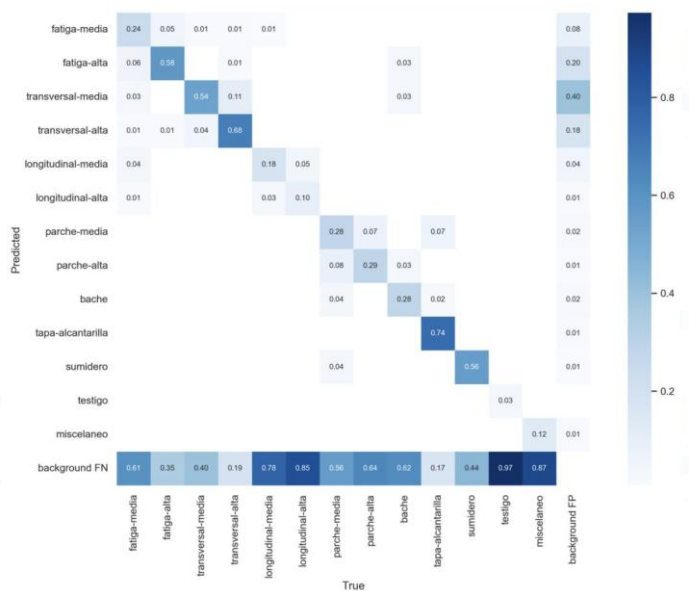


Figura 9.7: Matriz de confusión del conjunto test de Split 2 entrenado con YOLOv7

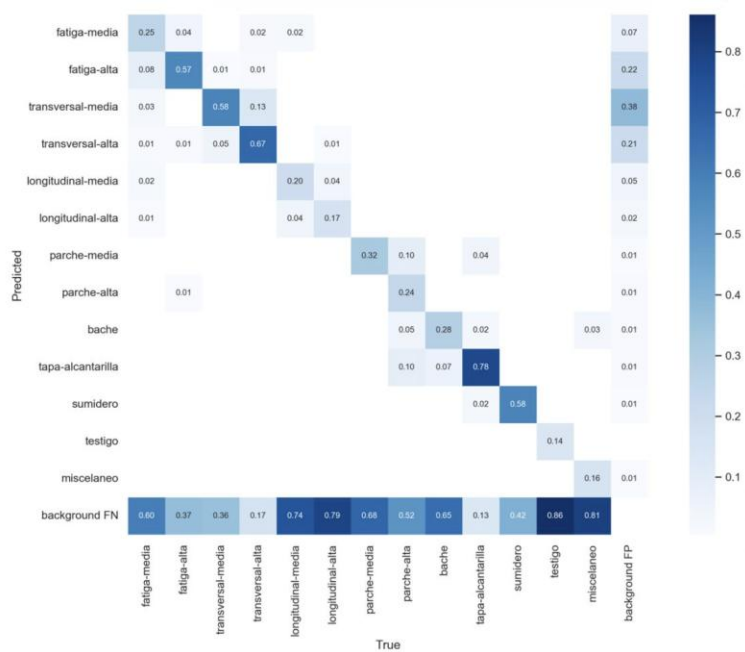
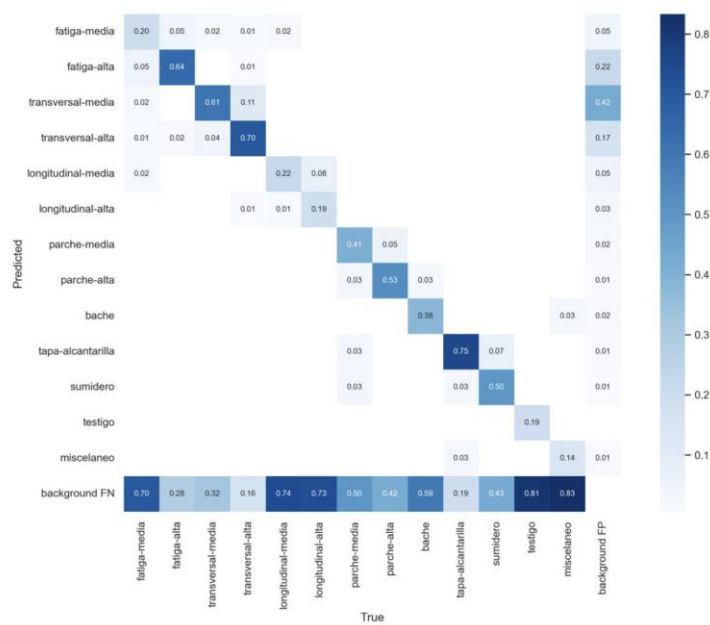
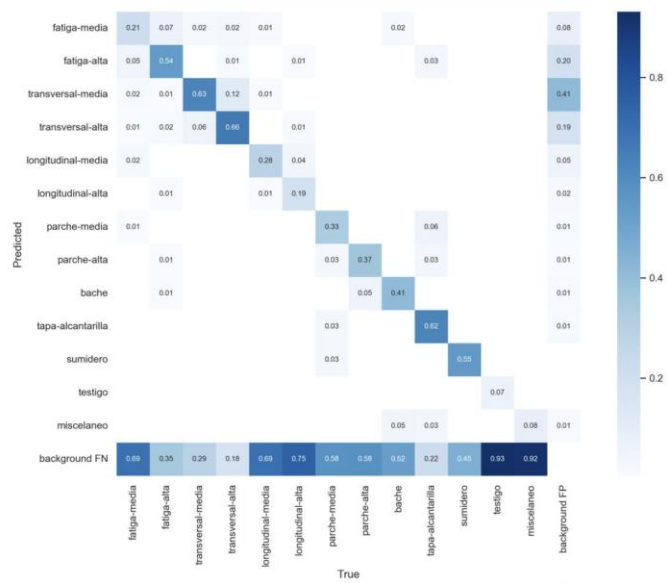


