Search and Rescue UAV missions with object recognition using convolutional neural networks

Fabio Andrade
Research Scientist, Norut
Assistant Professor, CEFET/RJ
PhD Candidate, NTNU

Supervisors: Rune Storvold and Tor Arne Johansen

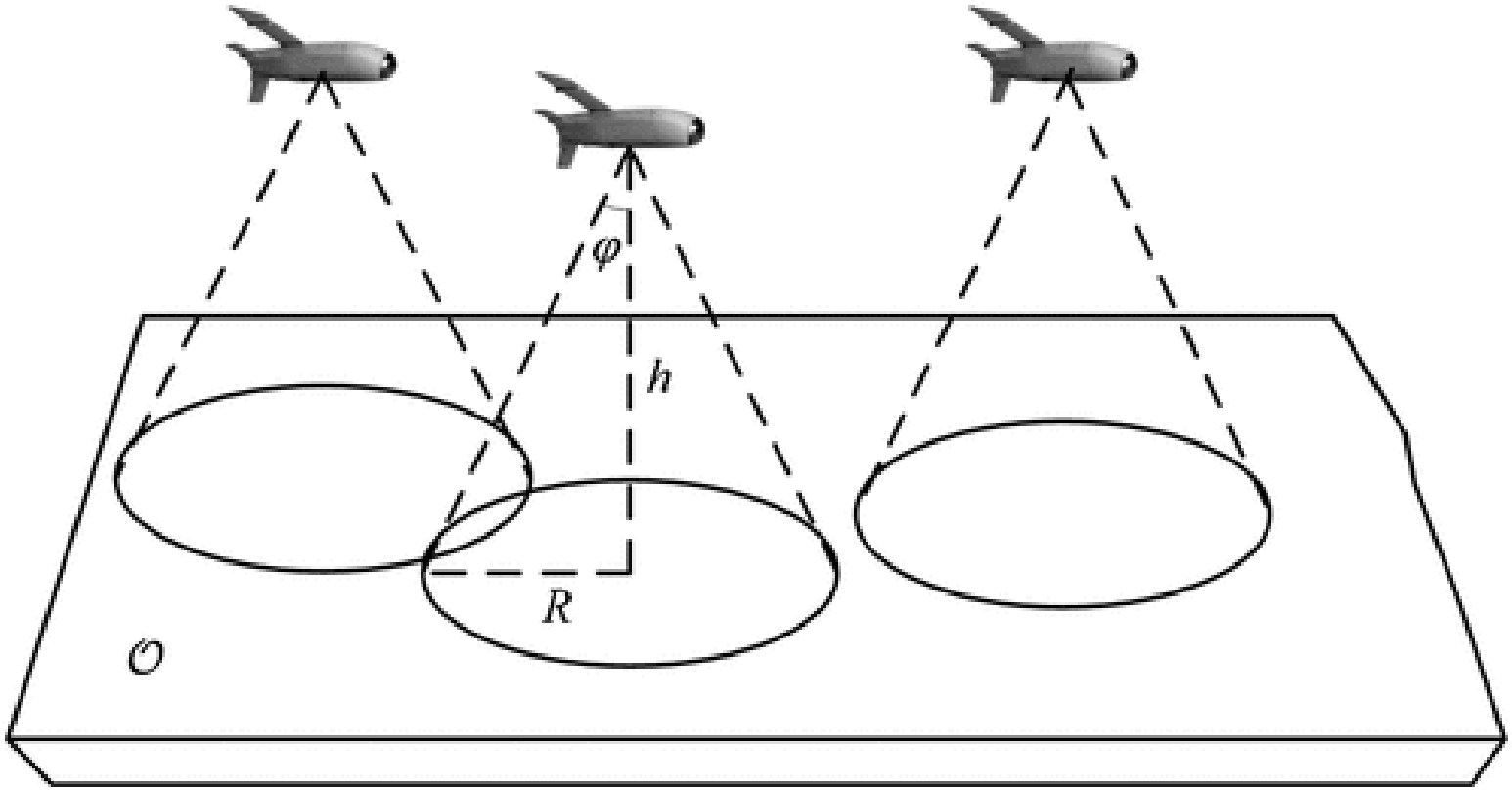
This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 642153.

