



TECNOLOGICO DE ESTUDIOS SUPERIORES DE ECATEPEC

DIVISION DE INGENIERIA EN SISTEMAS COMPUTACIONALES

**SISTEMA DE DETECCIÓN DE MOTOCICLETAS Y DEL USO DE
CASCO DE SEGURIDAD POR MEDIO DE ALGORITMOS DE
INTELIGENCIA ARTIFICIAL.**

T E S I S

**QUE PARA OBTENER EL TÍTULO DE
INGENIERO EN SISTEMAS COMPUTACIONALES**

P R E S E N T A :

GERARDO MARTINEZ MORALES

**DIRECTOR: M. en A.D.N. JUAN MANUEL STEIN CARRILLO
CODIRECTOR: DR. ADOLFO MELENDEZ RAMIREZ
REVISOR: DR. FRANCISCO JACOB AVILA CAMACHO**

ECATEPEC DE MORELOS, EDO. DE MÉXICO, 20 DE OCTUBRE DE 2022.

OFICIO

 GOBIERNO DEL ESTADO DE MÉXICO	DICTAMEN DE LIBERACIÓN Y CUMPLIMIENTO DE REQUISITOS ACADÉMICOS PARA TITULACIÓN FO-TESE-DA-78 TÍTOS	
--	---	---

DIRECCIÓN ACADÉMICA
DIVISIÓN INGENIERÍA EN SISTEMAS COMPUTACIONALES

TES_PRF-20211-006
03 DE OCTUBRE 2022

Datos generales

Nombre del egresado:
MARTINEZ MORALES GERARDO
Apellido paterno Apellido materno Nombres (s)

Matrícula: 201410221 Teléfono: 55-5711-6671 Móvil: 55-3056-6634 Correo Electrónico: fmtes.tese10@gmail.com 201410221@tese.edu.mx

Opción de titulación: **TESIS PROFESIONAL**
Únicamente en las opciones I, II, VII y X

PROYECTO: **DETECCIÓN DE MOTOCICLETAS Y DEL USO DEL CASCO DE SEGURIDAD, POR MEDIO DE ALGORITMOS DE INTELIGENCIA ARTIFICIAL.**

ASESOR (ES): **M. EN A.D.N. JUAN MANUEL STEIN CARRILLO, DR. ADOLFO MELÉNDEZ RAMÍREZ, DR. FRANCISCO JACOB ÁVILA CAMACHO**

Dictamen:

Con base en la solicitud de opción de titulación No. (1) correspondiente al periodo escolar (2022 - 1), presentada por el alumno de referencia en fecha **10 DE SEPTIEMBRE DEL 2021** y después de haber sido desarrolladas las actividades requeridas en dicha opción, el Consejo Académico Divisional, conformado por los integrantes que firman al calce, ha determinado que SI (X) NO () han sido cubiertos en su totalidad los requisitos que permitan a esta División Académica liberar al alumno de los trámites académicos para tal efecto e iniciar el trámite de titulación en la Unidad de Registro y Certificación de este Tecnológico de Estudios Superiores de Ecatepec.

Observaciones:

Consejo Académico Divisional	
M. EN C. IRVING ARDIEL AEBECER GUILLERMO Presidente de la Academia de Ciencias Básicas	LIC. MARIA REBECA GARCÍA PONCE Presidenta de la Academia de Ciencias Sociales
M. EN D. DULCE ANAHÍ FLORES NUÑEZ Presidenta de la Academia de Ciencias Económicas Administrativas	M. EN C. ROSA MARÍA HERNÁNDEZ ROJAS Presidenta de la Academia de Ciencias de la Ingeniería
I.S.C. FRANCISCO JAVIER HERNÁNDEZ BARAJAS Presidente de la Academia de Diseño de la Ingeniería	M. EN C. XÓCHITL-RAQUEL WONG COHÉN Academia de Ingeniería Aplicada

Elaboró: M. EN C. MERCEDES FLORES FLORES SECRETARIO DEL CONSEJO ACADÉMICO DIVISIONAL	Revisó y Visto Bueno: ING. JOSÉ LUIS CHÁVEZ ROSAS ENCARGADO DE LA DIVISIÓN ISC
--	--

DEDICATORIAS

Para mi papa:

Por todos los consejos, platicas y ánimos que me brindaste durante mi carrera universitaria. Por la confianza que me tuviste desde un inicio hasta el final de este camino. Por todo lo que me has dado para brindarme un mejor futuro.

Para mi mama:

Por todo lo que me brindaste a lo largo de mi carrera universitaria. Por todas las palabras de aliento y pláticas que me diste para que yo pudiera ser mejor persona y lograr esta hermosa meta de ser ingeniero en sistemas. Por todo en lo que me apoyaste para que fuera una persona ejemplar para mis hermanos.

Para mi hermana:

Por todas las pláticas que me diste a lo largo de mi carrera que me ayudaron a ser una mejor persona dentro y fuera de la escuela. Por todo el apoyo brindado en diferentes ocasiones para que pudiera sobresalir entre los demás.

Para mi hermano:

Por todos los consejos que me diste en mi carrera que me dieron la fortaleza para seguir estudiando y superarme como persona. Por todo el apoyo dado hacia mi persona para poder sobresalir entre mis compañeros de clase.

A mi Asesor y excelente M en A.D.N Juan Manuel Stein Carrillo:

Porque desde que estoy trabajando con usted me brindó su apoyo, su bondad y sobre todo su amistad. Por ser un maestro EJEMPLAR no solo para mí, sino para muchos compañeros en la carrera de ISC, y por apoyarme en todo momento en mi estancia en la carrera. ¡MUCHAS GRACIAS!

AGRADECIMIENTOS

Quiero agradecer principalmente a mis padres, por todo su apoyo, cariño y comprensión a lo largo de mi carrera universitaria. También agradecerles por su paciencia que me han tenido en todo mi camino en la universidad.

Agradecerles a mis hermanos, por todo su apoyo y todos sus consejos que me ayudaron a ser mejor persona y a desenvolverme mejor dentro y fuera de la universidad.

El agradecimiento también va hacia mi familia en general, quienes fueron un pilar muy importante en toda mi carrera universitaria, quienes con sus palabras de aliento me ayudaron a salir adelante.

Agradecerle principalmente al profesor M en ADN. Juan Manuel Stein Carrillo quien, con su apoyo fundamental, pude llevar a cabo la elaboración de la tesis, así como también agradecerles a los profesores DR. Adolfo Meléndez Ramírez y DR. Francisco Jacob Ávila Camacho quienes me apoyaron a lo largo de mi carrera universitaria.

Agradecerles también a mis amigos que tuve en la carrera universitaria, quienes con sus consejos y palabras de motivación me ayudaron a ser mejor persona y superarme en muchos ámbitos de mi vida.

índice

OFICIO	2
AGRADECIMIENTOS	4
RESUMEN.....	8
INTRODUCCION.....	10
CAPITULO I ANTECEDENTES.....	11
1.1 Justificación	11
1.2 Objetivos.....	13
1.2.1 Objetivo General	13
1.2.2 Objetivos Específicos	13
1.3 Alcances y limitaciones	14
1.3.1 Alcances.....	14
1.3.2 Limitaciones	14
1.4 Hipótesis.....	14
1.5 Planteamiento del problema.....	15
CAPITULO II MARCO TEORICO	20
2.1 Tránsito, Transporte y Movilidad	20
2.1.1 Sistemas de tráfico.....	21
2.2 Tecnologías de Detección de motocicletas	22
2.3 Machine Learning y Deep Learning	26
2.4 Aprendizaje Profundo.....	33
2.5 Binarización y Umbralización	38
2.6 Clasificador KNN.....	39

2.7 Detección de objetos.....	41
CAPITULO III METODOLOGÍA	44
3.1 Uso de MatLab	45
3.3 Redes neuronales profundas pre-entrenadas	51
3.4 MatLab y AlexNet	51
3.6 Aprendizaje de características, capas y clasificación	56
CAPITULO IV ANÁLISIS E INTERPRETACIÓN DE RESULTADOS.....	57
4.1 Carga de datos	57
4.2 Dividir datos	58
CONCLUSIONES.....	64
RECOMENDACIONES	65
BIBLIOGRAFIA.....	66
ANEXOS.....	71
GLOSARIO DE TERMINOS DE IA	71

Índice de Figuras

Figura 1. Defunciones por tipo de usuario.....	8
Figura 2. Equipo de protección de un motociclista.....	9
Figura 3. Identificación de la problemática por parte del autor.....	18
Figura 4. Técnicas basadas en sensores Activos	23
Figura 5. Técnicas basadas en sensores Pasivos	23
Figura 6. Redes neuronales en un cerebro humano	26
Figura 7. Aprendizaje supervisado	28
Figura 8. Aprendizaje no supervisado	29
Figura 9. Modelo por refuerzo	30
Figura 10. Machine Learning vs Deep Learning.....	31
Figura 11. Aprendizaje profundo por capas	33
Figura 12. Modelo Raster vs Modelo Vectorial	34
Figura 13. Frameworks más utilizados.....	35
Figura 14. Modelo de la red neuronal AlexNet.....	36
Figura 15. Comparación de CPU vs GPU	36
Figura 16. Capas de la red neuronal AlexNet	37
Figura 17. Ejemplos de imágenes capturadas según la propuesta	45
Figura 18. Deep Network Designer	48
Figura 19. Reuse Pretrained Network	51
Figura 20. Entradas y salidas de la red neuronal	52
Figura 21. Capas de AlexNet.....	54
Figura 22. Metodología de clasificación de imágenes en AlexNet.....	57
Figura 23. Análisis de las imágenes.....	60
Figura 24. Graficas de resultados arrojados por la red neuronal AlexNet	61
Figura 25. Resultado del análisis de las imágenes.	62
Figura 26. Identificación de autos, motociclistas y cascos.....	64

RESUMEN

En el Estado de México, los motociclistas y las muertes de peatón son las primeras víctimas de accidentes de tránsito como se puede ver en la siguiente

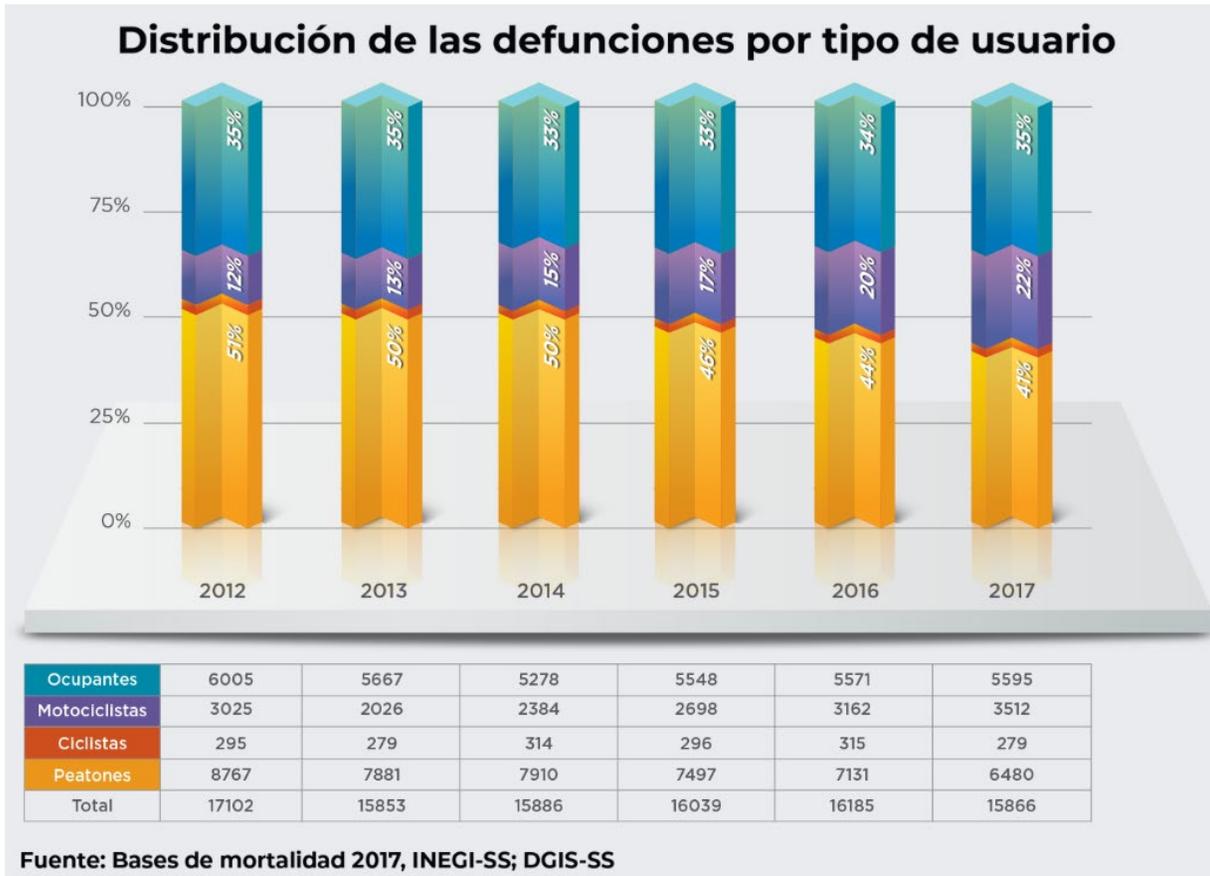


tabla.

Figura 1. Defunciones por tipo de usuario

Entre 2012 y 2017, en 2018 se contabilizaron 16807 motociclistas fallecidos, la cantidad ha venido incrementándose derivado al mayor uso de este medio de transporte en el Estado de México convirtiéndose en un problema que ha ido en aumento derivado de que más personas tienen la facilidad de adquirir una motocicleta.

Cuando un motociclista se ve involucradas en un accidente y este no cuente

con su equipo de protección, puede verse incapacitado temporal o en el peor de los casos, en ausencia de este, derivado de su fallecimiento.



Figura 2. Equipo de protección de un motociclista.

Para solucionar este problema, se han propuesto algoritmos para detectar el uso de motocicletas y cascos de motocicleta utilizando técnicas de procesamiento de imágenes combinadas con modelos de inteligencia artificial.

Los modelos de visión artificial no pueden detectar, reconocer y clasificar imágenes sin la ayuda de la tecnología de reconocimiento de imágenes. Por lo tanto, el software de reconocimiento de imágenes basado en IA debe poder decodificar imágenes y realizar análisis predictivos. Para hacer esto, los modelos de IA se entrenan utilizando conjuntos de datos masivos para producir predicciones precisas.

De estos elementos expuestos surge la necesidad de poder incentivar a los motociclistas a utilizar el equipo de protección y de esa manera reducir el número de accidentes viales, haciendo uso de indicadores en los semáforos que permitan hacer que los motociclistas porten adecuadamente por lo menos el casco.

INTRODUCCION

Según la Organización Mundial de la Salud (OMS), los accidentes de tráfico se cobran la vida de 1,2 millones de personas cada año y son la principal causa de muerte entre los jóvenes de 15 a 29 años en todo el mundo. De estas muertes, el 23% fueron motociclistas, el 22% peatones y el 4% ciclistas.

