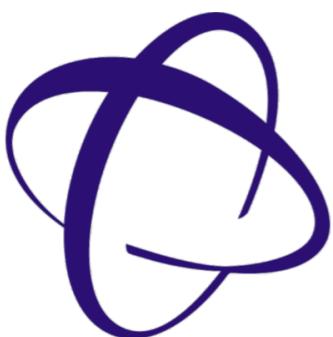


MODELOS DE ATENCIÓN VISUAL MULTIESPECTRAL Y RGB BASADOS EN DEEP LEARNING



Facultad de Ciencias

Trabajo Fin de Grado:

Autora: María José Rueda Montes

Tutora: Eva M. Valero Benito

Co-tutor: Juan L. Nieves Gómez

*Departamento de Óptica
Grado en Óptica y Optometría*



UNIVERSIDAD
DE GRANADA

ÍNDICE

1. INTRODUCCIÓN	3
2. OBJETIVOS.....	4
3. SALIENCIA	5
4. DEEP LEARNING.....	10
5. MODELOS ACTUALES DE SALIENCIA.....	14
6. EVALUACIÓN CUANTITATIVA DE LOS MODELOS.....	21
7. CONCLUSIONES.....	24
8. REFERENCIAS	25

1. INTRODUCCIÓN



(URL 1, en Referencias)



Atención selectiva a las localizaciones importantes



Regiones con mayor saliencia

Predicción de las localizaciones con mayor saliencia como objetivo en Visión Artificial

**Modelos de atención o
Modelos de saliencia**

2. OBJETIVOS

Revisión de las
aportaciones de cada
modelo

Abordar modelos que
utilizan imágenes
multiespectrales



Plantear las Redes
Neuronales
Convolucionales aplicadas
a saliencia



3. SALIENCIA



3. 1. MEDIDA DE LA SALIENCIA

“Eye Trackers” dispositivos que registran los seguimientos oculares



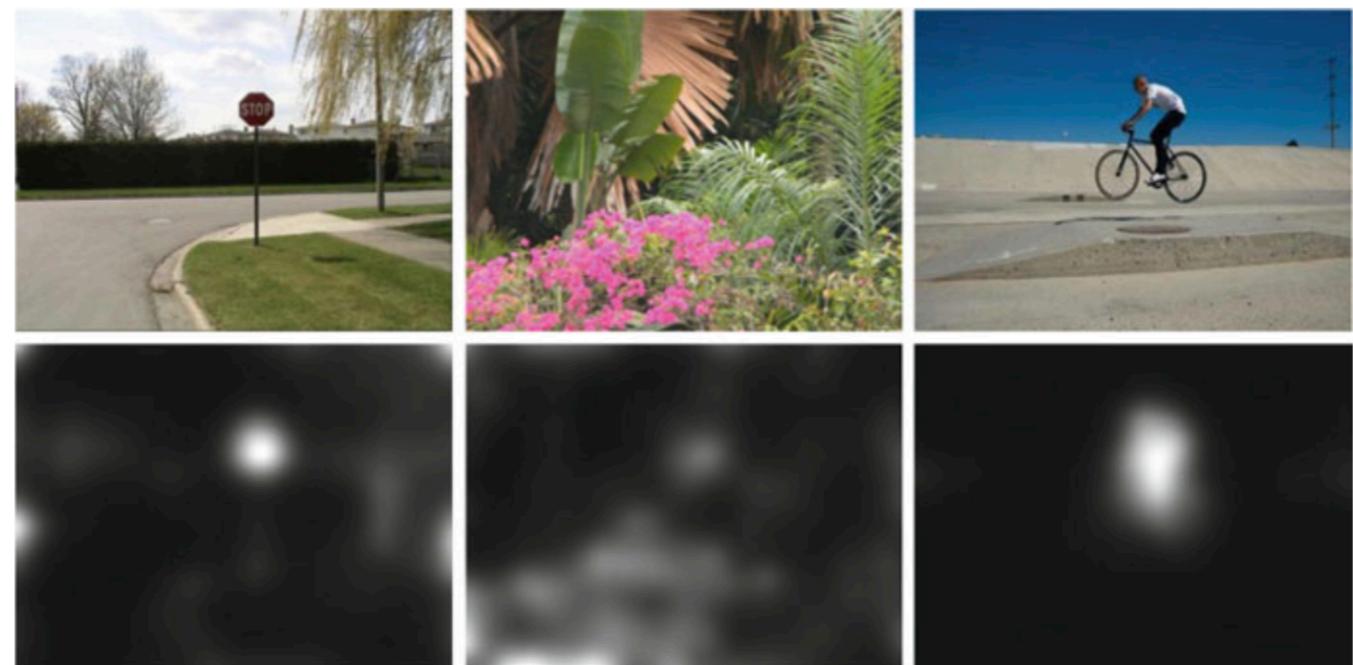
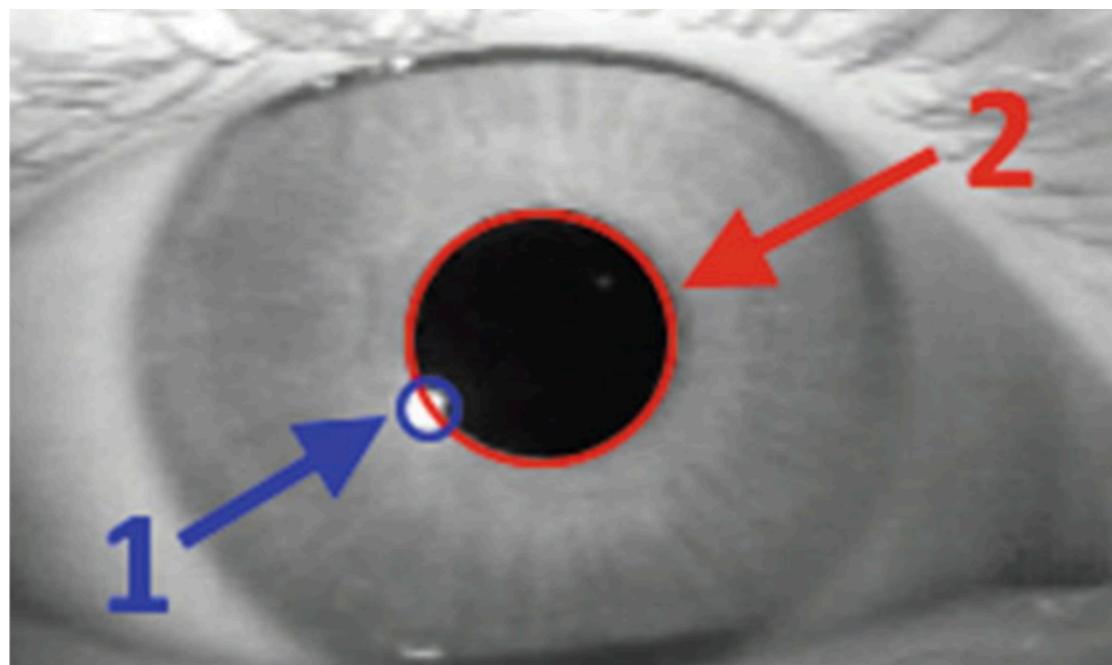
El software de los “Eye Trackers” genera



Referencia para los modelos

Mapa de densidad o mapa de calor

.....
“Ground Truth”

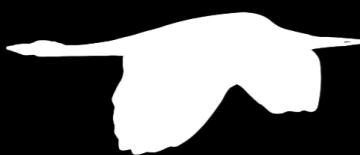


Reflejos corneal 1 y pupilar 2

Imágenes con su respectiva “Ground truth” (GT)



Imagen original



Mapa de predicción

Mapa de detección

3. 2. TIPOS DE MODELOS DE SALIENCIA

.....

Modelos de predicción de saliencia



Intentan localizar las fijaciones oculares

.....