Agenda

- Search and Rescue
- Convolutional Neural Networks
- Examples
- Our dataset

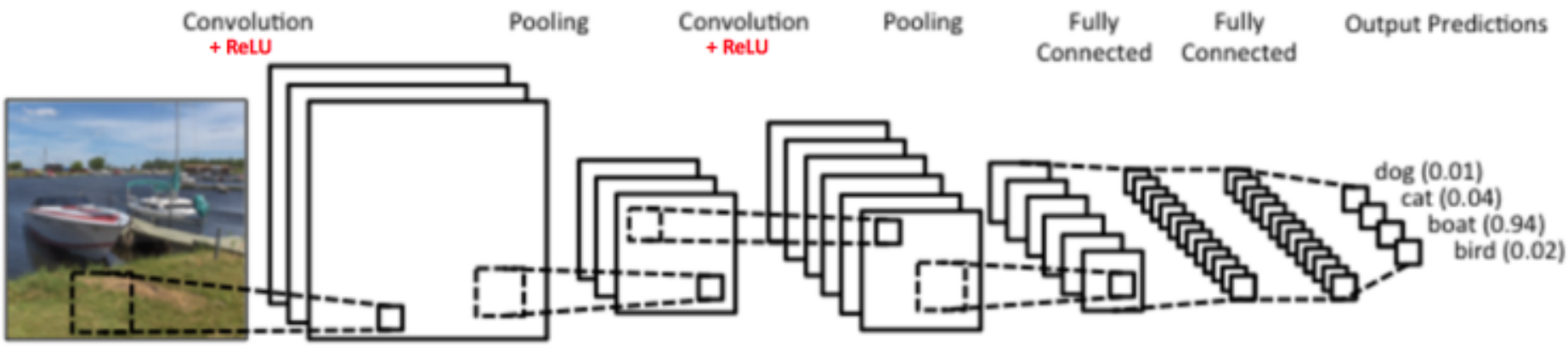
Source: Norut

Source: Norut

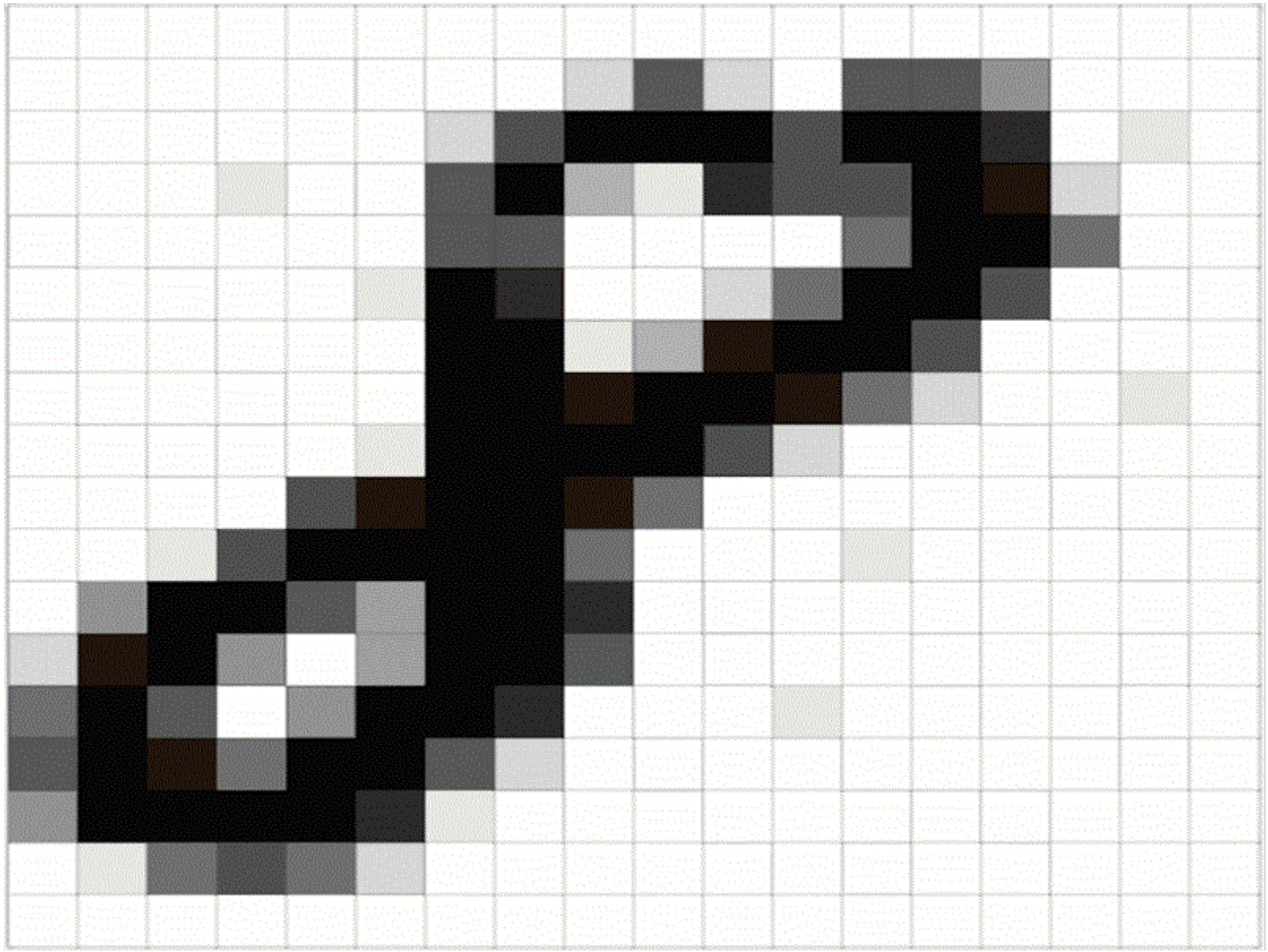


Source: Minqiang Zhang

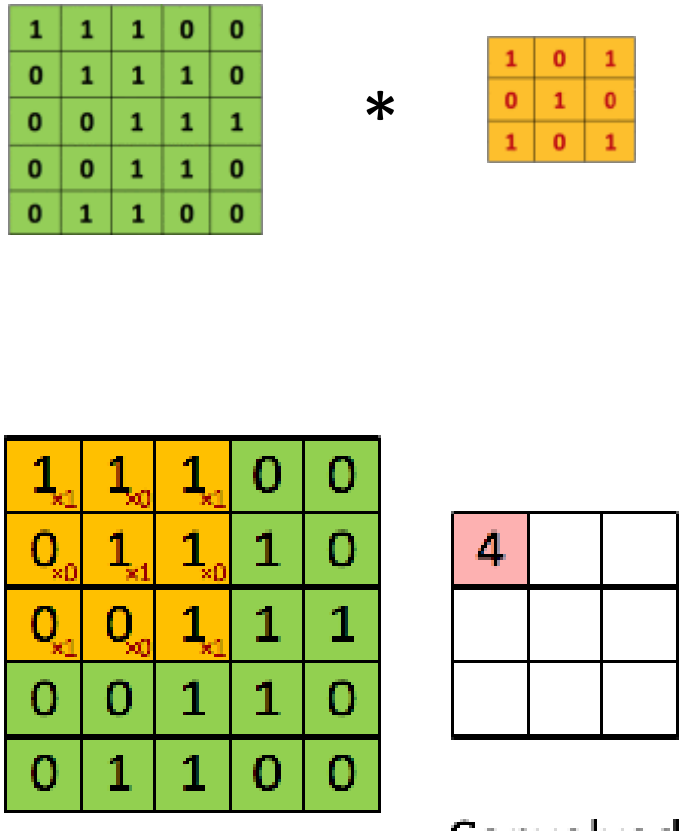
Convolutional Neural Networks?



