Università degli Studi dell'Insubria
Varese, Italy

LEARNING OBJECT SEGMENTATION USING A MULTI NETWORK SEGMENT CLASSIFICATION APPROACH

Simone Albertini
Ignazio Gallo, Angelo Nodari, Marco Vanetti

albertini.simone@gmail.com

February 25th, 2012

The problem

Object of interest segmentation

Prior Knowledge

Object Segmentation

The class the object belongs to

Useful for various high level tasks

The context

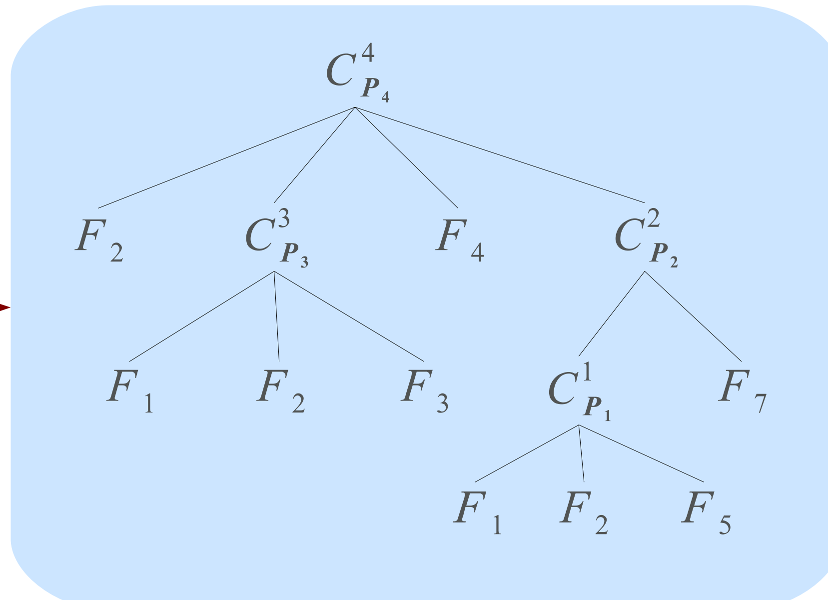
Commercial products images:

Low resolution

Artifacts

The proposed solution

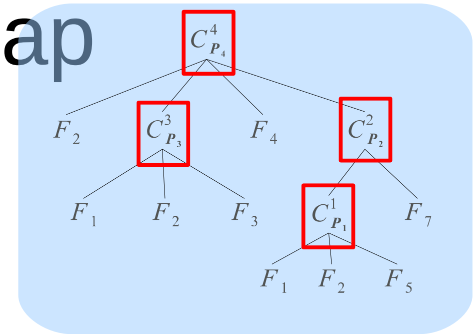
Multi-net for Object Segmentation (MNOS)

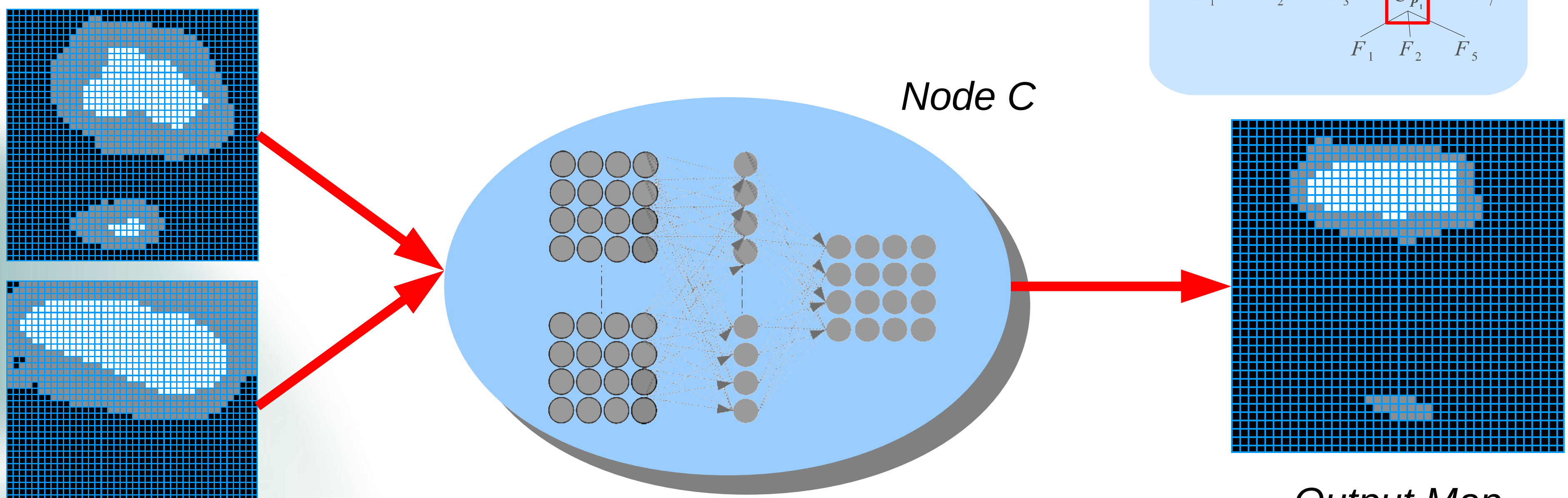


- Inner node *C*: *MLP* network with its configuration *P*
- Leaf *F*: feature node

The proposed solution

Each inner node C takes as input several map-images and predicts an output map





Output Map

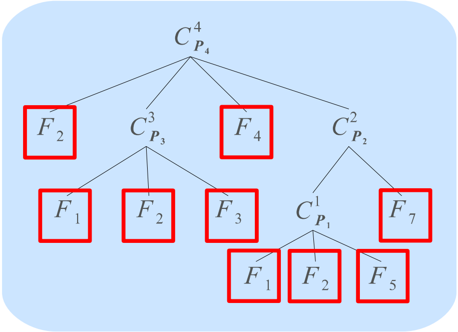
Input Maps
(one for each child node)