Modelos de detección de saliencia



Detectan y segmentan los objetos salientes

3. 3. BASES DE DATOS Y MÉTRICAS

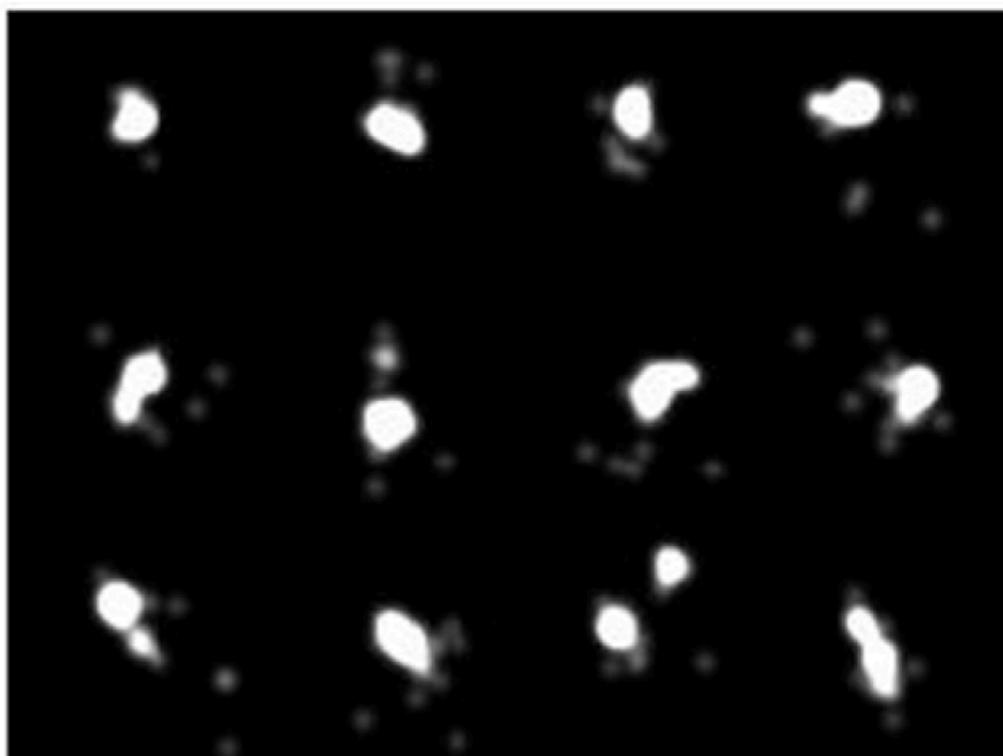
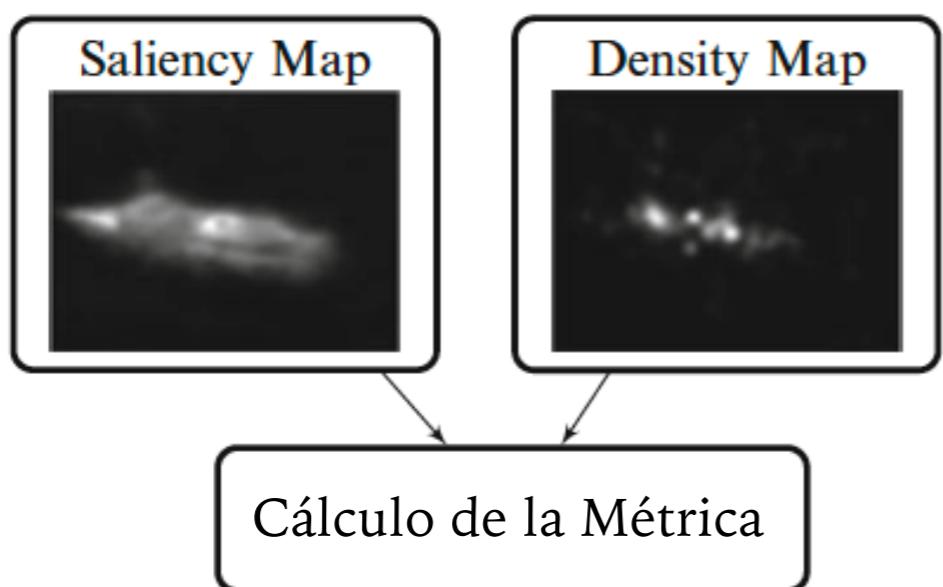
Bases de datos:

Imágenes + “Ground truth”

(CAT2000, MIT300, DUT-OMRON, ETC.)

Métricas: evalúan los modelos

(AUC, F-Score, EMD, NSS, ETC.)

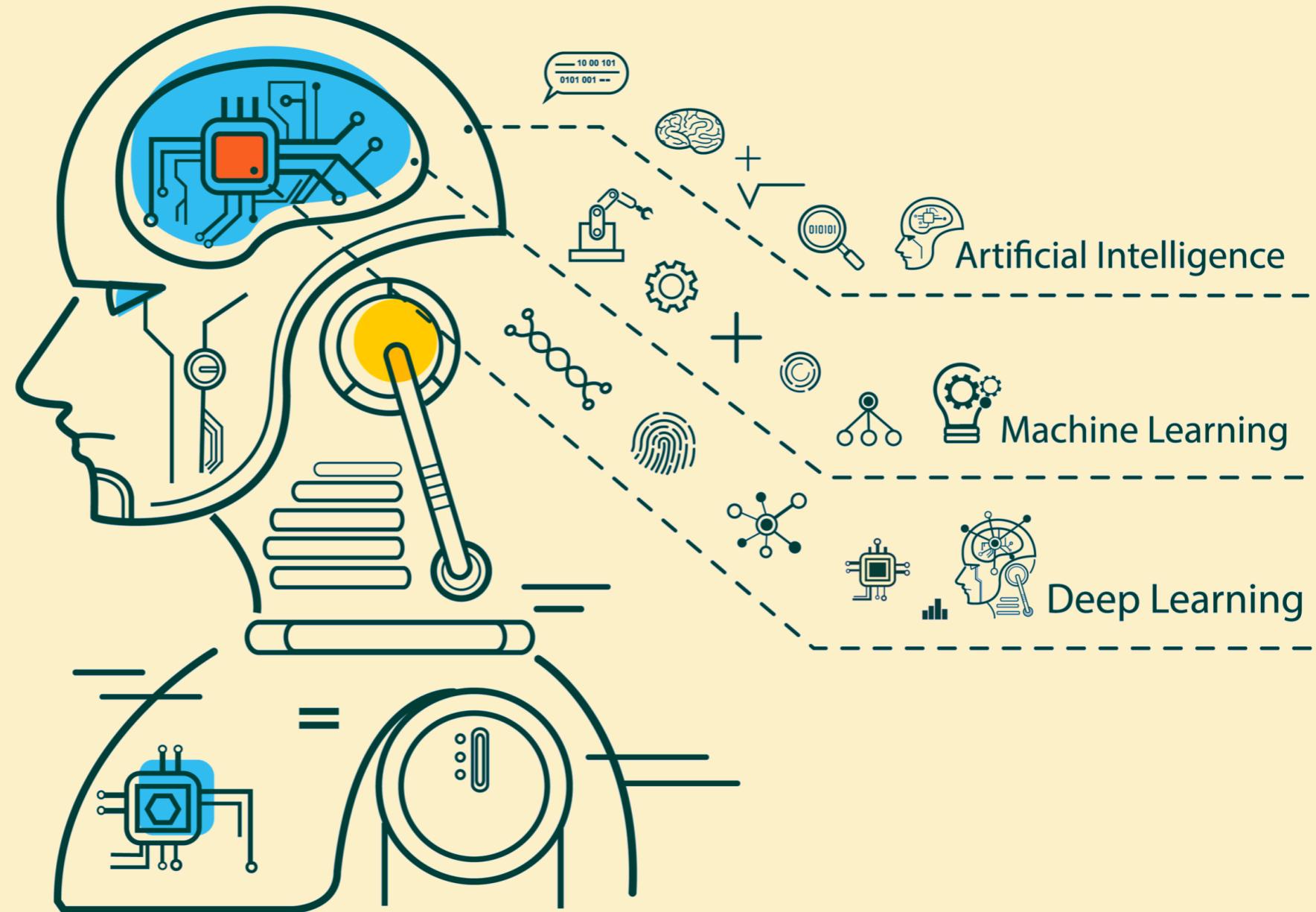


3. 4. SALIENCY BENCHMARK

Servicios encargados de calificar los modelos

mit300	cat2000												
Model Name	Published	Code	AUC-Judd [?]	SIM [?]	EMD [?]	AUC-Borji [?]	sAUC [?]	CC [?]	NSS [?]	KL [?]	Date tested [key]	Sample [img]	
Baseline: infinite humans [?]			0.90	1	0	0.84	0.62	1	2.85	0	Complete results		
Saliency Attentive Model (SAM- ResNet)	Marcella Cornia, Lorenzo Baraldi, Giuseppe Serra, Rita Cucchiara. Predicting Human Eye Fixations via an LSTM-based Saliency Attentive Model [IEEE TIP 2018]	python	0.88	0.77	1.04	0.80	0.58	0.89	2.38	0.56	first tested: 30/10/2016 last tested: 03/03/2017 maps from authors Complete results		
Saliency Attentive Model (SAM- VGG)	Marcella Cornia, Lorenzo Baraldi, Giuseppe Serra, Rita Cucchiara. Predicting Human Eye Fixations via an LSTM-based Saliency Attentive Model [IEEE TIP 2018]	python	0.88	0.76	1.07	0.79	0.58	0.89	2.38	0.54	first tested: 30/10/2016 last tested: 03/03/2017 maps from authors Complete results		
CEDNS	Chunhuan Lin, Fei Qi, Guangming Shi, Hao Li		0.88	0.73	1.27	0.74	0.58	0.85	2.39	0.34	first tested: 24/06/2018 last tested: 24/06/2018 maps from authors Complete results		
MSI-Net	Alexander Kroner, Mario Senden, Kurt Driessens, Rainer Goebel. Contextual Encoder-Decoder Network for Visual Saliency Prediction [arXiv 2019]		0.88	0.75	1.07	0.82	0.59	0.87	2.30	0.36	first tested: 06/12/2018 last tested: 06/12/2018 maps from authors Complete results		
DeepFix	Srinivas S S Kruthiventi, Kumar Ayush, R. Venkatesh Babu. DeepFix: A Fully Convolutional Neural Network for predicting Human Eye Fixations [arXiv 2015]		0.87	0.74	1.15	0.81	0.58	0.87	2.28	0.37	last tested: 02/10/2015 maps from authors Complete results		
EML-NET	Sen Jia. EML-NET: An Expandable Multi-Layer NETwork for Saliency Prediction [arXiv 2018]		0.87	0.75	1.05	0.79	0.59	0.88	2.38	0.96	first tested: 20/03/2018 last tested: 20/03/2018 maps from authors Complete results		

