

UNIVERSIDAD MAYOR DE SAN ANDRÉS
FACULTAD DE CIENCIAS PURAS Y NATURALES
CARRERA DE INFORMÁTICA



TESIS DE GRADO

**MODELO DE CONDUCCIÓN AUTÓNOMA BASADO EN
APRENDIZAJE PROFUNDO Y ALGORITMOS DE VISIÓN
COMPUTACIONAL**

PARA OPTAR AL TÍTULO DE LICENCIATURA EN INFORMÁTICA
MENCIÓN: CIENCIAS DE LA COMPUTACIÓN

POSTULANTE RAFAEL VILLCA POGGIAN
TUTOR METODOLÓGICO: M.Sc. ALDO VALDEZ ALVARADO
ASESOR: Lic. BRIGIDA ALEXANDRA CARVAJAL BLANCO

LA PAZ - BOLIVIA

2021



**UNIVERSIDAD MAYOR DE SAN ANDRÉS
FACULTAD DE CIENCIAS PURAS Y NATURALES
CARRERA DE INFORMÁTICA**



LA CARRERA DE INFORMÁTICA DE LA FACULTAD DE CIENCIAS PURAS Y NATURALES PERTENECIENTE A LA UNIVERSIDAD MAYOR DE SAN ANDRÉS AUTORIZA EL USO DE LA INFORMACIÓN CONTENIDA EN ESTE DOCUMENTO SI LOS PROPÓSITOS SON Estrictamente Académicos.

LICENCIA DE USO

El usuario está autorizado a:

- a) visualizar el documento mediante el uso de un ordenador o dispositivo móvil.
- b) copiar, almacenar o imprimir si ha de ser de uso exclusivamente personal y privado.
- c) copiar textualmente parte(s) de su contenido mencionando la fuente y/o haciendo la referencia correspondiente respetando normas de redacción e investigación.

El usuario no puede publicar, distribuir o realizar emisión o exhibición alguna de este material, sin la autorización correspondiente.

TODOS LOS DERECHOS RESERVADOS. EL USO NO AUTORIZADO DE LOS CONTENIDOS PUBLICADOS EN ESTE SITIO DERIVARA EN EL INICIO DE ACCIONES LEGALES CONTEMPLADOS EN LA LEY DE DERECHOS DE AUTOR.

DEDICATORIA

Dedico la presente tesis a mi familia por todo el apoyo, amor y paciencia durante todos estos años de estudio.

AGRADECIMIENTOS

A mi familia por el apoyo durante todos los años cursando la carrera.

A mi tutor M.Sc. Aldo Valdez Alvarado por la guía durante el desarrollo de la presente tesis,
y a mi asesora Lic. Brigida Carvajal Blanco por los consejos, sugerencias y apoyo.

RESUMEN

La creciente popularidad de vehículos de distintas marcas con funcionalidades autónomas, ocasionó la creación de simuladores y conjuntos de datos para el entrenamiento de modelos de conducción mediante técnicas de visión artificial y aprendizaje profundo.

El problema a abordar en las soluciones que se desarrollan es crear implementaciones eficientes para realizar inferencias rápidas y reaccionar a las distintas situaciones ambientales mediante cámaras.

Se propone un modelo compuesto por redes neuronales y algoritmos de visión computacional que se complementen, y así obtener una conducción autónoma básica y eficiente computacionalmente.

Palabras clave: Aprendizaje profundo, Visión computacional, Redes neuronales, Carla simulator

ABSTRACT

The increasing popularity of different car manufacturers with autonomous features, made possible the creation of simulators and data sets to train self driving models using Computer Vision techniques and Deep Learning.

The main focus of these solutions is to develop efficient implementations to perform fast inference and react to different environmental situations through cameras.

In this work a model composed of deep neural networks and computer vision algorithms which work together is proposed, in order to obtain an efficient and basic self driving capability

Keywords: Deep learning, Computer vision, Neural networks, Carla simulator

Índice general

DEDICATORIA	III
AGRADECIMIENTOS	IV
RESUMEN	V
ABSTRACT	VI
ÍNDICE	XI
ÍNDICE DE FIGURAS	XIII
ÍNDICE DE TABLAS	XIV
1 MARCO REFERENCIAL	1
1.1 INTRODUCCIÓN	1
1.2 ANTECEDENTES	2
1.2.1 ESTADO DEL ARTE	2
1.2.1.1 DETECCIÓN DE OBJETOS	2
1.2.1.2 SEGMENTACIÓN SEMÁNTICA	5
1.2.1.3 SISTEMAS DE CONDUCCIÓN AUTÓNOMA	6
1.2.2 TRABAJOS SIMILARES	7
1.3 PLANTEAMIENTO DEL PROBLEMA	8
1.3.1 PROBLEMA CENTRAL	9
1.3.2 PROBLEMAS SECUNDARIOS	10
1.4 DEFINICIÓN DE OBJETIVOS	10
1.4.1 OBJETIVO GENERAL	10
1.4.2 OBJETIVOS ESPECÍFICOS	10
1.5 HIPÓTESIS	11
1.5.1 OPERACIONALIZACIÓN DE VARIABLES	11
1.6 JUSTIFICACIÓN	11
1.6.1 JUSTIFICACIÓN ECONÓMICA	11
1.6.2 JUSTIFICACIÓN SOCIAL	12
1.6.3 JUSTIFICACIÓN CIENTÍFICA	12