Figura 9.8: Matriz de confusión del conjunto test de Split 3 entrenado con YOLOv7



**Figura 9.9: Matriz de confusión del conjunto test de Split 4 entrenado con YOLOv7**



**Figura 9.10: Matriz de confusión del conjunto test de Split 5 entrenado con YOLOv7**

### 9.4.2 Resultados por Tipo de Deterioro

**Tabla 9.13: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 1 obtenido con YOLOv5**

Etiqueta	P	R	F1	mAP50	mAP50-95
Todas	0.414	0.433	0.423	0.374	0.189
Fatiga Media	0.277	0.23	0.251	0.161	0.0516
Fatiga Alta	0.489	0.59	0.535	0.522	0.264
Transversal Media	0.402	0.621	0.488	0.491	0.222
Transversal Alta	0.497	0.738	0.594	0.696	0.415
Longitudinal Media	0.362	0.286	0.320	0.254	0.0945
Longitudinal Alta	0.383	0.273	0.319	0.282	0.121
Parche Media	0.335	0.406	0.367	0.254	0.183
Parche Alta	0.397	0.421	0.409	0.313	0.198
Bache	0.29	0.254	0.271	0.164	0.0728
Tapa Alcantarilla	0.608	0.778	0.683	0.7	0.372
Sumidero	0.583	0.75	0.656	0.693	0.323
Testigo	0.215	0.0769	0.113	0.0932	0.0594
Miscelaneo	0.547	0.202	0.295	0.238	0.0874

**Tabla 9.14: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 1 obtenido con YOLOv7**

Labels	P	R	F1	mAP50	mAP50-95
All	0.189	0.364	0.249	0.379	0.214
Fatiga Media	0.0516	0.121	0.072	0.159	0.0739
Fatiga Alta	0.264	0.554	0.358	0.557	0.289
Transversal Media	0.222	0.494	0.306	0.52	0.247
Transversal Alta	0.415	0.703	0.522	0.672	0.413
Longitudinal Media	0.0945	0.17	0.121	0.242	0.101
Longitudinal Alta	0.121	0.16	0.138	0.246	0.13
Parche Media	0.183	0.423	0.255	0.197	0.129
Parche Alta	0.198	0.194	0.196	0.228	0.153
Bache	0.0728	0.25	0.113	0.281	0.139
Tapa Alcantarilla	0.372	0.767	0.501	0.701	0.475
Sumidero	0.323	0.667	0.435	0.735	0.489
Testigo	0.0594	0.16	0.087	0.333	0.12
Miscelaneo	0.0874	0.0741	0.080	0.0557	0.022

