DRIVES

Development and Research on Innovative Vocational Education Skills

# COMPUTER VISION FUNDAMENTALS

## U2.E6. IMAGE SEGMENTATION

Computer Vision Expert

**May 2021, Version 1**

Co-funded by the Erasmus+ Programme of the European Union

The Development and Research on Innovative Vocational Educational Skills project (DRIVES) is co-funded by the Erasmus+ Programme of the European Union under the agreement 591988-EPP-1-2017-1-CZ-EPPKA2-SSA-B. The European Commission support for the production of this publication does not constitute endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

# LEARNING OBJECTIVES

| The student is able to … | |
|---|---|
| CVE.U2.E6.PC1 | The student understands the definition, purpose and need of image segmentation. |
| CVE.U2.E6.PC2 | The student knows the different types of image segmentation algorithms. |
| CVE.U2.E6.PC3 | The student can define region-based segmentation. |
| CVE.U2.E6.PC4 | The student knows the differences between threshold segmentation and regional growth segmentation. |
| CVE.U2.E6.PC5 | The student defines edge detection segmentation. |
| CVE.U2.E6.PC6 | The student knows and differentiates the sobel and Laplacian operators. |
| CVE.U2.E6.PC7 | The student is able to define segmentation based on clustering. |
| CVE.U2.E6.PC8 | The student defines segmentation based on weakly-supervised learning in CNN. |
| CVE.U2.E6.PC9 | The student knows the different use-cases of each image segmentation algorithm. |
| CVE.U2.E6.PC10 | The student is able to select the most fitting image segmentation algorithm and assess its behaviour. |

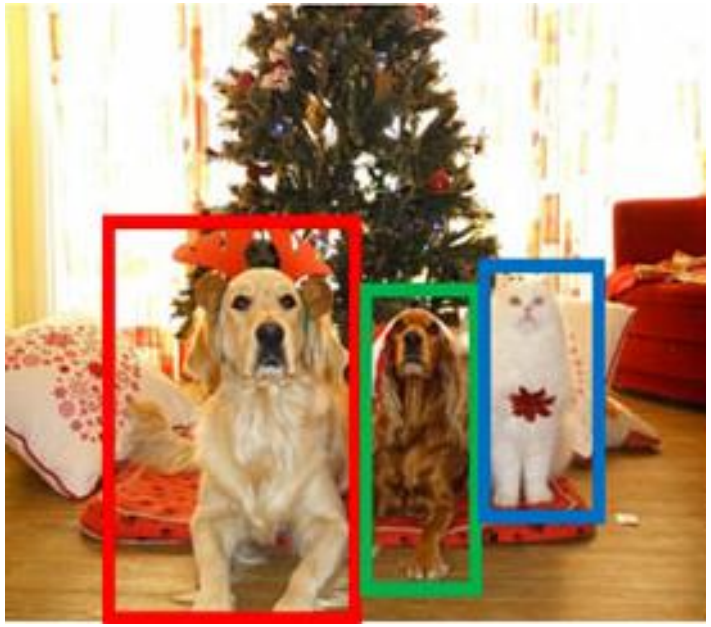
# IMAGE SEGMENTATION

Before detecting the objects and before classifying the image, we need to understand what the image consists of. So this is where segmentation comes in.

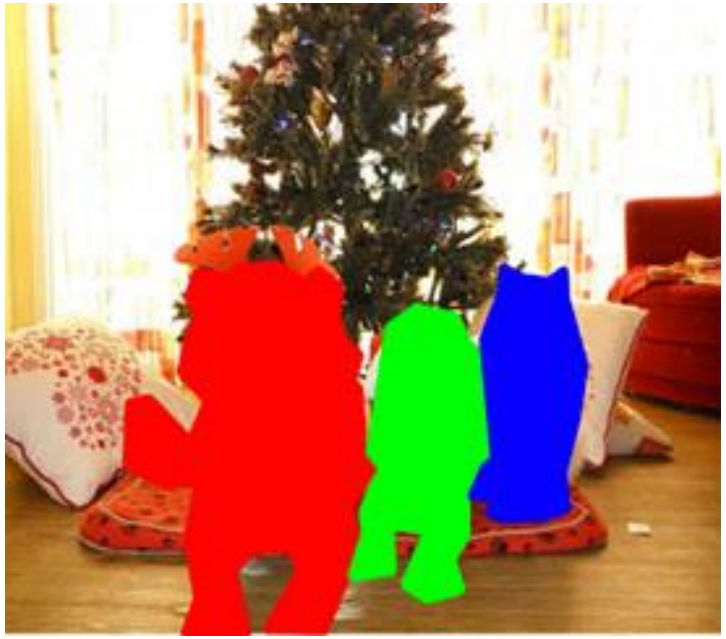
Image segmentation is based on certain criteria to divide an input image into a number of the same nature of the category in order to extract the area which people are interested in.