1.7	ALCANCES Y LÍMITES	12
1.7.1	ALCANCES	12
1.7.2	LÍMITES	12
1.8	APORTES	13
1.8.1	PRÁCTICO	13
1.8.2	TEÓRICO	13
1.9	METODOLOGÍA	13
2	MARCO TEÓRICO	15
2.1	SISTEMAS DE CONDUCCIÓN AUTÓNOMA	15
2.1.1	TAREAS DE LA CONDUCCIÓN AUTÓNOMA	15
2.1.2	ARQUITECTURA DE LA CONDUCCIÓN AUTÓNOMA	15
2.1.3	CARLA	16
2.1.4	NIVELES DE CONDUCCIÓN AUTÓNOMA	16
2.2	VISIÓN COMPUTACIONAL	17
2.2.1	PROCESAMIENTO DE IMÁGENES	17
2.2.1.1	REPRESENTACIÓN DE IMÁGENES EN UNA COMPUTADORA . . .	17
2.2.1.2	CONVOLUCIÓN	18
2.2.1.3	DETECCIÓN DE CONTORNOS	19
2.2.2	TRANSFORMACIONES MORFOLÓGICAS	19
2.2.2.1	DILATACIÓN	19
2.2.2.2	EROSIÓN	20
2.2.2.3	APERTURA	21
2.2.3	K-MEANS	21
2.2.4	FLOOD FILL	23
2.3	APRENDIZAJE AUTOMÁTICO	23
2.3.1	APRENDIZAJE SUPERVISADO	23
2.3.1.1	APRENDIZAJE CORRECTO PROBABLEMENTE APROXIMADO . . .	24
2.3.2	REGRESIÓN	25

2.3.2.1	REGRESIÓN LINEAL	25
2.3.2.2	REGRESIÓN GENERALIZADA	27
2.3.2.3	MEDIAS MÓVILES EXPONENCIALES	27
2.3.3	CLASIFICACIÓN	27
2.3.3.1	REGRESIÓN SOFTMAX	27
2.3.4	DESCENSO DEL GRADIENTE ESTOCÁSTICO	29
2.3.5	ADAM	30
2.4	APRENDIZAJE PROFUNDO	31
2.4.1	PERCEPTRÓN MULTICAPA	31
2.4.1.1	FUNCIONES DE ACTIVACIÓN	33
2.4.1.2	RETROPROPAGACIÓN DE LOS ERRORES	33
2.4.2	REDES NEURONALES CONVOLUCIONALES	34
2.4.2.1	STRIDES	34
2.4.2.2	POOLING	35
2.4.2.3	MOBILENET V2	36
2.4.2.4	FAST DEPTH	37
2.4.3	APRENDIZAJE DE REPRESENTACIONES PROFUNDAS	38
2.5	MÉTRICAS DE ERROR	40
2.5.1	ERROR CUADRÁTICO MEDIO	40
2.5.2	ERROR ABSOLUTO MEDIO	40
2.5.3	PRECISIÓN Y EXHAUSTIVIDAD	41
2.5.4	ÍNDICE JACCARD	42
2.6	COMPILACIÓN EN TIEMPO DE EJECUCIÓN	42
3	MARCO APLICATIVO	43
3.1	COMPRENSIÓN DEL PROYECTO	43
3.2	COMPRENSIÓN DE LOS DATOS	44
3.2.1	ACELERADOR Y DIRECCIÓN	44
3.2.2	PROFUNDIDAD	45

