

# Aplicación de CNNs en la conducción autónoma

---

Rafael Villca Poggian

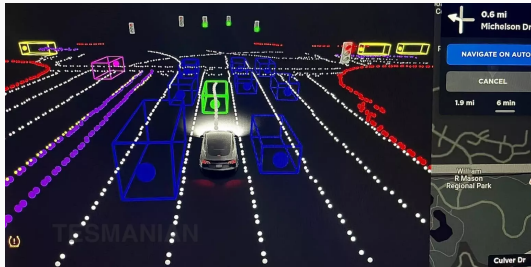
Club de Ciencia de Datos

1. Introducción
2. Proyecto
3. Resultados

# Introducción

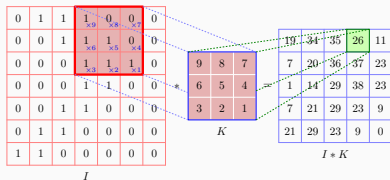
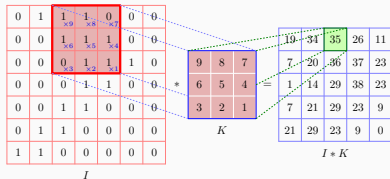
---

# Sistemas de Conducción Autónoma



# Procesamiento de Imágenes

107	98	100	105	131	113	113
18	9	11	18	46	33	42
62	55	57	61	88	71	78
97	85	76	77	131	142	120
16	13	8	14	63	61	35
59	61	59	59	104	94	66
99	88	79	83	102	127	118
21	13	6	16	34	56	43
60	56	51	57	71	86	74
122	97	94	87	110	156	128
46	15	14	16	43	87	57
82	53	53	48	74	116	91
140	127	126	115	139	162	119
60	43	42	41	66	88	44
95	76	75	68	95	121	83
152	136	135	138	133	124	110
68	56	55	62	54	42	32
102	85	82	90	86	80	74



# Operaciones Morfológicas

Original

Historically, certain computer programs were written using only two digits rather than four to define the applicable year. Accordingly, the company's software may recognize a date using "00" as 1900 rather than the year 2000.

Dilatación

Historically, certain computer programs were written using only two digits rather than four to define the applicable year. Accordingly, the company's software may recognize a date using "00" as 1900 rather than the year 2000.

Original



Erosión



Original



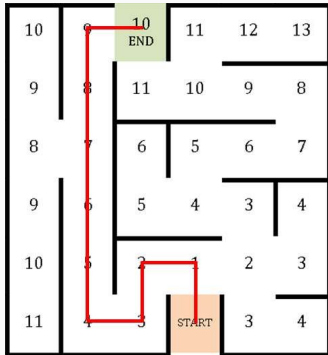
Erosión



Dilatación (Apertura)



# Flood Fill



---

## Algoritmo 1: *Flood Fill*

---

**Requiere:** *mat* : Matriz de valores

**Requiere:** *y* : coordenada vertical actual

**Requiere:** *x* : coordenada horizontal actual

**Requiere:** *d* : vector de direcciones

```
if mat[y][x]  $\neq$  val & x, y dentro de mat then
    mat[y][x] = marca ;           // se marca como visitado
    x1, y1  $\leftarrow$  min(x1, j), min(y1, i)
    x2, y2  $\leftarrow$  max(x2, j), max(y2, i)
    for i  $\rightarrow$  length d do
        | flood_fill(mat, y+d[i].y, x+d[i].x)
    end
end
/* si ya está visitado finaliza la rama de recursión */
```