We can divide or partition the image into various parts called segments. By dividing the image into segments, we can make use of the important segments for processing the image.

# IMAGE SEGMENTATION



**Object Detection**



**Instance Segmentation**

- **Object detection** builds a bounding box corresponding to each class in the image, but not tell anything about the shape of the object. We only get the set of bounding box coordinates.
- **Image segmentation** creates a pixel-wise mask for each object in the image. This technique gives us a better understanding of the object(s) in the image.

# TYPES OF IMAGE SEGMENTATION ALGORITHMS

DRIVES
Development and Research on Innovative Vocational Education Skills

**Region-based Segmentation**

- Threshold Segmentation
- Regional Growth Segmentation

**Edge Detection Segmentation**

- Sobel Operator
- Laplacian Operator

**Segmentation based on clustering**

**Segmentation based on weakly -supervised learning in CNN**

# REGION-BASED SEGMENTATION

Use pixel values from objects it could be a simple way to segment these different objects. Note that the pixel values will be different for the objects and the image's background if there's a sharp contrast between them.

In this case, we can set a threshold value. The pixel values falling below or above that threshold can be classified accordingly (as an object or the background). This technique is known as **Threshold Segmentation**.

# DIFFERENCES BETWEEN THRESHOLD SEGMENTATION AND REGIONAL GROWTH SEGMENTATION

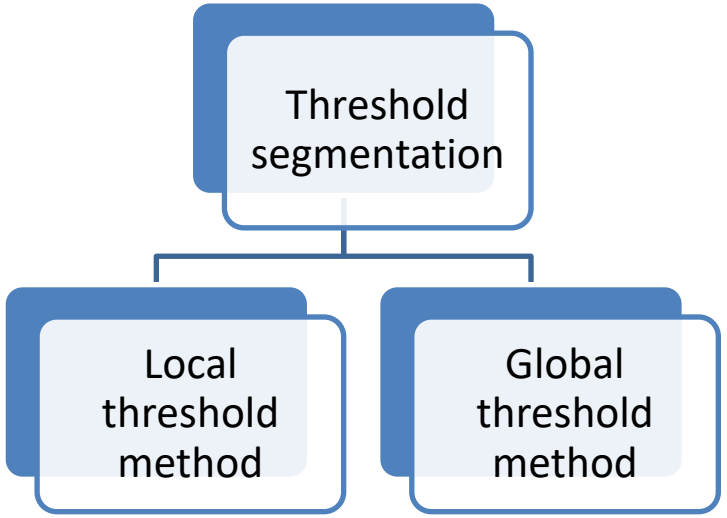
**Threshold Segmentation** essence is to determine automatically the optimal threshold according to a certain criterion, and use these pixels according to the gray level to achieve clustering.

The basic idea of the **regional growth** algorithm is to combine the pixels with similar properties to form the region, which is, for each region to be divided first to find a seed pixel as a growth point, and then merge the surrounding neighborhood with similar properties of the pixel in its area.

# THRESHOLD SEGMENTATION

It is a segmentation algorithm which directly divides the image gray scale information processing based on the gray value of different targets. Can be divided into local threshold method and global threshold method.

The **global threshold method** divides the image into two regions of the target and the background by a single threshold. The local threshold method needs to select multiple segmentation thresholds and divides the image into multiple target regions and backgrounds by multiple thresholds.