3.2.3	SEGMENTACIÓN SEMÁNTICA	45
3.3	PREPARACIÓN DE DATOS	47
3.3.1	UNIÓN DE DATAFRAMES	47
3.3.2	ETIQUETADO DE INTERSECCIONES	47
3.3.3	CONCATENACIÓN DE ATRIBUTOS	49
3.4	MODELADO	49
3.4.1	RED DE CONDUCCIÓN	50
3.4.2	RED DE PROFUNDIDAD	51
3.4.3	RED DE SEGMENTACIÓN SEMÁNTICA	52
3.4.4	SUAVIZADO DE DIRECCIÓN	52
3.4.5	CAJA DELIMITADORA	53
3.4.6	CUANTIFICACIÓN DIGITAL DEL COLOR	54
3.4.7	MODELO PARA LA CONDUCCIÓN AUTÓNOMA	55
3.5	EVALUACIÓN	56
3.5.1	DRIVENET	56
3.5.2	DEPTHNET	57
3.5.3	SEMSEGNET	58
3.5.4	CENTROIDES K-MEANS	58
3.6	DESPLIEGUE	59
4	RESULTADOS Y ANÁLISIS	61
4.1	RENDIMIENTO DE LOS MÓDULOS	61
4.1.1	ACELERACIÓN Y GIRO	61
4.1.2	ESTIMACIÓN DE PROFUNDIDAD	62
4.1.3	SEGMENTACIÓN SEMÁNTICA	63
4.1.3.1	MATRIZ DE CONFUSIÓN	64
4.1.3.2	PRECISIÓN Y EXHAUSTIVIDAD	65
4.1.3.3	CAJAS DELIMITADORAS	67
4.1.3.4	ÍNDICE JACCARD	67

4.1.4	DETECCIÓN Y CLASIFICACIÓN DE SEMÁFOROS	68
4.2	REPRESENTACIONES APRENDIDAS	69
4.2.1	DISTANCIAS EN LA REGIÓN DE INTERÉS	69
4.2.2	VISUALIZACIÓN DE ZONAS DE INTERÉS	71
4.3	PRUEBAS EN SIMULADOR	71
4.3.1	CONDUCCIÓN	71
4.3.2	FALLOS	73
4.4	PRUEBA DE LA HIPÓTESIS	74
5	CONCLUSIONES Y RECOMENDACIONES	76
5.1	CONCLUSIONES	76
5.2	RECOMENDACIONES	78
6	BIBLIOGRAFÍA	79
	ANEXOS	83

Índice de figuras

1	Contornos de objetos	2
2	Características HAAR	3
3	Vectores direccionales del HOG	3
4	Flujo predicción YOLO	4
5	Partición y recuadros de detección SSD	4
6	Segmentación de carretera por MCRF	5
7	Segmentación de una célula del dataset PhC-U373	6
8	Segmentación semántica de distintas clases	6
9	DAVE	7
10	Diagrama CRISP-DM	14
11	Intensidad de los canales de color RGB	18
12	Convolución del filtro K con la imagen I	19
13	Dilatación sobre un conjunto de letras con discontinuidades	20
14	Erosión de una letra	20
15	Apertura aplicada a una huella digital	21
16	Ajuste iterativo de k-means	22
17	Ajuste de regresión lineal	26
18	Ajuste de regresión logística	28
19	Pasos de los parámetros W en cada iteración camino al mínimo error	29
20	Red neuronal con una capa oculta	32
21	Red neuronal convolucional para la clasificación de dígitos manuscritos	34
22	Convolución con $s = 2$	35
23	Pooling con con $s = 1$	35
24	Cuello de botella residual	36
25	Arquitectura FastDepth	38
26	Extracción de las características profundas aprendidas	39
27	Algoritmo ScoreCam	39

28	Arquitectura del proyecto	43
29	Diagrama de secuencia de extracción de datos	44
30	Etiquetas de profundidad imagen	45
31	Segmentación semántica con códigos de color	46
32	Etapas del etiquetado de intersecciones	48
33	Entrenamiento de tres redes neuronales	50
34	Arquitectura DriveNet basada en Mobilenet V2	50
35	Arquitectura FastDepth	51
36	SemsegNet, FastDepth para clasificación 2D	52
37	Aplicación de medias exponenciales móviles	52
38	Cuantificación digital del color	54
39	Delimitado del área de interés para la detección de objetos	55
40	Modelo de conducción autónoma	56
41	Errores de entrenamiento y validación de la DriveNet	57
42	Errores de entrenamiento y validación de la DepthNet	57
43	Errores de entrenamiento y validación de la SemsegNet	58
44	Espacios de color y centroides para dos semáforos distintos	59
45	Predicción Media Móvil Exponencial	62
46	Predicción vs valor esperado de profundidad	63
47	Matriz de confusión de segmentación semántica	64
48	Predicción vs valor esperado de segmentación	66
49	Cajas delimitadoras	67
50	Cuantización y Clasificación del Color	69
51	Distancia a de objetos	70
52	Regiones de Interés en la Inferencia de Dirección	71
53	Comparación de Dirección Real y Estimada	72
54	Charco erroneamente Clasificado	73
55	Colisión con Poste	74

Índice de tablas

1	Niveles de Conducción Autónoma	16
2	MobileNet V2	37
4	Correspondencia numérica de clases de segmentación	46
5	Errores en el conjunto de validación de los parámetros seleccionados	61
6	Métricas de error de clasificación	65
7	Índice de Jaccard	68