En otras palabras, según el Informe Mundial de Seguridad Vial de la OMS, el 49% de todas las muertes en carretera ocurren en las vías públicas más vulnerables. En 2017, las autoridades registraron 6.480 accidentes con peatones y 279 ciclistas. A modo de comparación, hubo 6.759 accidentes que no involucraron motocicletas.

De un análisis previamente publicado, se puede concluir que los accidentes entre ciclistas y ciclistas en el Estado de México son un grave problema, provocando externalidades negativas a nivel socioeconómico de la sociedad, reduciendo el ingreso y riqueza de los hogares de clase media a nivel micro, el grado en que el miembro más productivo de la familia queda incapacitado temporal o permanentemente de por vida, generalmente como resultado de un accidente de tránsito o, en el peor de los casos, si no hay accidente de tránsito. falleció en un accidente por incumplimiento.

Los seres humanos tienen una capacidad innata para distinguir e identificar con precisión objetos, personas, animales y lugares a partir de fotografías. Sin embargo, las computadoras no están equipadas para clasificar imágenes. Sin embargo, pueden ser entrenados para interpretar información visual utilizando aplicaciones de visión por computadora y técnicas de reconocimiento de imágenes.

El reconocimiento de imágenes, un subcampo de la IA y la visión artificial incorpora técnicas de aprendizaje profundo para admitir muchos casos de uso del mundo real. La inteligencia artificial se basa en la visión artificial para percibir con precisión el mundo.

CAPITULO I ANTECEDENTES

1.1 Justificación

En el caso de México, los ciclistas, peatones y motociclistas representan el 60% de todas las fatalidades de tránsito. La Organización Mundial de la Salud ha identificado cinco factores principales que aumentan el riesgo de accidentes de tránsito:

- Exceso de velocidad
- Conducir ebrio
- Los motociclistas no usan cascos
- No use cinturones de seguridad
- No utilizar asientos para niños

Para los motociclistas, el uso correcto de un casco certificado (según los estándares DOT y ECE) reduce el riesgo de muerte en un accidente en un 40% y reduce las lesiones graves en aproximadamente un 70%. Por lo tanto, es importante no solo usar un casco, sino también asegurarse de que su calidad esté certificada.

La Nom-206-scfi/ssa2-2018, nos sugiere "Cascos de seguridad para la prevención y atención inmediata de lesiones en la cabeza de motociclistas

acciones de promoción de la salud especificaciones de seguridad y métodos de prueba, información comercial y etiquetado".

Según la Organización Mundial de la Salud, los cascos cumplen tres funciones:

- Reduce la desaceleración del cráneo, reduciendo así el movimiento del cerebro, absorbiendo el impacto. El material blando integrado en el casco absorbe parte del impacto y frena la cabeza. Esto significa que el cerebro no golpea el cráneo con tanta fuerza.
- Distribuye la fuerza del impacto sobre un área más grande para que no se concentre en un área específica del cráneo.
- Actúa como una barrera mecánica entre la cabeza y el objeto, evitando el contacto directo entre el cráneo y el objeto de impacto.

Los motociclistas que no usan casco corren un riesgo mucho mayor de sufrir lesiones en la cabeza o ambas cosas. En su guía de seguridad vial para formuladores de políticas y profesionales del casco, la OMS señala que los cascos agregan una capa adicional de protección a la cabeza, previniendo algunos de los peores tipos de lesiones cerebrales traumáticas. "Es probable que los motociclistas sean un grupo con una morbilidad y mortalidad relativamente alta en el futuro.

Por lo tanto, la Organización Mundial de la Salud ha llamado la atención sobre las necesidades especiales de los usuarios de la carretera más vulnerables", dicen los investigadores del INSP. Para prevenir lesiones y accidentes de tráfico: Fomentamos el uso de cascos homologados, el cumplimiento de los límites de velocidad y el no beber y utilizar el móvil durante la conducción. Usa tu cabeza, usa tu casco.

1.2 Objetivos

1.2.1 Objetivo General

Realizar un sistema de visión artificial que permita identificar a motociclistas que no porten casco como elemento principal de seguridad pasiva con el fin de alertar al motociclista al llegar a un semáforo a que lo utilice y se reduzcan las lesiones en caso de que se produzca un incidente.

1.2.2 Objetivos Específicos

- Detectar objetos en el flujo de tráfico.
- Clasificación binaria para la identificación de motocicletas.
- Encuentra y extrae motocicletas de imágenes en bruto.
- Extraiga el ROI (Región de interés) del motociclista y cree un contorno del área de la cabeza del conductor y la coordenada H del ROI en la pila HSB (Tono, Saturación y Brillo).
- Clasificación binaria de estas imágenes para identificar el objetivo del casco.
- Identificación de redes neuronales.
- Creación de un entorno de retorno de la inversión.
- Reconocimiento de casco mediante entrenamiento de redes neuronales.

1.3 Alcances y limitaciones

1.3.1 Alcances

- Identificar los elementos del sistema de reconocimiento facial y el método de entrenamiento de redes neuronales utilizados en el proyecto.
- Implementar un prototipo que detecte cuando un motociclista lleva casco.
- Usar bases de datos de imágenes para volver a entrenar una red neuronal para que aprenda a reconocer objetos de manera correcta y precisa.
- Realizar las pruebas necesarias para publicar y validar los datos más precisos.
- Activa la alarma si el motociclista no lleva casco.

1.3.2 Limitaciones

- Este método no funcionará si las personas usan sombreros, gorras u otros artículos que no cambien drásticamente la tez.
- La distancia debe estar normalizada para trabajar a cierta distancia antes del punto de recepción.
- Se debe utilizar un conjunto de datos de al menos 1000 imágenes para mejorar la precisión de la identificación de motociclistas y cascos.

1.4 Hipótesis

La identificación de motociclistas que utilizan casco a través de técnicas de visión y reconocimiento artificial permitirá que se reduzcan el índice de mortalidad y accidentes graves en el Estado de México.

1.5 Planteamiento del problema

Resulta necesaria la aplicación de nuevas tecnologías para el análisis del tráfico a nivel local, con el fin de obtener resultados más confiables a partir de algoritmos más robustos. El análisis de tráfico mediante visión artificial comprende un área de investigación de importancia significativa para el desarrollo sostenible de la movilidad en zonas urbanas.

Para tal fin, se han creado los Sistemas Inteligentes de Transporte (ITS). Estos son aplicaciones avanzadas que combinan sistemas electrónicos, de comunicaciones, de computadores y sensores. Dichos ITS integran vehículos, personas e información de la vía con estrategias de gestión del tráfico para proveer información en tiempo real para aumentar la seguridad, la eficiencia y el confort en los sistemas de tráfico tanto públicos, como privados.

La Inteligencia Artificial (IA) es la combinación de algoritmos planteados con el propósito de crear máquinas que presenten las mismas capacidades que el ser humano. Una tecnología que todavía nos resulta lejana y misteriosa, pero que desde hace unos años está presente en nuestro día a día a todas horas. Los algoritmos son un conjunto de instrucciones o reglas definidas y no-ambiguas, ordenadas y finitas que permiten, típicamente, solucionar un problema, realizar un cómputo, procesar datos y llevar a cabo otras tareas o actividades.

En su forma más básica, el aprendizaje automático utiliza algoritmos programados que reciben y analizan datos de entrada para predecir los valores de salida dentro de un rango aceptable. A medida que se introducen nuevos datos en estos algoritmos, aprenden y optimizan sus operaciones para mejorar el rendimiento, desarrollando “inteligencia” con el tiempo.

Existen una gran variedad de técnicas para la detección, el reconocimiento y clasificación de Motocicletas en el tráfico, la mayoría de ellas se soporta en la extracción de características propias de las motocicletas, tales como elementos circulares, HOG (histograma de gradientes) para el color o la forma, la relación de aspecto y la estimación de las áreas dentro de regiones de interés.

En su tesis doctoral de 1963, Lawrence Roberts analizó la posibilidad de extraer información 3D a partir de vistas 2D de poliedros (Roberts, 1963). Este trabajo más tarde ganó reconocimiento como uno de los primeros defensores de la visión artificial (Huang, 2003). Poco después, varios investigadores del Instituto Tecnológico de Massachusetts (MIT) se basaron en el trabajo de Roberts y utilizaron un método simple para aplicar la visión a bloques geométricos.

Los investigadores involucrados tardaron meses en descubrir que era necesario estudiar imágenes capturadas del mundo real en lugar de mirar vistas de caras poliédricas. Esto marcaría el comienzo oficial de lo que ahora se llama visión artificial y su uso para resolver diversos problemas de fabricación e investigación.

La tarea de reconocimiento y detección de objetos ha sido y es un área importante de investigación en visión artificial. El primer detector de rostros exitoso fue presentado en 2001 por James Viola y Michael Jones (Viola, 2001). En 2005, el trabajo de investigación de Navneet y Triggs revolucionó la detección de peatones mediante histogramas orientados a gradientes (HOG) (Dalal & Triggs, n.d.).

A través de esta investigación se demuestra el potencial de las aplicaciones de visión artificial en temas relacionados con el análisis de tráfico. Con la llegada

de los algoritmos de aprendizaje profundo, aumenta la necesidad de detectar objetos (principalmente personas y vehículos) de forma más rápida y precisa. Introducido en 2012, AlexNET es un algoritmo para reconocer objetos en una escena con una precisión de análisis del 85 % (Krizhevsky & Hinton, n.d.).

En los años siguientes, aparecieron algoritmos con diferente velocidad de procesamiento y precisión de análisis, como RCCNN, Fast RCNN, Mask RCNN, SSD y YOLO. Constituyen la tendencia actual como herramienta para el desarrollo de sistemas de detección y reconocimiento mediante visión artificial en diversos campos de investigación, especialmente en temas relacionados con el tráfico de vehículos.

El uso de algoritmos de aprendizaje profundo ha permitido el desarrollo de varios sistemas de análisis de tráfico en tiempo real y fuera de línea. Algunos sistemas en tiempo real utilizan dos unidades principales de procesamiento. Un ejemplo es el análisis conjunto entre redes neuronales de alta fidelidad y filtros Kalman para detección y conteo de vehículos (Yang & Qu, 2017).

Adicionalmente se han presentado también sistemas tolerantes a perturbaciones, oclusiones y cambios en enfoque mediante uso de arquitecturas en red (Kim & Sim, 2017), o sistemas enfocados al estudio de tendencias de flujo de tráfico en zonas conflictivas con el fin de aumentar la seguridad vial (Battiato, Farinella, & Giudice, 2013).

A nivel local se ha realizado una serie de estudios relacionados con el análisis del tráfico vehicular, más no a la utilización de visión artificial y algoritmos de aprendizaje profundo para tratar la temática de estudio. El presente trabajo de investigación propone un sistema de análisis de tráfico basado en visión artificial. La determinación de parámetros tales como el flujo vehicular total

constituyen, un requisito imprescindible al diseño y ejecución de obras de vialidad, a su vez, del desarrollo ordenado y sistemático de una urbe.

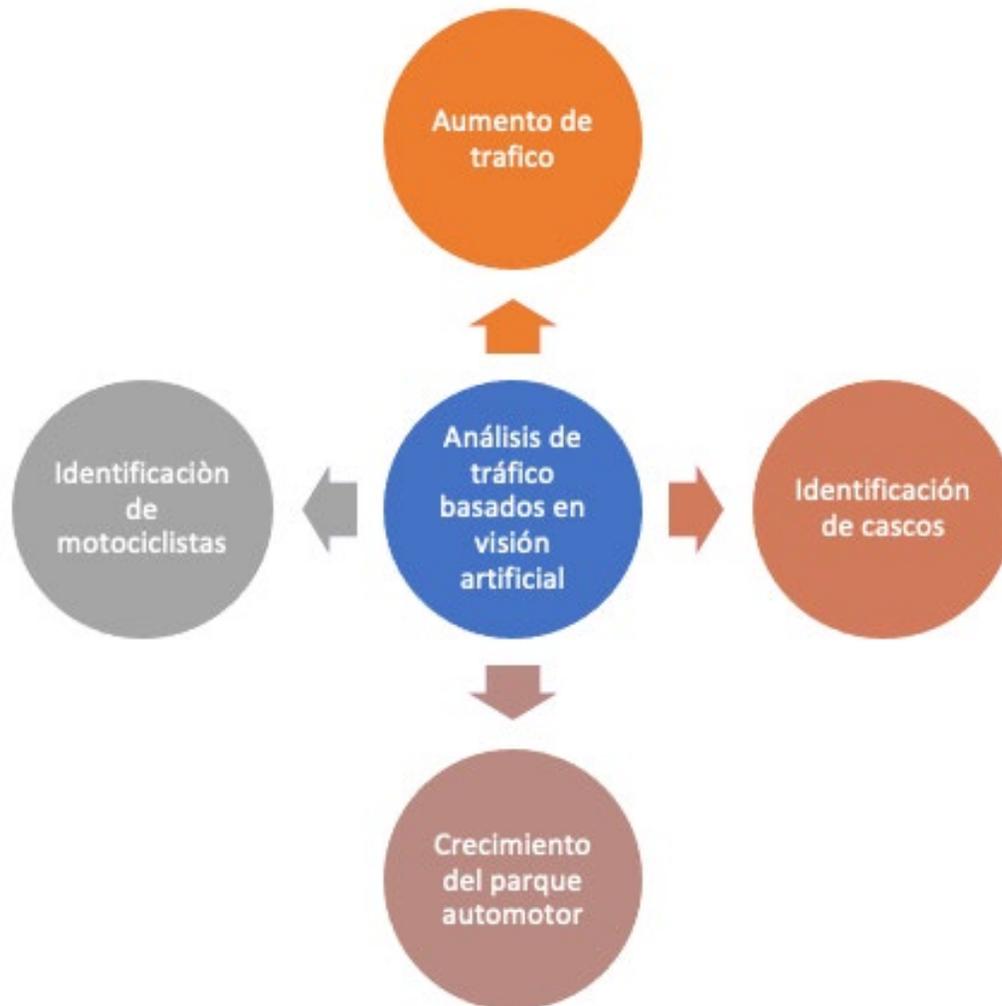


Figura 3. Identificación de la problemática por parte del autor.

Por lo que la metodología planeada consiste en los siguientes pasos para poder llevar a cabo la detección de una motocicleta y de quien la maneje use casco:

- i. Detección de objetos en el flujo vehicular.
- ii. Clasificación binaria para identificación de las motocicletas.
 - i. La clasificación binaria es la tarea de clasificar los elementos de un conjunto

- ii. en dos grupos sobre la base de una regla de clasificación.
- iii. Ubicación y Extracción de la moto desde la imagen original.
- iv. Extracción del ROI (región de interés) del motociclista y construcción del contorno de la zona de la cabeza del conductor.
- v. Clasificación binaria de estas imágenes para identificación del uso del casco.

CAPITULO II MARCO TEORICO

2.1 Tránsito, Transporte y Movilidad

El tráfico es uno de los problemas asociados al desarrollo urbano en el Estado de México, razón por la cual las autoridades locales se interesaron en este estudio. Las soluciones y los esfuerzos para mejorar la movilidad en las ciudades deben abordarse mediante un análisis de tráfico confiable, y el uso de sistemas de visión artificial proporciona una solución menos disruptiva y altamente confiable para este problema.