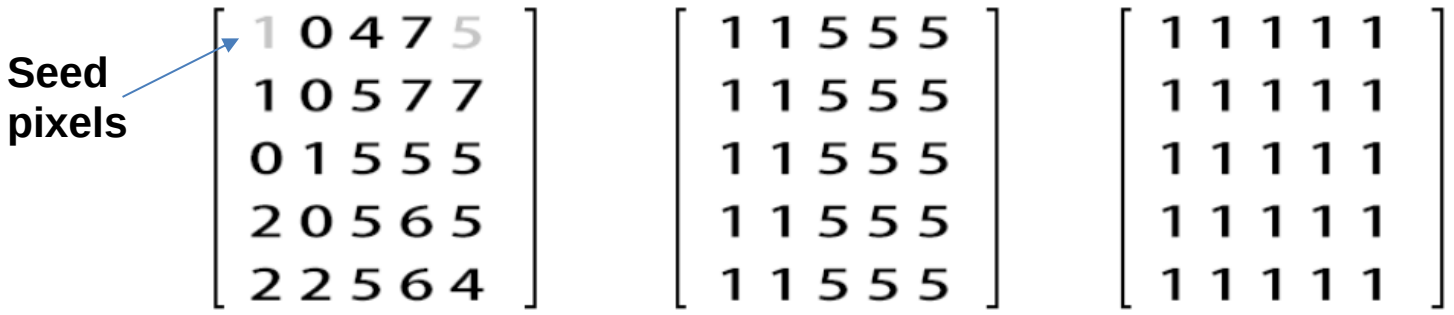
# ADVANTAGES AND DISADVANTAGES OF THRESHOLD SEGMENTATION

The **advantage** of the threshold method is the simple calculation and faster operation speed. The segmentation effect can be achieve when the target and the background have high contrast.

The **disadvantage** is the difficulty in obtaining accurate results for image segmentation problems where there is no significant gray scale difference or a large overlap of the gray scale values in the image. It is sensitive to noise and grayscale unevenness because only takes into account the gray information of the image without considering the spatial information of the image.

# REGIONAL GROWTH SEGMENTATION

The method requires first selecting a seed pixel, and then merging the similar pixels around the seed pixel into the region where the seed pixel is located. The figure below shows an example of a known seed point for region growing.




(a) shows the need to split the image.

(b) shows the regional growth results at T = 3, and the whole plot is well divided into two regions.

(c) shows the results of the region growth at T = 6 and the whole plot is in an area. Thus the choice of threshold is very important

# ADVANTAGES AND DISADVANTAGES OF REGIONAL GROWTH SEGMENTATION

The **advantage** of regional growth is the separation of the connected regions with the same characteristics and the provision of good boundary information and segmentation results. The idea of regional growth is simple and requires only a few seed points to complete. And the growth criteria in the growing process can be freely specified. It can pick multiple criteria at the same time.

The **disadvantage** is the large computational cost. The noise and grayscale unevenness can lead to voids and over-division. The last disadvantage is the shadow effect on the image is often not very good.

# EDGE DETECTION SEGMENTATION

**Edge detection segmentation** algorithm refers to the use of different regions of the pixel gray or color discontinuity detection area of the edge in order to reach image segmentation.

The edge of the object is in the form of discontinuous local features of the image, which is, the most significant part of the image changes in local brightness, like gray value of the mutation, color mutation, texture changes and so on.

There is a case where the gray value is not continuous. This discontinuity can often be detected using derivative operations, and derivatives can be calculated using differential operators. The widely first-order differential operators are Prewitt operator, Roberts operator and Sobel operator. The second-order differential operator has nonlinear operators such as Laplacian, Kirsch operator and Wallis operator.

# THE SOBEL OPERATOR

It is technically a discrete differential operator used to calculate the approximation of the gradient of the image luminance function, and it is based on the first derivative.

The influence of the Sobel operator on the position of the pixel is weighted, which is better than the Prewitt operator and the Roberts operator. The Sobel operator consists of two sets of 3x3 matrices, which are transverse and longitudinal templates, and are plotted with the image plane, respectively, to obtain the difference between the horizontal and the longitudinal difference.

# THE SOBEL OPERATOR

The following two figures are used to detect the edges of the image.

$$Gx = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

Detect horizontal edge

(transverse template)

$$Gy = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

Detect vertical edge

(longitudinal template)

The horizontal and vertical gradient approximations of each pixel of the image can be combined to calculate the size of the gradient using the following formula:

$$G = \sqrt[2]{G_x^2 + G_y^2}$$

The gradient can then be calculated using the following formula:

$$\Theta = \arctan\left(\frac{G_y}{G_x}\right)$$

if the above angle Θ is equal to zero, that is, the image has a longitudinal edge, and the left is darker than the right.

