

A survey on deep learning techniques in image and video semantic segmentation

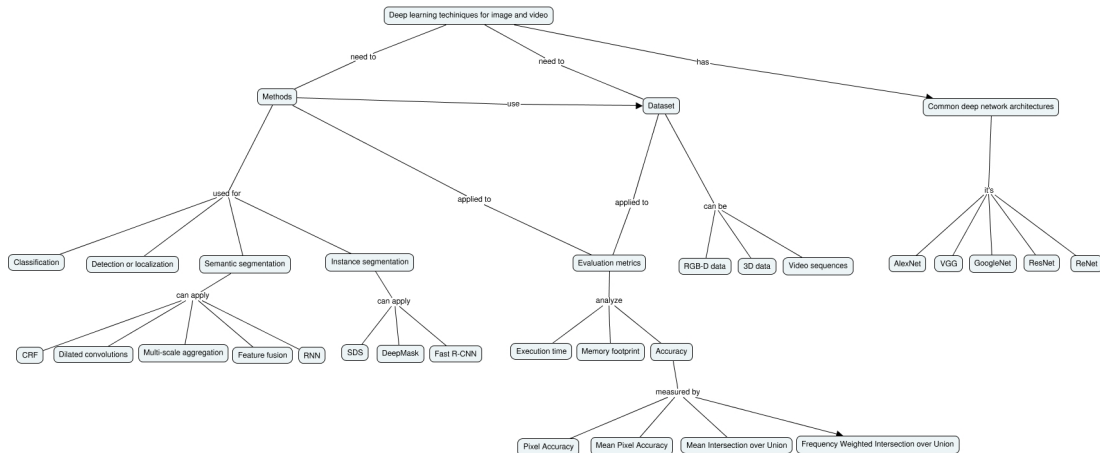
(paper analysis)

Jessica Motta

SENAI CIMATEC

June 2, 2020

Concept Map



CNN- How it works?

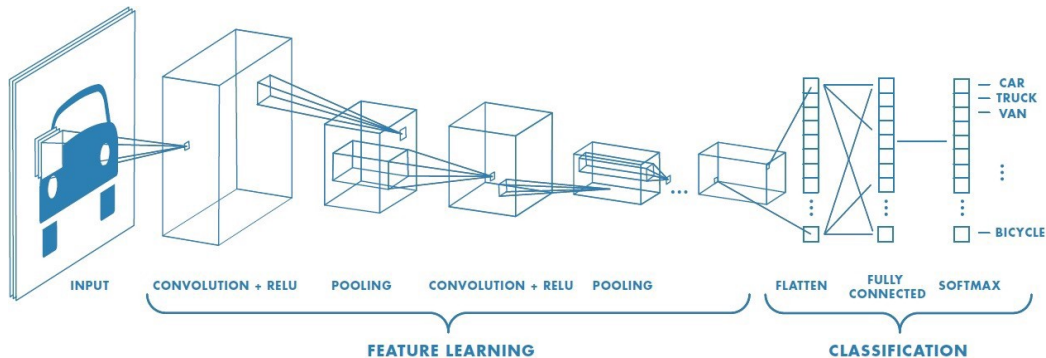


Figure: Convolutional Neural Network [Sha19]

Common deep networks architecture

COMPARATIVE FOR COMMON DEEP NETWORK ARCHITECTURES			
Network	Year champion ILSVRC*	Number of Layers	Accuracy
AlexNet	2012	3	84.6%
VGG	2013	16	92.7%
GoogleNet	2014	22	93.3%
ResNet	2016	152	96.4%

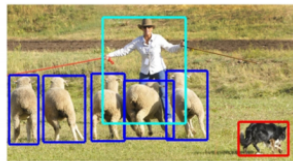
Table: Deep network architectures. [GGOE⁺18]

*ILSVRC (ImageNet Large Scale Visual Recognition Challenge)

Methods to image analysis



(a) Image classification



(b) Object localization



(c) Semantic segmentation



(d) Instance segmentation

Figure: Methods to image analysis. [LMB⁺14]

Segmentation apply in video

Evaluation Metrics

Execution time

Memory footprint

Accuracy



IoU calculation

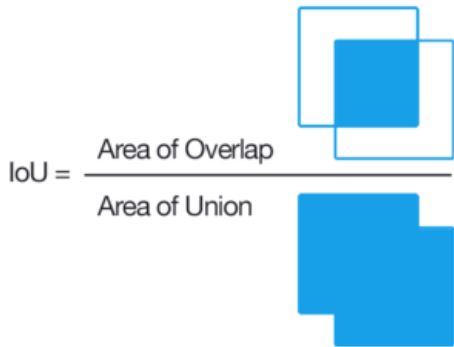


Figure: IoU calculation visualized. [Tiu19]

Accuracy



Figure: Accuracy evaluation. [Ros16]

Accuracy results

ACCURACY RESULTS (METHODS AND DATASETS) (%)									
Method / Dataset	PASCAL VOC-2012	Pascal- Person- Part	CamVid	CityScapes	Stanford Background	SiftFlow	SUN3D	ShapeNet Part	Youtube- Objects
PSPNet	85,4								
DeepLab		64,94							
DAG-RNN			91,60						
rCNN					80,20				
LSTM-CF							58,50		
PointNet								83,70	
PointNet++								85,10	
DGCNN								85,10	
Clockwork Convnet									68,50
SegmPred				59,40					

Table: Accuracy results for the most relevant methods and dataset. [GGEO⁺18]

Wich cases doesn't apply deep learning?

For **high performance**, deep networks require **extremely large** datasets.

It's **expensive** to get data, computer power and hiring researchers.

Deep networks **aren't easily interpreted** as classical Machine Learning algorithms.



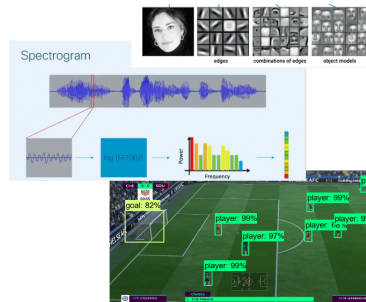
Advantages to use Deep learning against to Classical methods

Better performance (accuracy results)

More data works to DP **but not** with CML

DP **no need** for feature engineering

It's **adaptable** to differents domains and applications



Advantages to use Classical methods against to Deep learning

Woks better with **small dataset**

Low computational and financial cost

The algorithms it's **easier** to understand and interpret

Conclusion

A survey to **help** researchers to choose the **best method and dataset** it's appropriate for their application

Semantic segmentation has been approached with many **success** stories but has a **problem** whose solution would be useful for real-world applications

And **deep learning** has proved to be extremely **powerful** to solve this problem

Future research directions

3D datasets

Sequence datasets

Point cloud segmentation using Graph Convolutional Networks (GCN)

Context knowledge






Real-time segmentation

Memory

Temporal coherency on sequences

Multi-view integration

References I

-  Alberto Garcia-Garcia, Sergio Orts-Escolano, Sergiu Oprea, Victor Villena-Martinez, Pablo Martinez-Gonzalez, and Jose Garcia-Rodriguez, *A survey on deep learning techniques for image and video semantic segmentation*, Applied Soft Computing **70** (2018), 41–65.
-  Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick, *Microsoft coco: Common objects in context*, European conference on computer vision, Springer, 2014, pp. 740–755.
-  Adrian Rosebrock, *Intersection over union (iou) for object detection*, 2016.
-  Shashikant, *Convolutional neural network: A step by step guide*, 2019.
-  Ekin Tiu, *Metrics to evaluate your semantic segmentation model*, 2019.