Los sistemas tradicionales de conteo de vehículos basados en el uso de regiones de interés y la sustracción de antecedentes están siendo reemplazados por algoritmos de aprendizaje profundo. La versatilidad de este algoritmo permite sistemas de conteo de vehículos en tiempo real (Yang y Qu, 2017). Otras obras utilizan detectores y drones para identificar y rastrear peatones o vehículos (Saribas & Cevikalp, 2018).

Sin embargo, en algunos casos especiales, los métodos tradicionales de conteo de vehículos funcionan junto con las redes neuronales para reducir los efectos no deseados del cruce de carril (o el ingreso parcial a las zonas de detección) (Zheng & Wang, 2015).

A nivel local, se deben aplicar nuevos métodos de análisis de tráfico para obtener resultados más confiables a partir de algoritmos más robustos. El análisis del tráfico mediante visión artificial es un área de investigación que tiene grandes implicaciones para el desarrollo sostenible de la movilidad en las zonas urbanas. (Yang y Qu, 2017).

2.1.1 Sistemas de tráfico

El sistema de control de tráfico consiste principalmente en la emisión de permisos

Un centro de control que detecta condiciones anormales recibe información de los detectores y toma las decisiones apropiadas y los medios para comunicar esas decisiones al conductor, generalmente utilizando variables y señales de control remoto.

Estos sistemas forman parte de los denominados ITS (Sistemas Inteligentes de Transporte). ITS se define como la combinación de tecnologías de la información y la comunicación con vehículos y redes que mueven personas y mercancías. Inteligente porque proporciona a los viajeros y operadores conocimientos adicionales. Los sistemas ITS en los automóviles ayudan a los conductores a navegar y evitar atascos y colisiones. En trenes y autobuses, permiten a los conductores optimizar las operaciones de la flota y proporcionar a los pasajeros servicios automatizados de emisión de billetes e información en tiempo real. En las redes viales, los ITS coordinan las señales de tráfico, registran y procesan los accidentes, transmiten información, dirigen y guían a los conductores.

Para ello se crearon los Sistemas Inteligentes de Transporte (ITS). Estas son aplicaciones avanzadas que combinan electrónica, comunicaciones, computadoras y sistemas de sensores. Estos ITS combinan información sobre vehículos, personas y carreteras con estrategias de gestión del tráfico para proporcionar información en tiempo real para mejorar la seguridad, la eficiencia y la comodidad de los sistemas de transporte público y privado. (Roess, Prassas y McShane, 2004). En particular, los diferentes niveles de desarrollo tecnológico de los ITS hacen posible la detección de motocicletas mediante varios métodos.

Entre ellos, radar, sensor lidar, sensor de sonido y sensor óptico. Cabe señalar que identificar motocicletas es una tarea bastante difícil, ya que estos vehículos tienen grandes variaciones en forma, color y tamaño.

A esto se suma el entorno de tráfico altamente congestionado, los cambios de iluminación y las interacciones aleatorias entre los diversos usuarios de la vía que participan en el tráfico diario de la ciudad. Por otro lado, diversas tecnologías de detección e identificación de motocicletas enfrentan dos desafíos principales: la potencia informática (dado que la mayoría de las motocicletas se montan en la carretera) y la confiabilidad de la detección. (Amir, Likun y Tang, 2013).

2.2 Tecnologías de Detección de motocicletas

Las tecnologías para la detección de motocicletas se pueden dividir en dos grupos (Herbert, 2000):

1. Técnicas basadas en sensores Activos
2. Técnicas basadas en sensores Pasivos

Dentro de estas técnicas se encuentran los sensores que utilizan tecnología acústica, de radar o de laser -Lidar. Se llaman activas en la medida que los sensores aquí descritos identifican los objetos con la reflexión de la señal emitida por ellos. LiDAR son las siglas en inglés de *Light Detection and Ranging*, un término similar al radar (*Radio Detection and Ranging*, en inglés)



Figura 4. Técnicas basadas en sensores Activos

Los sensores ópticos por lo general son los que se conocen como sensores pasivos, sensores **fotográficos**, óptico-electrónicos.



Figura 5. Técnicas basadas en sensores Pasivos

Para la detección de elementos a través de inteligencia artificial se incluyen sensores que utilizan tecnología acústica, de radar o láser: lidar. Se denominan activos porque los sensores aquí descritos reconocen objetos reflejando la señal emitida por ellos. Su principal ventaja es que no requieren una gran potencia informática para medir directamente determinados indicadores (distancia, velocidad relativa).

La identificación y clasificación de motocicletas se realiza calculando los volúmenes de los reflectores utilizando técnicas acústicas. La tecnología LiDAR

utiliza las características del sensor, la geometría del punto de conexión del sensor y la información de seguimiento del vehículo para inferir si el vehículo identificado es una motocicleta.

El sensor ha demostrado su eficacia en la detección precisa de motocicletas, a partir de las cuales se pueden obtener las medidas de tráfico correspondientes. (Ling, Gibson y Middleton, 2013) (Blanc, Aufrere, Malaterre y Alison, 2004) (Nashashibi y Bargeton, 2008).

Además, esta técnica se ha utilizado para identificar cascos de moto. Este es el caso de (Sasikala, Padol, Katekar, & Dhanasekaran, 2015) donde se utiliza un sistema de identificación por radiofrecuencia (RFID) que se fija en el interior del casco para el seguimiento de las autoridades de tráfico y se instalan lectores a lo largo del mismo. Sin embargo, este enfoque no es una solución rentable para las autoridades de tránsito que intentan controlar el uso del casco.

Los sensores ópticos a menudo se denominan sensores pasivos porque recopilan datos sin emitir ninguna señal. (Herbert, 2000). Los sensores ópticos implican el uso de cámaras, que en algunos casos son menos costosas que los sensores de tecnología activa. Una de las principales ventajas de este tipo de tecnologías es que la detección se realiza independientemente de los cambios que se puedan realizar en la infraestructura, lo que puede prolongar la vida útil de dichos dispositivos.

Hace mucho tiempo que la inteligencia artificial dejó el reino de la ciencia ficción y entró en nuestras vidas. Aunque todavía está en sus primeras etapas, está destinado a causar una revolución comparable a la creada por Internet. Su uso en campos tan diversos como la salud, las finanzas, el transporte o la educación ha llevado a la UE a desarrollar su propia legislación sobre robots.

La inteligencia artificial (IA) es una combinación de algoritmos propuestos para crear máquinas con habilidades similares a las humanas. Una tecnología que

nos sigue siendo lejana y misteriosa, pero que lleva varios años en nuestra vida cotidiana.

La IA se encuentra en el reconocimiento facial móvil, asistentes de voz virtuales como Apple Siri, Amazon Alexa o Microsoft Cortana, y se integra en nuestros dispositivos cotidianos a través de bots (abreviatura de robot) o aplicaciones móviles. Por ejemplo: Lyli Enlace externo, se abre en ventana nueva, personal shopper versión digital; Parla Enlace externo, se abre en ventana nueva., para ayudarnos a aprender idiomas; Ems Enlace externo, se abre en ventana nueva., a Hacer más llevadera la difícil tarea de encontrar un nuevo piso; o Gyant Enlace externo, se abre en ventana nueva., un asistente virtual de Facebook que emite "diagnósticos" médicos.

El objetivo de todos ellos: facilitar la vida de las personas. Los avances en inteligencia artificial ya están impulsando el uso de big data, ya que su capacidad para procesar grandes cantidades de datos y brindar beneficios operativos, comerciales y de comunicación la convierte en una tecnología crítica en las próximas décadas.

Los algoritmos son un conjunto de instrucciones o reglas definidas y no-ambiguas, ordenadas y finitas que permiten, típicamente, solucionar un problema, realizar un cómputo, procesar datos y llevar a cabo otras tareas o actividades.

Nuestras redes neuronales naturales nos ayudan a reconocer, clasificar e interpretar imágenes en función de nuestras experiencias pasadas, el conocimiento aprendido y la intuición. De la misma manera, una red neuronal artificial ayuda a las máquinas a identificar y clasificar imágenes. Pero primero necesitan ser entrenados para reconocer objetos en una imagen.

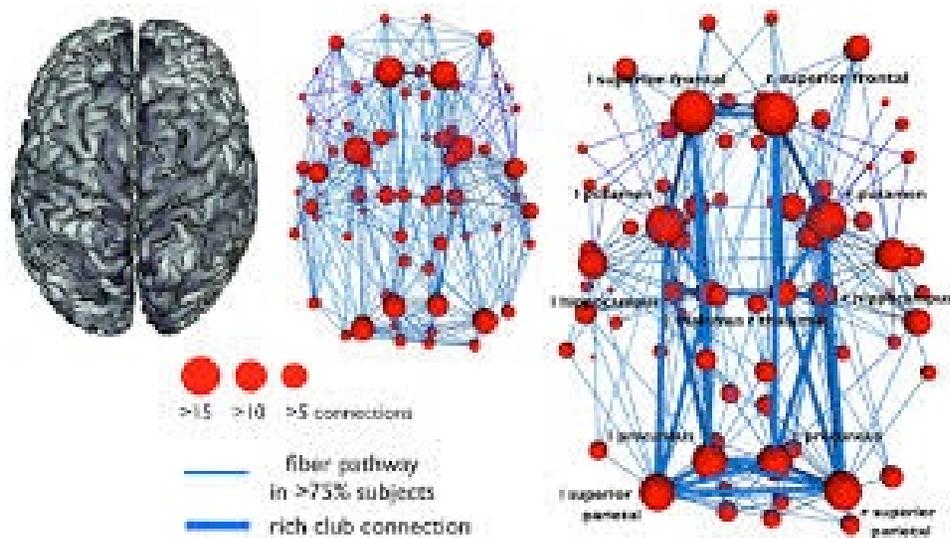


Figura 6. Redes neuronales en un cerebro humano

La red neuronal tiene que aprender a reconocer la variedad de objetos en las imágenes, y para ello necesitamos muchas imágenes. La visión artificial es un campo de investigación que lleva más de 40 años, por lo que hemos desarrollado varias aplicaciones técnicas.

Estas aplicaciones incluyen el Reconocimiento óptico de caracteres (OCR): implica reconocer automáticamente símbolos o caracteres pertenecientes a un alfabeto particular de una imagen y luego almacenarlos como datos. Inspección Robótica: Inspección rápida de piezas usando visión estereoscópica con iluminación especializada para asegurar la calidad de los componentes fabricados.

2.3 Machine Learning y Deep Learning

Una gran diversidad de técnicas de visión se ha propuesto para la detección de motocicletas. (Chung-Cheng, Min-Yu, & Hung-Tsung, 2007), aunque los términos se utilizan a veces como sinónimos, el Deep Learning (aprendizaje

profundo) y el Machine Learning (aprendizaje automático), no son lo mismo, siendo el primero un tipo particular del segundo.

El machine learning es la ciencia de desarrollo de algoritmos y modelos estadísticos que utilizan los sistemas de computación con el fin de llevar a cabo tareas sin instrucciones explícitas, en vez de basarse en patrones e inferencias. Los sistemas de computación utilizan algoritmos de machine learning para procesar grandes cantidades de datos históricos e identificar patrones de datos. Esto les permite generar resultados con mayor precisión a partir de un conjunto de datos de entrada. Por ejemplo, los científicos de datos pueden entrenar una aplicación médica para diagnosticar el cáncer con imágenes de rayos X a partir del almacenamiento de millones de imágenes escaneadas y diagnósticos correspondientes.

La idea central del machine learning es la existencia de una relación matemática entre cualquier combinación de datos de entrada y salida. El modelo de machine learning no conoce de antemano esta relación, pero puede adivinarla si se le dan suficientes conjuntos de datos. Esto significa que cada algoritmo de machine learning se crea en torno a una función matemática modificable. El principio subyacente puede entenderse así:

Entrenamos el algoritmo al darle las siguientes combinaciones de entrada y salida (e,s): (2,10), (5,19) y (9,31)

El algoritmo calcula que la relación entre la entrada y la salida es: $o=3*i+4$

A continuación, le damos la entrada 7 y le pedimos que prediga la salida. Puede determinar automáticamente la salida como 25.

Si bien se trata de conocimientos básicos, el machine learning se centra en el principio de que los sistemas de computación pueden relacionar

matemáticamente todos los puntos de datos complejos, siempre y cuando tengan suficientes datos y potencia de computación para procesarlos. Por lo tanto, la precisión de la salida está relacionada directamente con la magnitud de la entrada dada.

Los tres grupos de algoritmos del machine Learning son los siguientes:

1. Aprendizaje supervisado
2. Aprendizaje sin supervisión
3. Aprendizaje semisupervisado
4. Aprendizaje por refuerzo

En el aprendizaje supervisado, la máquina se enseña con el ejemplo, de este modo, el operador proporciona al algoritmo de aprendizaje automático un conjunto de datos conocidos que incluye las entradas y salidas deseadas, y el algoritmo debe encontrar un método para determinar cómo llegar a esas entradas y salidas.



Figura 7. Aprendizaje supervisado

Mientras el operador conoce las respuestas correctas al problema, el algoritmo identifica patrones en los datos, aprende de las observaciones y hace predicciones. El algoritmo realiza predicciones y es corregido por el operador, y este proceso sigue hasta que el algoritmo alcanza un alto nivel de precisión y rendimiento.

El entrenamiento se puede lograr mediante la repetición. Las computadoras deben recibir tantas imágenes identificadas o etiquetadas como sea posible. Por ejemplo, si quisiera enseñar a una computadora a identificar motocicletas, se le mostraría numerosas imágenes con motocicletas etiquetadas. Para etiquetar la motocicleta la computadora reconocería qué píxeles específicos tienen patrones o características de una motocicleta y luego asociaría esa estructura de píxeles con motocicletas. Este caso se conoce como Aprendizaje Supervisado, dentro del área del Aprendizaje Automático.

En un proceso de aprendizaje no supervisado, se deja que el algoritmo de aprendizaje automático interprete grandes conjuntos de datos y dirija esos datos en consecuencia. Así, el algoritmo intenta organizar esos datos de alguna manera para describir su estructura. Esto podría significar la necesidad de agrupar los datos en grupos u organizarlos de manera que se vean más organizados.

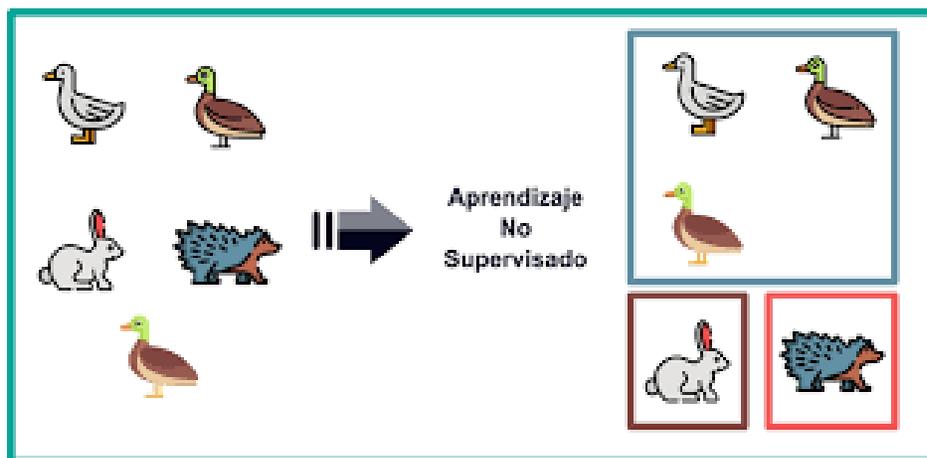


Figura 8. Aprendizaje no supervisado

En el aprendizaje semisupervisado, como su nombre indica, este método combina el aprendizaje supervisado y el no supervisado. Para entrenar los sistemas, esta técnica se basa en el uso de una pequeña cantidad de datos

etiquetados y de una gran cantidad de datos sin etiquetar. En primer lugar, los datos etiquetados se utilizan para entrenar parcialmente el algoritmo de machine learning. Después, el propio algoritmo entrenado parcialmente etiqueta los datos no etiquetados. Este proceso se denomina pseudoetiquetado. A continuación, el modelo se vuelve a entrenar con la mezcla de datos resultante sin programarlo explícitamente.