source: https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/

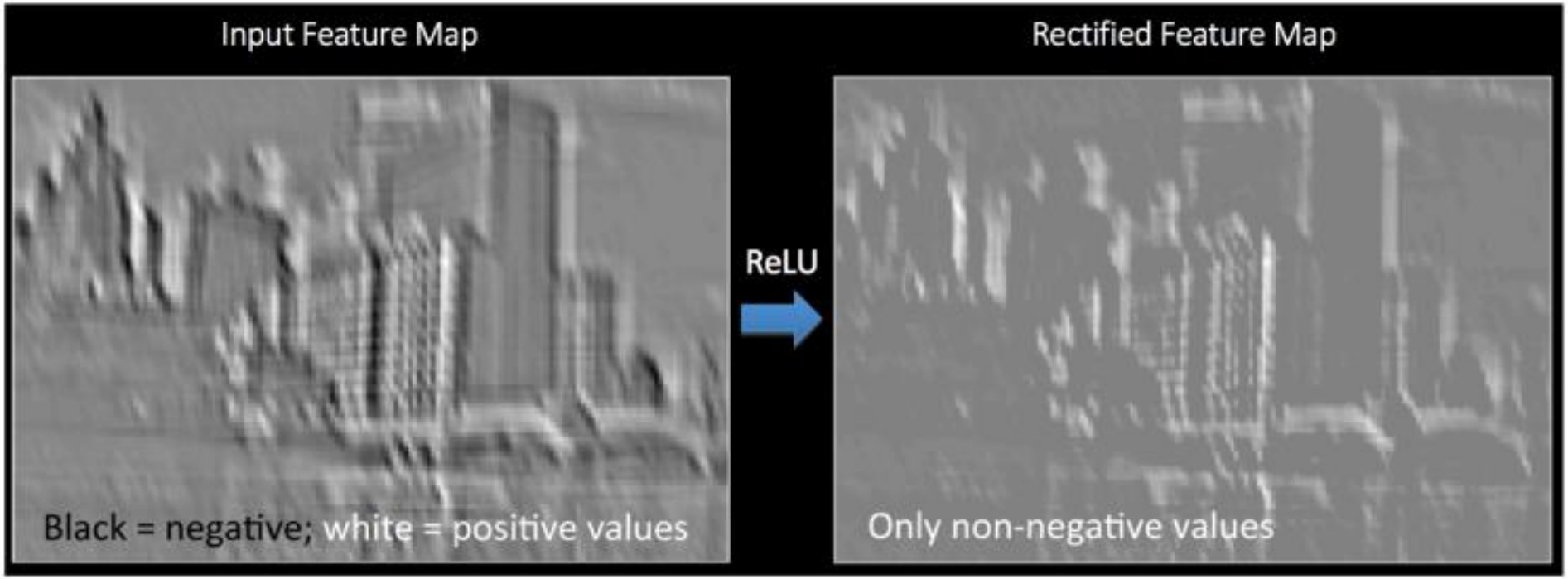


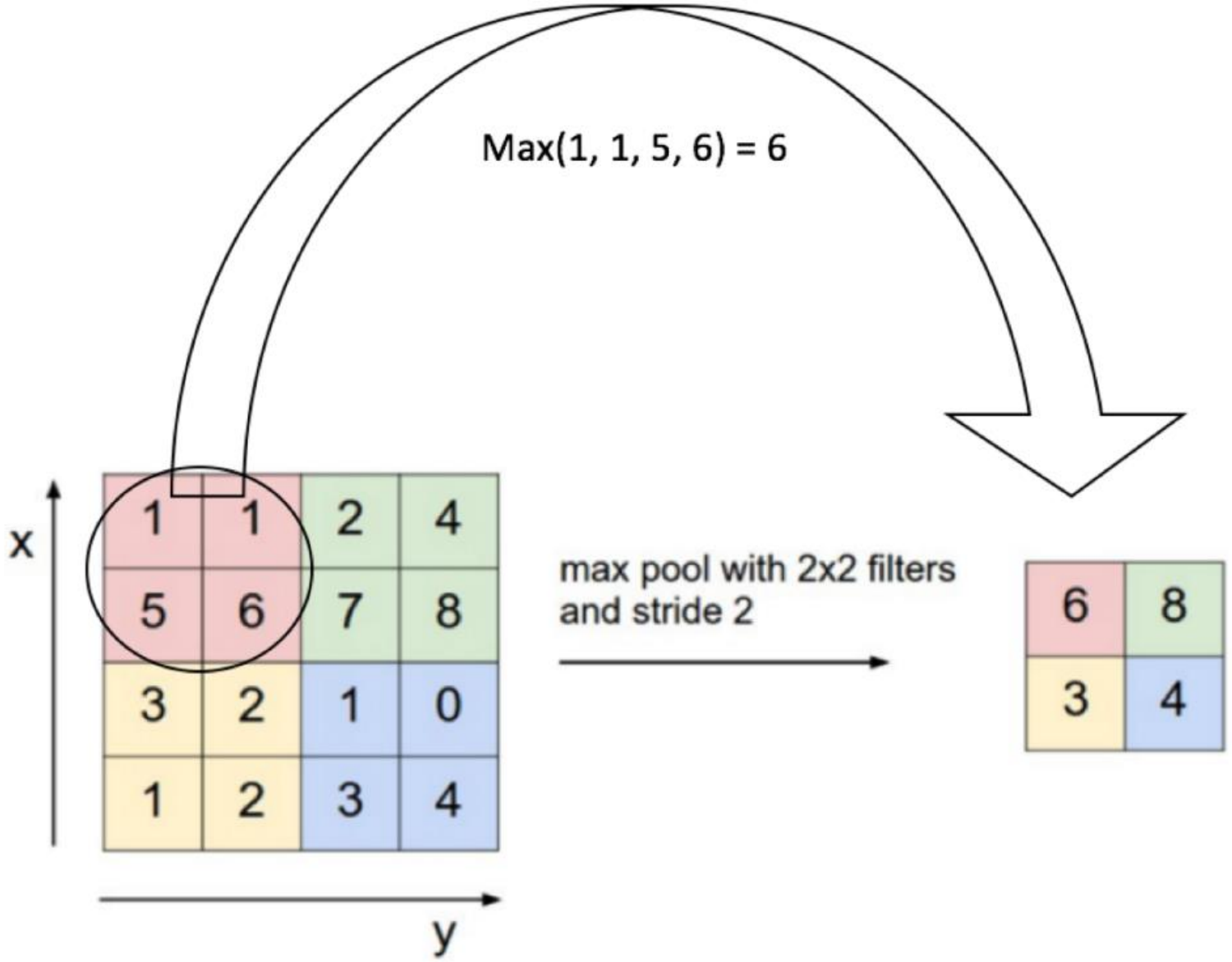
The Convolution Step



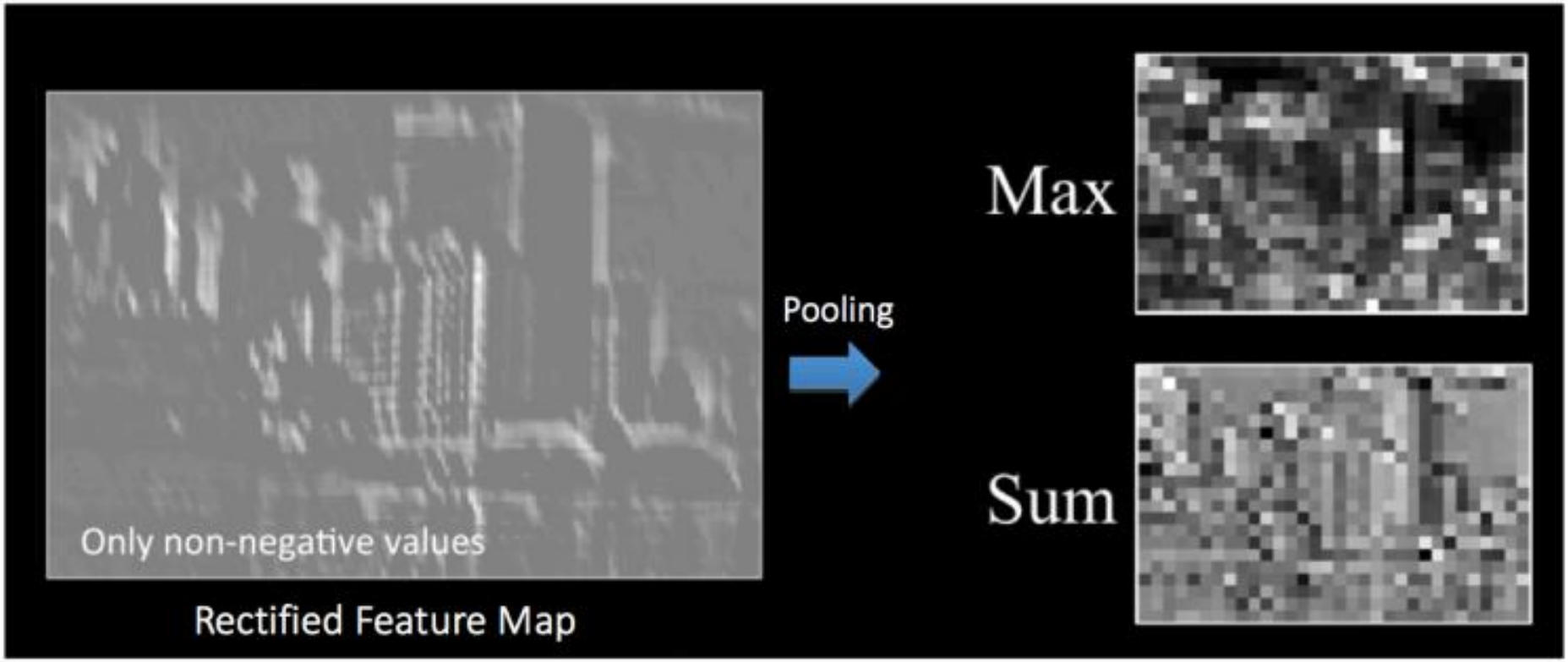
Operation	Filter	Convolved Image
Identity	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	
Edge detection	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$	
	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$	
	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	
Sharpen	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
Box blur (normalized)	$\frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	
Gaussian blur (approximation)	$\frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$	

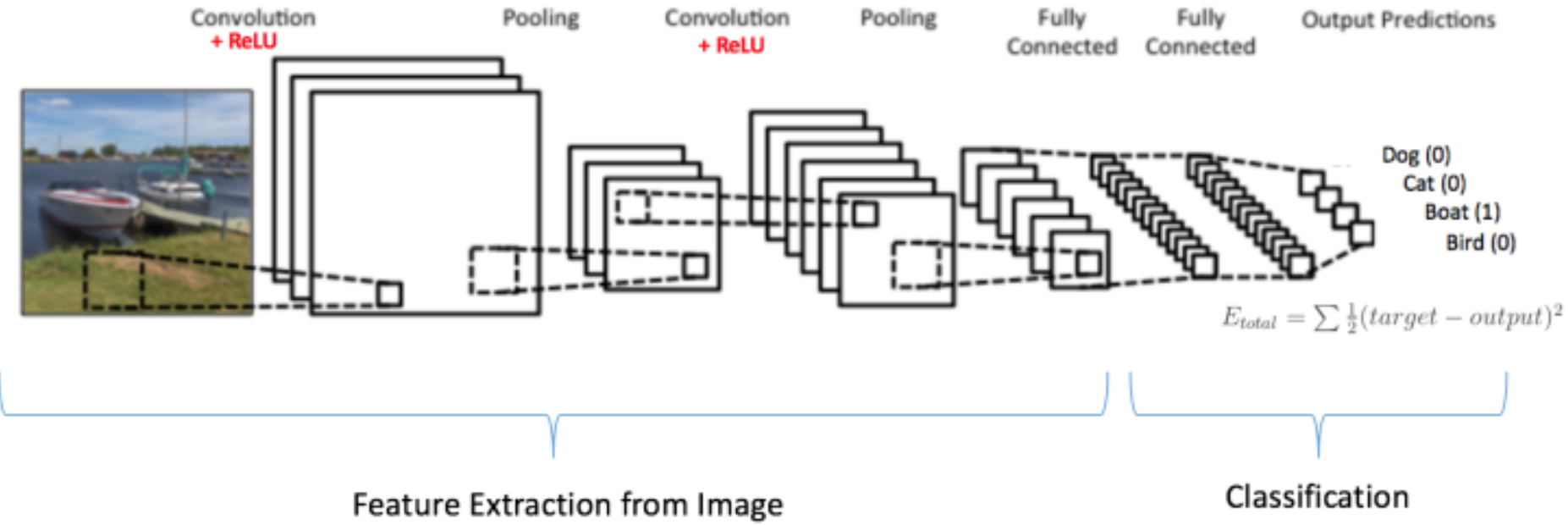
Input

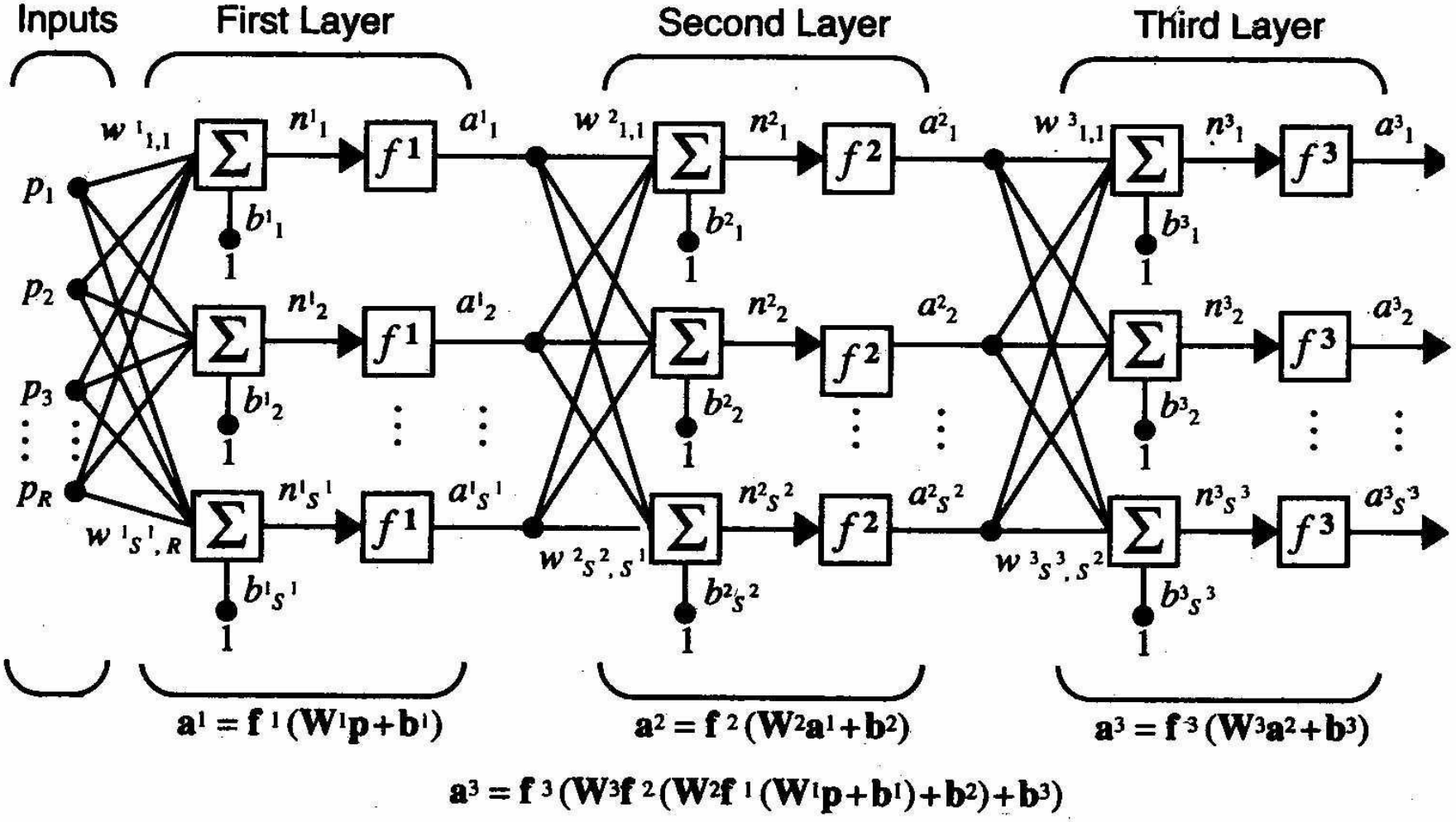

Input Feature Map
Rectified Feature Map
ReLU
Black = negative; white = positive values
Only non-negative values