---



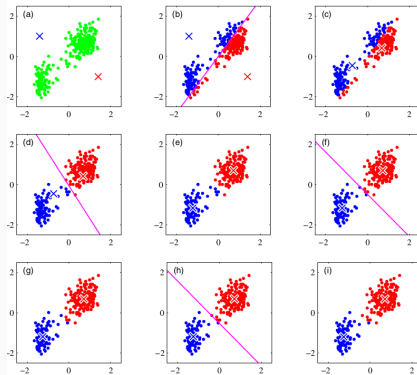
# K-Means

Sea  $R$  de dimensiones  $m \times k$  y un elemento  $r_{ij}$

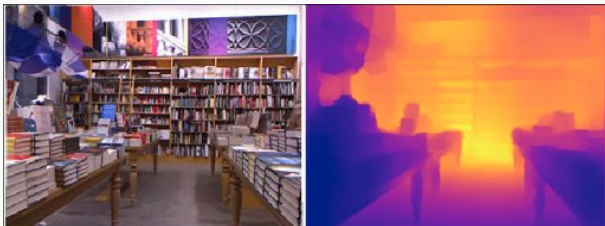
$$J = \sum_{i=1}^m \sum_{j=1}^k r_{ij} \|x_i - \mu_j\|^2$$

$$r_{ij} = \begin{cases} 1 & \text{si } j = \underset{p}{\operatorname{argmin}} \|x_i - \mu_p\| \\ 0 & \text{e.o.c} \end{cases}$$

$$\mu_j = \frac{\sum_i r_{ij} x_i}{\sum_i r_{ij}}$$



# Visión Computacional



Classification



Cat

Object Detection



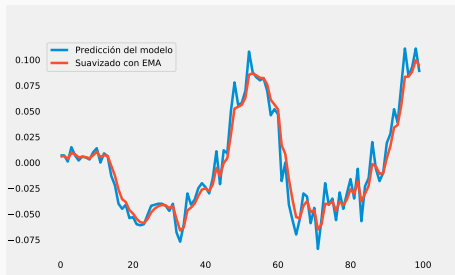
Semantic Segmentation



# Medias Móviles Exponenciales

- promedia los valores en un tiempo  $t$  de manera ponderada exponencialmente por los términos anteriores de la serie.
- Se aplica para encontrar tendencias y suavizar oscilaciones dadas por valores atípicos de poca duración en la señal

$$S_t = \begin{cases} y_0, & t = 0 \\ \alpha y_t + (1 - \alpha)S_{t-1}, & t \geq 1 \end{cases}$$

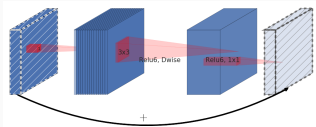


# Convolución Separable

- Se separa la convolución en dos etapas para obtener velocidad a costa de exactitud.
- **Depthwise:** entrada de dimensiones  $h_i \times w_i \times c$  con  $c$  filtros de  $k \times k$ , para obtener una salida de dimensiones  $h_o \times w_o \times c$ .
- **Pointwise:** sobre el anterior paso aplica  $d$  filtros de dimensiones  $1 \times 1 \times c$ , para apilar las salidas y obtener una capa de  $h_o \times w_o \times d$
- Este mismo resultado se puede obtener mediante una capa estándar, aplicando  $d$  filtros de dimensión  $k \times k \times c$
- El número de operaciones es mayor, ya que se realizan  $h_i \cdot w_i \cdot c \cdot k^2 \cdot d$
- Esta modificación realiza  $h_i \cdot w_i \cdot c \cdot k^2 + h_o \cdot w_o \cdot c \cdot 1^2 \cdot d$  operaciones,  $h \cdot w \cdot c \cdot (k^2 + d)$  si la entrada y salida tienen las mismas dimensiones.

# Cuello de Botella Residual

- En base esta nueva convolución, se proponen bloques denominados cuellos de botella residuales

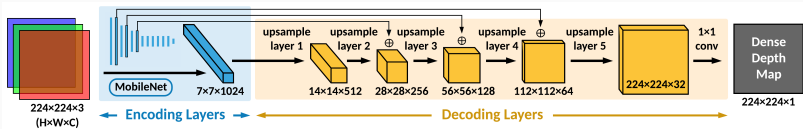


- Consisten en una capa PointWise con no linealidad Relu truncada con valor máximo 6 llamada Relu6 para obtener  $t \cdot k$  filtros
- Seguido de una capa DepthWise  $3 \times 3$  con stride  $s$  y Relu6
- Finalmente otra capa PointWise sin activación no lineal y que devuelve  $d$  filtros