La ventaja de este método es que no necesita grandes cantidades de datos etiquetados. Resulta útil cuando se trabaja con datos como documentos largos que los humanos tardarían mucho en leer y etiquetar.

El aprendizaje por refuerzo se centra en los procesos de aprendizajes reglamentados, en los que se proporcionan algoritmos de aprendizaje automáticos con un conjunto de acciones, parámetros y valores finales.

Al definir las reglas, el algoritmo de aprendizaje automático intenta explorar diferentes opciones y posibilidades, monitorizando y evaluando cada resultado para determinar cuál es el óptimo. En consecuencia, este sistema enseña la máquina a través del proceso de ensayo y error.



Figura 9. Modelo por refuerzo

Ahora puede haber millones de estos patrones o funciones a los que queremos aplicar inteligencia artificial, en este caso redes neuronales, uno de los tipos de aprendizaje automático más populares en inteligencia artificial. En cuanto a la visión artificial, es un algoritmo que procesa y simula la corteza visual del ojo humano para reconocer varias características de entrada (esto ya lo hemos visto en el procesamiento de imágenes) y finalmente es capaz de reconocer objetos y finalmente "ver". Detecta puntos, líneas, curvas y se especializa hasta que puede reconocer formas complejas como rostros o contornos de animales.

El Deep Learning lleva a cabo el proceso de Machine Learning usando una red neuronal artificial que se compone de un número de niveles jerárquicos. En el nivel inicial de la jerarquía la red aprende algo simple y luego envía esta información al siguiente nivel. El siguiente nivel toma esta información sencilla, la combina, compone una información algo un poco más compleja, y se lo pasa al tercer nivel, y así sucesivamente.

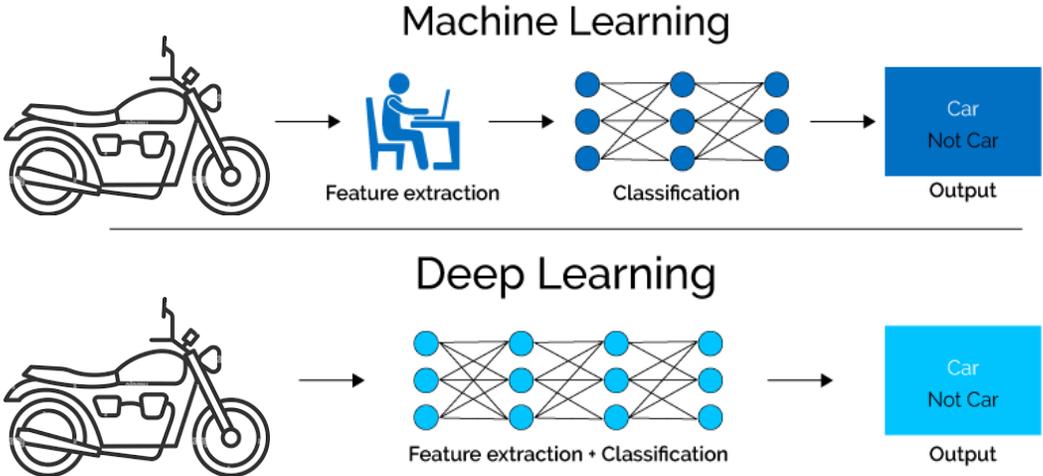


Figura 10. Machine Learning vs Deep Learning

Las redes neuronales que operan en dos o tres capas de neuronas conectadas se conocen como redes neuronales superficiales. Las redes de deep learning pueden tener muchas capas, incluso cientos de ellas. Ambas son técnicas de machine learning que aprenden directamente de los datos introducidos.

El deep learning atrae mucha atención últimamente, y hay razones de peso para ello. Está consiguiendo resultados que antes no eran posibles. El deep learning se adapta especialmente bien a aplicaciones de identificación complejas, tales como el reconocimiento facial, la traducción de textos y el reconocimiento de voz. También es una tecnología clave utilizada en sistemas y tareas avanzados de conducción asistida, que incluyen la asignación de carriles y el reconocimiento de señales de tráfico.

Si el resultado de un sistema es predecible, se considera determinista. La mayoría de las aplicaciones de software responden de forma predecible a la acción del usuario, por lo que se puede decir: "Si el usuario hace esto, obtiene aquello". Sin embargo, los algoritmos de machine learning aprenden mediante la observación y las experiencias. Por lo tanto, son de naturaleza probabilística. El enunciado cambia ahora a: "Si el usuario hace esto, hay un X % de posibilidades de que ocurra".

En el machine learning, el determinismo es una estrategia que se utiliza al aplicar los métodos de aprendizaje descritos con anterioridad. Cualquiera de los métodos de entrenamiento supervisados, no supervisados y otros se pueden hacer deterministas en función de los resultados deseados por la empresa. La pregunta de investigación, la recuperación de datos, la estructura y las decisiones de almacenamiento determinan si se adopta una estrategia determinista o no determinista.

2.4 Aprendizaje Profundo

A diferencia de Machine Learning, donde los datos de entrada se analizan mediante algoritmos, el Deep Learning (aprendizaje profundo) utiliza una red neuronal en capas. Hay tres tipos de capas involucradas que son: entrada, oculta y salida. La entrada de información es recibida por la capa de entrada, procesada por la capa oculta y los resultados generados por la capa de salida. Como las capas están interconectadas, cada capa depende de los resultados de la capa anterior.

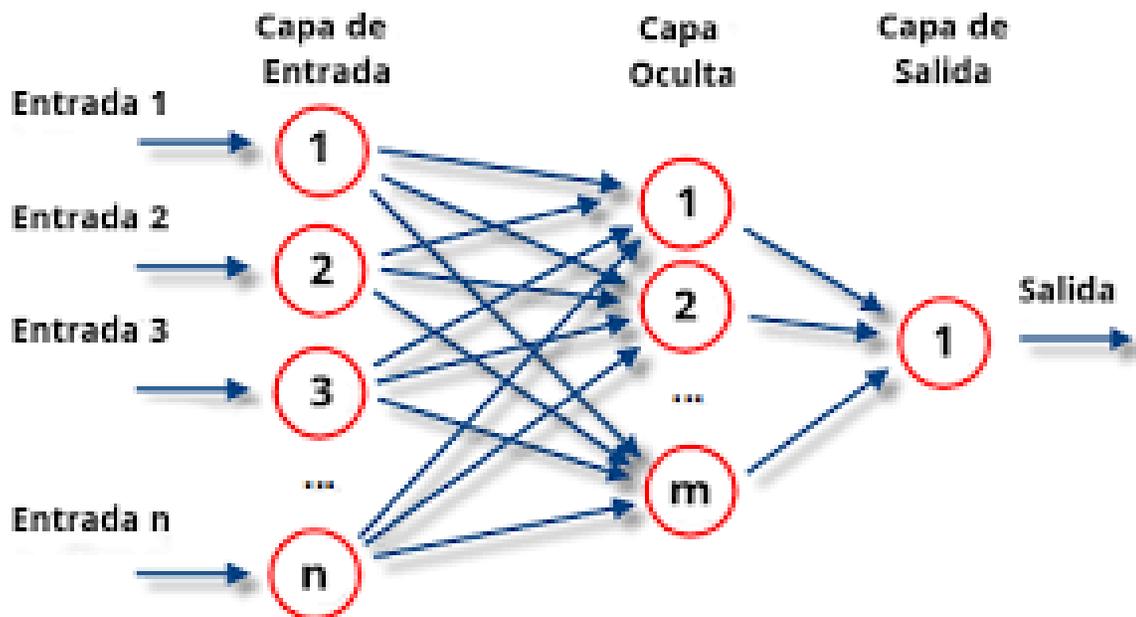


Figura 11. Aprendizaje profundo por capas

Por lo tanto, un gran conjunto de datos es esencial para entrenar una red neuronal para que el sistema de aprendizaje profundo se incline a imitar el proceso de razonamiento humano y continúe aprendiendo.

Una computadora ve y procesa una imagen de manera muy diferente a los humanos. Una imagen, para una computadora, es solo un montón de píxeles, ya sea como una imagen vectorial o rasterizada. En las imágenes de trama, cada píxel se organiza en forma de cuadrícula, mientras que en una imagen vectorial se organizan como polígonos de diferentes colores.

El aprendizaje profundo es un tipo de técnica de machine learning que se basa en el cerebro humano. Los algoritmos de aprendizaje profundo analizan los datos con una estructura lógica similar a la que utilizan los humanos. El aprendizaje profundo utiliza sistemas inteligentes, denominados redes neuronales artificiales, para procesar información por capas. Los datos fluyen desde la capa de entrada a través de varias capas de redes neuronales “profundas” ocultas antes de llegar a la capa de salida. Las capas adicionales ocultas permiten un aprendizaje mucho más eficaz que el de los modelos estándar de machine learning.

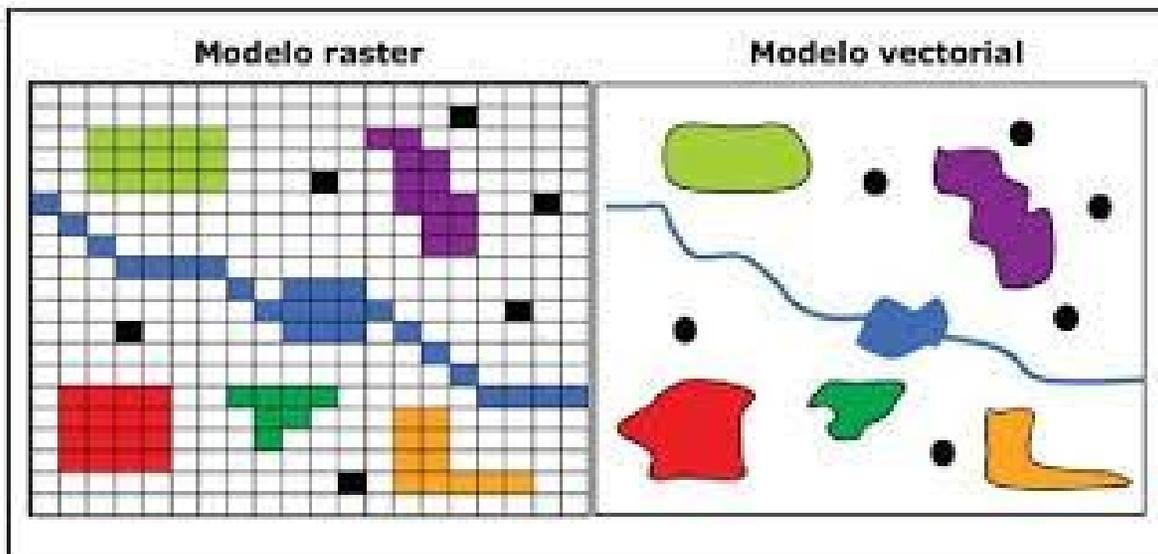


Figura 12. Modelo Raster vs Modelo Vectorial

Durante la organización de los datos, cada imagen se categoriza y se extraen las características físicas. Finalmente, la codificación geométrica se transforma en etiquetas que describen las imágenes. Esta etapa – reunir, organizar, etiquetar y anotar imágenes – es fundamental para el rendimiento de los modelos de visión artificial. Para el proceso del sistema de reconocimiento de imágenes se cuenta con tres procesos:

Proceso 1: conjuntos de datos de entrenamiento

Proceso 2: Entrenamiento de redes neuronales

Proceso 3: Prueba

Algunos Frameworks utilizados en la detección de objetos:

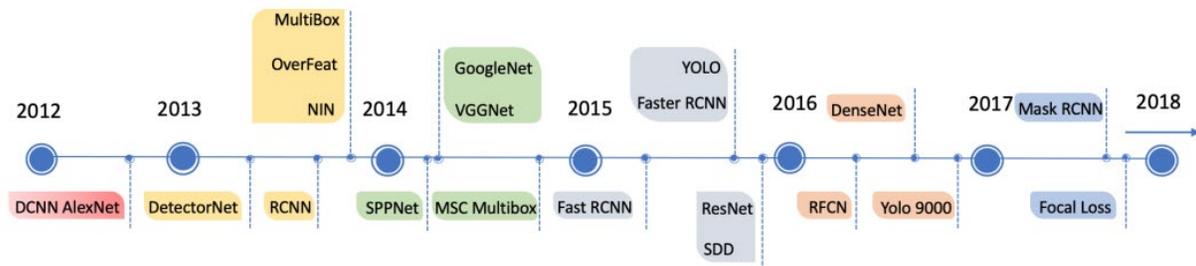


Figura 13. Frameworks más utilizados

En el año 2012, Alex Krizhevsky propuso, junto con su director de tesis Geoffrey Hinton, un nuevo modelo de red convolucional que logró reducir ese porcentaje hasta el 17%, a este nuevo modelo se conocería como AlexNet. La arquitectura de la red se compone de cinco capas convolucionales y tres capas densas. El *kernel* tiene dimensiones de 11×11 en la primera capa, 5×5 en la segunda, y 3×3 en el resto.

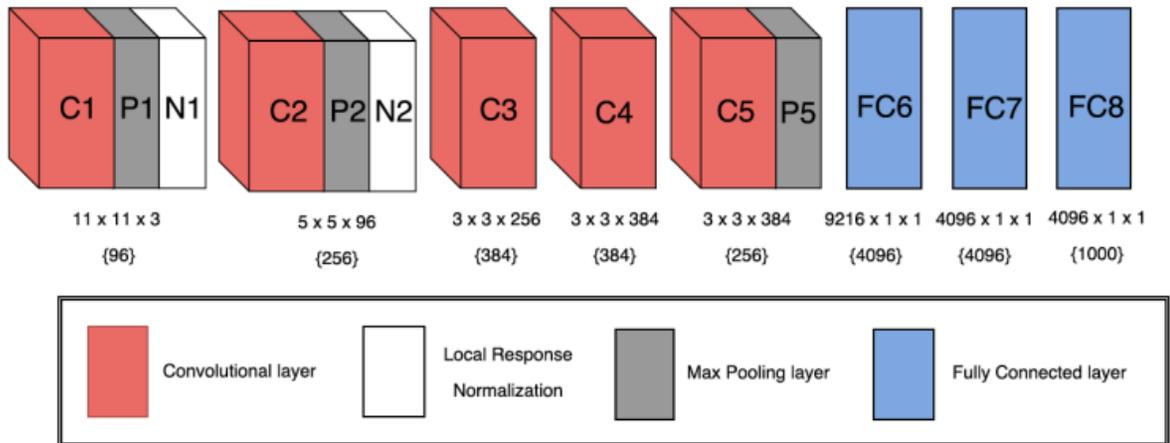


Figura 14. Modelo de la red neuronal AlexNet

La red neuronal es simulada por dos GPUs, GPU es el acrónimo de Graphics Processing Unit y representa precisamente el corazón de una tarjeta gráfica al igual que la CPU lo hace en un PC.