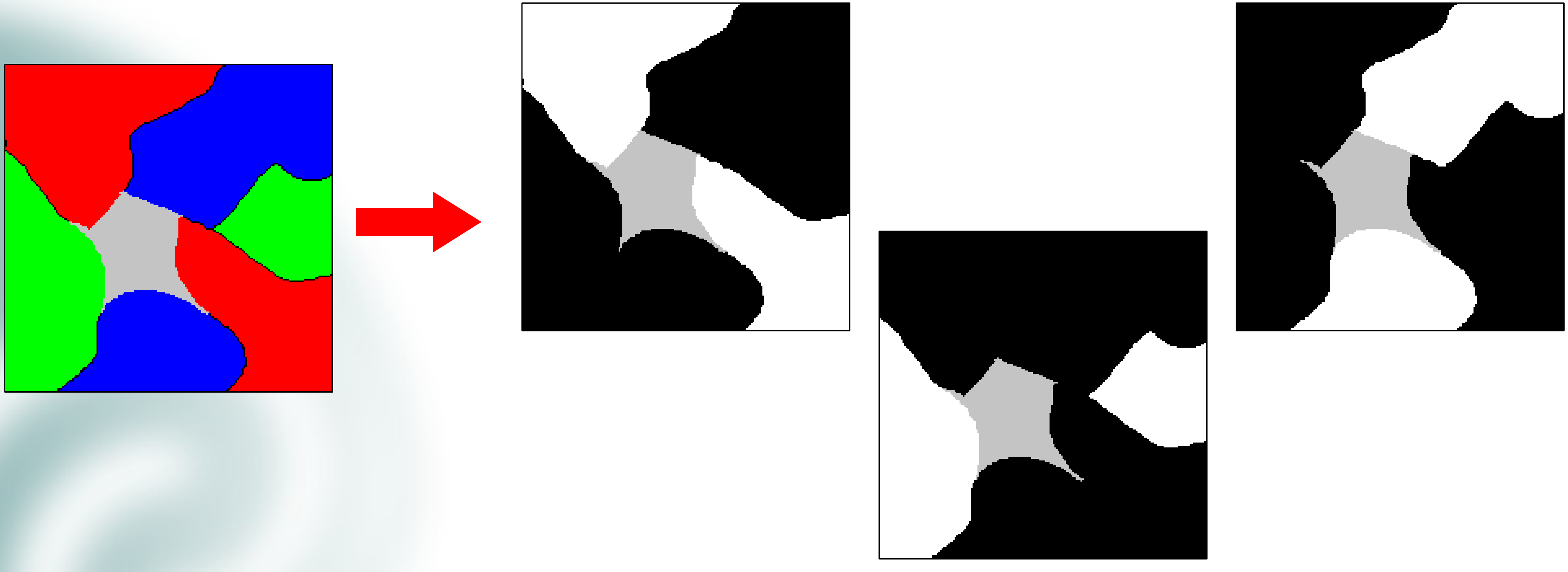
The Network predicts the probability each pixel belongs to the Object of Interest

The proposed solution

A leaf node only applies operators and transformations to the original image:

- Extracts color channels (RGB or HSV)

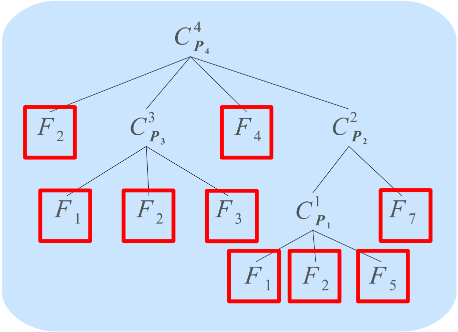


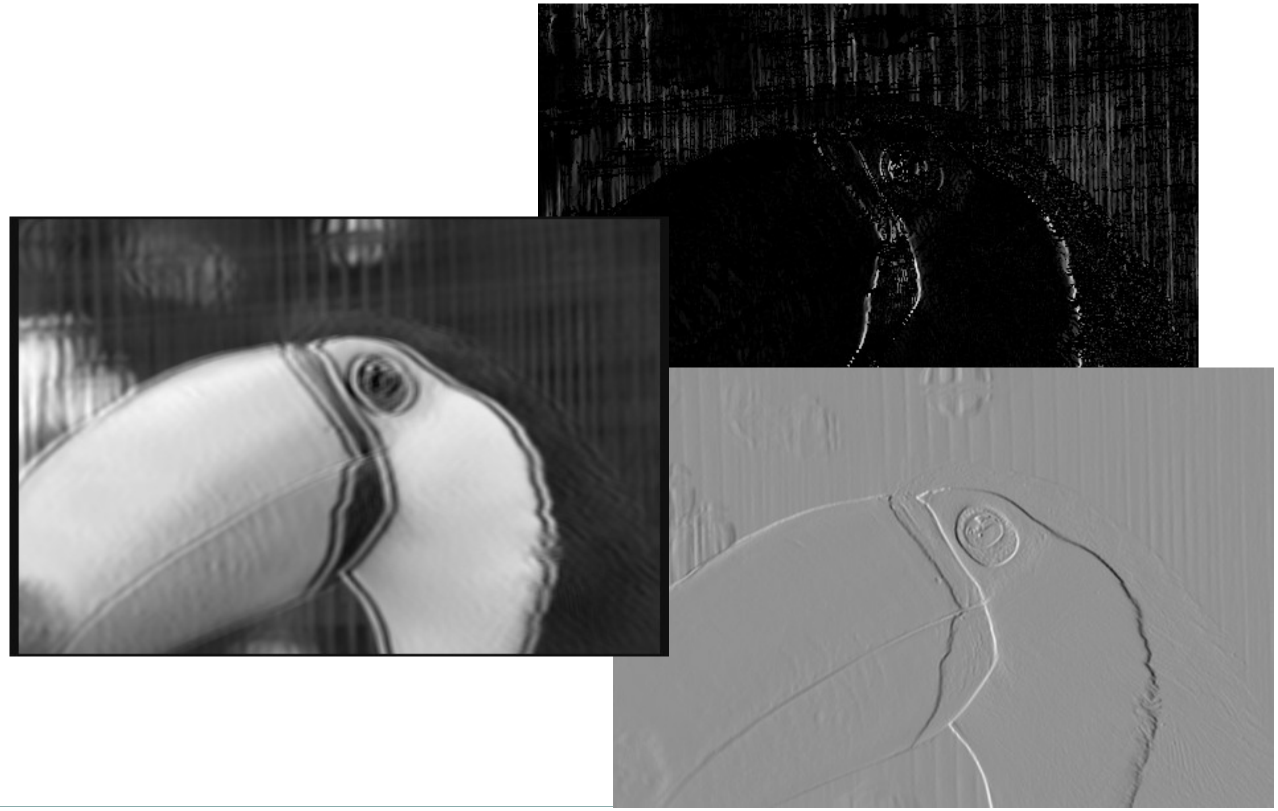


The proposed solution

A leaf node only applies operators and transformations to the original image:

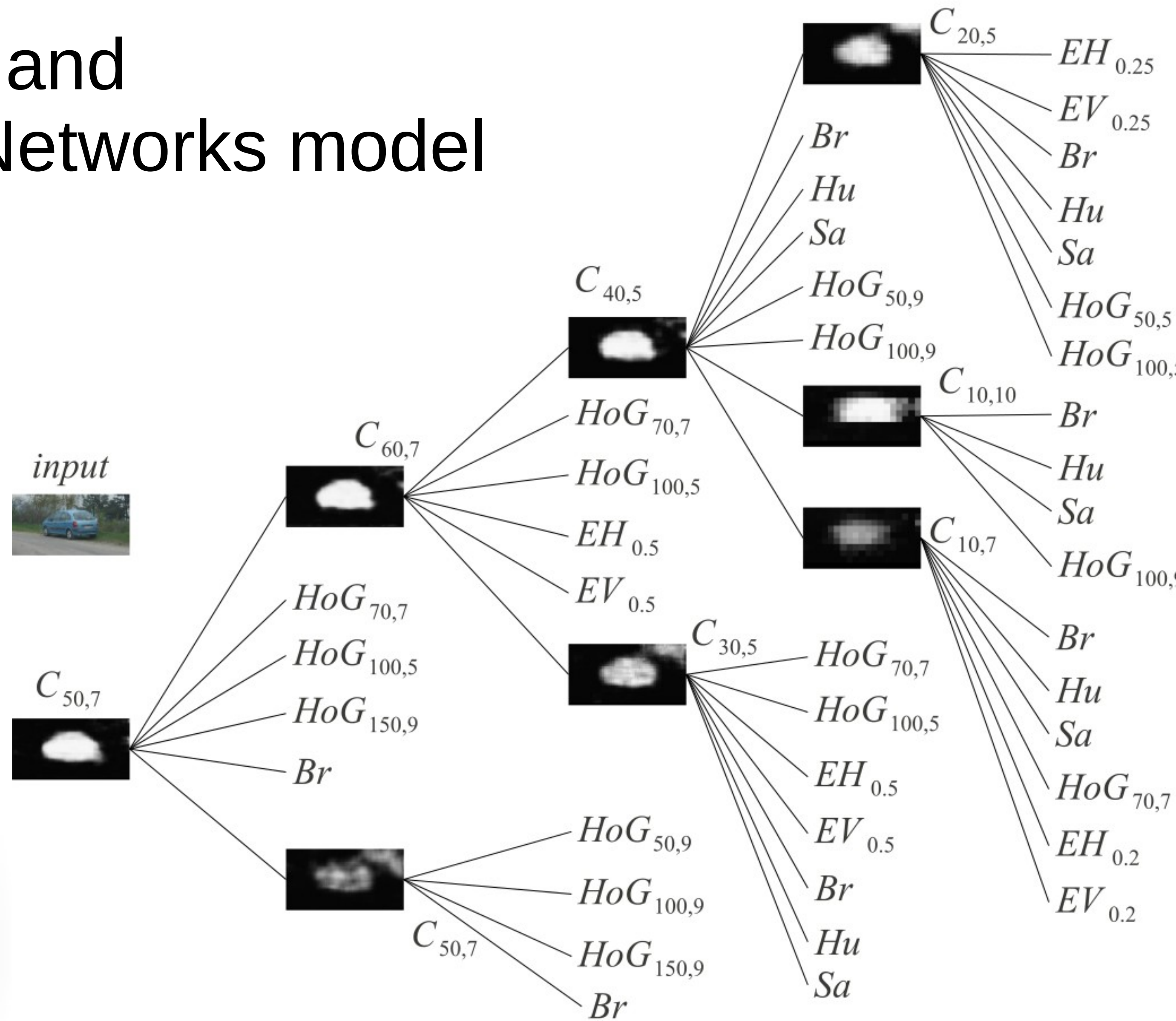
- Haar, Hog, High Frequencies





Multi-Net for Object Detection

Feed-forward and
Multi-Neural Networks model



Two possible internal nodes

- Sliding window approach:

  Pattern generated from the raw intensity values of the input images

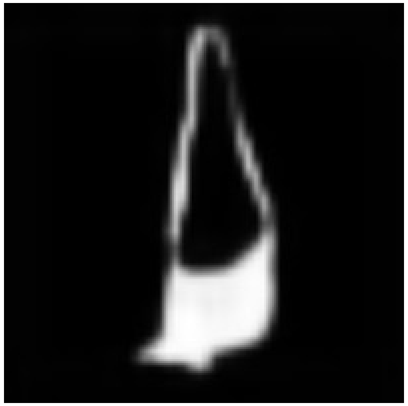
Gallo, I. and Nodari, A. (2011).
Learning object detection using multiple neural netwoks.
In VISAPP 2011. INSTICC Press.

- Segment-based approach:

  Input images are partitioned and features are extracted from each segment

Two possible internal nodes

Nodes with sliding windows operate at pixel level

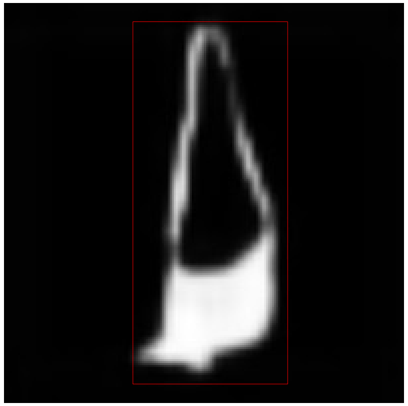


Two possible internal nodes

Nodes with sliding windows operate at pixel level

Two possible internal nodes

Nodes with sliding windows operate at pixel level

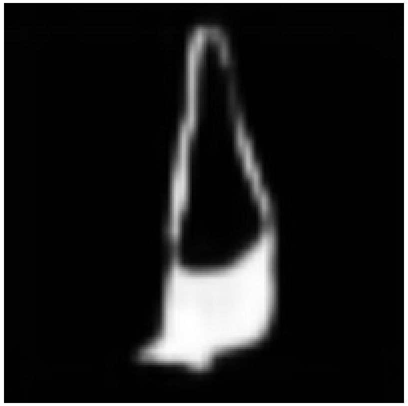

Good for detection

Objective:

sharp and neat edges to segment the object of interest

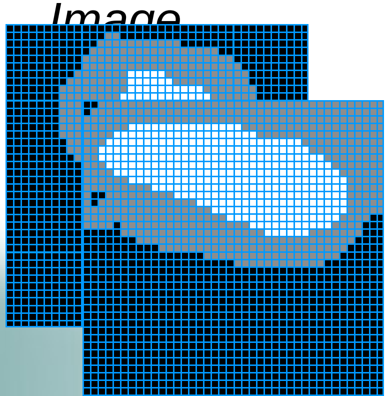
Two possible internal nodes

Nodes with sliding windows operate at pixel level

Segments

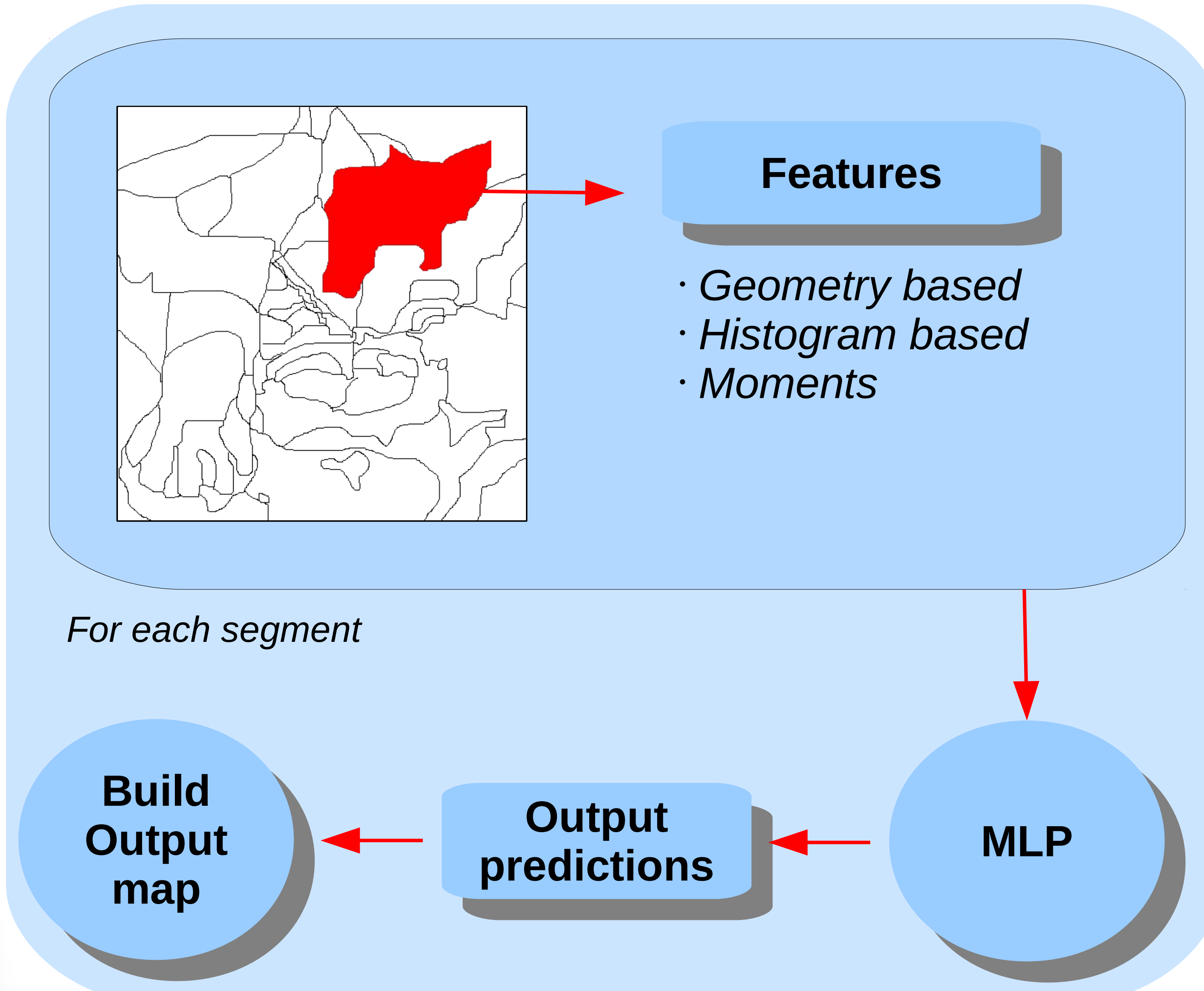
Original Image



Input maps

Output map

Internal node



Segment Features

- Geometric:

  Area, Perimeter

  Perimeter over area

  Bounding box locations and dimensions

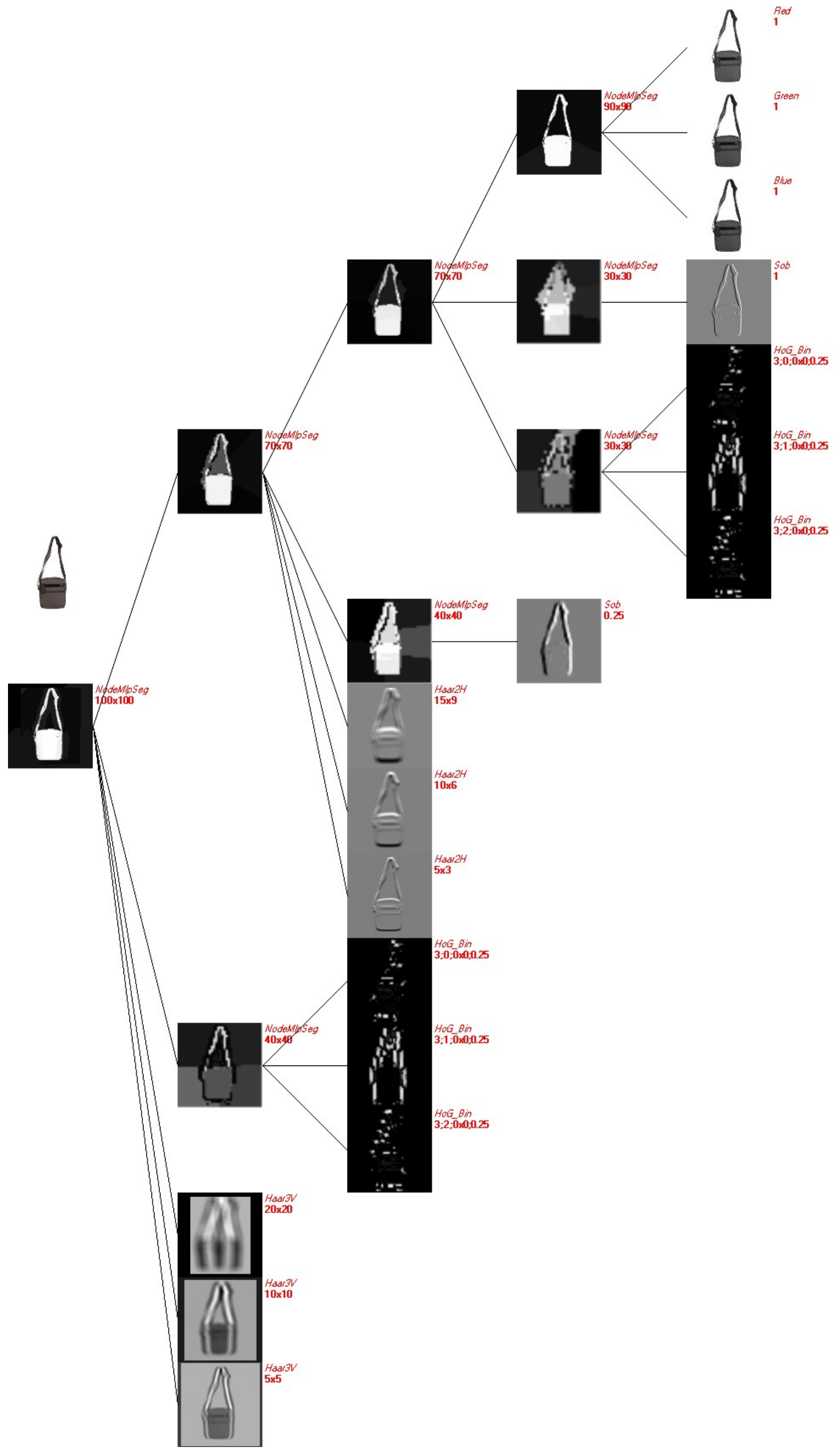
- Intensity histogram
- Hu Moments

The MNOS tree

- We should feed the majority of the node with at least some leaf

- High frequency filters seems to be very effective

  … Usage of sliding window nodes in the low levels of the tree



Experiments

Custom Dataset: Drezzy

8 classes:

- Bags (285)
- Shoes (400)
- Hats (158)
- Ties (203)
- Man Clothing (150)
- Man Underwear (278)
- Woman Clothing (355)
- Woman Underwear (239)

Jpeg images
Resolution: 100 x 100 or 200 x 200 pixels

VOC Challenge metrics adopted

First experiment

Only nodes based on segments classification

Dataset	Obj Acc Train	Obj Acc Test
Bags	83,92	79,13
Shoes	80,05	77,76
Hats	68,63	64,25
Ties	90,42	77,76
Man Clothing	73,09	68,06
Man Underwear	40,97	38,75
Woman Clothing	52,53	57,84
Woman Underwear	36,33	35,13

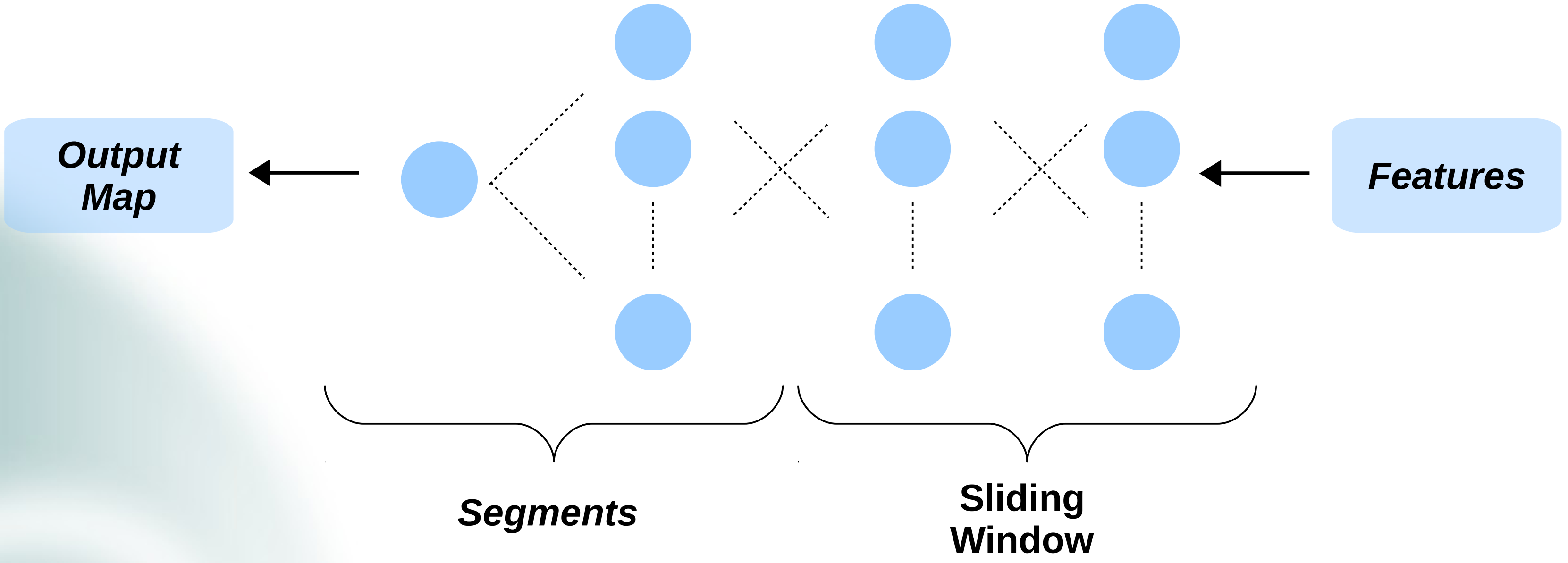
First experiment

The network cannot solve the problem:

- with complex backgrounds
- when we have other objects with high contrast

Hybrid Model

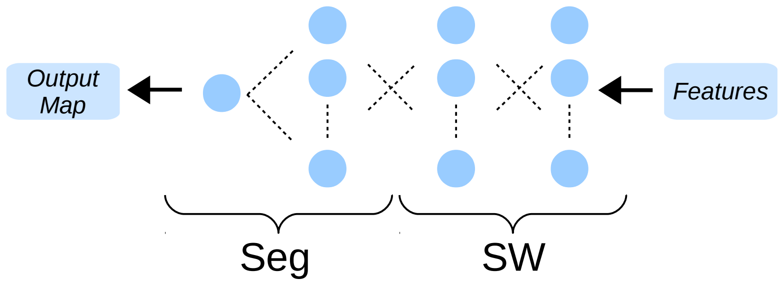
Hybrid model: Sliding window and segment classification

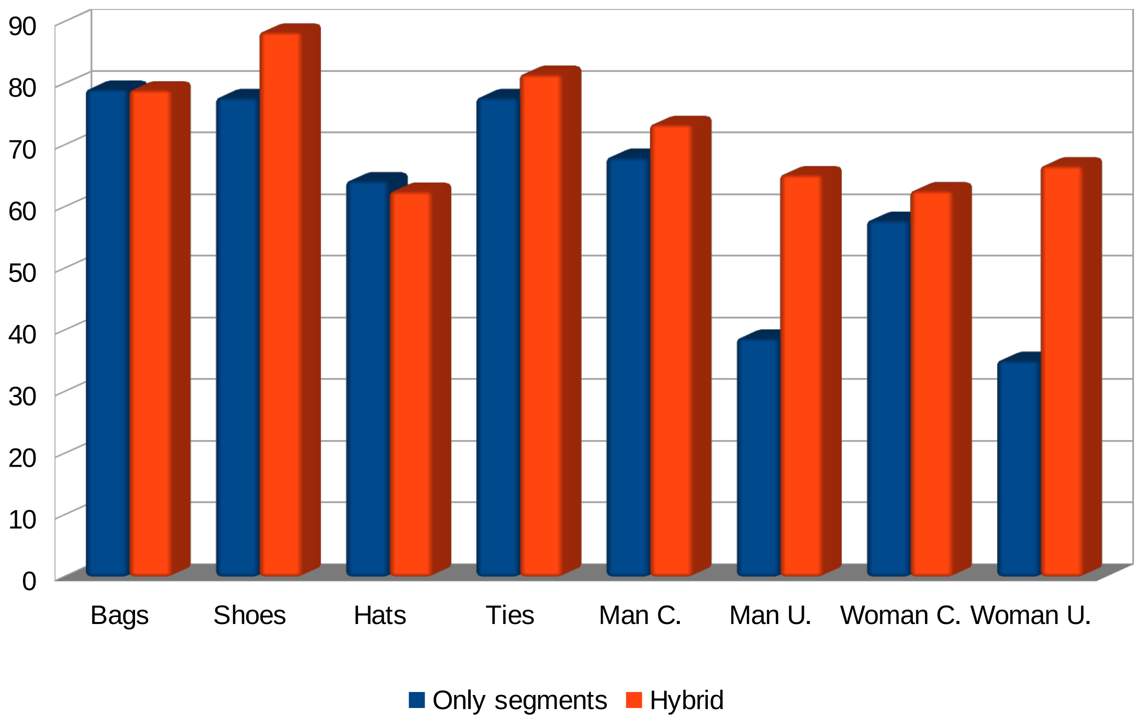


Hybrid Model

Hybrid model: Sliding window and segment classification

Dataset	**Acc test**	***Diff***
Bags	79,00	- 0,13
Shoes	88,39	+ 10,63
Hats	62,55	- 1,70
Ties	81,52	+ 3,76
Man Clothing	73,40	+ 5,34
Man Underwear	65,25	+ 26,50
Woman Clothing	62,64	+ 4,80
Woman Underwear	66,68	+ 31,55