“Mit Saliency Benchmark”



4. DEEP LEARNING



4. 1. CAPA DE CONVOLUCIÓN

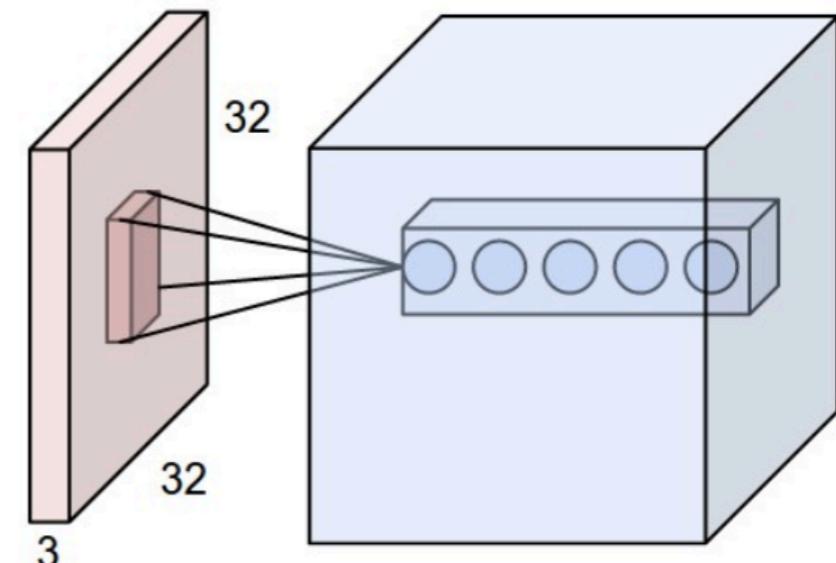
Convolución:

1	1	1	0	0
0	1	1	1	0
0	0	1 _{x1}	1 _{x0}	1 _{x1}
0	0	1 _{x0}	1 _{x1}	0 _{x0}
0	1	1 _{x1}	0 _{x0}	0 _{x1}

Imagen de entrada y
Filtro

4	3	4
2	4	3
2	3	4

Mapa de
características s



Varias neuronas en diferentes capas



Campo receptivo

0	0	0	0	0	0	30
0	0	0	0	50	50	50
0	0	0	20	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0

Representación en píxeles

*

0	0	0	0	0	0	30	0
0	0	0	0	30	0	0	0
0	0	0	30	0	0	0	0
0	0	0	30	0	0	0	0
0	0	0	30	0	0	0	0
0	0	0	30	0	0	0	0
0	0	0	0	0	0	0	0

Filtro en píxeles

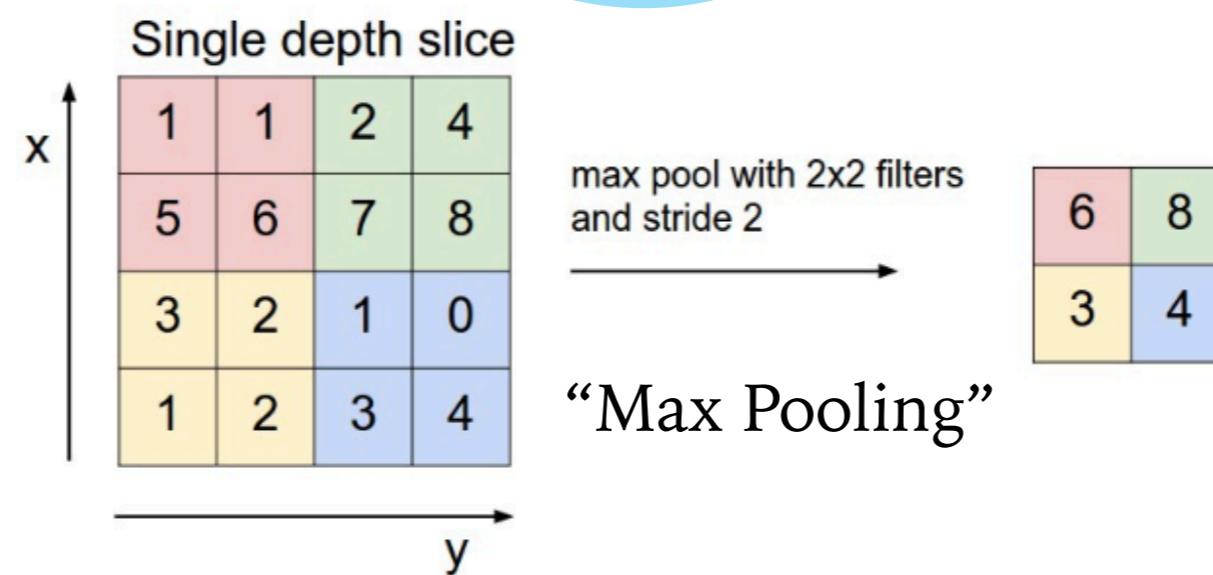
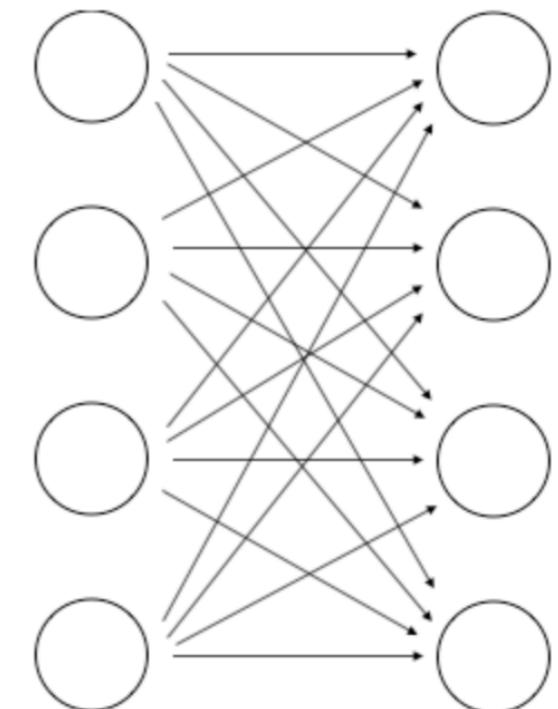
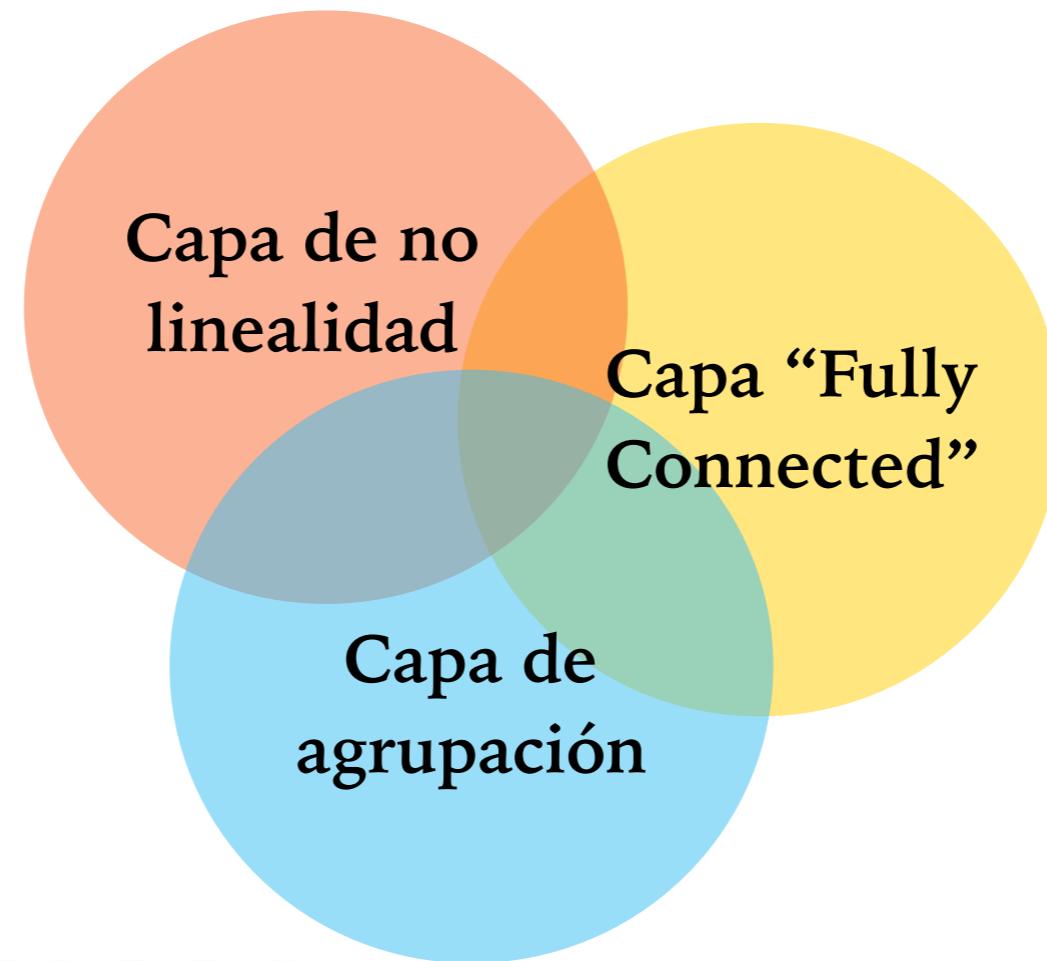
Multiplication and Summation = $(50*30)+(50*30)+(50*30)+(20*30)+(50*30) = 6600$ (A large number!)