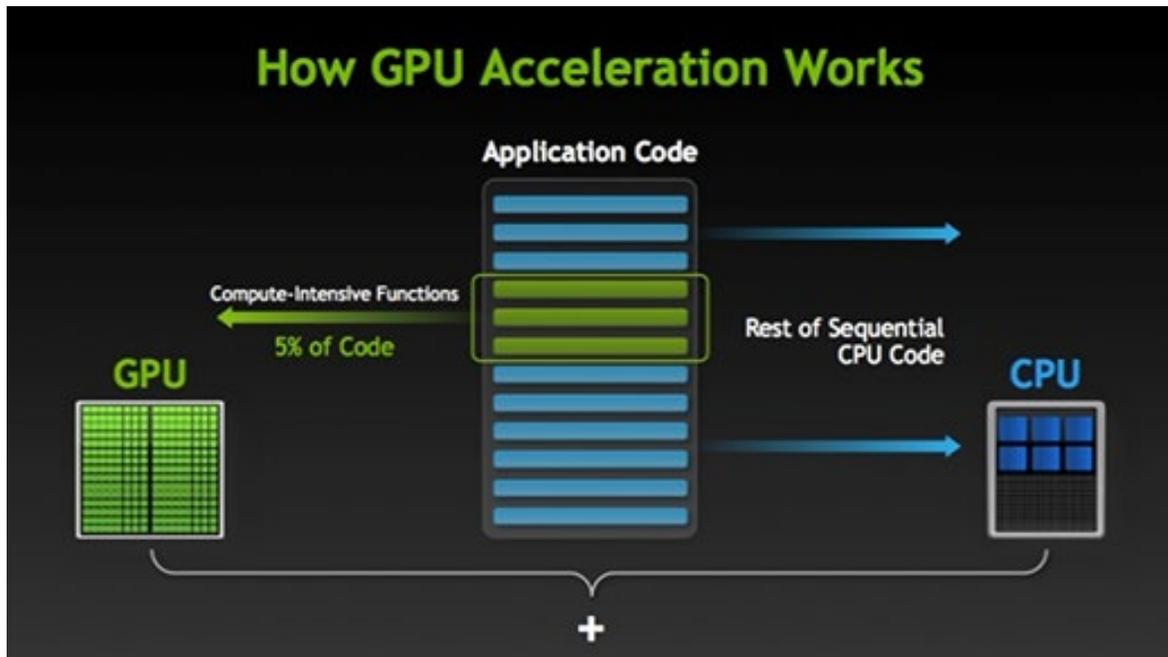


Figura 15. Comparación de CPU vs GPU

A cada una de ellas corresponde el cálculo de la mitad de los mapas de características de la capa convolucional de que se trate.

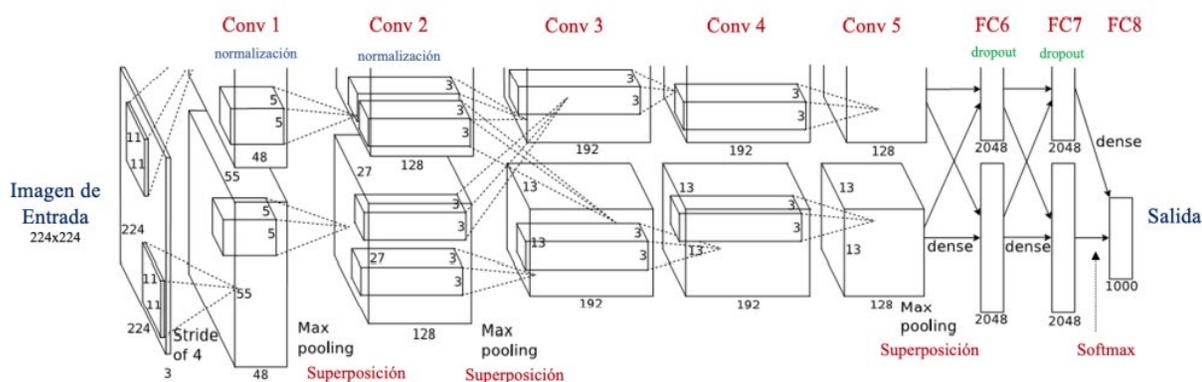


Figura 16. Capas de la red neuronal AlexNet

Como se describe en (Amir, Likun y Tang, 2013), existen varios métodos que se pueden utilizar para detectar, identificar y clasificar motocicletas en el tráfico. La mayoría de ellos están respaldados por la extracción de elementos de motocicletas, como elementos circulares, color o forma HOG (histogramas de gradiente), relación de aspecto y cálculo de área en la región de interés.

La literatura descrita cubre el período de 1985 a 2016, cuando se utilizan redes neuronales convolucionales para detectar tráfico en situaciones densas (Huynh & Hamamoto, 2016). En cierto sentido, muchos de estos métodos son inseguros porque no se pueden combinar con otras heurísticas de procesamiento de imágenes para el conteo de pasajeros y la detección del uso del casco para detectar y clasificar correctamente las motocicletas. Cabe señalar que solo la literatura que permite detectar y monitorear motocicletas con situaciones de oclusión es relevante en este artículo.

Hasta cierto punto, el conjunto de datos recopilados para la tesis contiene algunas muestras que otros usuarios de la vía, como automóviles u otras motocicletas, ocultan. (Chiu, Ku, & Chen, 2007) Chiu, propone un sistema donde se realice la segmentación del objeto en movimiento, sustrayendo la imagen del fondo. Hecho esto, se crean un “componente conectado” que permite la formación de cajas que encierran los objetos que se clasifican como motocicletas.

Luego determine si el objeto está ocluido por la longitud visual, el ancho y la relación de aspecto. Si realmente se produce un proceso de oclusión, la motocicleta se segmentará en modo de oclusión. Si un objeto cubierto es una motocicleta se determina buscando un descriptor específico de casco.

Si se detecta lo mismo, se utiliza un algoritmo para detectar y rastrear todos los objetos (oscurecidos y no oscurecidos) en la imagen que se clasifican como motocicletas. Para ello se utiliza un filtro de Kalman. El sistema identificó y clasificó correctamente las motocicletas con una precisión del noventa por ciento.

2.5 Binarización y Umbralización

Cabe mencionar que observar a dos o más pasajeros en un mismo vehículo es un fenómeno común en el Estado de México, siendo esta una de las actividades de seguridad vial más peligrosas que enfrenta esta urbe.

El trabajo presentado (Waranusast, Timtong, Bundon, & Tangnoi, 2014) es uno de los pocos que intenta contar el número de pasajeros por vehículo. Para detectar objetos en movimiento, primero procesan las imágenes en el conjunto

de entrenamiento con un filtro de paso bajo. Luego crean una imagen de fondo que se sustrae de los elementos en movimiento mediante operaciones simples de binarización y umbralización. Seguimos cerrando morfológicamente las imágenes para reducir el ruido.

La binarización es utilizada en el análisis de imágenes, mediante esta técnica se pueden separar objetos o regiones, que nos pueden interesar, del resto de la imagen. Es decir, que se puede separar el fondo de la imagen de la imagen de los objetos que deseamos analizar. (Aleph, 2021)

La binarización de una imagen digital consiste en convertir la imagen digital en una imagen en blanco y negro, de tal manera que se preserven las propiedades esenciales de la imagen. (Aleph, 2021)

La **binarización** consiste en una reducción de información en la **que** los únicos valores posibles son verdadero y falso (1 y 0). En caso de una imagen digital los valores verdadero y falso se corresponden a dos colores: blanco y negro. (Aleph, 2021)

La umbralización es uno de los más importantes métodos de segmentación. El objetivo es convertir una imagen en escala de grises a una nueva con sólo dos niveles, de manera que los objetos queden separados del fondo. (Aleph, 2021)

2.6 Clasificador KNN

K-Nearest-Neighbor es un algoritmo basado en instancia de tipo supervisado de Machine Learning. Puede usarse para clasificar nuevas muestras (valores discretos) o para predecir (regresión, valores continuos). Al ser un método

sencillo, es ideal para introducirse en el mundo del Aprendizaje Automático. Sirve esencialmente para clasificar valores buscando los puntos de datos “más similares” (por cercanía) aprendidos en la etapa de entrenamiento (ver 7 pasos para crear tu ML) y haciendo conjeturas de nuevos puntos basado en esa clasificación.

Ventajas:

- El algoritmo es simple y fácil de aplicar.
- No es necesario crear un modelo, configurar varios parámetros o formular hipótesis suplementarias.
- El algoritmo es polivalente. Puede ser utilizado para la clasificación o la regresión.

Desventajas:

- El algoritmo se vuelve más lento a medida que el número de observaciones aumenta y las variables independientes aumentan.

Al tratarse de uno de los algoritmos más simples de Machine Learning, es muy implementado por los desarrolladores de sistemas basados en el aprendizaje, intuitivos e inteligentes que pueden efectuar y tomar pequeñas decisiones solos. Esto hace que sea aún más práctico para el aprendizaje y el desarrollo y puede servir para toda industria que utilice sistemas, soluciones o servicios inteligentes.

Este es un método que simplemente busca en las observaciones más cercanas a la que se está tratando de predecir y clasifica el punto de interés basado en la mayoría de datos que le rodean. Como dijimos antes, es un algoritmo:

- Supervisado: esto -brevemente- quiere decir que tenemos etiquetado nuestro conjunto de datos de entrenamiento, con la clase o resultado esperado dada “una fila” de datos.
- Basado en Instancia: Esto quiere decir que nuestro algoritmo no aprende explícitamente un modelo (como por ejemplo en Regresión Logística o árboles de decisión). En cambio, memoriza las instancias de entrenamiento que son usadas como “base de conocimiento” para la fase de predicción.

2.7 Detección de objetos

Podemos tener la errónea intuición de que la detección de imágenes sea una tarea sencilla, pero veremos que realmente no lo es y de hecho es un gran problema a resolver. Nosotros los humanos podemos ver una foto y reconocer inmediatamente cualquier objeto que contenga de un vistazo rápido, si hay objetos pequeños o grandes, si la foto es oscura ó hasta algo borrosa. Imaginemos un niño escondido detrás de un árbol donde apenas sobresale un poco su cabeza ó un pie.

Para la detección de imágenes mediante Algoritmos de Machine Learning esto implica una red neuronal convolucional que detecte una cantidad limitada (ó específica) de objetos, no pudiendo detectar objetos que antes no hubiera visto, ó si están en tamaños que logra discernir y todas las dificultades de posibles “focos”, rotación del objeto, sombras y poder determinar en qué posición -dentro de la imagen- se encuentra.

Un algoritmo de Machine Learning de detección, para considerarse como tal deberá:

- Detectar multiples objetos.
- Dar la posición X e Y del objeto en la imagen (o su centro) y dibujar un rectángulo a su alrededor.
- Otra alternativa es la segmentación de imágenes (no profundizaremos en este artículo).
- Detectar “a tiempo” ... o puede que no sirva el resultado. Esta es una característica que debemos tener en cuenta si por ejemplo queremos hacer detección en tiempo real sobre video.

Entonces para entrenar nuestra máquina de manera supervisada deberemos indicar la clase del objeto (por ejemplo perro ó gato) y además la posición dentro de la imagen, X, Y el ancho y alto del objeto.

Y por si esto fuera poco, podrían ser múltiples objetos en la misma imagen, con lo cual, para detectar 2 perros en una foto, necesitamos como salida 10 neuronas.

Este es un gran cambio, pues en clasificación de imágenes veníamos acostumbrados a devolver un array con por ejemplo Perro = [1 0] y Gato = [0 1].

La nueva salida deberá contener adicionalmente la posición (por ej. 54,45) y dimensión (por ej. 100,100) de cada clase, resultando en algo mínimo como

[1 0 100 100 54 45] pudiendo detectar sólo 1 objeto ó

[1 0 100 100 54 45 0 1 200 200 30 25] para 2 objetos.

Chiverton utiliza la sustracción de imágenes de fondo para detectar objetos en movimiento. Los métodos descritos se mencionaron anteriormente en este documento. La relación de aspecto (ancho a largo) de la región de interés se usa luego usando el rectángulo delimitador de la imagen. (Valbuena, 2012)

Esta relación permite clasificar el objeto como una motocicleta o de otra clase. Nuevamente, este método es ampliamente utilizado en la literatura descrita en este documento (Valbuena, 2012). Se aísla la región de interés donde se encuentra la cabeza del motociclista y se usa un proceso de extracción de características para determinar si se está usando un casco o no. Donde el histograma en escala de grises calculado en la región de interés se divide en histogramas en escala de grises de los cuatro cuadrantes en los que se puede ubicar el corpus, formando así un vector. (Valbuena, 2012).

Podemos partir de este punto: tenemos una red CNN entrenada para detectar perros y gatos y supongamos que tiene una muy buena tasa de aciertos. A esta red le pasamos una imagen nueva y nos devuelve “perro” ó “gato”. Agregaremos una tercera salida “otros” por si le pasamos la foto de algo que no sepa reconocer.

Si a nuestra red pre-entrenada, le pasamos una imagen con 2 perros será incapaz de detectarlos, puede que no detecte ni siquiera a uno. Si le pasamos una imagen con perros y gatos, tampoco los podrá identificar y mucho menos localizar.

Entonces lo que el “sentido común de ingenieros” nos dice es: “vamos a iterar”. Es decir, iteremos un “área reducida” dentro de la foto de izquierda a derecha y de arriba abajo y le aplicamos la CNN pre-entrenada para ver si detecta algo.

CAPITULO III METODOLOGÍA

La metodología propuesta consiste en los siguientes pasos:

1. Acceder a los datos y prepararlos
2. Crear la red neuronal
3. Configurar las entradas y salidas de la red
4. Ajustar los parámetros de la red (las ponderaciones y tendencias) para optimizar el rendimiento
5. Entrenar la red
6. Validar los resultados de la red
7. Integrar la red en un sistema de producción

Cabe aclarar que tanto el procesamiento de imágenes como los modelos de IA utilizados en este estudio se ejecutaron en una computadora portátil convencional (Core i7, 2.1GHZ, 32GB RAM).

El marco metodológico, así como el problema a resolver, a menudo determinan el tipo de imagen que se utilizará para el estudio correspondiente. En este caso, se desarrolló una trampa fácilmente repetible a bajo costo con el potencial de aumentar significativamente los puntos de captura en un futuro cercano para cubrir una mayor proporción de la población de estudio. En ese sentido, las fotografías fueron tomadas de frente y arriba de los vehículos en la ciudad de Ecatepec. Estas fotos fueron tomadas entre las 15:00 y las 16:00 en un puente peatonal aproximadamente cuatro metros sobre la carretera asfaltada.



Figura 17. Ejemplos de imágenes capturadas según la propuesta

- Sin embargo, estas fotografías no fueron suficientes para poder llevar a cabo la identificación a través de una red neuronal, por lo que se optó por utilizar un paquete de 1000 fotografías de motociclistas encontradas en internet, donde se dividieron los datos en conjuntos de datos de entrenamiento y de validación, donde se utilizó el 70% de las imágenes (350 imágenes) para el entrenamiento y el 30% (150 imágenes) para la validación.