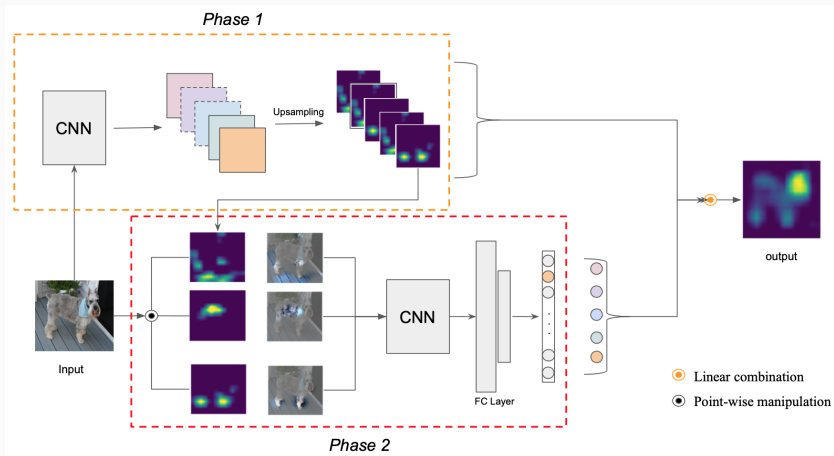
# MobileNet V2

Entrada	Operador	Factor $t$	Canales $c$	Repeticiones $n$	Stride $s$
$224^2 \times 3$	conv2d	-	32	1	2
$112^2 \times 32$	bottleneck	1	16	1	1
$112^2 \times 16$	bottleneck	6	24	2	2
$56^2 \times 24$	bottleneck	6	32	3	2
$28^2 \times 32$	bottleneck	6	64	4	2
$14^2 \times 64$	bottleneck	6	96	3	1
$14^2 \times 96$	bottleneck	6	160	3	2
$7^2 \times 160$	bottleneck	6	320	1	1
$7^2 \times 320$	conv2d $1 \times 1$	-	1280	1	1
$7^2 \times 1280$	avg pool $7 \times 7$	-	-	1	-
$1^2 \times 1280$	conv2d $1 \times 1$	-	#clases	-	-

# FastDepth



# SCORECAM





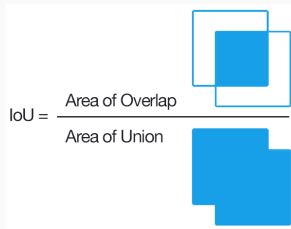
$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y} - y)^2$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y} - y|$$