4. 2. CAPA DE NO LINEALIDAD, DE AGRUPACIÓN Y FULLY CONNECTED

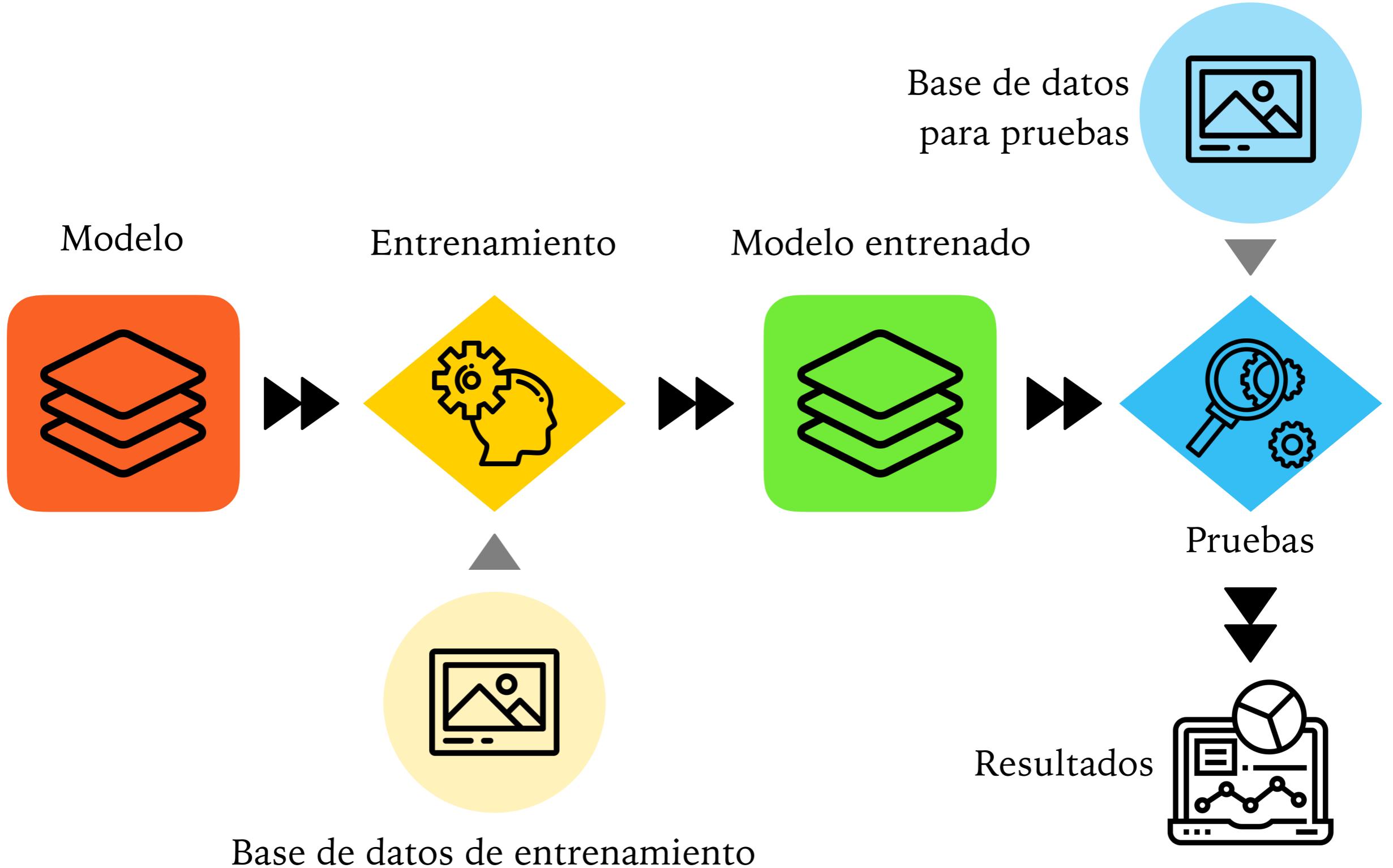
9	3	5	-8
-6	2	-3	1
1	3	4	1
3	-4	5	1

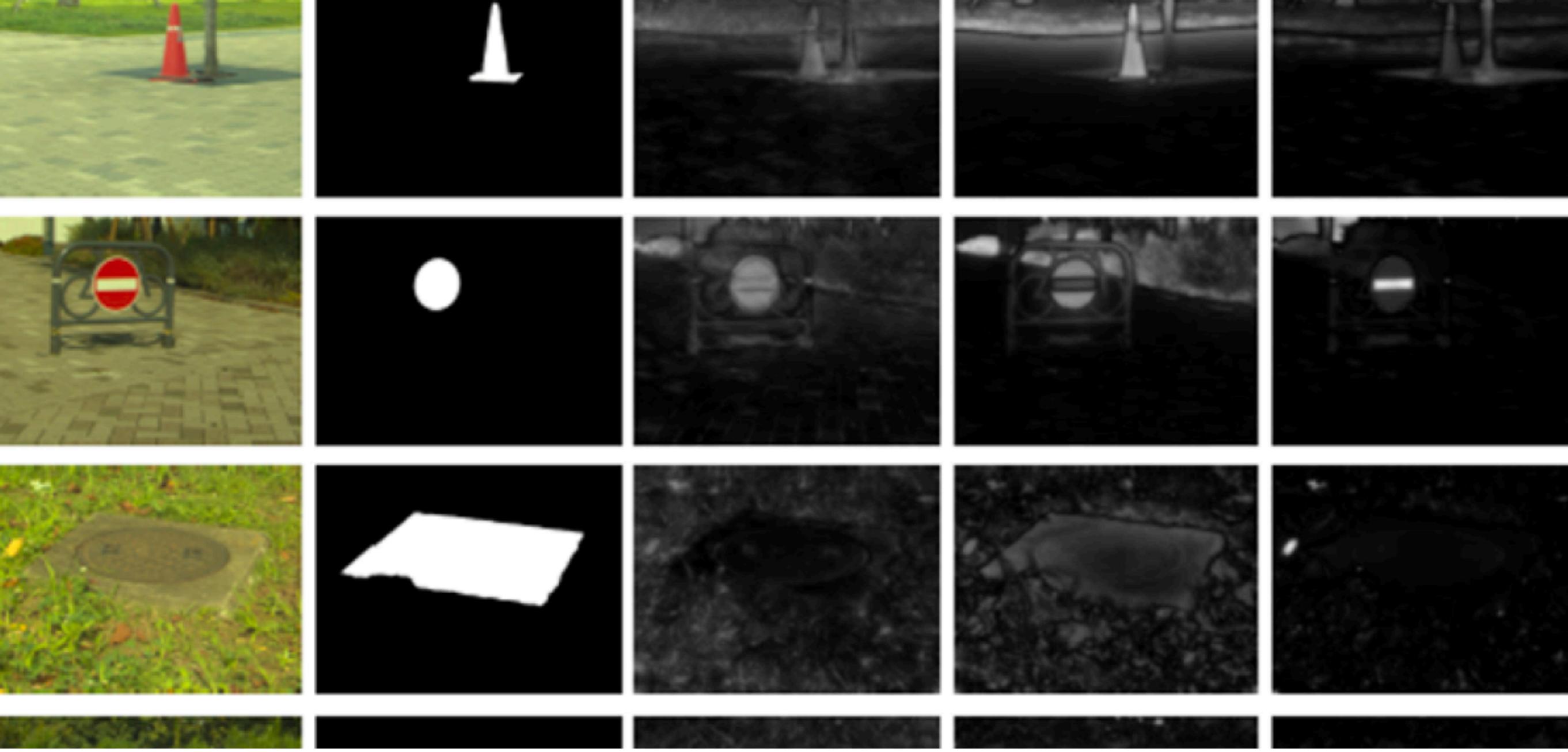
$$f(x) = \max(0, x)$$

9	3	5	0
0	2	0	1
1	3	4	1
3	0	5	1



4. 3. ENTRENAMIENTO Y PRUEBA



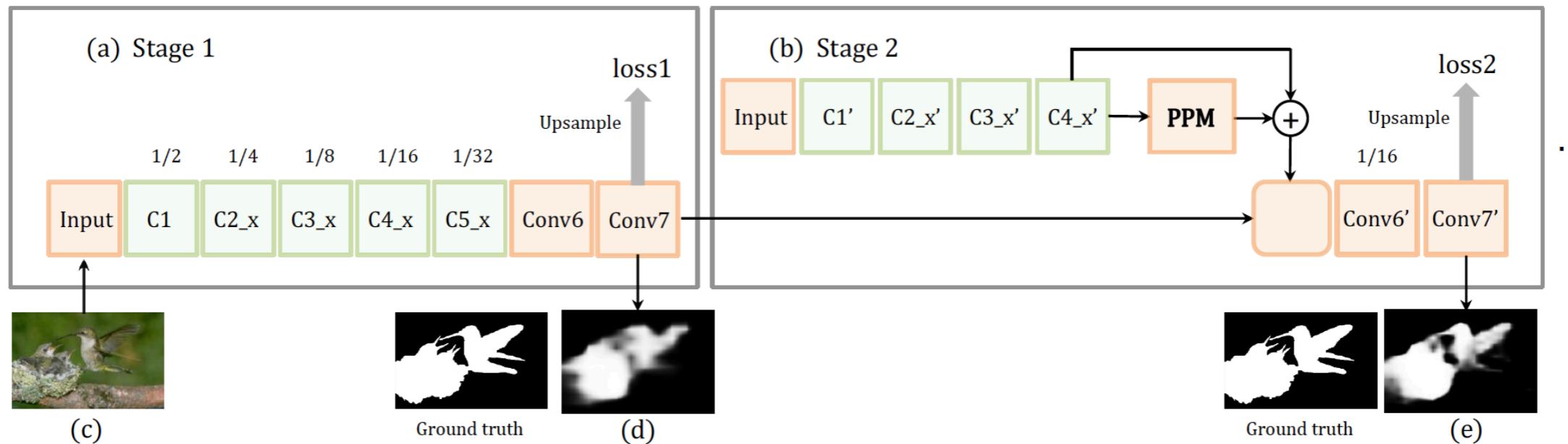