3.1 Uso de MatLab

MATLAB es un sistema de cómputo numérico que ofrece un entorno de desarrollo integrado con un lenguaje de programación propio. Proporciona

capacidades de IA similares a las de las herramientas dedicadas de IA como Caffe y TensorFlow.

MATLAB y Deep Learning Toolbox proporcionan funciones de línea de comandos y apps para crear, entrenar y simular redes neuronales superficiales. Las apps facilitan el desarrollo de redes neuronales para tareas tales como la clasificación, la regresión (incluida la regresión de series temporales) y el clustering. Tras crear las redes con estas herramientas, es posible generar automáticamente código de MATLAB para capturar el trabajo y automatizar las tareas.

El equipo de IBM Research tomó este reto y, a través de innovadores métodos de agrupación, ha construido una biblioteca de “Distributed Deep Learning” (DDL) que se conecta a los entornos populares de machine learning de código abierto como TensorFlow, Caffe, Torch y Chainer. DDL permite que estos marcos se adapten a decenas de servidores IBM aprovechando cientos de GPUs. Efectivamente, IBM Research ha inventado el motor de reacción del deep learning. (IBM, 2022)

Con la biblioteca DDL, nos tomó sólo 7 horas entrenar ImageNet-22K usando ResNet-101 en 64 servidores IBM Power Systems, que tienen un total de 256 aceleradores GPU NVIDIA P100 en ellos. Bajar de 16 días a 7 horas cambia el flujo de trabajo de los científicos de datos. ¡Eso es un incremento en velocidad de 58x! (IBM, 2022)

La biblioteca del distributed deep learning (DDL) está disponible como una vista previa de la tecnología en nuestra última versión 4 de la distribución de software PowerAI de deep learning. DDL presenta una interfaz de programación de aplicaciones (API) que cada uno de los marcos de deep learning puede

conectarse, para escalar a través de varios servidores. PowerAI pone a disposición estas características de cluster scaling a las organizaciones que utilizan deep learning para formar sus modelos de IA. (IBM, 2022)

Rosenblatt diseñó la primera red neuronal artificial en 1958 (el perceptrón). Desafortunadamente dos compañeros de clases, muy respetados y reconocidos, publicaron un análisis de matemático que explicada las deficiencias de las redes neuronales en 1969. Debido a esto y a la muerte de Ronsenblatt en 1971, los fondos para investigación en redes neuronales se terminaron. En 1986, se demostró que las redes neuronales podían ser usadas para resolver problemas prácticos y fue entonces que éstas comenzaron a ser importantes.

Las redes neuronales descomponen las entradas en capas de abstracción. Se pueden entrenar con muchos ejemplos para que reconozcan patrones de voz o en imágenes, por ejemplo, igual que el cerebro humano. Su comportamiento está definido por la forma en que se conectan sus elementos individuales, así como por la importancia (o ponderación) de dichas conexiones. Estas ponderaciones se ajustan automáticamente durante el entrenamiento de acuerdo con una regla de aprendizaje especificada hasta que la red neuronal lleva a cabo la tarea deseada correctamente.

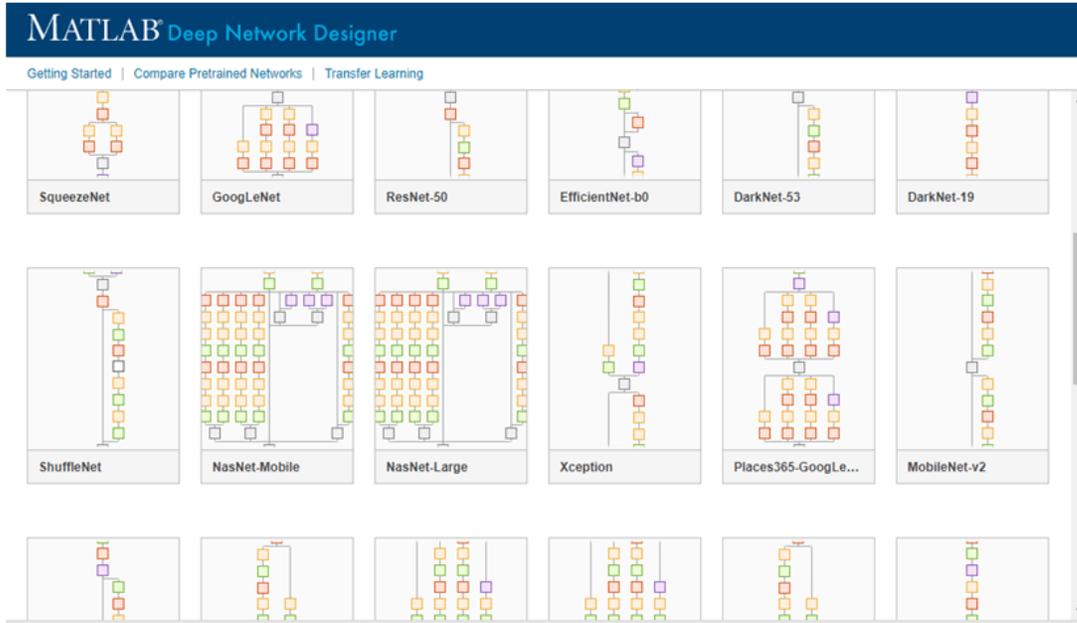


Figura 18. Deep Network Designer

Deep Learning Toolbox™ proporciona un marco para diseñar e implementar redes neuronales profundas con algoritmos, modelos previamente entrenados y aplicaciones. Puede utilizar redes neuronales convolucionales (ConvNet, CNN) y redes de memoria a corto plazo (LSTM) para realizar la clasificación y la regresión de datos de imágenes, series temporales y texto.

Puede crear arquitecturas de red como redes adversarias generativas (GAN) y redes siamesas mediante la diferenciación automática, bucles de entrenamiento personalizados y pesos compartidos. Con la aplicación Deep Network Designer, puede diseñar, analizar y entrenar redes gráficamente. La aplicación Experiment Manager lo ayuda a administrar varios experimentos de aprendizaje profundo, realizar un seguimiento de los parámetros de capacitación, analizar resultados y comparar código de diferentes experimentos. Puede visualizar activaciones de capas y monitorear gráficamente el progreso del entrenamiento.

Deep Network Designer se usa para realizar transferencias del aprendizaje para la clasificación de imágenes siguiendo estos pasos:

- Abra la app Deep Network Designer y elija una red preentrenada.
- Importe el nuevo conjunto de datos.
- Reemplace las capas finales por otras adaptadas al nuevo conjunto de datos.
- Establezca las tasas de aprendizaje para aprender más rápido en las nuevas capas que en las capas transferidas.
- Entrene la red con Deep Network Designer o expórtela para entrenarla en la línea de comandos.

El preprocesamiento de las entradas y los objetivos de la red aumenta la eficiencia del entrenamiento de redes neuronales superficiales. El posprocesamiento permite el análisis detallado del rendimiento de la red. MATLAB y Simulink proporcionan herramientas para ayudarle a:

- Reducir las dimensiones de los vectores de entrada mediante el análisis de componentes principales
- Llevar a cabo análisis de regresión entre la respuesta de la red y los objetivos correspondientes
- Escalar las entradas y los objetivos para que se ajusten al rango $[-1, 1]$
- Normalizar la media y la desviación estándar del conjunto de datos de entrenamiento
- Utilizar el preprocesamiento de datos y la división de datos automatizados al crear sus redes

El perfeccionamiento de la capacidad de la red para generalizar contribuye a evitar el sobreajuste, un problema habitual en el diseño de redes neuronales. El sobreajuste se produce cuando una red ha memorizado el conjunto de entrenamiento, pero no ha aprendido a generalizar cuando hay entradas nuevas. El sobreajuste produce un error relativamente pequeño en el conjunto de entrenamiento, pero un error mucho mayor cuando se presentan datos nuevos a la red.

Deep Learning Toolbox proporciona un conjunto de bloques para crear redes neuronales superficiales en Simulink. Todos los bloques son compatibles con Simulink Coder™. Estos bloques se dividen en cuatro librerías:

- Bloques de función de transferencia, que admiten un vector de entrada de red y generan un vector de salida correspondiente.
- Bloques de función de entrada de red, que admiten cualquier número de vectores de entrada ponderados, vectores de salida de capa de ponderación y vectores de tendencias, y devuelven un vector de entrada de red.
- Bloques de función de ponderación, que aplican un vector de ponderación neuronal a un vector de entrada (o un vector de salida de capa) a fin de obtener un valor de entrada ponderado para una neurona.
- Bloques de preprocesamiento de datos, que asignan datos de entrada y salida dentro de los rangos más adecuados para que la red neuronal los gestione directamente.

Si lo prefiere, puede crear y entrenar sus redes en el entorno de MATLAB y generar automáticamente bloques de simulación de red para usarlos con Simulink. Este enfoque también permite ver las redes de forma gráfica.

3.3 Redes neuronales profundas pre-entrenadas

Una red neuronal convolucional (CNN o ConvNet) es una arquitectura de red para deep learning que aprende directamente de los datos, sin necesidad de extraer características manualmente.

Estas redes son particularmente útiles para encontrar patrones en imágenes para reconocer objetos, caras y escenas. También resultan eficaces para clasificar datos sin imágenes, tales como datos de audio, series temporales y señales.

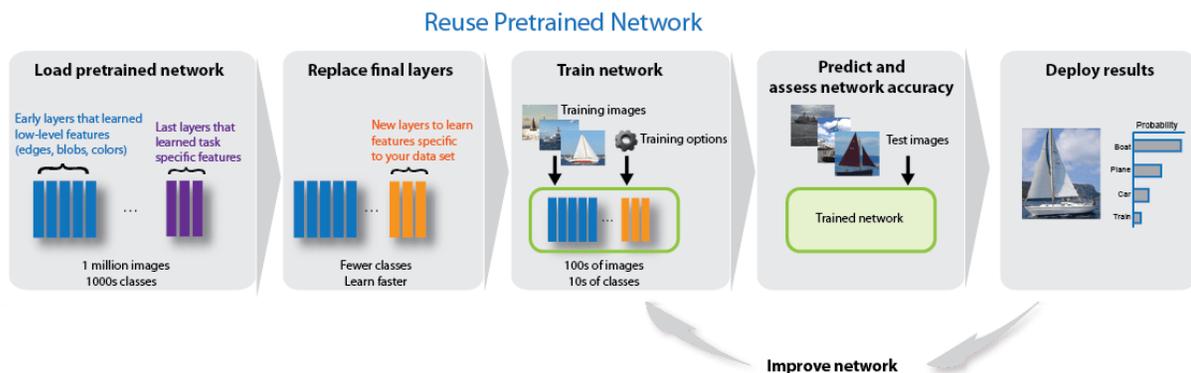


Figura 19. Reuse Pretrained Network

3.4 MatLab y AlexNet

Se puede usar una red pre-entrenada como punto de partida para aprender una nueva tarea. Ajustar una red con transferencia del aprendizaje suele ser más rápido y fácil que entrenarla con pesos inicializados al azar de cero. Puede transferir de forma rápida las características aprendidas a una nueva tarea con menos imágenes de entrenamiento.

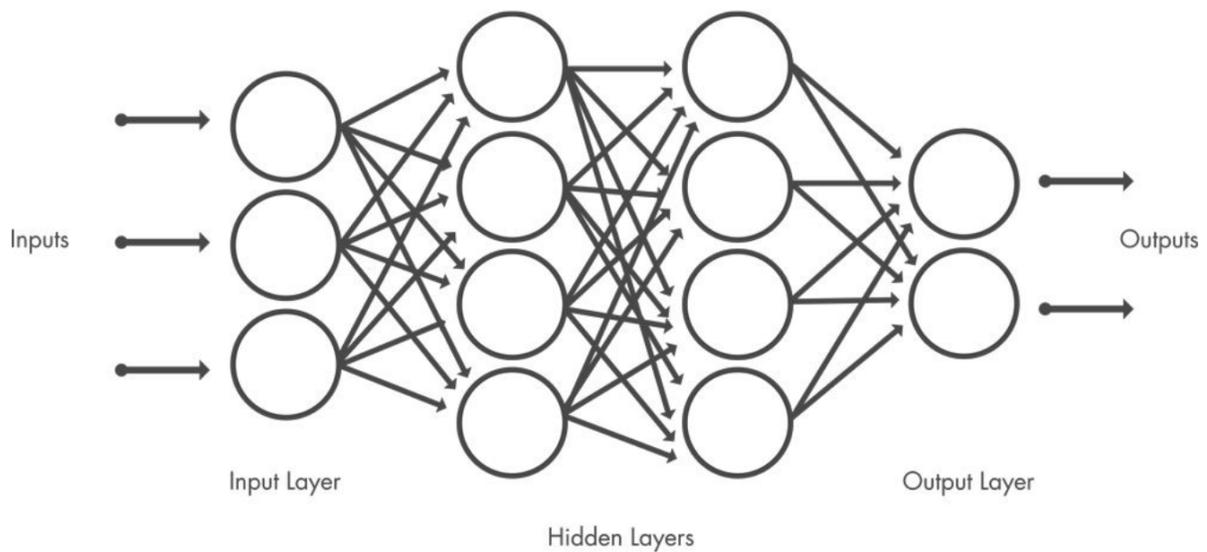


Figura 20. Entradas y salidas de la red neuronal

Antes de explorar AlexNet, es esencial comprender qué es una red neuronal convolucional. Las redes neuronales convolucionales son una de las variantes de las redes neuronales donde las capas ocultas consisten en capas convolucionales, capas de agrupación, capas totalmente conectadas y capas de normalización.

La convolución es el proceso de aplicar un filtro sobre una imagen o señal para modificarla. Ahora, ¿qué es la agrupación? Es un proceso de discretización basado en muestras. La razón principal es reducir la dimensionalidad de la entrada. Por lo tanto, permite hacer suposiciones sobre las características contenidas en las subregiones agrupadas.

Una pila de capas distintas que transforman el volumen de entrada en volumen de salida con la ayuda de una función diferenciable se conoce como Arquitectura CNN. (por ejemplo, mantener los puntajes de la clase)

En otras palabras, se puede entender que una arquitectura CNN es una disposición específica de las capas mencionadas anteriormente. A lo largo de los años se han desarrollado numerosas variaciones de dichos arreglos, lo que ha dado como resultado varias arquitecturas de CNN. Los más comunes entre ellos son:

- LeNet-5 (1998)
- Alex Net (2012)
- ZFNet (2013)
- GoogleNet / Origen (2014)
- VGGNet (2014)
- ResNet (2015)

Desde el año 2010, el proyecto ImageNet ha lanzado anualmente un concurso de reconocimiento de imágenes, ILSVRC –ImageNet Large Scale Visual Recognition Challenge-. Uno de los parámetros utilizados para medir el éxito de los concursantes es el error de las cinco primeras (top-5 error). La red debe proponer cinco etiquetas diferentes a cada imagen. Si ninguna de ellas se corresponde con su descripción, se considera que se ha producido un fallo. Con el objetivo de batir el récord existente hasta entonces (top-5 error del 25%), en el año 2012, Alex Krizhevsky propuso, junto con su director de tesis Geoffrey Hinton, un nuevo modelo de red convolucional que logró reducir ese porcentaje hasta el 17%. Este nuevo modelo se conocería como AlexNet.