$$P = \frac{T_p}{T_p + F_p}$$

$$R = \frac{T_p}{T_p + F_n}$$

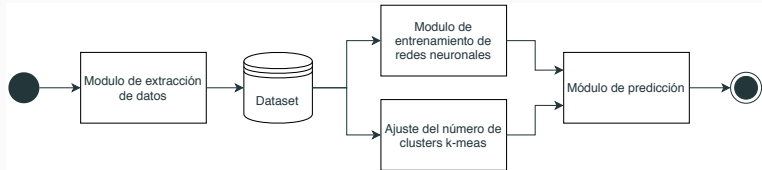
$$F = 2 \frac{P \cdot R}{P + R}$$



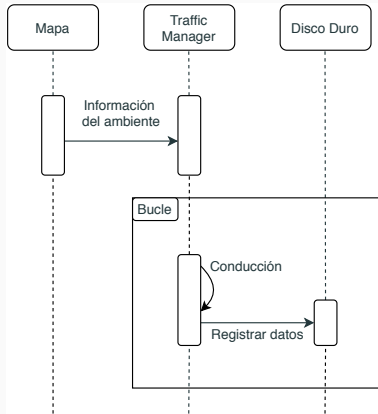
# Proyecto

---

# Estructura del Proyecto



# Extracción de datos

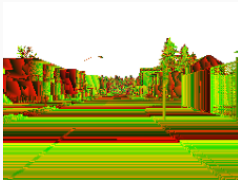


# Profundidad

RGB



Profundidad



Normalizado

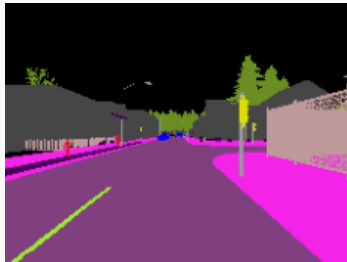


# Segmentación

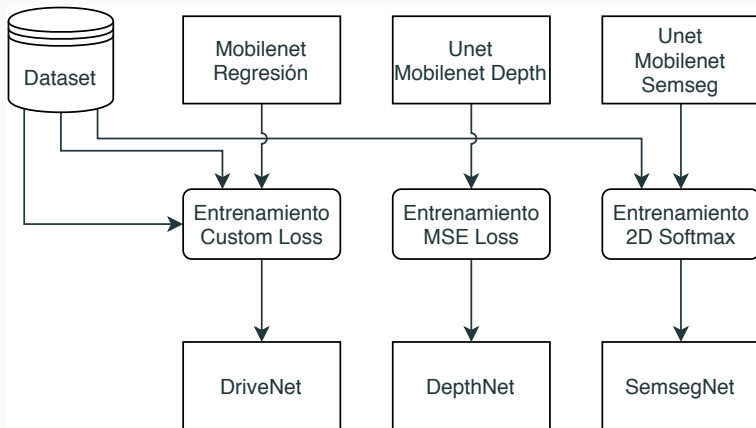
RGB

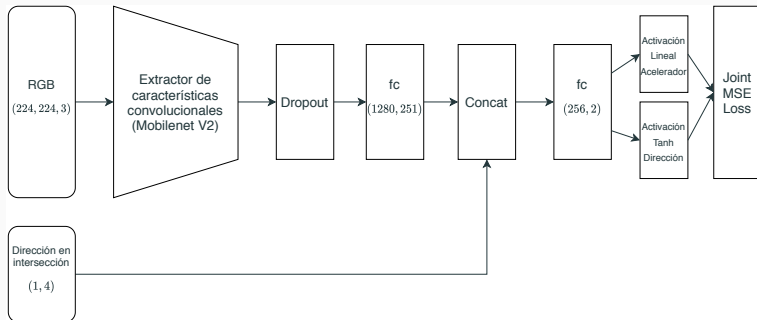


Segmentación semántica



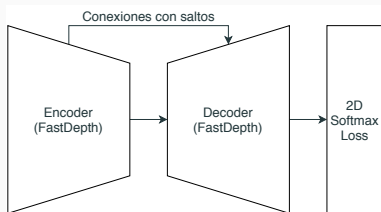
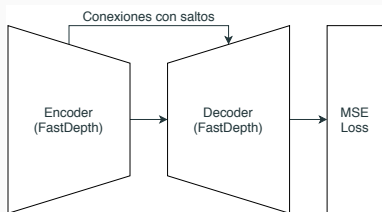
# Entrenamiento de las Redes Neuronales