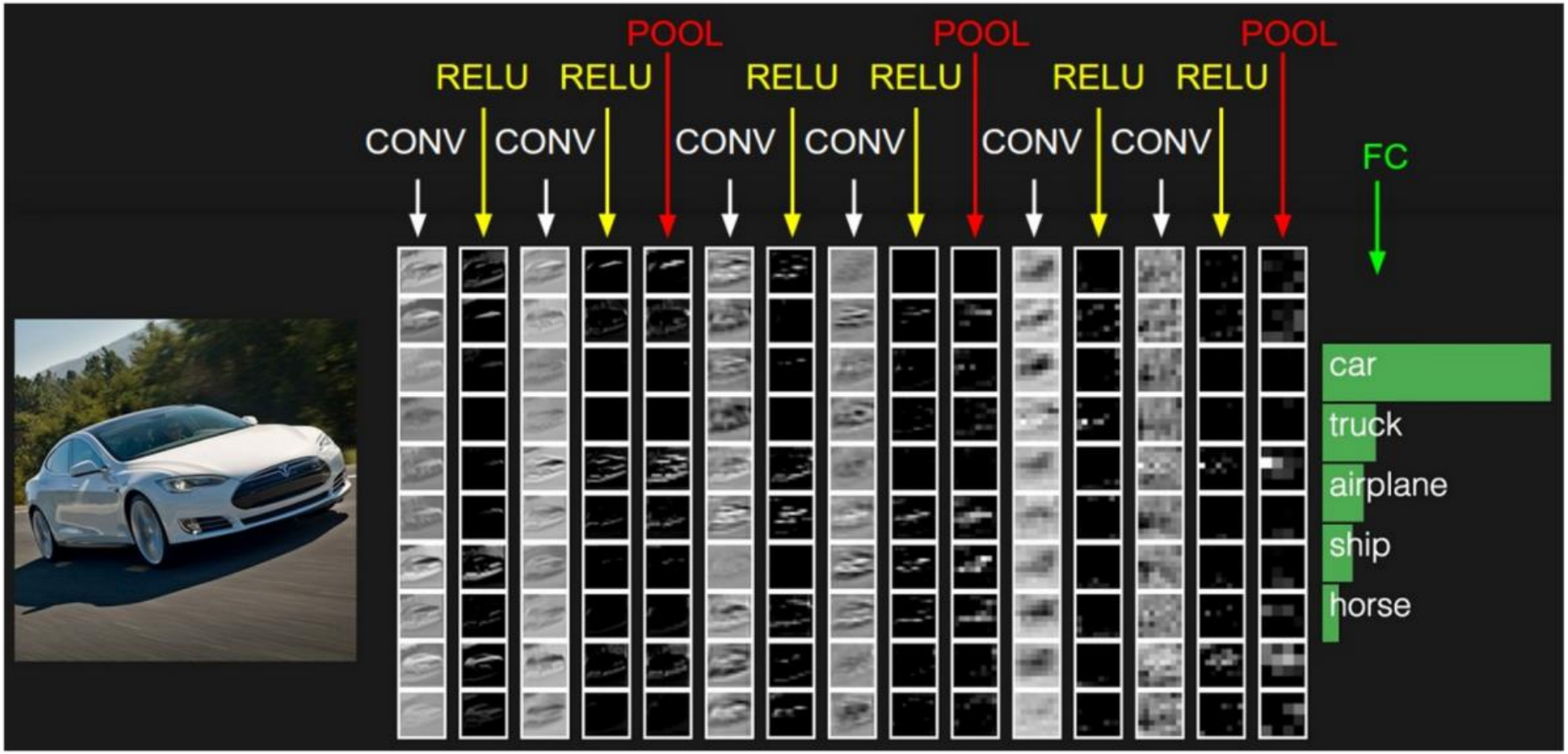
Rectified Feature Map

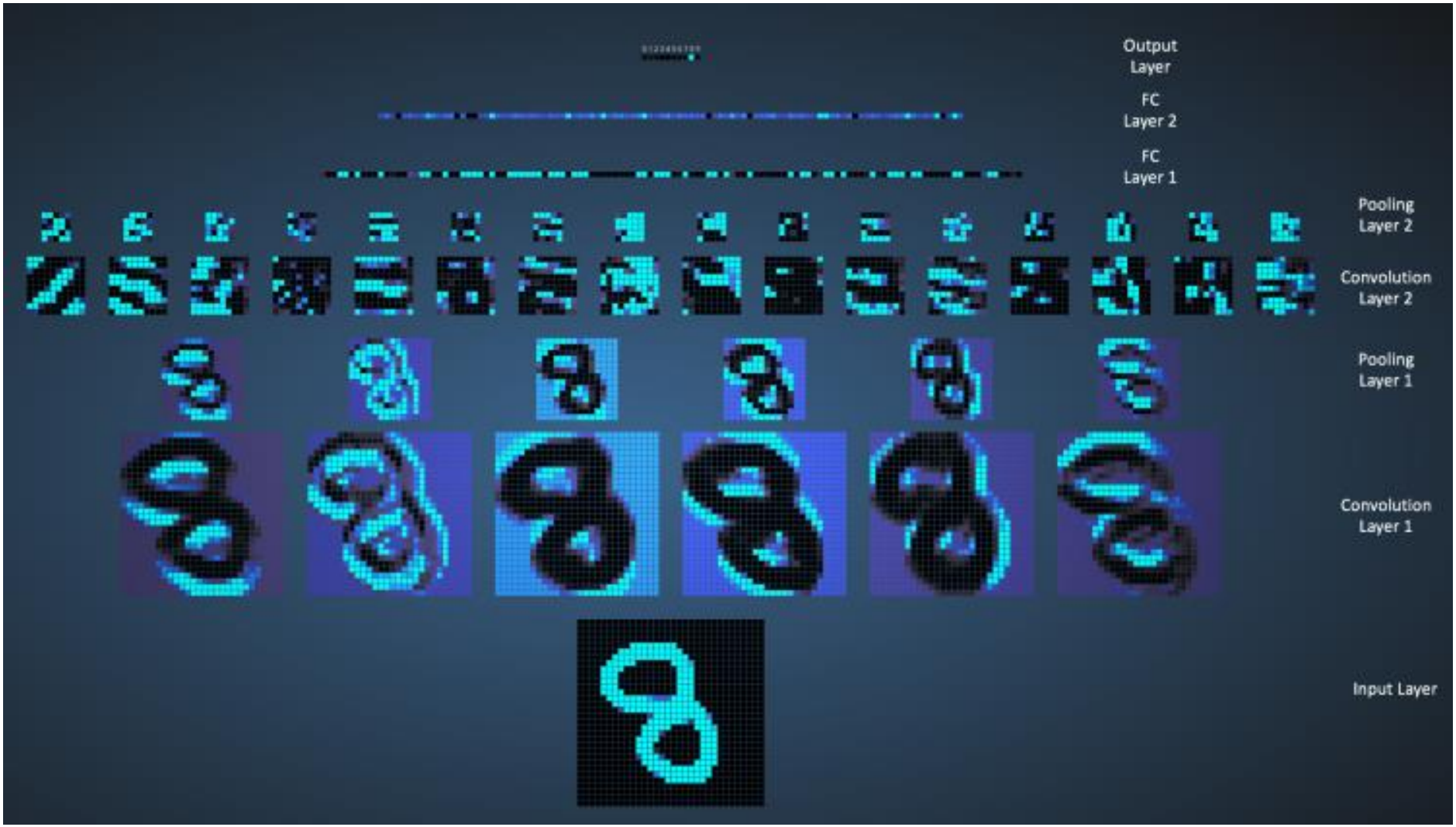

Only non-negative values
Rectified Feature Map
Pooling
Max
Sum


Convolution
+ ReLU
Pooling
Convolution
+ ReLU
Pooling
Fully
Connected
Fully
Connected
Output Predictions
Dog (0)
Cat (0)
Boat (1)
Bird (0)
$E_{total} = \sum \frac{1}{2}(target - output)^2$
Feature Extraction from Image
Classification