AlexNet AlexNet es el nombre de una arquitectura de red neuronal convolucional (CNN), se ha entrenado con más de un millón de imágenes y puede clasificarlas en 1000 categorías de objetos (como teclado, taza de café, lápiz y muchos animales).

La red ha aprendido representaciones ricas en características para una amplia gama de imágenes. La red toma una imagen como entrada y, a continuación, emite una etiqueta para el objeto en la imagen junto con las probabilidades para cada una de las categorías de objetos.

AlexNet es el nombre de una arquitectura de red neuronal convolucional (CNN)

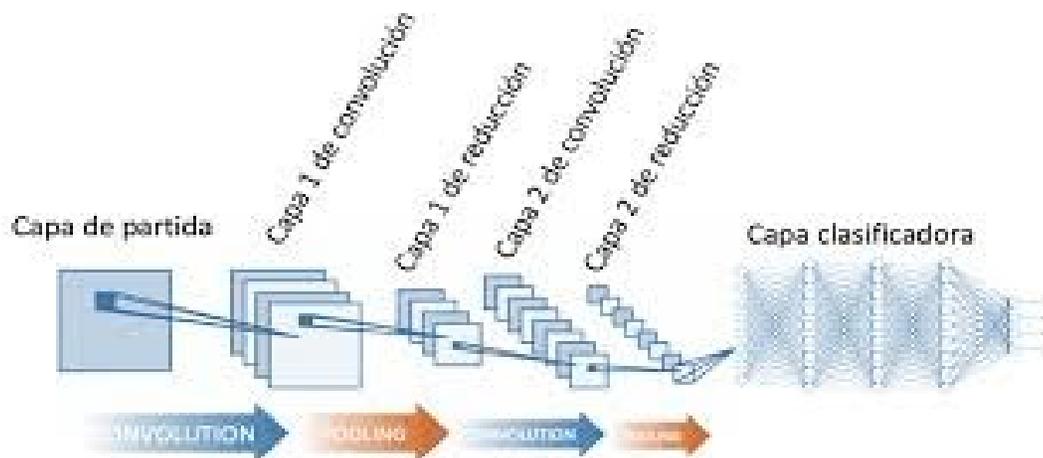


Figura 21. Capas de AlexNet

La arquitectura de la red se compone de cinco capas convolucionales y tres capas densas.

Las salidas de las capas convolucionales son normalizadas por lotes (aunque el esquema de normalización del artículo original era ligeramente diferente). El kernel tiene dimensiones de 11×11 en la primera capa, 5×5 en la segunda, y 3×3 en el resto. La primera, cuarta y quinta capas convolucionales vienen seguidas cada una de ellas por una capa de agrupación (max-pooling) de 3×3 , en la que existe un desplazamiento (stride) de dos, lo que provoca que haya un solapamiento de las celdas en la misma. Las capas convolucionales y de pooling vienen seguidas de tres capas capas densas. Las dos primeras tienen 4.096

neuronas cada una. La última es la de salida. Está compuesta con 1.000 neuronas dotadas de una función de activación softmax. Son ellas las encargadas de realizar la clasificación de la imagen.

Como suele ser usual en los sistemas de clasificación, la función de pérdida utilizada es la entropía cruzada. Las derivadas de esta última se retropropagan del modo usual.

El funcionamiento interno de los mecanismos de convolución de AlexNet (y en general, de todas las redes convolucionales) fue objeto de un interesante estudio del año 2013 de los profesores de la Universidad de Nueva York Matthew D. Zeiler y Rob Fergus. Hasta este momento, la mayoría de los especialistas en aprendizaje profundo consideraban que el contenido de las capas de convolución ocultas, al no ser directamente visualizable, no podía ser interpretado por observadores externos. Y constituía lo que se conoce como un sistema de caja negra –black box-.

Para analizar el comportamiento de AlexNet, los autores introdujeron el concepto de “convolución inversa”, que se expone detenidamente en nuestro artículo de las GANs, y que constituye, en ciertos casos, una suerte de reversión del proceso de convolución.

En la versión 2010 del desafío ImageNet, AlexNet superó con creces al segundo mejor modelo con un 37,5 % de error entre los 1 primeros frente al 47,5 % de error entre los 1 primeros y un 17,0 % de error entre los 5 primeros frente a un 37,55 de error entre los 5 primeros. AlexNet pudo reconocer objetos descentrados y la mayoría de sus 5 clases principales para cada imagen fueron razonables. AlexNet ganó la competencia de 2012 con una tasa de error entre

los 5 primeros del 15,3 % en comparación con la tasa de error entre los 5 primeros del segundo lugar del 26,2 %.

El éxito de AlexNet se atribuye principalmente a su capacidad para aprovechar la GPU para el entrenamiento y poder entrenar esta enorme cantidad de parámetros.

En las siguientes capas, hubo múltiples mejoras sobre AlexNet que dieron como resultado modelos como VGG, GoogleNet y, últimamente, ResNet .

3.6 Aprendizaje de características, capas y clasificación

Estas capas realizan operaciones que alteran los datos con el objetivo de aprender características específicas de dichos datos. Las 3 capas más frecuentes son: convolución, activación o ReLU, y pooling.

- **Convolución:** somete las imágenes de entrada a un conjunto de filtros convolucionales, cada uno de los cuales activa ciertas características de las imágenes.
- **Unidad lineal rectificadora (ReLU):** permite un entrenamiento más rápido y eficaz al asignar los valores negativos a cero y mantener los valores positivos. También se lo denomina activación, dado que solo las características activadas pasan a la siguiente capa.
- **Pooling:** simplifica la salida al disminuir la tasa de muestreo no lineal, reduciendo así el número de parámetros que la red necesita aprender.

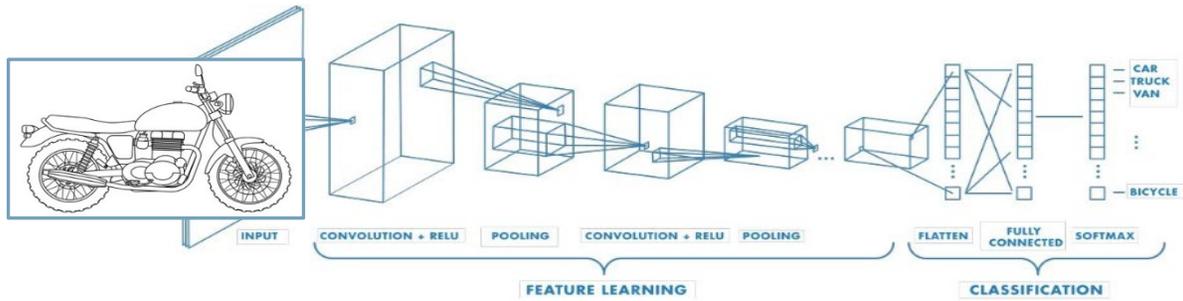


Figura 22. Metodología de clasificación de imágenes en AlexNet

CAPITULO IV ANÁLISIS E INTERPRETACIÓN DE RESULTADOS

4.1 Carga de datos

- Se descomprimió y cargo las 500 imágenes como un almacén de datos.
- imageDatastore etiqueta de forma automática las imágenes basándose en los nombres de las carpetas y almacena los datos como un objeto ImageDatastore.
- Un almacén de datos de imágenes permite almacenar un gran volumen de datos de imágenes, incluidos los que no caben en la memoria, y leer eficazmente lotes de imágenes durante el entrenamiento de una red neuronal convolucional.

```
unzip('MerchData.zip');
imds = imageDatastore('MerchData', ...
    'IncludeSubfolders',true, ...
    'LabelSource','foldernames');
```

4.2 Dividir datos

- Se dividieron los datos en conjuntos de datos de entrenamiento y de validación.
- Se utilizó el 70% de las imágenes (350 imágenes) para el entrenamiento y
- el 30% (150 imágenes) para la validación.
- `splitEachLabel` divide el almacén de datos de imágenes en dos nuevos almacenes de datos.

```
[imdsTrain,imdsValidation] = splitEachLabel(imds,0.7,'randomized');
```

Visualizar Datos

```
numTrainImages = numel(imdsTrain.Labels);
idx = randperm(numTrainImages,16);
figure
for i = 1:16
    subplot(4,4,i)
    I = readimage(imdsTrain,idx(i));
    imshow(I)
end
```

4.3 Cargar una red pre-entrenada

```
net = alexnet;
```

Utilizar `analyzeNetwork` para obtener una visualización interactiva de la arquitectura de la red e información detallada sobre sus capas.

```
analyzeNetwork(net)
```

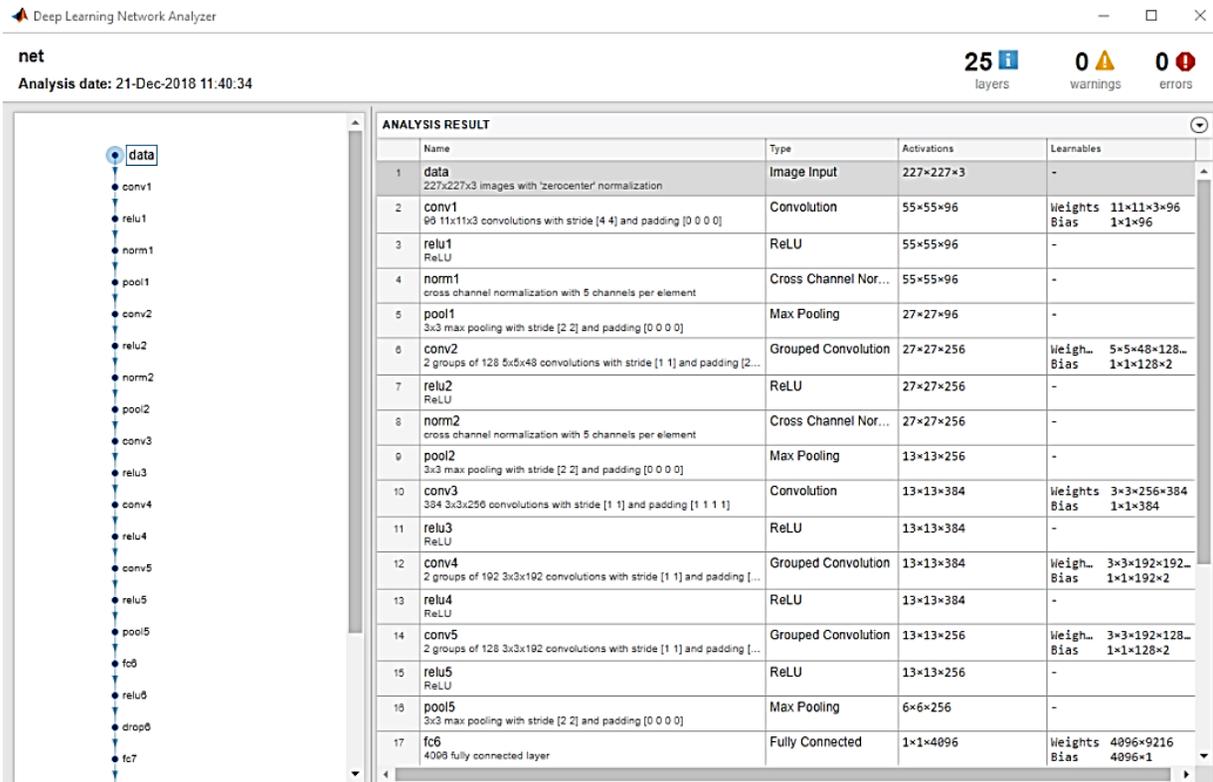


Figura 23. Análisis de las imágenes

La primera capa, la de entrada de imágenes, requiere imágenes de entrada de un tamaño de 227 por 227 por 3, donde 3 es el número de canales de color.

```
inputSize = net.Layers(1).InputSize
```

Sustituir capas finales

- Las últimas tres capas de la red net pre-entrenada están configuradas para 1000 clases.
- Es preciso ajustar las tres capas para el problema de clasificación nuevo.
- Para ello de extraí todas las capas, excepto las tres últimas, de la red pre-entrenada.

```
layersTransfer = net.Layers(1:end-3);
```

4.4 Entrenar la red

```
pixelRange = [-30 30];  
imageAugmenter = imageDataAugmenter( ...  
    'RandXReflection',true, ...  
    'RandXTranslation',pixelRange, ...  
    'RandYTranslation',pixelRange);  
augimdsTrain = augmentedImageDatastore(inputSize(1:2),imdsTrain, ...  
    'DataAugmentation',imageAugmenter);
```

```
netTransfer = trainNetwork(augimdsTrain,layers,options);
```

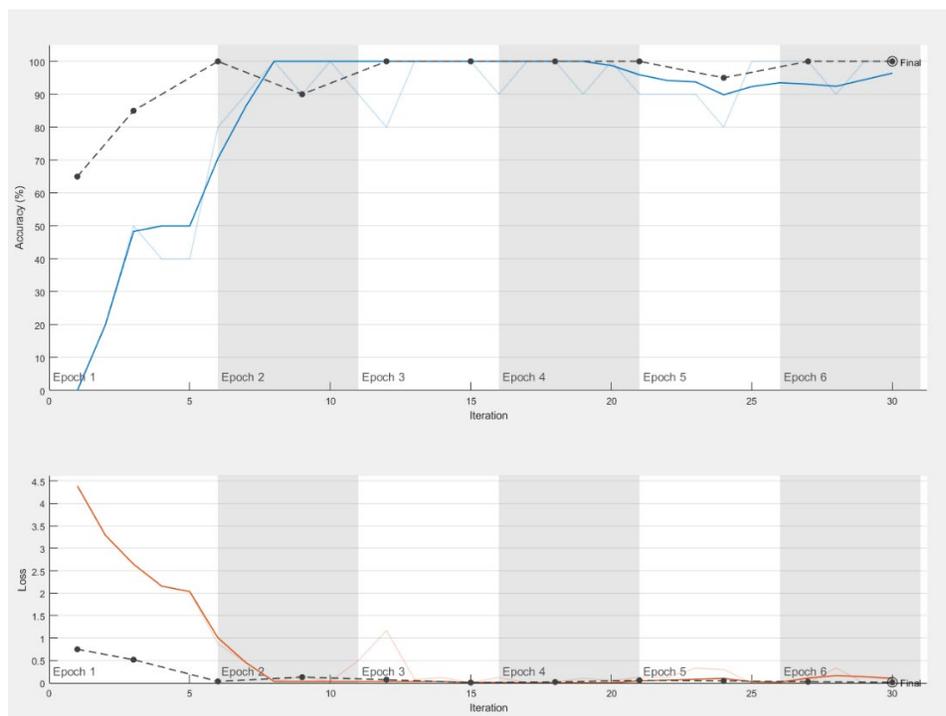


Figura 24. Graficas de resultados arrojados por la red neuronal AlexNet

Results	
Validation accuracy:	100.00%
Training finished:	Reached final iteration
Training Time	
Start time:	21-Oct-2022 10:23:47
Elapsed time:	3 sec
Training Cycle	
Epoch:	6 of 6
Iteration:	30 of 30
Iterations per epoch:	5
Maximum iterations:	30
Validation	
Frequency:	3 iterations
Patience:	Inf
Other Information	
Hardware resource:	Single GPU
Learning rate schedule:	Constant
Learning rate:	0.0001

Figura 25. Resultado del análisis de las imágenes.