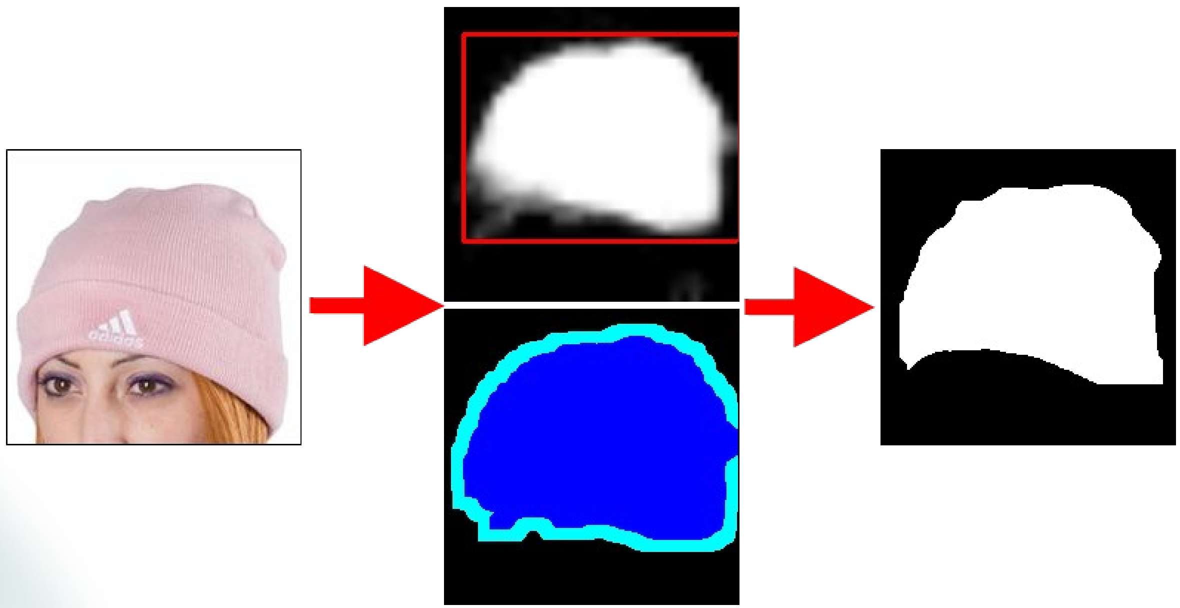


Hybrid Model

Hybrid model: Sliding window and segment classification

Post Processing with GrabCut

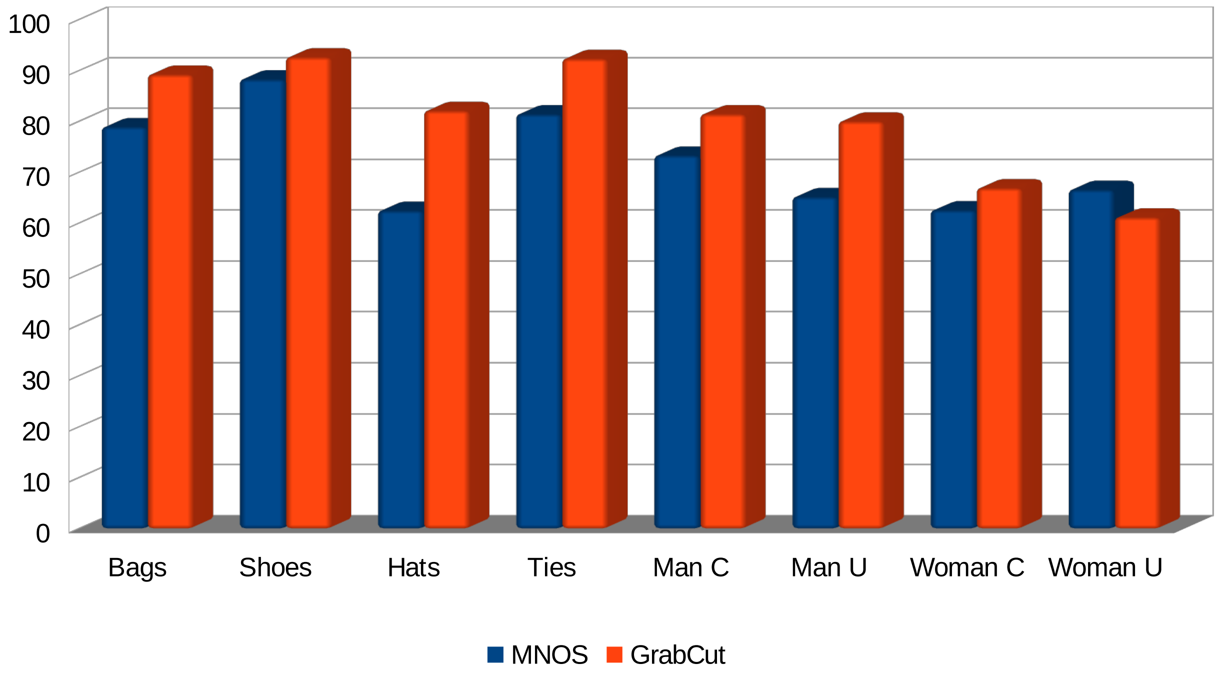
Two approaches: Bounding Box vs Region Mask



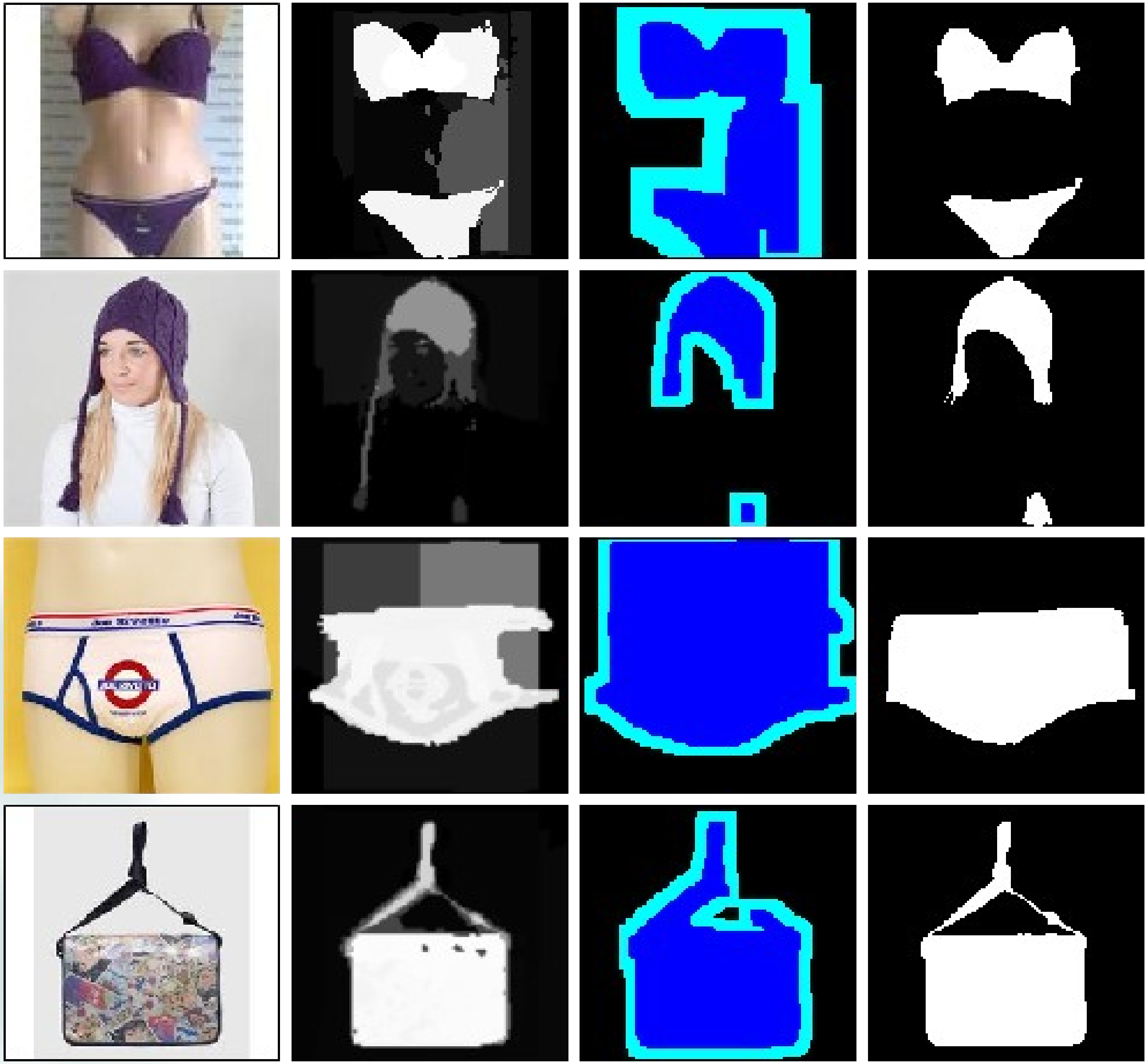
The MNOS segmentation mask is labeled in order to initialize the GrabCut