5. MODELOS ACTUALES DE ATENCIÓN VISUAL

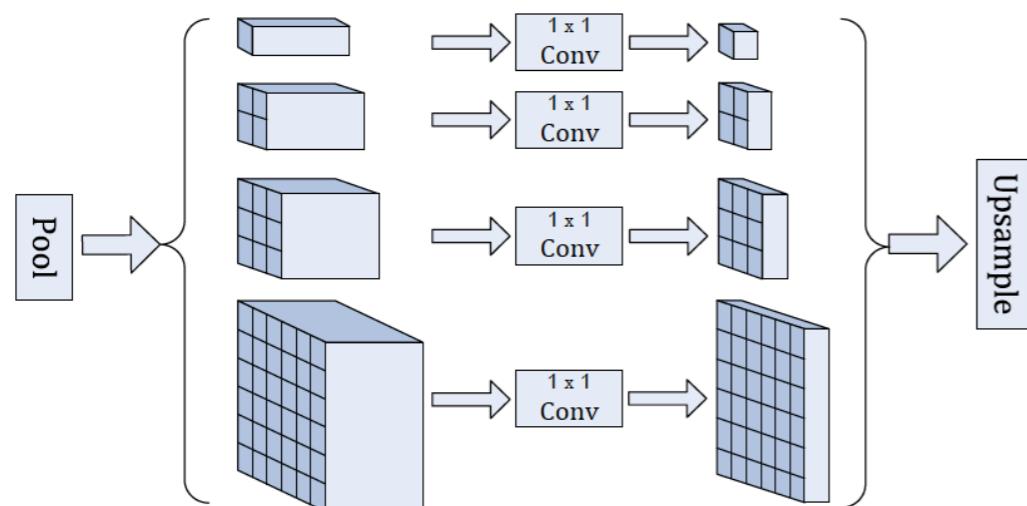


5. 1. MODELOS DE SALIENCIA RGB

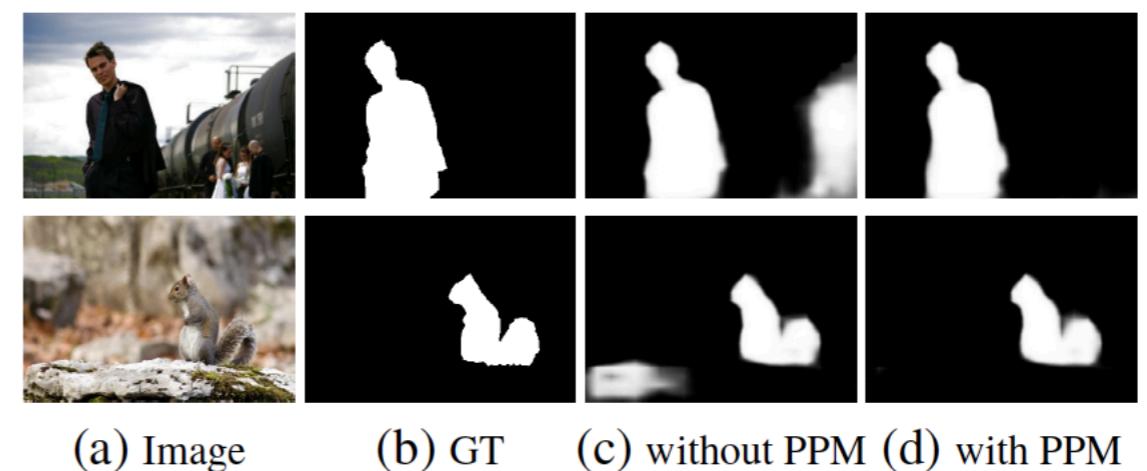
5. 1. a. Modelo de detección de saliencia “Stagewise Refinement”(Wang *et al.*, 2017)



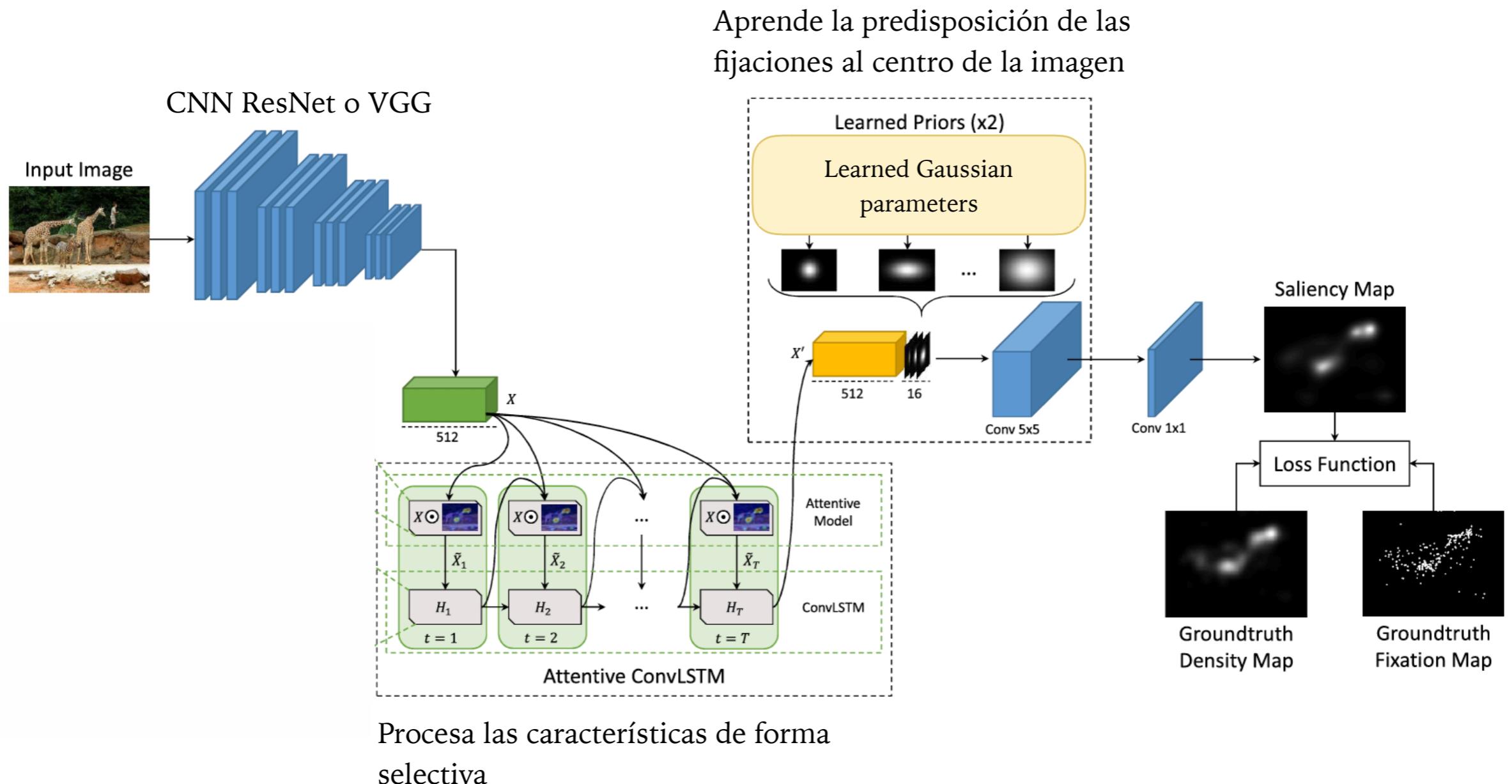
Módulo de agrupación piramidal (PMM):