# DepthNet y SemsegNet



# Clasificación del Color

Verde RGB



Verde Cuantizado



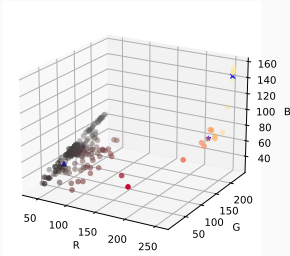
Rojo RGB



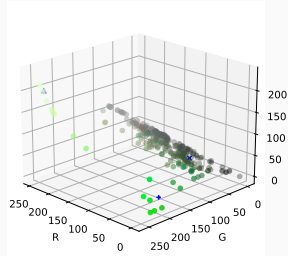
Rojo Cuantizado

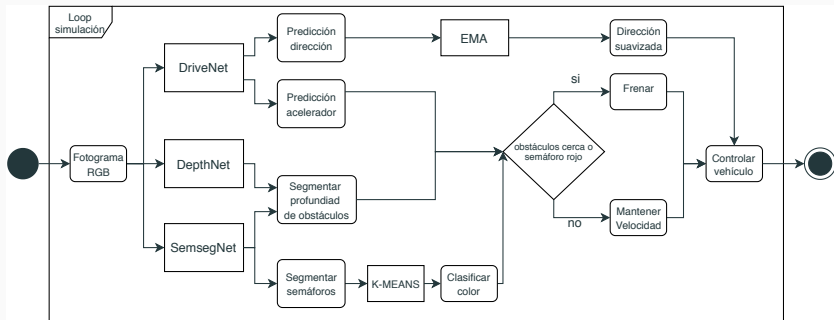


Semáforo Rojo y clusters



Semáforo Verde y clusters

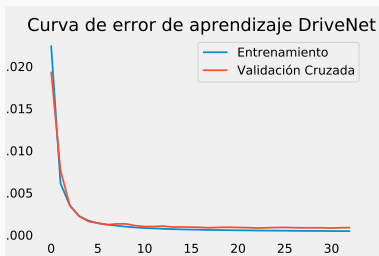
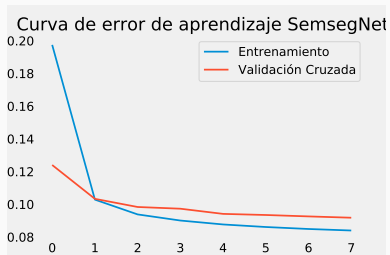
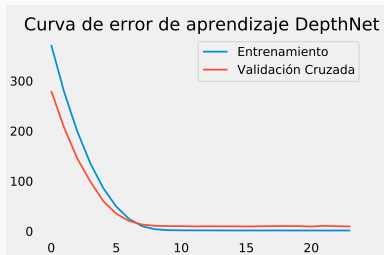




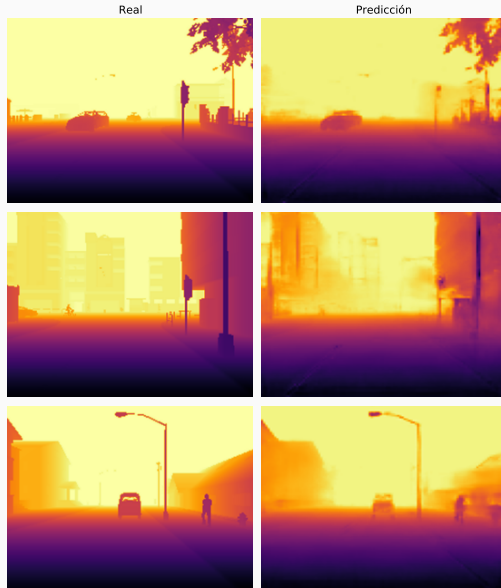
## Resultados

---

# Curvas de Aprendizaje



# Resultados Profundidad



# Resultados Semseg

Matriz de confusión porcentual

Nada	97.87	0.69	0.06	0.03	0.01	0.11	0.00	0.00	0.12	1.03	0.03	0.05	0.01
Edificios	0.66	97.38	0.28	0.18	0.02	0.25	0.00	0.00	0.29	0.62	0.09	0.15	0.08
Cercas	0.29	2.32	90.40	0.47	0.10	0.34	0.00	0.48	4.42	0.40	0.72	0.05	0.01
Otro	0.92	8.39	1.10	80.77	0.18	0.34	0.00	0.24	5.57	0.42	0.46	1.59	0.02
Peatones	2.00	12.47	3.64	3.30	48.25	0.50	0.00	6.50	6.75	0.14	13.76	2.70	0.00
Postes	10.14	16.13	1.31	1.22	0.03	54.04	0.00	0.52	7.89	6.40	0.52	0.59	1.22
Carriles	37.90	0.01	0.00	0.00	0.13	0.00	0.00	61.64	0.02	0.00	0.21	0.00	0.08
Caminos	0.01	0.00	0.03	0.01	0.01	0.01	0.00	99.13	0.58	0.00	0.22	0.00	0.00
Aceras	0.32	0.46	0.58	0.46	0.04	0.28	0.00	1.41	96.07	0.01	0.25	0.12	0.00
Vegetación	3.27	2.19	0.21	0.05	0.00	0.23	0.00	0.00	0.08	93.82	0.02	0.08	0.05
Vehículos	0.19	1.33	0.51	0.19	0.24	0.03	0.00	3.72	0.98	0.16	92.54	0.04	0.07
Paredes	1.37	4.90	0.13	0.85	0.22	0.29	0.00	0.09	1.64	0.52	0.60	89.36	0.04
Señales	0.83	24.90	1.02	0.88	0.00	3.22	0.00	0.00	0.00	4.14	0.02	0.03	64.96
	Nada	Edificios	Cercas	Otro	Peatones	Postes	Carriles	Caminos	Aceras	Vegetación	Vehículos	Paredes	Señales