# LAPLACIAN OPERATORS

The Laplace operator's response to isolated pixels is stronger than the edge or line, and therefore applies only to noise-free images. In the presence of noise, the Laplacian operator needs to perform low-pass filtering before detecting the edge.

The Laplace transform of a two-dimensional image function is an isotropic second derivative, which is more suitable for digital image processing, and the pull operator is expressed as a discrete form:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

# LAPLACIAN OPERATORS

The Laplace operator can be expressed in the form of a template.

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

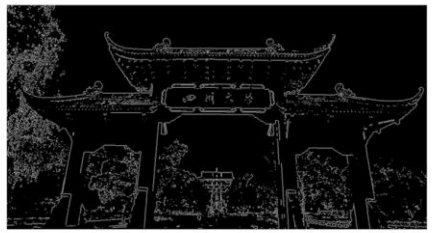
Discrete Laplacian garlic template

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

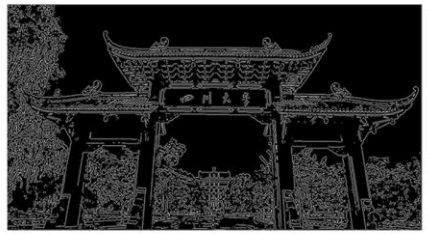
Extended template

It is used to improve the blurring effect due to the blurring effect, since it conforms to the descent model. In the imaging process diffusion effect is often occurring .

original



sobel



laplacian

# SEGMENTATION BASED ON CLUSTERING

The feature space clustering method is used to segment the pixels in the image space with the corresponding feature space points. K-means is one of the most commonly used clustering algorithm.

The algorithm based on clustering is based on the similarity between things as the criterion of class division, which is, it is divided into several subclasses according to the internal structure of the sample set, so that the same kind of samples are similar, and the different are not similar.

The **implementation** process of K-means:

- Select randomly K initial clustering centers;
- Calculate the distance from each sample to each cluster center, and return each sample to the nearest clustering center;
- For each cluster, with the mean of all samples as the cluster of new clustering centers;
- Repeat steps (2) to (3) until the cluster center no longer changes or reaches the set number of iterations.

# ADVANTAGES AND DISADVANTAGES OF K-MEANS CLUSTERING

The advantage of K-Means clustering algorithm is the simplicity and speed of the algorithm, and it is highly efficient and scalable for large data sets. And its time complexity is close to linear, and suitable for mining large-scale data sets.

The disadvantage of K-means is the non-explicit selection criteria of the clustering number K, and is difficult to estimate. It can be seen from the K-means algorithm framework that every iteration of the algorithm traverses all the samples, so the time of the algorithm is very expensive. Lastly, the K-means algorithm is a distance-based partitioning method. It is only applicable to the data set that is convex and not suitable for clustering nonconvex clusters.

# SEGMENTATION BASED ON WEAKLY-SUPERVISED LEARNING IN CNN

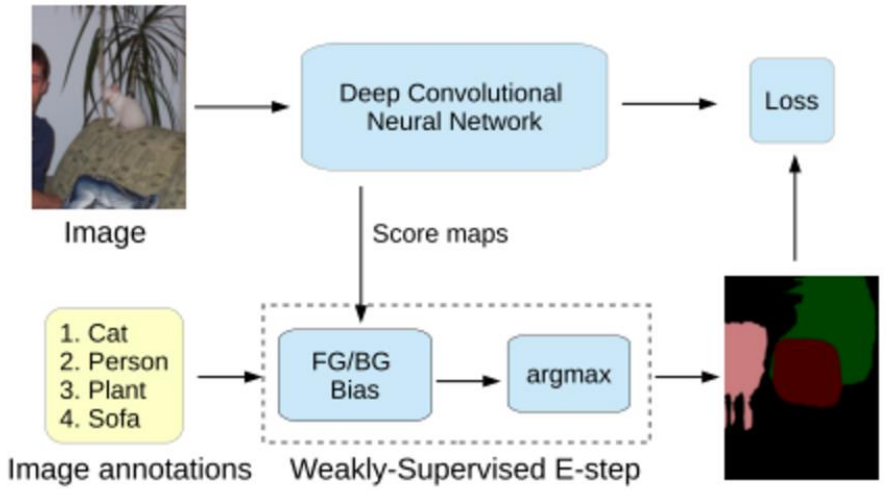
It refers to the problem of assigning a semantic label to every pixel in the image and consists of three parts.

1) Give an image which contains which objects.