**Tabla 9.15: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 2 obtenido con YOLOv5**

Etiqueta	P	R	F1	mAP50	mAP50-95
Todas	0.475	0.403	0.436	0.369	0.184
Fatiga Media	0.42	0.201	0.272	0.224	0.103
Fatiga Alta	0.511	0.584	0.545	0.529	0.273
Transversal Media	0.444	0.633	0.522	0.506	0.214
Transversal Alta	0.551	0.782	0.646	0.743	0.43
Longitudinal Media	0.404	0.254	0.312	0.231	0.0788
Longitudinal Alta	0.388	0.274	0.321	0.252	0.0769
Parche Media	0.334	0.355	0.344	0.257	0.187
Parche Alta	0.307	0.467	0.370	0.365	0.261
Bache	0.393	0.18	0.247	0.164	0.0757
Tapa Alcantarilla	0.607	0.675	0.639	0.64	0.322
Sumidero	0.671	0.632	0.651	0.56	0.236
Testigo	0.551	0.136	0.218	0.206	0.0899
Miscelaneo	0.597	0.0609	0.111	0.119	0.0401

**Tabla 9.16: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 2 obtenido con YOLOv7**

Etiqueta	P	R	F1	mAP50	mAP50-95
All	0.385	0.426	0.404	0.368	0.201
Fatiga Media	0.264	0.319	0.289	0.192	0.0752
Fatiga Alta	0.44	0.652	0.525	0.563	0.284
Transversal Media	0.411	0.663	0.507	0.495	0.224
Transversal Alta	0.509	0.782	0.617	0.723	0.422
Longitudinal Media	0.298	0.286	0.292	0.23	0.0848
Longitudinal Alta	0.395	0.232	0.292	0.246	0.1
Parche Media	0.16	0.28	0.204	0.234	0.176
Parche Alta	0.331	0.349	0.340	0.292	0.199
Bache	0.231	0.323	0.269	0.219	0.113
Tapa Alcantarilla	0.614	0.833	0.707	0.804	0.506
Sumidero	0.322	0.474	0.383	0.432	0.298
Testigo	0.687	0.135	0.226	0.19	0.0737
Miscelaneo	0.349	0.21	0.262	0.162	0.0522

**Tabla 9.17: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 3 obtenido con YOLOv5**

Etiqueta	P	R	F1	mAP50	mAP50-95
Todas	0.426	0.44	0.433	0.37	0.191
Fatiga Media	0.326	0.328	0.327	0.204	0.0908
Fatiga Alta	0.462	0.614	0.527	0.543	0.289
Transversal Media	0.37	0.648	0.471	0.471	0.214
Transversal Alta	0.486	0.785	0.600	0.718	0.409
Longitudinal Media	0.328	0.277	0.300	0.226	0.0725
Longitudinal Alta	0.36	0.397	0.378	0.29	0.0999
Parche Media	0.276	0.382	0.320	0.238	0.181
Parche Alta	0.344	0.421	0.379	0.33	0.21
Bache	0.227	0.207	0.217	0.119	0.0692
Tapa Alcantarilla	0.623	0.683	0.652	0.65	0.353
Sumidero	0.764	0.667	0.712	0.637	0.293
Testigo	0.603	0.15	0.240	0.19	0.0891
Miscelaneo	0.371	0.162	0.226	0.184	0.105

**Tabla 9.18: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 3 obtenido con YOLOv7**

Etiqueta	P	R	F1	mAP50	mAP50-95
Todas	0.417	0.449	0.432	0.378	0.197
Fatiga Media	0.281	0.293	0.287	0.211	0.0993
Fatiga Alta	0.481	0.598	0.533	0.537	0.284
Transversal Media	0.429	0.641	0.514	0.5	0.224
Transversal Alta	0.456	0.756	0.569	0.679	0.393
Longitudinal Media	0.343	0.277	0.306	0.219	0.0833
Longitudinal Alta	0.328	0.208	0.255	0.211	0.0916
Parche Media	0.308	0.343	0.325	0.272	0.179
Parche Alta	0.213	0.263	0.235	0.206	0.161
Bache	0.403	0.324	0.359	0.264	0.104
Tapa Alcantarilla	0.624	0.804	0.703	0.763	0.517
Sumidero	0.572	0.769	0.656	0.587	0.257
Testigo	0.655	0.254	0.366	0.318	0.106
Miscelaneo	0.324	0.3	0.312	0.146	0.0556