Source: http://ffden-2.phys.uaf.edu


POOL
POOL
POOL
RELU
RELU
RELU
RELU
RELU
RELU
CONV
CONV
CONV
CONV
CONV
CONV
FC
car
truck
airplane
ship
horse


Output Layer
FC Layer 2
FC Layer 1
Pooling Layer 2
Convolution Layer 2
Pooling Layer 1
Convolution Layer 1
Input Layer

Vision-based Vehicle Detecting and Counting for Traffic Flow Analysis

Zhang et al

Count vehicles

Fast R-CNN

Set of object proposals as input

Produces bounding boxes

Fine-tune a pre-trained Fast R-CNN with traffic videos

Real-time Scene Understanding for UAV Imagery based on Deep Convolutional Neural. Networks

Clay Sheppard and Maryam Rahnemoonfar

UAV vertical images
High altitude
3864 images
4 classes: building, ground, road, tree

Generating binary tags for fast medical image retrieval based on convolutional nets and radon transform

Xinran Liu, H.R.Tizhoosh, J.Kofman

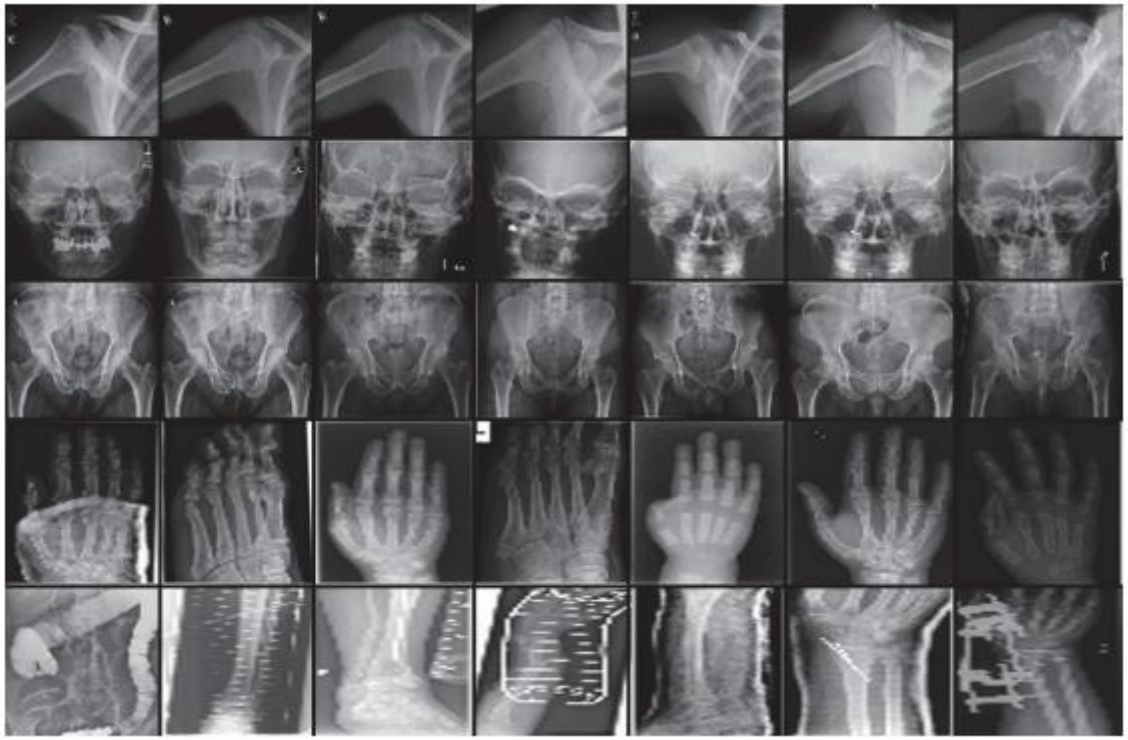


Image retrievel

Gray level (no color)

Fast animal detection in uav images using convolutional neural networks

Benjamin Kellenberger, Michele Volpi, Devis Tuia

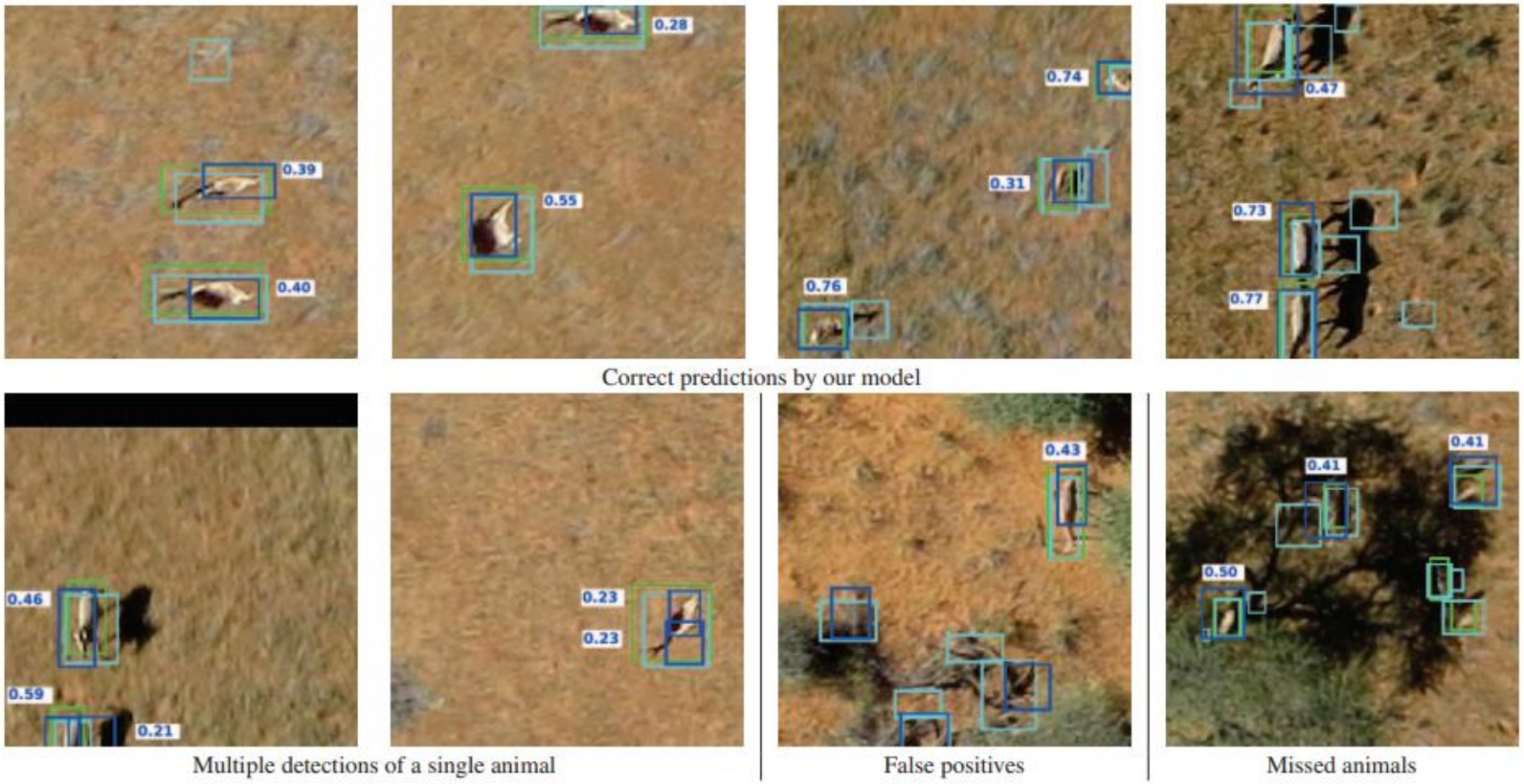


Fig. 3. Detection examples on the test set (blue; with IoU scores for predictions) and Fast R-CNN baseline (cyan). Ground truth is in green. Top row shows correct detections for our model, while bottom row show failure cases.