4.5 Clasificar imágenes de validación utilizando la red ajustada

```
[YPred,scores] = classify(netTransfer, augimdsValidation);
```

Se calcula la precisión de la clasificación en el conjunto de validación. La precisión es la fracción de etiquetas que la red predice correctamente.

```
YValidation = imdsValidation.Labels;  
accuracy = mean(YPred == YValidation)
```

accuracy = 1

La exactitud o “Accuracy” (AC) se refiere a la dispersión del conjunto de valores obtenidos a partir de mediciones repetidas de una magnitud. Cuanto menor es la dispersión mayor la precisión.

En términos estadísticos, la exactitud está relacionada con el sesgo de una estimación. También se conoce como Verdadero Positivo (o “True positive rate”). Se representa por la proporción entre los positivos reales predichos por el algoritmo y todos los casos positivos

4.6 Resultados del Proyecto

Se logro identificar de acuerdo con las imágenes cargadas y seleccionadas poder identificar y distinguir a un vehículo motocicleta entre autos

De igual manera se realizó el procedimiento para la identificación de imágenes de cascos de motocicleta.

Sin embargo, distinguir a más de una persona en una motocicleta haciendo uso de casco no fue posible,



Figura 26. Identificación de autos, motociclistas y cascos

CONCLUSIONES

El método descrito en este artículo le permite detectar motocicletas en el tráfico, además, la geometría utilizada para las imágenes facilita la repetición de este experimento en diferentes lugares para ayudar a identificar la población de motocicletas en el Estado de México.

Cabe mencionar que la metodología propuesta incorpora varias limitaciones, entre ellas, es importante tomar en cuenta la calidad de las imágenes, ya que como ejemplo se puede decir que, para poder validar el efecto del sol sobre los objetos a detectar, así como condiciones climáticas como el nublado y/o lluvia hace que esas imágenes no ayudan a identificar adecuadamente a una persona que porte el casco en una motocicleta.

Así como también tomar en cuenta que para más de una persona en una motocicleta es necesario realizar el conteo de objetos para poder identificar sobre ellos el uso del casco. El modelo debe validarse también en circunstancias de alta densidad vehicular. Elementos como sombreros y capuchas limitan en

gran medida la precisión del modelo. Además, se recomienda aumentar la cantidad de imágenes con la finalidad de obtener una mejor precisión al estar realizando la clasificación de la detección de motocicletas y cascos.

Realizar este proyecto me deja satisfacción por el aprendizaje obtenido sobre las redes neuronales desde la teoría hasta la práctica, entender el funcionamiento de una red al estar utilizando MatLab, así como el poder analizar los resultados obtenidos.

RECOMENDACIONES

Es importante tener en cuenta que este método se puede utilizar para muchas aplicaciones de transporte, pero es necesario tomar en cuenta incrementar en gran medida el número de fotografías para poder entrenar la red neuronal y pueda realizar un mejor recuento y clasificación de vehículos, con la finalidad de poder resolver el problema de determinar con precisión el número de pasajeros en una moto.

Para ello, es necesario realizar algunos cambios en el método propuesto, incluida la automatización de algunos procesos y el ajuste del modelo de acuerdo con las restricciones identificadas, lo que se puede hacer casi en tiempo real utilizando en la detección tenga un alto rendimiento.

BIBLIOGRAFIA

1. Alvares, E. (s.f.). Computer hoy. Obtenido de <https://computerhoy.com/noticias/tecnologia/raspberry-pi-4-guerra-ofertas-amazon-aliexpress-ofrecer-mejor-precio-532141>
2. A.M., L. X., & T. B. (2013). On-Road Approaching Motorcycle Detection and Tracking Techniques: A Survey. IEEE International Conference on Control System, Computing and Engineering, (págs. 63-68). Penang.
3. ANDI. (2012). INDUSTRIA DE VEHÍCULOS- Caracterización del Sector Automotor . Bogota: Andi.
4. Bradski, G., & Kaehler, A. (2008). Contours. Sebastapol: O'Reilly.
5. Chiu, C.-C., Ku, M.-y., & Chen, H.-T. (2007). Motorcycle detection and Tracking Systema with Occlusion Segmentation. 8th International Workshop on Image Analysis for Multimedia Interactive Services (págs. 21-27). Taiwan: IEEE.
6. Chiverton, J. (2012). Helmet presence classification with motorcycle detection and Tracking. IET Intelligent Transport Systems vol:6, 259-269.
7. Efrain. (19 de Marzo de 2020). Course Hero. Obtenido de <https://www.coursehero.com/file/100807445/covid-19docx/>
8. Esquivvel, L. (17 de Agosto de 2020). Dinero en imagen. Obtenido de <https://www.dineroenimagen.com/economia/los-sectores-ganadores-y-los-mas-golpeados-por-la-pandemia-en-mexico/125854>
9. Fiter, M. (7 de Junio de 2020). El independiente. Obtenido de <https://www.elindependiente.com/futuro/2020/06/07/la-lucha-del-reconocimiento-facial-contra-el-covid-asi-aprende-a-reconocer-mascarillas/>
10. Forero, L. (26 de Febrero de 2015). Las 2 Orillas. Obtenido de Colombia va sin freno hacia la accidentalidad vial:

<http://www.las2orillas.co/colombia-va-sin-freno-hacia-la-accidentalidad-vial/>

11. Friedman, J., Hastie, T., & Tibshirani, R. (2000). Additive logistic regression a statistical view of boosting. *The annals of statistics*, 337-407.
12. Goodnight, J. (20 de Agosto de 2021). SAS (Software y Soluciones de Analítica). Obtenido de https://www.sas.com/es_mx/insights/analytics/what-is-artificial-intelligence.html
13. Guzman, A. M. (Mayo de 2020). colegio de sociologos peru. Obtenido de <https://colegiodesociologosperu.org.pe/wp-content/uploads/El-Coronavirus-y-su-impacto-en-la-sociedad-actual-y-futura-mayo-2020.pdf>
14. Huynh, C., & Hamamoto, K. (2016). Convolutional Neural Network for Motorbike Detection in Dense Traffic. *Communications and Electronics (ICCE), 2016 IEEE Sixth International Conference on* (págs. 27-29). IEEE.
15. Instituto Javeriano de Salud Pública. (2015). Patrón de mortalidad en motociclistas en Colombia 2000-2014. Bogota: Universidad Javeriana.
16. Integra, N. (2020). Nexus Integra. Obtenido de <https://nexusintegra.io/es/ventajas-y-desventajas-de-la-inteligencia-artificial/>
17. kaspersky. (20 de Agosto de 2020). kaspersky.es. Obtenido de <https://www.kaspersky.es/resource-center/definitions/what-is-facial-recognition>
18. Kapur, J., Sahoo, P., & Wong, A. (1985). "A New Method for Gray-Level Picture Thresholding Using the Entropy of the Histogram". *Graphical Models and Image Processing*, 273-285.
19. Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems 25 (NIPS 2012)*, 25-37.

20. Loaiza, C. (30 de Mayo de 2012). Motos: Accidentes fatales. Obtenido de Sura Blog: <http://www.sura.com/blogs/autos/motos-accidentes-fatales.aspx>
21. MAPFRE. (9 de Mayo de 2014). Institucional. Obtenido de La seguridad vial en América Latina y el Caribe: presente y desafíos: <http://www.mapfre.com.ar/institucional/novedades/novedades.html?id=3512>
22. MINISTERIO DE TRANSPORTE. (2004). RESOLUCIÓN 1737 DE 2004. En M. D. TRANSPORTE. Bogota: Legis.
23. Mathworks. (s.f.). mathworks. Obtenido de <https://es.mathworks.com/discovery/convolutional-neural-network-matlab.html>
24. Miranda, G. (Diciembre de 2019). scielo. Obtenido de http://www.scielo.org.mx/scielo.php?pid=S0035-00522019000600213&script=sci_arttext
25. Murillo, D. G. (03 de 2001). Academia . Obtenido de <https://d1wqtxts1xzle7.cloudfront.net/36957218/redesneuronales-with-cover-page-v2.pdf?Expires=1642952460&Signature=TiGfBHxY~fzda2opom8W5GYdRVkO4q6aJgWsagnpK0MtwntBhFGO31-SFGxAenY0cT4BKoSMCU625vPDYOCNS9qL343688iRy3UvqNWde2QFNNpt5aNjGBMx9cilZXh1U6hjrjUPatIE>
26. Salud, O. M. (s.f.). Organización Panamericana de Saludo. Obtenido de <https://www.paho.org/es/temas/coronavirus>
27. Santamaria, P. (6 de Abril de 2020). El Output. Obtenido de <https://eloutput.com/productos/gadgets/sistemas-operativos-raspberry-pi-usos-recomendaciones/>
28. Santamarina, P. (8 de Septiembre de 2021). El Output. Obtenido de <https://eloutput.com/productos/domotica/instalar-so-raspberry-pi-herramientas/>
29. Silva, R., Aires, K., & Veras, R. (2014). Helmet detection on Motorcyclists using image descriptors and Classifiers. 27th SIBGRAPHI

Conference on Graphics, Patterns and Images (págs. 141-148). Picos: IEEE.

30. Silva, R., Santos, T., Kalyf, A., Veras, R., Aires, K., & Soares, A. (2013). Automatic detection of motorcyclist without helmet. Latin American Computing Conference. Rio de Janeiro.
31. Suzuki, S., & Abe, K. (1985). Topological structural analysis of digital binary images by border following. Tokio.
32. Szegedy, C. (2015). Rethinking the Inception Architecture for Computer Vision. arXiv preprint arXiv:1512.00567. Obtenido de <https://github.com/dmlc/mxnet-model-gallery/blob/master/imagenet-1k-inception-v3.md>
33. Tsai, W. (1985). "Moment-preserving thresholding: a new approach",. Computer Vision, Graphics, and Image Processing, 377-393.
34. Turi. (2014). Modelo Iter_45. Obtenido de http://s3.amazonaws.com/GraphLab-Datasets/deeplearning/imagenet_model_iter45
35. UNECE. (2016). The United Nations Motorcycle Helmet Study. Geneva: United Nations.
36. Valbuena, S. (2012). Muertes y lesiones no fatales por accidentes de transporte, Colombia, 2011. Bogota: Instituto Nacional de Medicina Legal y Ciencias Forenses.
37. Waranusast, R., Bundon, N., & Pattanathaburt, P. (2013). Machine Vision Techniques for Motorcycle Safety Helmet Detection. International Conference on Image and Vision Computing , (págs. 13-18). New Zeland.
38. Waranusast, R., Timtong, V., Bundon, N., & Tangnoi, C. (2014). A computer Vision Approach for Detection and Counting of Motorcycles

Riders in University Campus. Proceedings of the IEEE Congress .
Phitsanulok: IEEE.

39. Wolpert, D., & Macready, G. (1997). No free lunch theorems for optimization. IEEE Transactions on Evolutionary Computation, 67 - 82.
40. Tolosa, A. (18 de Mayo de 2020). Genetica Medica News. Obtenido de https://genotipia.com/genetica_medica_news/coronavirus-estructura-infeccion-celulas/
41. World, C. (11 de Novimebre de 2020). ComputerWorld from IDG. Obtenido de <https://www.computerworld.es/tendencias/las-grandes-tecnologias-de-la-era-covid>

GLOSARIO DE TERMINOS DE IA

1. La inteligencia artificial (IA) es una colección de tecnologías que permiten que las computadoras utilicen conjuntos de instrucciones, llamados algoritmos, para percibir, aprender, razonar y actuar.
2. Un algoritmo es un conjunto de instrucciones paso a paso para completar una tarea.
3. Los datos son información que las computadoras almacenan y procesan.
4. Un conjunto de datos es un grupo organizado de datos.
5. Los macrodatos constituyen conjuntos de datos muy grandes que pueden incluir billones o trillones de entradas.
6. Un conjunto de datos de entrenamiento es un conjunto de datos que las máquinas procesan para aprender. Si algunas personas, algunos períodos de tiempo o algunas regiones quedan fuera de los conjuntos de datos de entrenamiento, el modelo de IA tendrá defectos que pueden afectar las vidas de las personas de manera desigual.
7. Un modelo hace referencia a un algoritmo que se ha entrenado en función de un conjunto de datos para reconocer ciertos tipos de patrones y hacer predicciones. Un modelo está diseñado para que una computadora se comporte como un sistema del mundo real.
8. Un programa es un conjunto de instrucciones que debe seguir una computadora y que se imparte en un lenguaje que la computadora puede comprender. Un programa puede indicar a una computadora cómo usar un algoritmo.
9. El aprendizaje automático es un tipo de IA que aprende, mediante ejemplos, a reconocer patrones y resolver problemas.

10. La ciencia de datos es un campo amplio que usa enormes cantidades de información disponible para proporcionar conocimientos significativos. Aunque puede usar herramientas de IA para el análisis, no siempre lo hace.
11. La IA débil, o inteligencia artificial estrecha, incluye aplicaciones de algoritmos de IA para realizar tareas específicas. La IA débil es la que tenemos ahora.
12. La IA avanzada, o inteligencia artificial general (IAG), es cuando una máquina puede razonar en su totalidad como un ser humano. La IA avanzada aún debe lograrse.
13. El aprendizaje supervisado es un método de entrenamiento que usa conjuntos de datos etiquetados en el que se dan ejemplos de pares de entrada-salida a un algoritmo. Un conjunto de imágenes que se identifica como “gato” o “no gato” es un ejemplo de un conjunto de datos etiquetado.
14. El aprendizaje no supervisado es un método de entrenamiento que no usa etiquetas ni salidas correctas. El algoritmo descubre la estructura de los datos por sí mismo.
15. La adaptabilidad es la habilidad de mejorar el desempeño aprendiendo a través de la experiencia.
16. La autonomía es la habilidad de realizar tareas en entornos complejos sin la guía constante de un usuario.
17. Una red neuronal es un sistema inspirado por la biología humana que usa una red de algoritmos para comprender datos de entrada y traducirlos en valores de salida. Las redes neuronales cuentan con conjuntos de datos de entrenamiento para aprender y mejorar con el tiempo.
18. El aprendizaje profundo es un subconjunto de aprendizaje automático que usa redes neuronales con un mínimo de tres capas.
19. Una capa es un paso del procesamiento de datos.

20. Un deepfake es una imagen, un video o un clip de audio alterado o completamente creado mediante el uso de herramientas de inteligencia artificial.
21. La robótica es la construcción y programación de los robots para que puedan operar en situaciones complejas del mundo real.
22. La visión computarizada es un campo de la IA que posibilita a las computadoras usar datos para reconocer imágenes, videos y otros estímulos visuales.
23. Un programa de reconocimiento de imágenes puede identificar si personas y ciertos objetos aparecen en imágenes o videos a través de algoritmos entrenados y un sistema de cámaras.
24. El reconocimiento facial es un programa de reconocimiento de imagen entrenado para verificar o confirmar la identidad de una persona a través de su rostro.
25. El sesgo es la tendencia, conocida o desconocida, a preferir una persona o un objeto a otro, lo cual puede influir en la comprensión o en los resultados.
26. La ética es un sistema de reglas o principios que afecta la forma de pensar y comportarse de las personas.