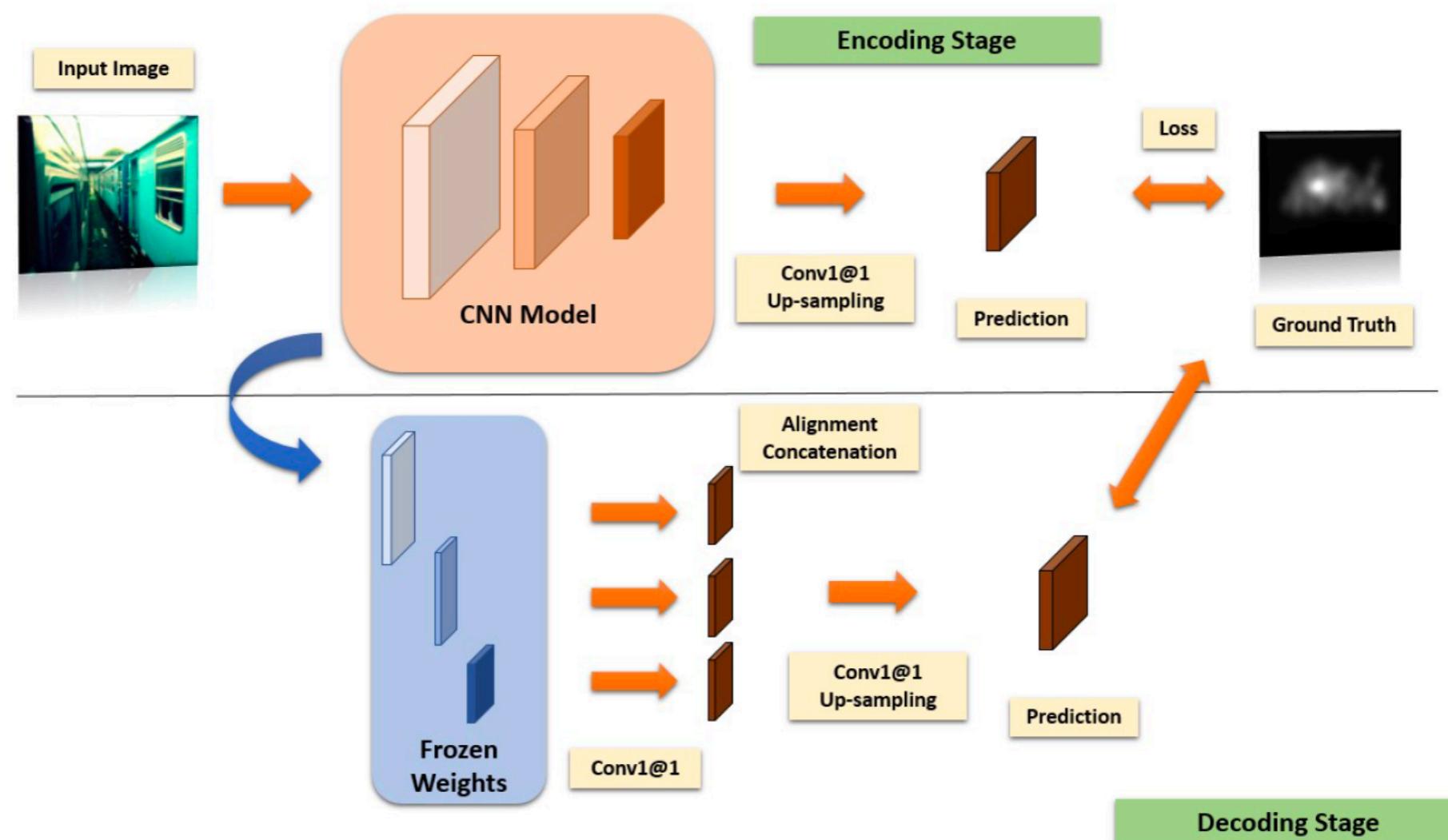
Resultados cualitativos:



5. 1. b. SAM-VGG, SAM-ResNet: “Saliency Attentive Model” basado en LSTM para la predicción de saliencia (Cornia *et al.*, 2018)

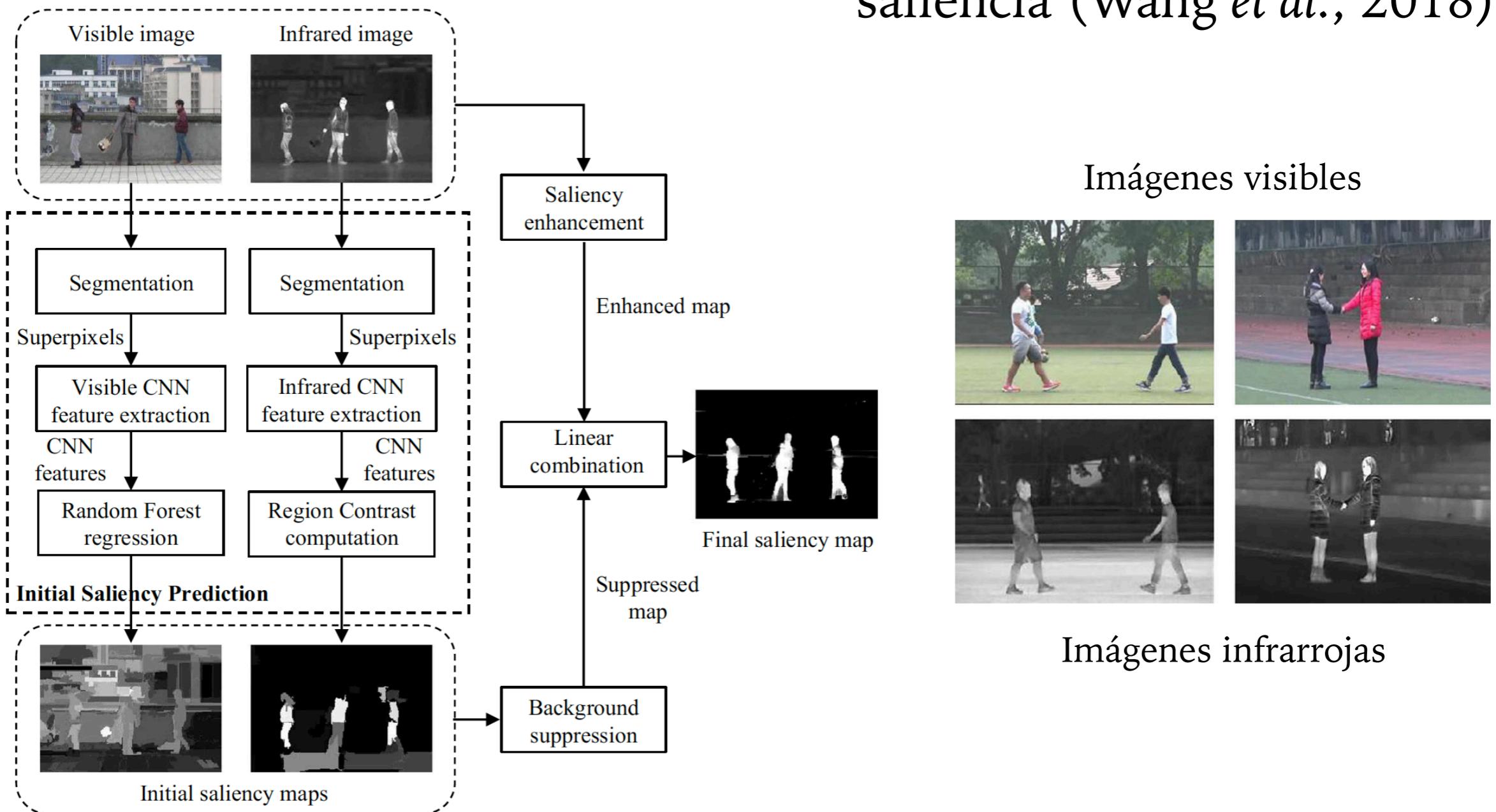


5. 1. c. EML-NET: “Expandable Multi-Layer NETwork” para la predicción de saliencia (Jia y Bruce, 2019)