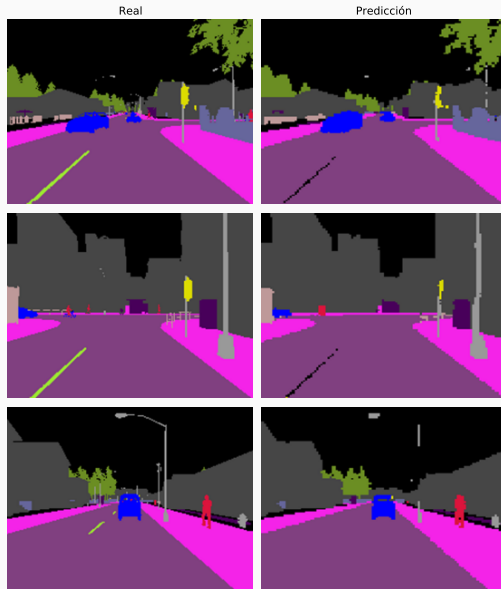
# Resultados Semseg

Clase	Precisión	Exhaustividad	Valor-F
Peatones	0.7141	0.4825	0.5759
Postes	0.7398	0.5404	0.6246
Vehículos	0.9402	0.9254	0.9327
Señales	0.7802	0.6496	0.7089

Clase	IoU
Peatones	45.72%
Postes	31.93%
Vehículos	72.46%
Señales	37.25%



# Resultados Semseg

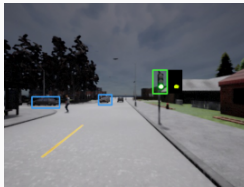


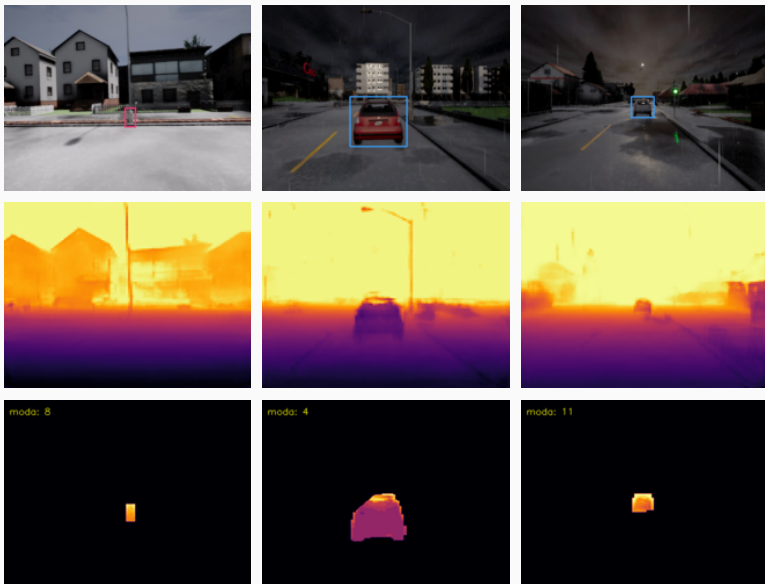
# Semáforos

Rojos



Verdes





# Zonas de Interés CNN

a) Poca iluminación



b) Sol frontal



c) Intersección