Post Processing with GrabCut

Dataset	**GrabCut**	***Diff***
Bags	89,29	+10,29
Shoes	92,70	+4,31
Hats	82,19	+19,64
Ties	92,39	+10,87
Man Clothing	81,50	+8,10
Man Underwear	80,10	+14,85
Woman Clothing	66,97	+4,33
Woman Underwear	61,22	-5,46



Post Processing with GrabCut



Results with VOC 2011 dataset

Class	MNOS	GC	Voc Best	Class	MNOS	GC	Voc Best
Aeroplane	36,04	55,60	54,3	DiningTable	26,01	26,42	30,1
Bicycle	14,99	13,43	23,9	Dog	34,56	38,16	33,9
Bird	24,18	37,06	46,0	Horse	31,11	42,15	49,1
Boat	31,05	36,11	35,3	Motorbike	51,13	49,44	54,4
Bottle	20,80	19,11	49,4	Person	32,73	35,82	46,4
Bus	50,93	57,14	66,2	Pottedplant	19,38	24,44	28,8
Car	37,27	37,52	56,2	Sheep	31,26	31,58	51,3
Cat	36,82	39,33	46,1	Sofa	26,45	28,02	26,4
Chair	4,78	9,13	15,0	Train	50,26	52,37	44,9
Cow	40,65	54,07	47,4	Tvmonitor	18,18	26,54	45,8

Mean MNOS: 30,93

Mean GC: 35,67

Best Voc Mean: 43,3

Results with VOC 2011 dataset

Class	MNOS	GC	Voc Best	Class	MNOS	GC	Voc Best
Aeroplane	36,04	55,60	54,3	DiningTable	26,01	26,42	30,1
Bicycle	14,99	13,43	23,9	Dog	34,56	38,16	33,9
Bird	24,18	37,06	46,0	Horse	31,11	42,15	49,1
Boat	31,05	36,11	35,3	Motorbike	51,13	49,44	54,4
Bottle	20,80	19,11	49,4	Person	32,73	35,82	46,4
Bus	50,93	57,14	66,2	Pottedplant	19,38	24,44	28,8
Car	37,27	37,52	56,2	Sheep	31,26	31,58	51,3
Cat	36,82	39,33	46,1	Sofa	26,45	28,02	26,4
Chair	4,78	9,13	15,0	Train	50,26	52,37	44,9
Cow	40,65	54,07	47,4	Tvmonitor	18,18	26,54	45,8

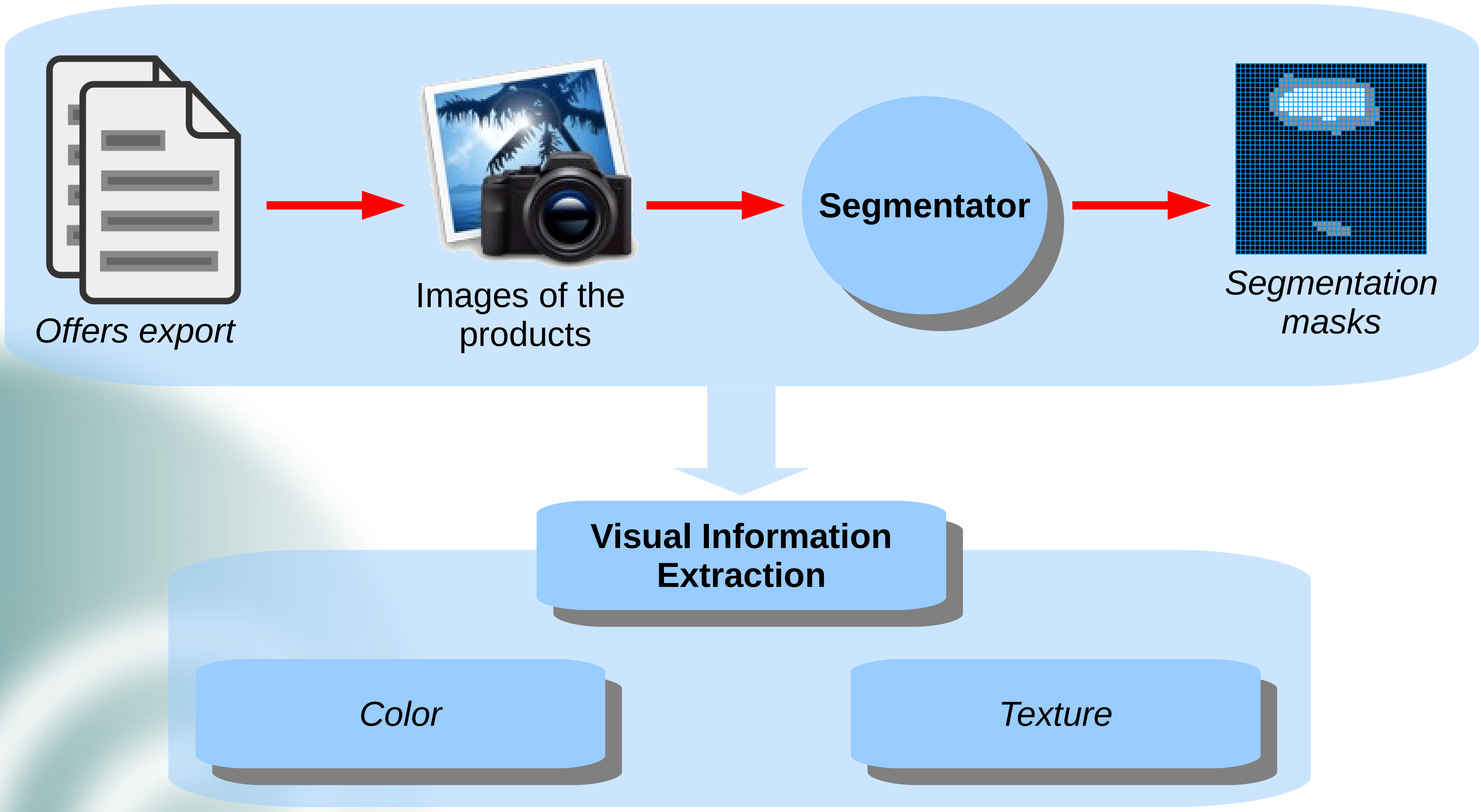
Mean MNOS: 30,93

Mean GC: 35,67

Best Voc Mean: 43,3

Commercial application

www.drezzy.it



Commercial application

www.drezzy.it