2) Give the border of an object.

3) The object area in the image is marked with a partial pixel

DeepLab method is divided into two steps, the first is still using the FCN to get the coarse score map and interpolate to the original image size, and then the second step began to borrow the fully connected CRF from the FCN to get the details of the segmentation refinement.

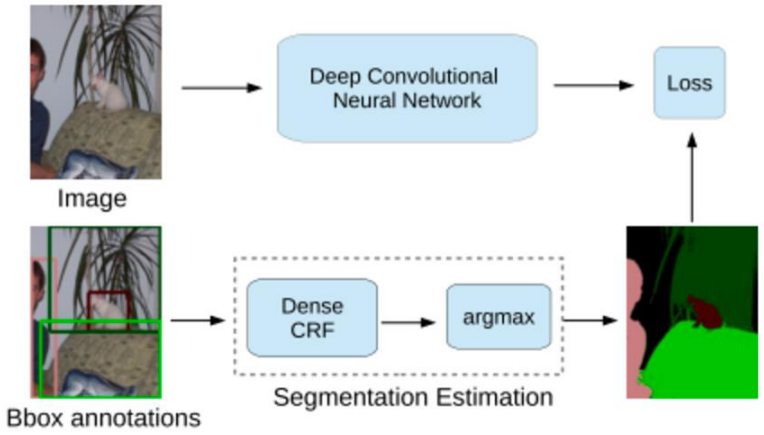


# SEGMENTATION BASED ON WEAKLY-SUPERVISED LEARNING IN CNN

For the image-level tagged data, the pixel value x of the image can be observed and as also the mark z of the image level, but do not know the label y for each pixel, so y is treated as a hidden variable. Use the probability graph mode:

$$P(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}; \boldsymbol{\theta}) = P(\boldsymbol{x}) \left( \prod_{m=1}^{M} P(y_m | \boldsymbol{x}; \boldsymbol{\theta}) \right) P(\boldsymbol{z} | \boldsymbol{y}) .$$

Use the EM algorithm to estimate $\boldsymbol{\theta}$ and y. E step is fixed $\boldsymbol{\theta}$ y expect the value of y, and M step is fixed y using SGD to calculate $\boldsymbol{\theta}$ .



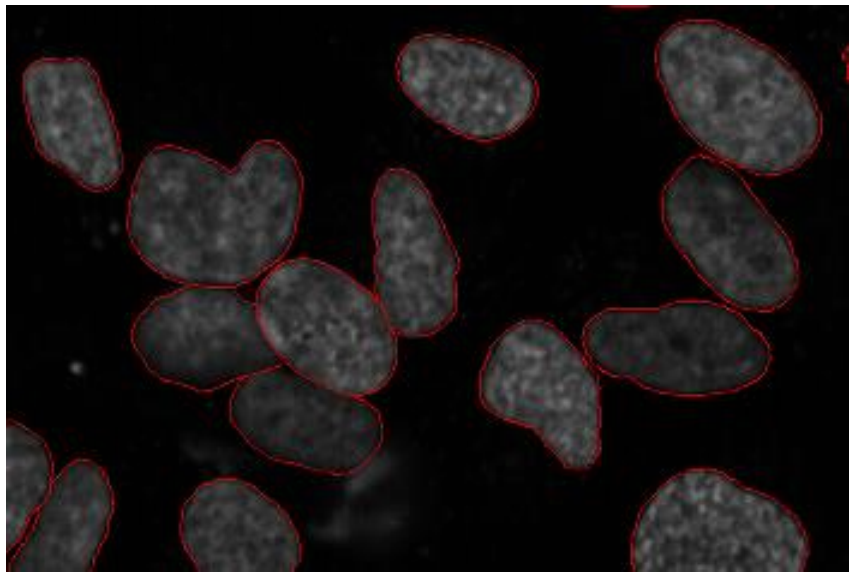

Experiments show that simply using the image level of the mark to get the segmentation effect is poor, but the use of bounding box training data can get better results

# USE-CASES OF EACH IMAGE SEGMENTATION ALGORITHM

Cancer has long been a deadly illness. The shape of the cancerous cells plays a vital role in determining the severity of the cancer.