Large animals in Namibia

Real time

654 RGB images, 1196 animals after fine tuning

Convolutional Neural Networks For Near Real-time Object Detection from Uav Imagery In Avalanche Search And Rescue Operations

Mesay Belete Bejiga, Abdallah Zeggada, and Farid Melgani

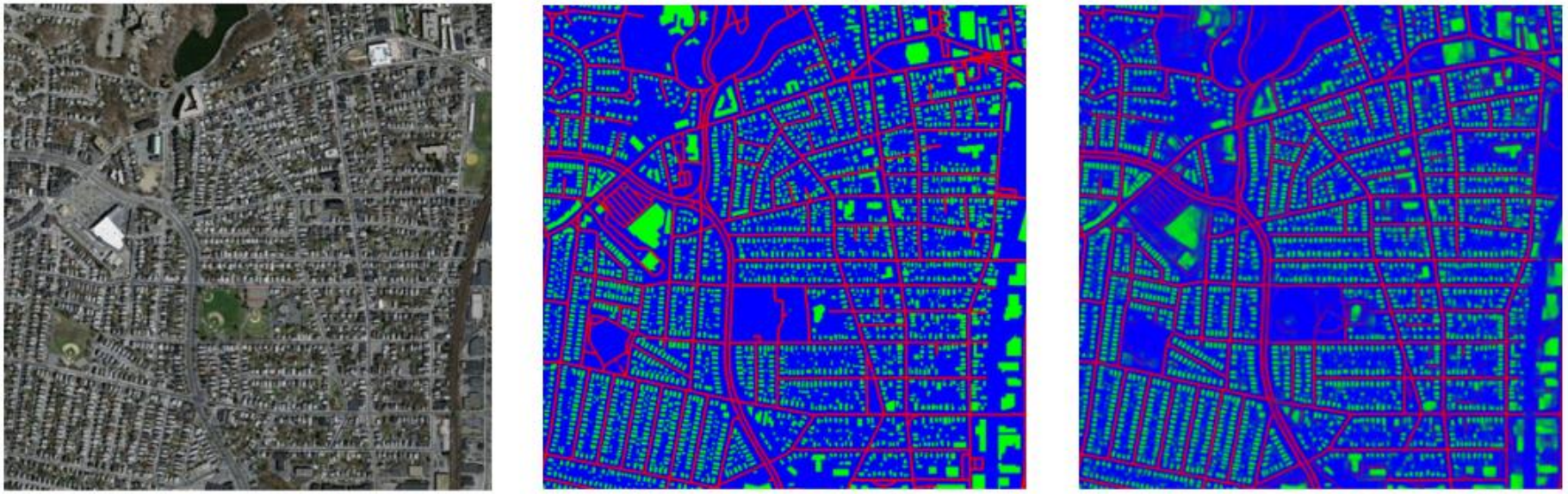
Fig. 3. Example of correctly classified negative (top left) and positive (top right) images, false positive object in yellow rectangle (bottom left), and false negative or undetected object (bottom right). Objects of interest are indicated by yellow arrow.

Pre-trained CNN to detect objects

SVM to classify

Multiple Object Extraction from Aerial Imagery with Convolutional Neural Networks

Shunta Saito, Takayoshi Yamashita, and Yoshimitsu Aoki



(a) Aerial imagery S (b) Ground truth $\tilde{M}$ (c) Predicted label image $\hat{M}$

Figure 2. An example of the resulting predicted label image.

Aerial images

3 classes: road, building, background

A Convolutional Neural Network for Automatic Analysis of Aerial Imagery

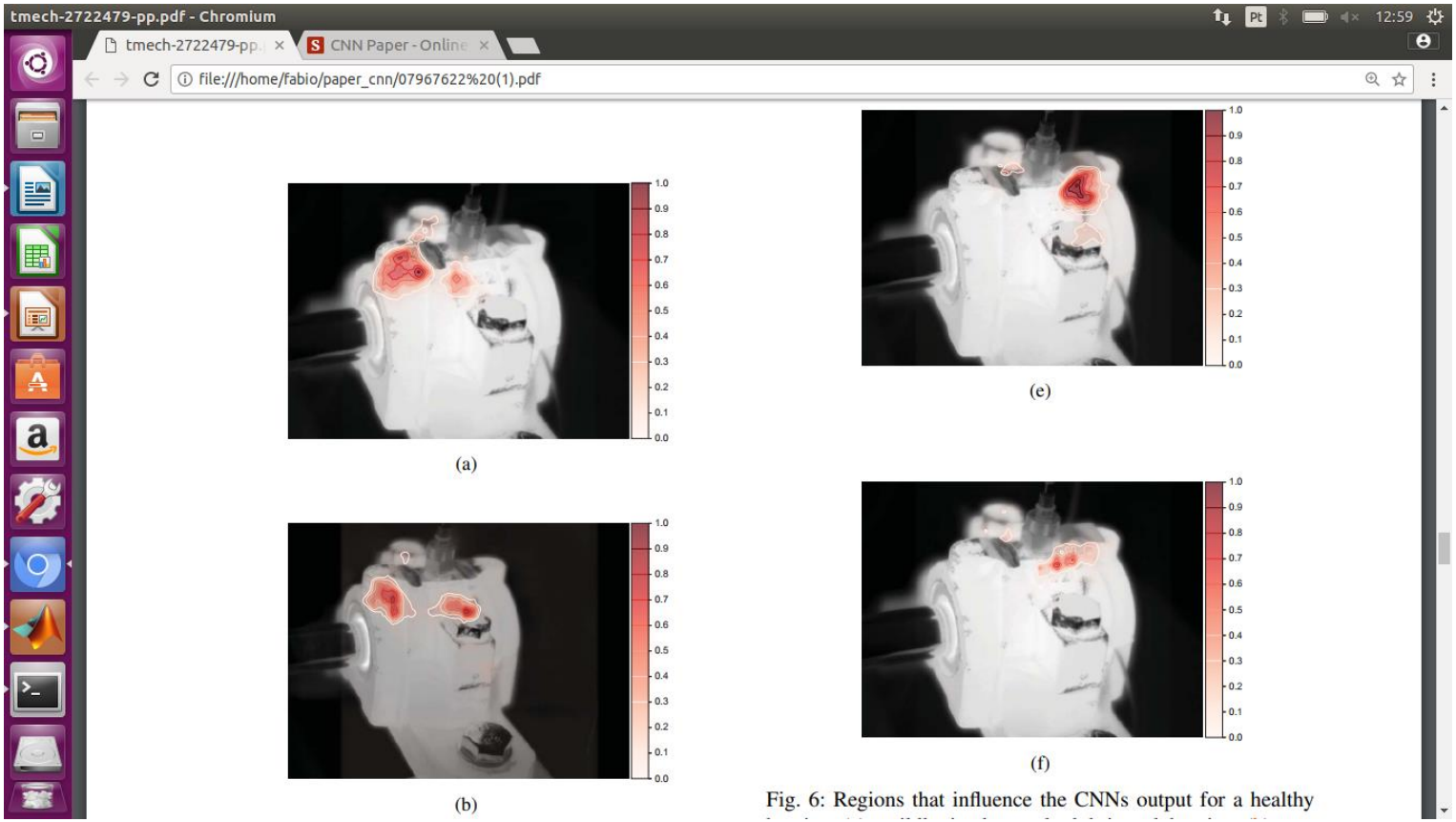
Frederic Maire, Luis Mejias and Amanda Hodgson

Marine species

Bboxes using blobs (confidence of the pixel being the center of a window containing a mammal

Deep Learning for Infrared Thermal Image Based Machine Health Monitoring

Olivier Janssens et al



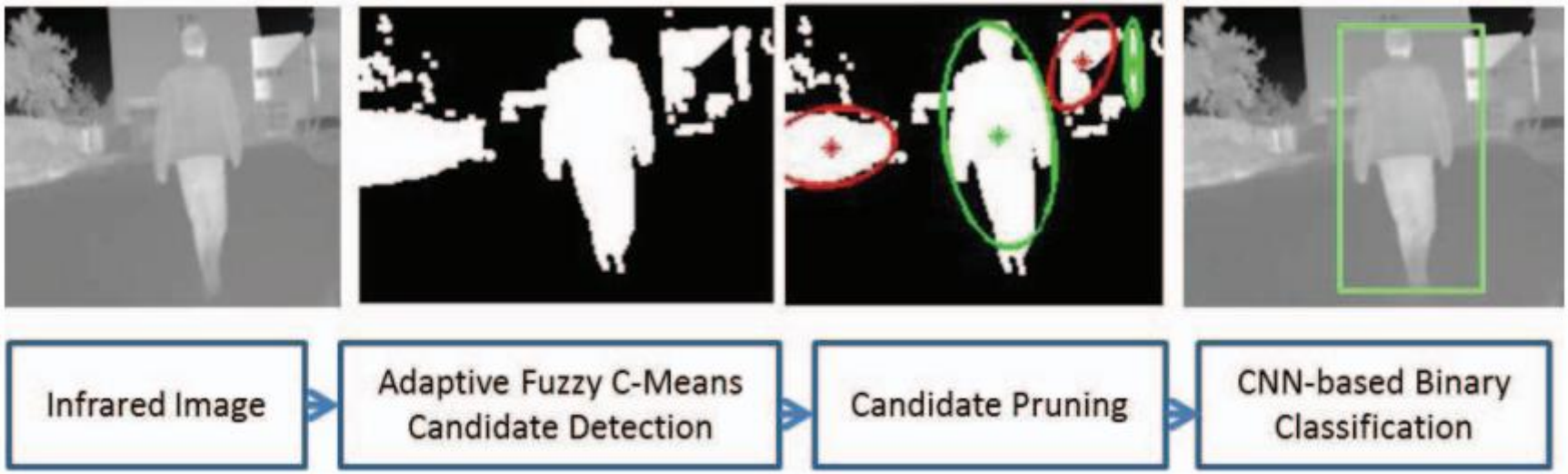
Infrared thermal video

Determine condition of the machine

Machine fault detection and oil level prediction

Pedestrian Detection in Thermal Images Using Adaptive Fuzzy C-Means Clustering and Convolutional Neural Networks