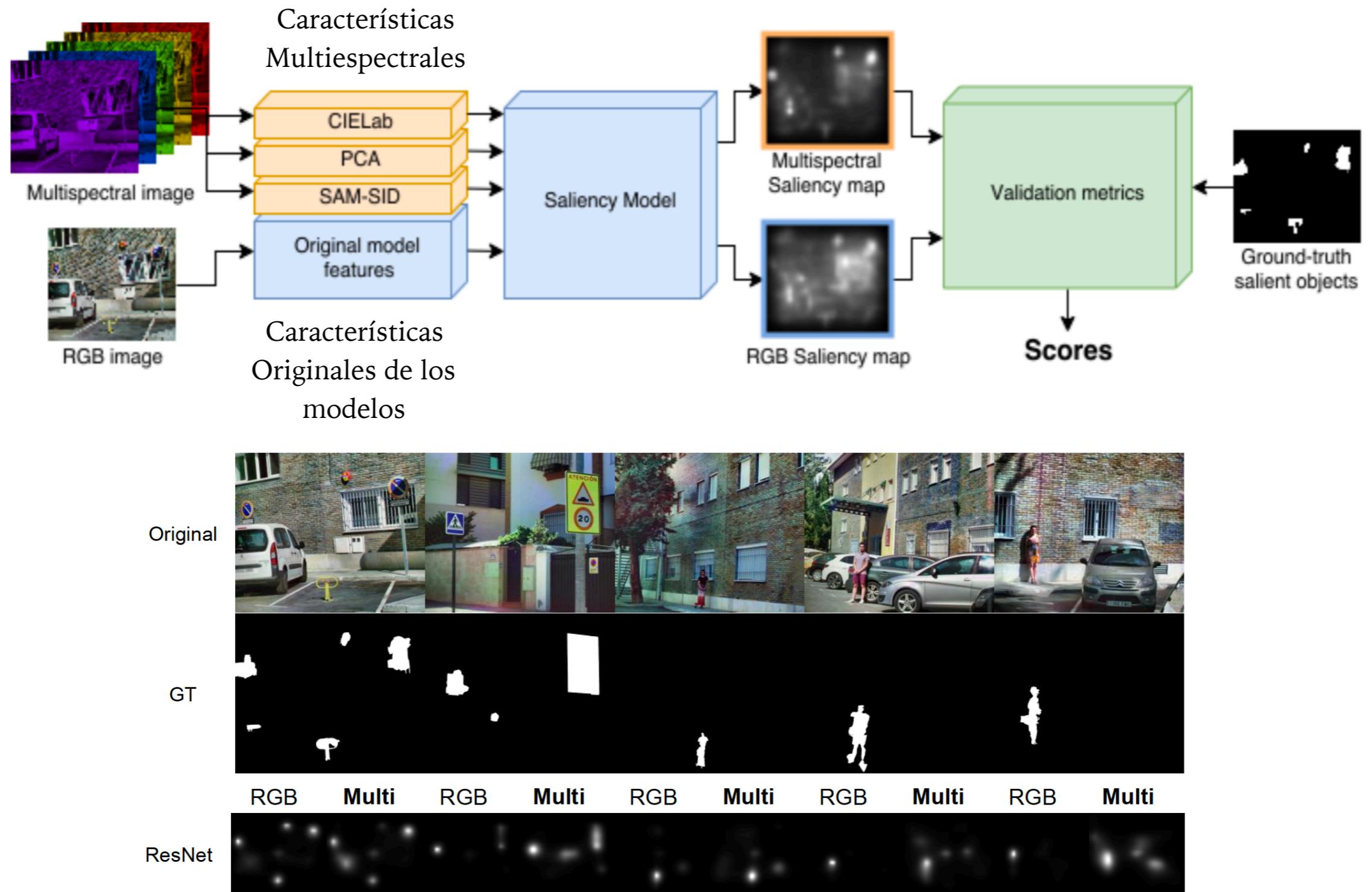


5. 2. MODELOS DE SALIENCIA MULTIESPECTRAL

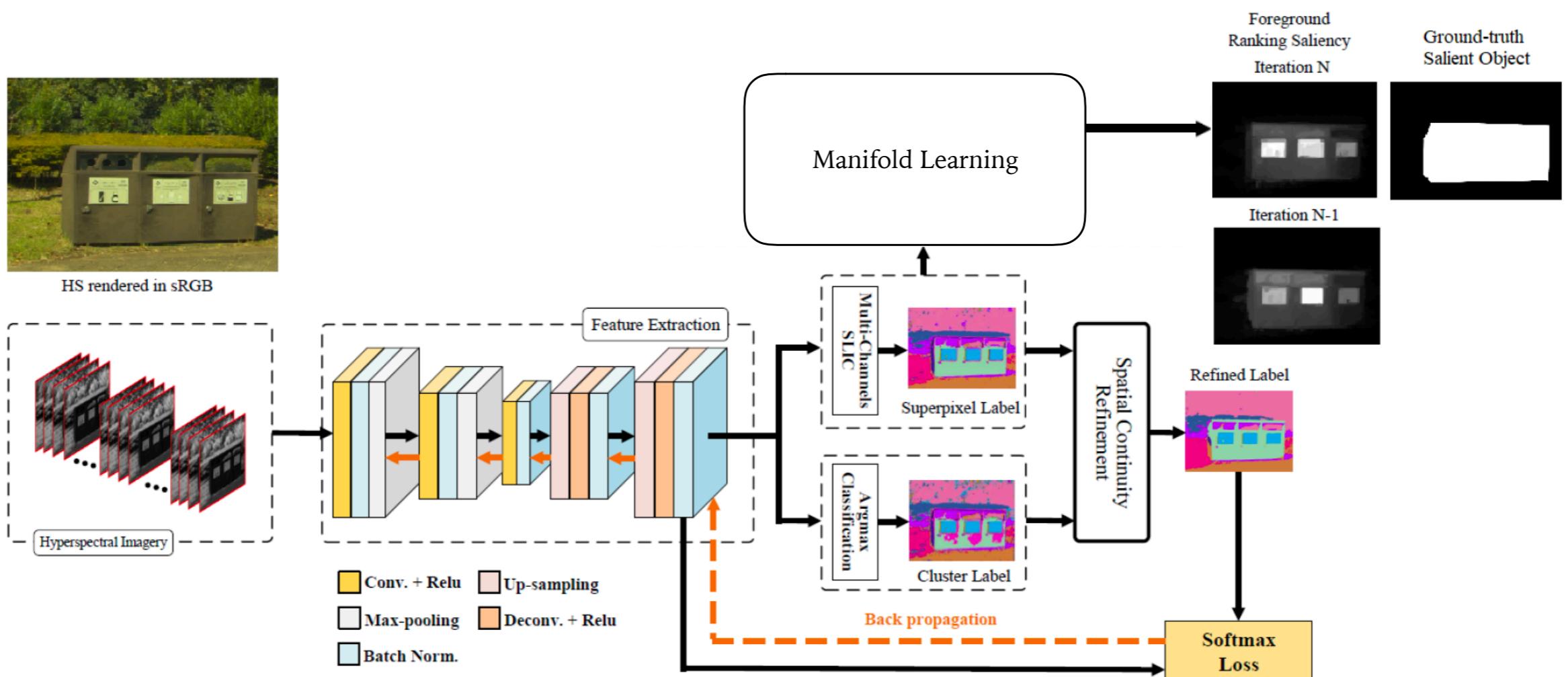
5. 2. a. “Semantic feature based multi-spectral” para la detección de
saliencia (Wang *et al.*, 2018)



5. 2. b. “Multispectral Saliency Validation” en el visible e infrarrojo cercano (Valero *et al.*, 2019)



5. 2. c. SUDF: “Saliency from Unsupervised Deep Features” para la detección de saliencia (Imamoglu *et al.*, 2019)





6. EVALUACIÓN CUANTITATIVA DE LOS MODELOS



6. 1. MODELOS DE PREDICCIÓN DE SALIENCIA

	AUC-Judd	EMD	NSS
EML-NET	0.88	1.84	2.47
SAM-ResNet	0.87	2.15	2.34
SAM-VGG	0.87	2.14	2.30
eDN	0.82	4.56	1.14

Base de datos MIT300

Mejor resultado en **rojo**, peor resultado en **azul**

	AUC-Judd	EMD	NSS
EML-NET	0.87	1.05	2.38
SAM-ResNet	0.88	1.04	2.38
SAM-VGG	0.88	1.07	2.38
eDN	0.85	2.64	1.30

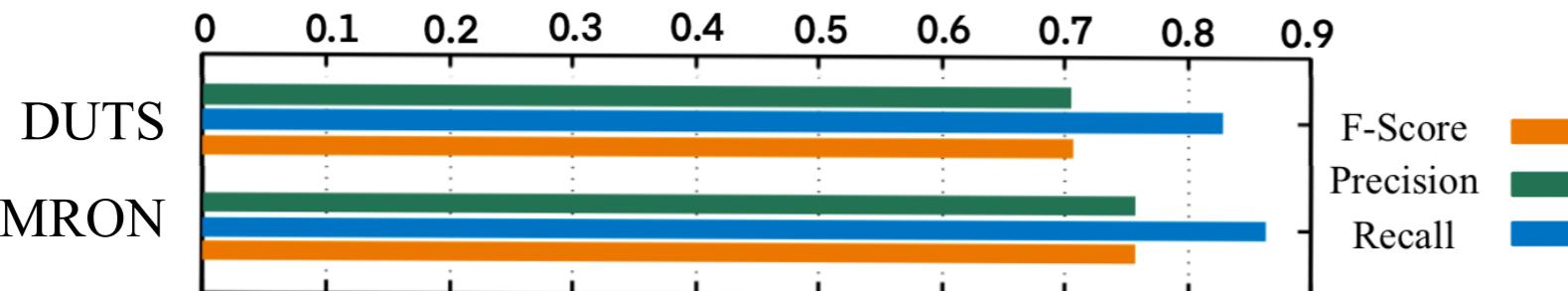
Base de datos CAT2000

		AUC		NSS	
SAM-ResNet	RGB	0.42	63.7 %	2.34	0.04 %
	Hiperespectral	0.69		2.44	

Base de datos del trabajo “Multispectral Saliency Validation”