**Tabla 9.19: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 4 obtenido con YOLOv5**

Etiqueta	P	R	F1	mAP50	mAP50-95
Todas	0.613	0.526	0.566	0.552	0.307
Fatiga Media	0.562	0.327	0.413	0.385	0.188
Fatiga Alta	0.632	0.685	0.657	0.705	0.394
Transversal Media	0.53	0.628	0.575	0.577	0.269
Transversal Alta	0.661	0.826	0.734	0.797	0.486
Longitudinal Media	0.484	0.262	0.340	0.324	0.145
Longitudinal Alta	0.543	0.244	0.337	0.345	0.131
Parche Media	0.588	0.71	0.643	0.634	0.422
Parche Alta	0.632	0.714	0.671	0.748	0.554
Bache	0.765	0.587	0.664	0.634	0.301
Tapa Alcantarilla	0.756	0.833	0.793	0.855	0.547
Sumidero	0.78	0.722	0.750	0.732	0.351
Testigo	0.399	0.0587	0.102	0.124	0.0535
Miscelaneo	0.641	0.246	0.356	0.323	0.148

**Tabla 9.20: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 4 obtenido con YOLOv7**

Etiqueta	P	R	F1	mAP50	mAP50-95
Todas	0.444	0.487	0.465	0.428	0.238
Fatiga Media	0.317	0.264	0.288	0.208	0.0913
Fatiga Alta	0.462	0.701	0.557	0.605	0.335
Transversal Media	0.409	0.666	0.507	0.524	0.241
Transversal Alta	0.536	0.793	0.640	0.719	0.447
Longitudinal Media	0.408	0.329	0.364	0.279	0.113
Longitudinal Alta	0.294	0.264	0.278	0.203	0.0901
Parche Media	0.379	0.571	0.456	0.308	0.228
Parche Alta	0.443	0.542	0.488	0.522	0.408
Bache	0.435	0.486	0.459	0.448	0.155
Tapa Alcantarilla	0.679	0.806	0.737	0.78	0.463
Sumidero	0.596	0.533	0.563	0.577	0.332
Testigo	0.562	0.213	0.309	0.264	0.121
Miscelaneo	0.254	0.162	0.198	0.126	0.0705

**Tabla 9.21: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 5 obtenido con YOLOv5**

Etiqueta	P	R	F1	mAP50	mAP50-95
Todas	0.665	0.578	0.618	0.621	0.357
Fatiga Media	0.673	0.386	0.491	0.499	0.243
Fatiga Alta	0.737	0.739	0.738	0.786	0.474
Transversal Media	0.6	0.67	0.633	0.647	0.31
Transversal Alta	0.723	0.863	0.787	0.847	0.522
Longitudinal Media	0.517	0.247	0.334	0.339	0.146
Longitudinal Alta	0.579	0.248	0.347	0.366	0.153
Parche Media	0.547	0.903	0.681	0.796	0.548
Parche Alta	0.708	0.786	0.745	0.803	0.594
Bache	0.729	0.662	0.694	0.701	0.345
Tapa Alcantarilla	0.902	0.88	0.891	0.908	0.627
Sumidero	0.804	0.778	0.791	0.767	0.382
Testigo	0.393	0.0588	0.102	0.182	0.0808
Miscelaneo	0.73	0.294	0.419	0.427	0.213

**Tabla 9.22: Resultados por tipo de etiqueta al evaluar el conjunto de prueba del Split 5 obtenido con YOLOv7**

Etiqueta	P	R	F1	mAP50	mAP50-95
Todas	0.424	0.443	0.433	0.392	0.215
Fatiga Media	0.281	0.269	0.275	0.189	0.0746
Fatiga Alta	0.447	0.593	0.510	0.508	0.278
Transversal Media	0.414	0.674	0.513	0.531	0.234
Transversal Alta	0.502	0.76	0.605	0.662	0.404
Longitudinal Media	0.41	0.405	0.407	0.306	0.151
Longitudinal Alta	0.406	0.325	0.361	0.296	0.123
Parche Media	0.307	0.353	0.328	0.206	0.147
Parche Alta	0.412	0.364	0.387	0.406	0.278
Bache	0.442	0.467	0.454	0.391	0.142
Tapa Alcantarilla	0.533	0.786	0.635	0.726	0.511
Sumidero	0.617	0.5	0.552	0.587	0.31
Testigo	0.582	0.145	0.232	0.204	0.0877
Miscelaneo	0.16	0.125	0.140	0.0826	0.049