Vijay John et al

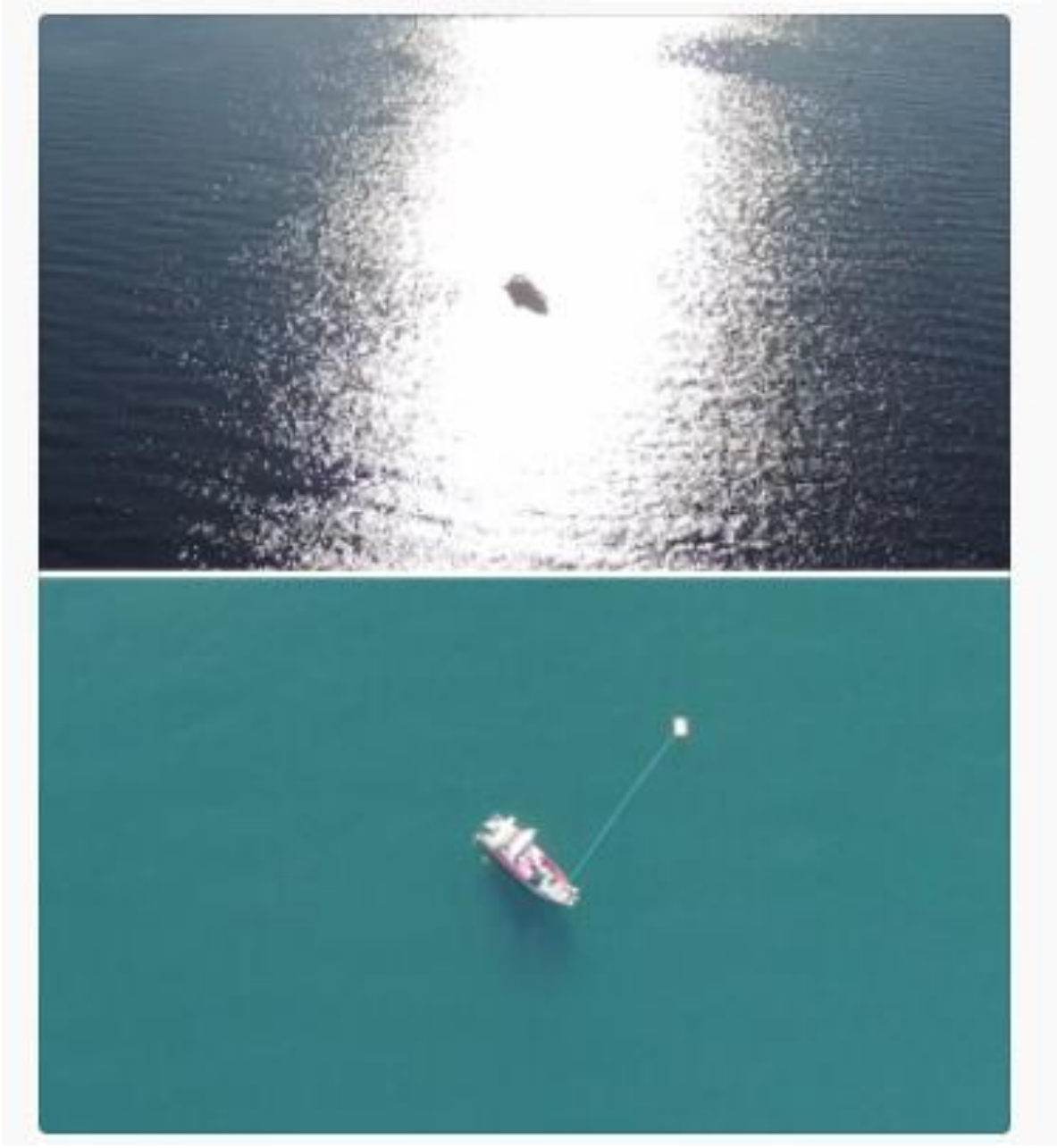


Infrared images

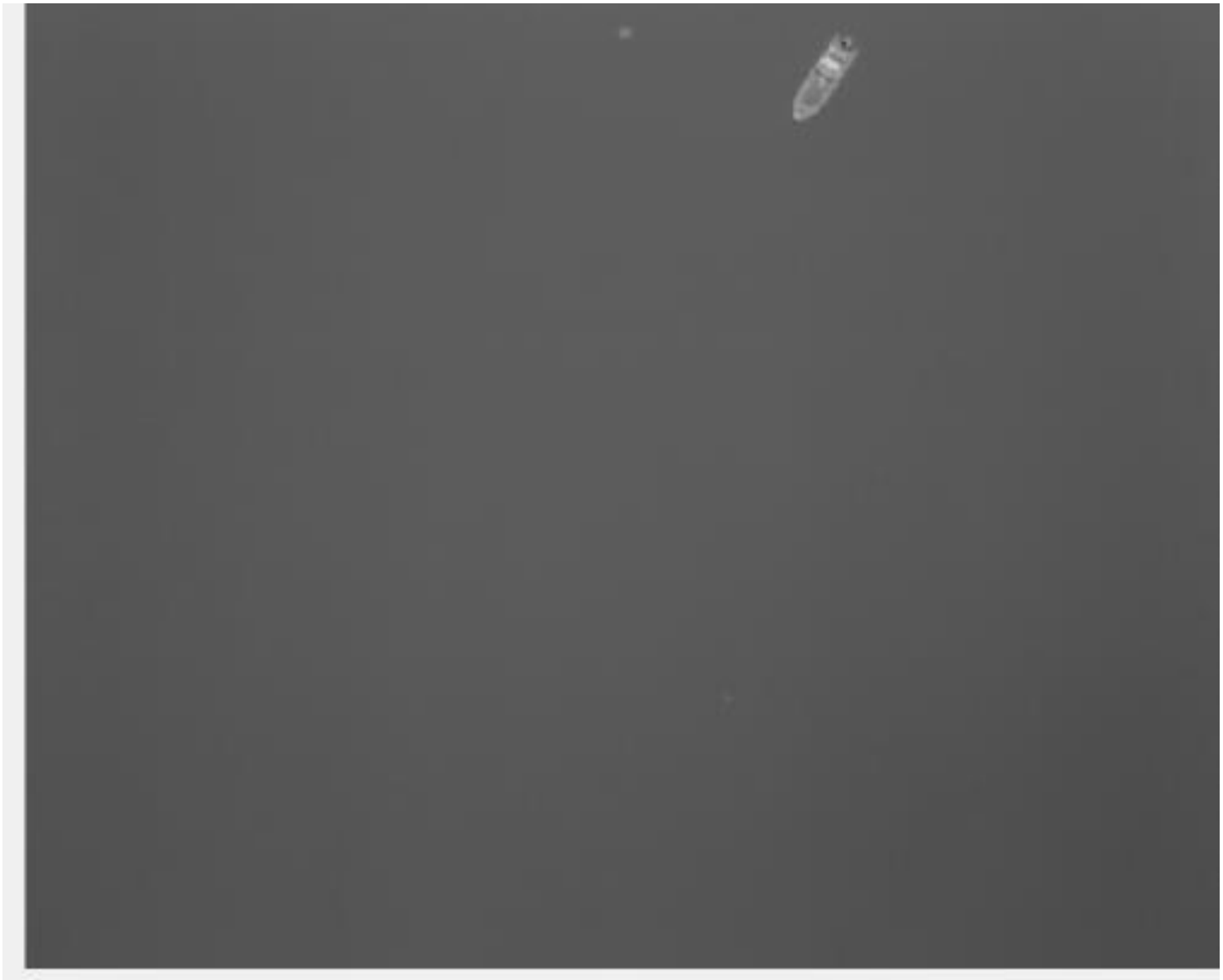
Candidate bbox is a pedestrian or not?