6. 2. MODELOS DE DETECCIÓN DE SALIENCIA

Modelo “Stagewise Refinement”

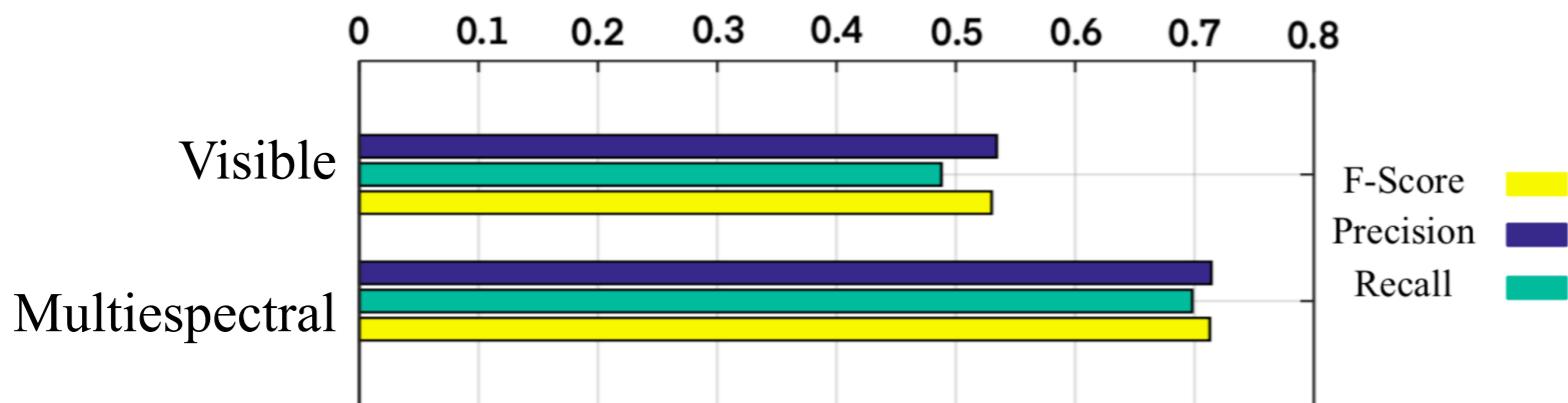


Base de datos DUTS y DUT-OMRON

Modelo “Saliency from Unsupervised Deep Features”

SUDF	
F-Score	0.47
Precision	0.54
Recall	0.48

Modelo “Semantic feature based multi-spectral”



Base de datos infAR

Base de datos HS-SOD

7. CONCLUSIONES

Las aportaciones de los modelos contribuyen al progreso

Labor importante de “Saliency Benchmark”



Campo en desarrollo

8. REFERENCIAS

8. 1. LIBROS

- Duchowski, AT. *Eye Tracking Methodology. Theory and Practice*. Londres: Springer; 2017.
- Goldstein, EB. *Sensation and Perception*. Internacional; Wadsworth; 2009.
- Li, J; Gao, W. Visual Saliency Computation. A machine Learning Perspective. Suiza: Springer; 2014.
- Mancas, M; Ferrera, VP; Riche, N; Taylor, JG. *From Human Attention to Computational Attention. A Multidisciplinary Approach*. Nueva York: Springer; 2016.
- Schanda, J. (Ed.). Colorimetry: understanding the CIE system. Panonia: John Wiley & Sons; 2007.

8. 2. ARTÍCULOS CIENTÍFICOS Y WEBS CIENTÍFICAS

- Albawi, S; Mohammed, T; Al-Zawi, S. Understanding of a convolutional neural network. IEEE. *2017 International Conference on Engineering and Technology (ICET)*; Agosto, 2017; Antalya, Turquía: p. 1-6.
- Borji A, Bylinskii Z, Judd T, Itti L, Durand F, Oliva A; Torralba A, “Mit saliency benchmark,” *Mit saliency benchmark*, consulta 1 de Julio de 2019 en: http://saliency.mit.edu/results_mit300.html
- Borji A; Itti L, “Cat2000: A large scale fixation dataset for boosting saliency research,” *arXiv preprint*, 14 de Mayo de 2015, en: <https://arxiv.org/abs/1505.03581>

- Borji, A. What is a salient object? A dataset and a baseline model for salient object detection. *IEEE Transactions on Image Processing* 2014; 24(2): 742-756.
- Cornia, M; Baraldi, L; Serra, G; Cucchiara, R. Predicting human eye fixations via an lstm- based saliency attentive model. *IEEE Transactions on Image Processing* 2018; 27(10): 5142-5154.
- Deshpande, A, “A Beginner’s Guide To Understanding Convolutional Neural Networks”, *CS Undergrad at UCLA* ('19), 2017, en: <https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/>
- Imamoglu N, Ding G, Fang Y, Kanezaki A, Kouyama T; Nakamura R, Salient object detection on hyperspectral images using features learned from unsupervised segmentation task. *arXiv preprint*, 28 de Febrero de 2019, en: <https://arxiv.org/abs/1902.10993>
- Imamoglu, N; Oishi, Y; Zhang, X; Ding, G; Fang, Y; Kouyama, T; Nakamura, R. Hyperspectral Image Dataset for Benchmarking on Salient Object Detection. IEEE. *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*; Mayo 2018; Cagliari, Italia: p. 1-3.
- E Itti, L; Koch, C; Niebur, E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence* 1998; (11): 1254-1259.
- F Jia S; Bruce N, “Eml-net: An expandable multi-layer network for saliency prediction. *arXiv preprint*, 2 de Mayo de 2018, en: <https://arxiv.org/abs/1805.01047>
- Jiang, M; Huang, S; Duan, J; Zhao, Q. Salicon: Saliency in context. IEEE. Proceedings of the IEEE conference on computer vision and pattern recognition; Junio, 2015; Boston, Massachusetts: p. 1072-1080.
- Judd T, Durand F; Torralba A, “A benchmark of computational models of saliency to predict human fixations.,” *MIT Libraries*, 13 de Enero de 2012, en: <https://dspace.mit.edu/handle/ 1721.1/68590>
- Judd, T; Ehinger, K; Durand, F; Torralba, A. Learning to predict where humans look. IEEE. *2009 IEEE 12th international conference on computer vision*; Septiembre, 2009; Kyoto, Japón: p. 2106-2113.
- Kruthiventi, S; Ayush, K; Babu, R. Deepfix: A fully convolutional neural network for predicting human eye fixations. *IEEE Transactions on Image Processing* 2017; 26(9): 4446-4456.
- Ng A; Ngiam J; Foo C; Mai Y; Suen C; Coates A; Maas A; Hannun A; Huval B; Wang T; Tandon S, “Feature Extraction Using Convolution.,” *UFLDL Tutorial*, 2 de Junio 2019, en: <http://ufldl.stanford.edu/tutorial/supervised/FeatureExtractionUsingConvolution/>

- Oliva, A; Torralba, A; Castelhano, M; Henderson, J. M. Top-down control of visual attention in object detection. IEEE. *Proceedings 2003 International Conference on Image Processing* (Cat. No. 03CH37429); Septiembre, 2003; Barcelona, España: p. 1-253
- Ruan X, Tong N; Lu H, “How far we away from a perfect visual saliency detection - DUT- OMRON: a new benchmark dataset,” *The DUT-OMRON Image Dataset*, 7 de Enero de 2019, en: <http://saliencydetection.net/dut-omron/>
- Valero, E; Nieves, JL; Etchebehere, S. Multispectral saliency validation in the visible and near-infrared. *Sometido a Color Research and Application*, 2019.
- Vig, E; Dorr, M; Cox, D. Large-scale optimization of hierarchical features for saliency prediction in natural images. IEEE. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2014; Columbus, Ohio: p. 2798-2805.
- Wang, T; Borji, A; Zhang, L; Zhang, P; Lu, H. A stagewise refinement model for detecting salient objects in images. IEEE. *Proceedings of the IEEE International Conference on Computer Vision*; 2017; Venecia, Italia: p. 4019-4028.
- Wang, L; Gao, C; Jian, J; Tang, L; Liu, J. Semantic feature based multi-spectral saliency detection. *Multimedia Tools and Applications* 2018; 77(3): 3387-3403.
- Zhai, Y; Shah, M. Visual attention detection in video sequences using spatiotemporal cues. ACM. *Proceedings of the 14th ACM international conference on Multimedia*; Octubre, 2006; California, EEUU: p. 815-824.

8. 3. ENLACES A LAS IMÁGENES DE LA PRESENTACIÓN

- URL 1: <https://images.freeimages.com/images/large-previews/7da/balloon-against-sky-1414259.jpg>
- URL 2: https://warioduckerman.com/wp-content/uploads/2019/02/AZ_fall-2018_shutterstock_1096779713_crp-for-web-1.jpg
- URL 3: <https://cdn.lynda.com/course/585239/585239-636374514847689553-16x9.jpg>