As we can see in the figure, the shapes of all the cancerous cells are clearly. There are many other applications where Image segmentation is transforming industries:

- Traffic Control Systems
- Self Driving Cars
- Locating objects in satellite images



# SUMMARY

- **Image segmentation** creates a pixel wise mask for each object in the image.This technique gives us a better understanding of the object(s) in the image.
- Types of image segmentation algorithms are **Region-based Segmentation, Edge Detection Segmentation, Segmentation based on clustering,** and **Segmentation based on weakly-supervised learning in CNN.**
- **Region-based segmentation u**se pixel values from objects it could be a simple way to segment these different objects.
- **Edge detection segmentation** refers to the use of different regions of the pixel gray or color discontinuity detection area of the edge in order to reach image segmentation.
- The feature space **clustering** method is used to segment the pixels in the image space with the corresponding feature space points.
- **Segmentation based on weakly-supervised learning in CNN** refers to the problem of assigning a semantic label to every pixel in the image and consists of three parts.

# SUMMARY

| Algorithm | Description | Advantages | Limitations |
|---|---|---|---|
| **Region-Based Segmentation** | Separates the objects into different regions based on some threshold value(s). | • Simple calculations<br>• Fast operation speed<br>• Well performance when the object and background have high contrast. | When there is no significant grayscale difference or an overlap of the grayscale pixel values, it becomes very difficult to get accurate segments. |
| **Edge Detection Segmentation** | Makes use of discontinuous local features of an image to detect edges and hence define a boundary of the object. | It is good for images having better contrast between objects. | Not suitable when there are too many edges in the image and if there is less contrast between objects. |
| **Segmentation based on Clustering** | Divides the pixels of the image into homogeneous clusters. | Works really well on small datasets and generates excellent clusters. | • Computation time is too large and expensive.<br>• k-means is a distance-based algorithm. It is not suitable for clustering non-convex clusters. |
| **Mask R-CNN** | Gives three outputs for each object in the image: its class, bounding box coordinates, and object mask | • Simple, flexible and general approach<br>• It is also the current state-of-the-art for image segmentation | High training time |

# REFERENCES

*Image Segmentation* | *Types Of Image Segmentation*. (n.d.). Retrieved from https://www.analyticsvidhya.com/blog/2019/04/introduction-image-segmentation-techniques-python/

Ryan, T. W. (1985). Image Segmentation Algorithms. *Architectures and Algorithms for Digital Image Processing II*, *0534*, 172. https://doi.org/10.1117/12.946577

# REFERENCE TO AUTHORS


**Regina Sousa**

- PhD student in Biomedical Engineering
- Research Collaborator of the Algoritmi Research Center

0000-0002-2988-196X



**Ana Luísa Sousa**

- PhD student in Information System and Tecnologies
- Research Collaborator of the Algoritmi Research Center

0000-0001-5731-3583



**Diana Ferreira**

- PhD student in Biomedical Engineering
- Research Collaborator of the Algoritmi Research Center

0000-0003-2326-2153

# REFERENCE TO AUTHORS


**António Abelha**

- Assistant Professor at the University of Minho
- Integrated Researcher of the Algoritmi Research Center

0000-0001-6457-0756



**José Machado**

- Associate Professor with Habilitation at the University of Minho
- Integrated Researcher of the Algoritmi Research Center

0000-0003-4121-6169



**Victor Alves**

- Assistant Professor at the University of Minho
- Integrated Researcher of the Algoritmi Research Center

0000-0003-1819-7051

# REFERENCE TO AUTHORS

This Training Material has been certified according to the rules of **ECQA – European Certification and Qualification Association.**

The Training Material was developed within the international job role committee **“Computer Vision Expert”:**

**UMINHO – University of Minho** (https://www.uminho.pt/PT)

The development of the training material was partly funded by the EU under Blueprint Project DRIVES.

DRIVES

Development and Research on Innovative Vocational Education Skills

# Thank you for your attention

DRIVES project is project under **The Blueprint for Sectoral Cooperation on Skills in Automotive Sector**, as part of New Skills Agenda.

The aim of the Blueprint is **to support an overall sectoral strategy and to develop concrete actions to address short and medium term skills needs.**

**Follow DRIVES project at:**

**More information at:**

www.project-drives.eu

Co-funded by the Erasmus+ Programme of the European Union

The Development and Research on Innovative Vocational Educational Skills project (DRIVES) is co-funded by the Erasmus+ Programme of the European Union under the agreement 591988-EPP-1-2017-1-CZ-EPPKA2-SSA-B. The European Commission support for the production of this publication does not constitute endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.