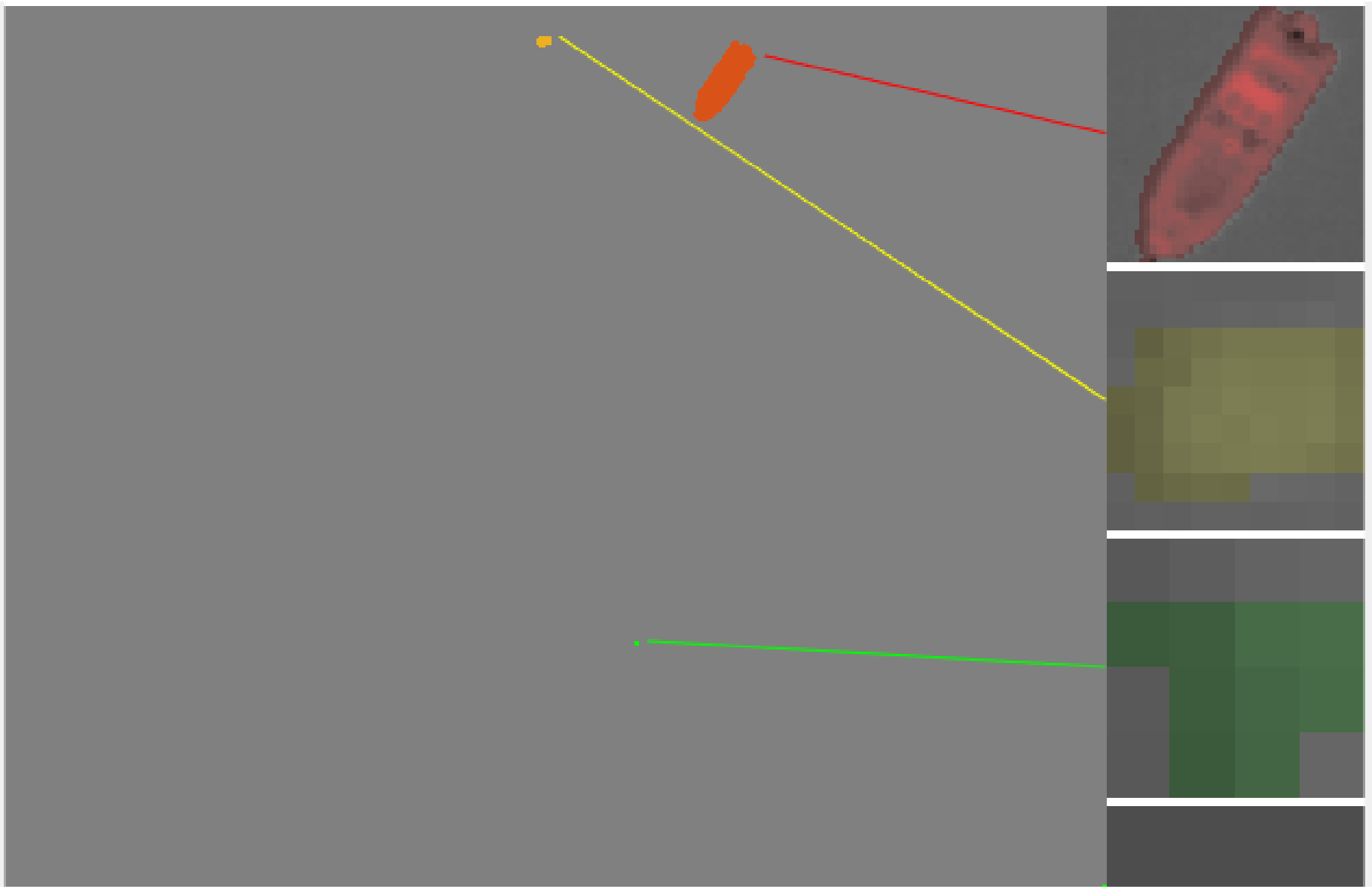
LSI Public dataset (8000 pedestrians and 8000 negative samples)

RGB

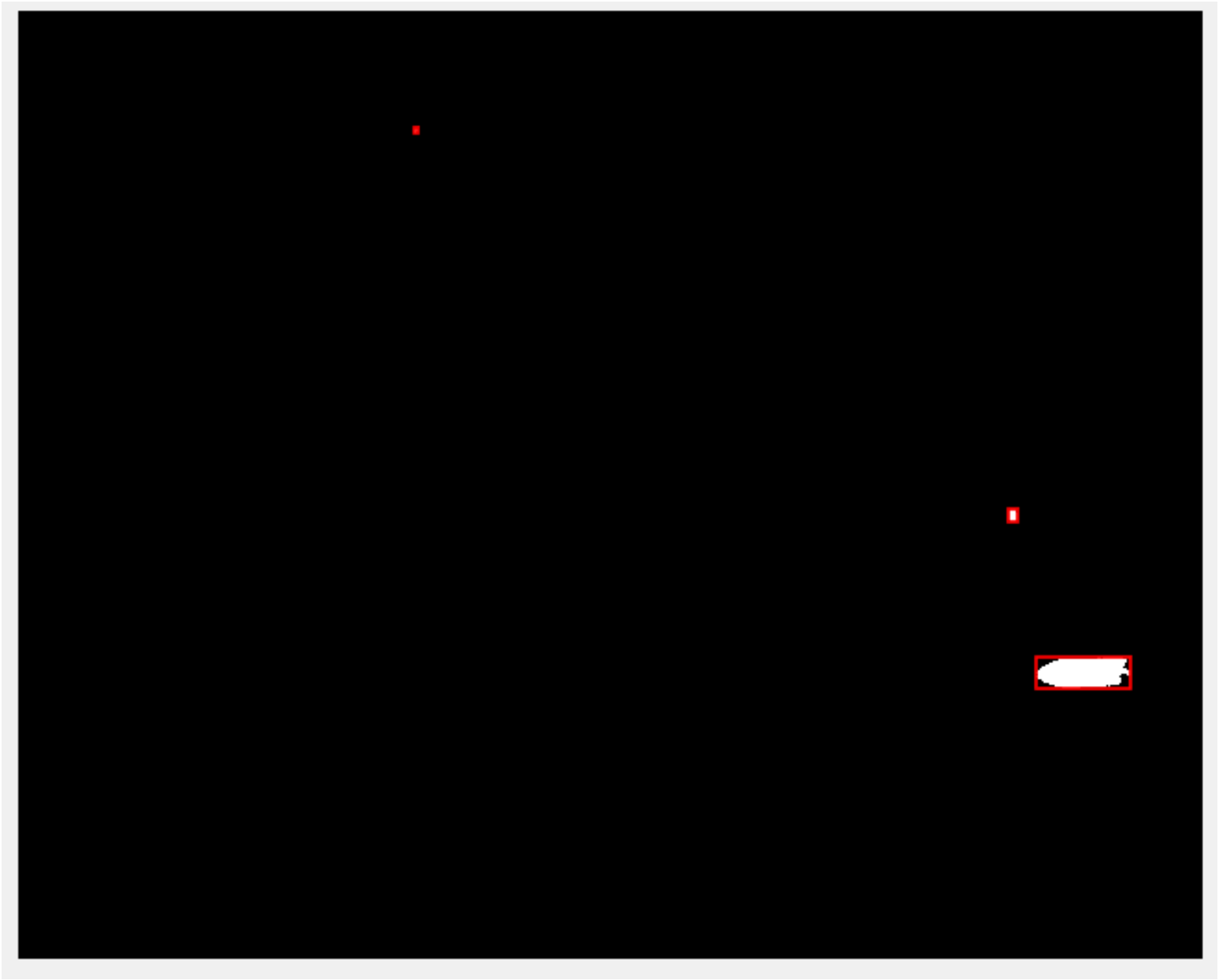


IR



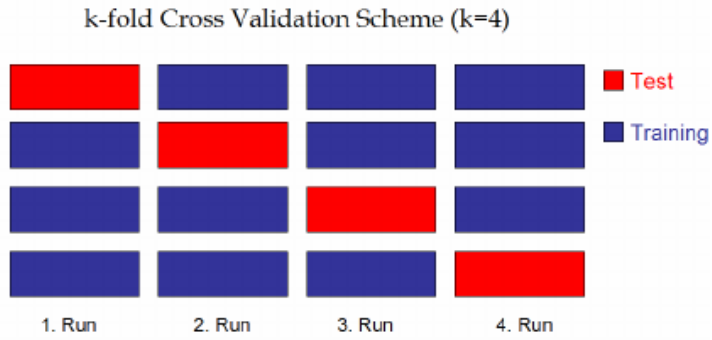


Bounding boxes

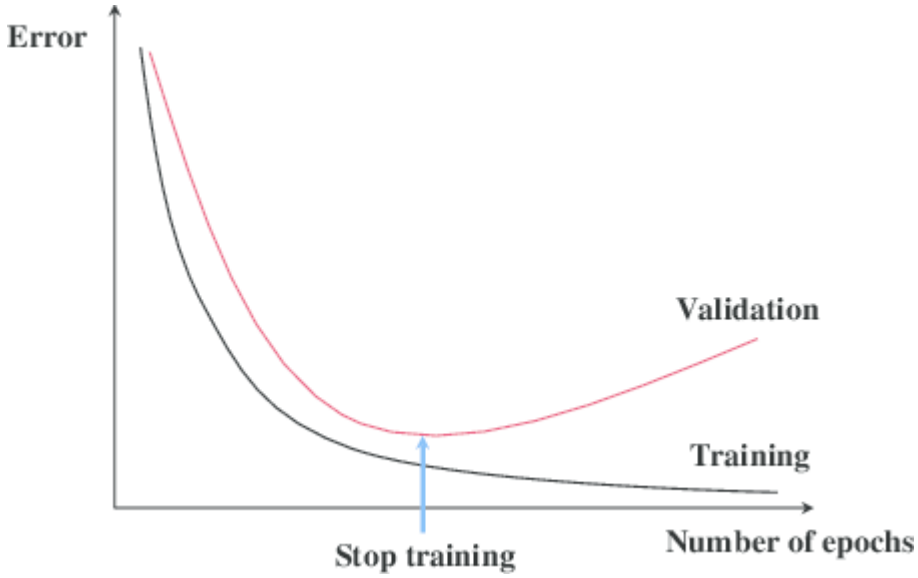


4 classes: boat, pallet, buoy, person

K-fold cross validation



Early stop by validation curve



Thank you!

fabio@ieee.org