



SLOVENSKI STANDARD
oSIST prEN 18288:2026
01-junij-2026

Umetna inteligenca - Taksonomija nalog umetne inteligence (UI) v računalniškem vidu

Artificial Intelligence - Taxonomy of AI tasks in computer vision

Künstliche Intelligenz - Taxonomie der KI-Aufgaben in der Computer Vision

Intelligence artificielle - Taxonomie des tâches de l'IA dans le domaine de la vision par ordinateur

Ta slovenski standard je istoveten z: prEN 18288

ICS:

35.240.01	Uporabniške rešitve informacijske tehnike in tehnologije na splošno	Application of information technology in general
-----------	---	--

oSIST prEN 18288:2026

en,fr,de

Sample Document

get full document from standards.iteh.ai

EUROPEAN STANDARD
NORME EUROPÉENNE
EUROPÄISCHE NORM

DRAFT
prEN 18288

April 2026

ICS

English version

Artificial Intelligence - Taxonomy of AI tasks in computer vision

Intelligence artificielle - Taxonomie des tâches de l'IA dans le domaine de la vision par ordinateur

Künstliche Intelligenz - Taxonomie der KI-Aufgaben in der Computer Vision

This draft European Standard is submitted to CEN members for enquiry. It has been drawn up by the Technical Committee CEN/CLC/JTC 21.

If this draft becomes a European Standard, CEN and CENELEC members are bound to comply with the CEN/CENELEC Internal Regulations which stipulate the conditions for giving this European Standard the status of a national standard without any alteration.

This draft European Standard was established by CEN and CENELEC in three official versions (English, French, German). A version in any other language made by translation under the responsibility of a CEN and CENELEC member into its own language and notified to the CEN-CENELEC Management Centre has the same status as the official versions.

CEN and CENELEC members are the national standards bodies and national electrotechnical committees of Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Republic of North Macedonia, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Türkiye and United Kingdom.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation. Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

Warning : This document is not a European Standard. It is distributed for review and comments. It is subject to change without notice and shall not be referred to as a European Standard.



CEN-CENELEC Management Centre:
Rue de la Science 23, B-1040 Brussels

Contents	Page
European foreword	4
1 Scope	5
2 Normative references	5
3 Terms and definitions	5
3.1 Computer vision related terms	5
4 Abbreviated terms	5
5 Tasks	5
5.1 General	5
5.2 Definition of computer vision task	6
5.3 Categories of computer vision tasks	6
5.4 Task variants and versions	6
5.5 Role of context on computer vision tasks	6
6 Core computer vision tasks	6
6.1 General	6
6.2 Image analysis	6
6.2.1 General	6
6.2.2 Object localization	6
6.2.3 Object classification	7
6.2.4 Object detection	8
6.2.5 Semantic segmentation	9
6.2.6 Instance segmentation	9
6.2.7 Panoptic segmentation	9
6.2.8 Image registration	10
6.2.9 Image similarity	11
6.2.10 Image classification	11
6.2.11 Face analysis	11
6.2.12 Saliency detection	12
6.2.13 Scene Graph Generation (SGG)	12
6.2.14 Instance identification	12
6.2.15 Feature extraction	13
6.2.16 Geometric feature detection	13
6.2.17 Image stitching	14
6.2.18 Image retrieval	14
6.3 Spatial analysis	15
6.3.1 General	15
6.3.2 Depth estimation	15
6.3.3 Object / 3D reconstruction	16
6.3.4 Object pose estimation	17
6.3.5 Camera pose estimation	17
6.3.6 Simultaneous Localization and Mapping (SLAM)	18
6.4 Temporal analysis	18
6.4.1 General	18
6.4.2 Multiple Object Tracking (MOT)	19
6.4.3 Action recognition	19
6.4.4 Event detection	19
6.4.5 Object re-identification (Re-ID)	20
6.4.6 Visual odometry	21
6.4.7 Change detection	21
6.4.8 Deepfake detection	21
6.4.9 Video summarization	22
6.4.10 Optical flow estimation	22
6.4.11 Motion detection	22

6.4.12	Single object tracking	23
6.5	Tasks producing image as an output	23
6.5.1	General	23
6.5.2	Image generation	23
6.5.3	Image denoising	24
6.5.4	Super-resolution	24
6.5.5	Image inpainting	24
6.5.6	Image-to-image translation	25
6.5.7	Style transfer	25
6.5.8	Object swapping	26
6.5.9	Novel view synthesis	26
6.6	Tasks producing video as an output	26
6.6.1	General	26
6.6.2	Video enhancement	27
6.6.3	Video generation	27
6.6.4	Video stabilization	27
6.7	Multi-modal tasks	28
6.7.1	General	28
6.7.2	Visual question answering (VQA)	28
6.7.3	Image captioning	28
6.7.4	Visual grounding	29
6.7.5	RGB-depth fusion	29
	Bibliography	30

Sample Document

get full document from standards.iteh.ai

prEN 18288 (E)

European foreword

This document (prEN 18288:2026) has been prepared by Technical Committee CEN/CENELEC JTC 21 "Artificial Intelligence", the secretariat of which is held by DS.

This document is currently submitted to the CEN Enquiry.

Sample Document

get full document from standards.iteh.ai

1 Scope

This document describes a taxonomy of the AI tasks related to computer vision, including tasks pertaining to either the analysis or generation of images and videos. It identifies existing and competing terminologies, as well as co-existing variants of the same tasks.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

- ISO Online browsing platform: available at <http://www.iso.org/obp>
- IEC Electropedia: available at <http://www.electropedia.org/>

3.1 Computer vision related terms

3.1.1

computer vision

artificial vision capability of a functional unit to acquire, process, and interpret visual data

Note 1 to entry: Computer vision involves the use of visual sensors to create an electronic or digital image of a visual scene.

Note 2 to entry: Not to be confused with machine vision.

Note 3 to entry: Computer vision; artificial vision: terms and definition standardized by ISO/IEC [ISO/IEC 2382-28:1995 [1]].

[SOURCE: ISO/IEC 2382:2015 [2]]

3.1.2

image

<digital> graphical content intended to be presented visually.

4 Abbreviated terms

MOT	Multiple Object Tracking
SGG	Scene Graph Generation
SLAM	Simultaneous Localization And Mapping
VQA	Visual Question Answering

5 Tasks

5.1 General

This clause provides an overview of the concept of a [computer vision \(3.1.1\)](#) task in vision-based systems. It establishes the foundational principles for defining, categorizing, and contextualizing tasks to ensure consistency and comparability across applications.

prEN 18288 (E)

5.2 Definition of computer vision task

A computer vision task is defined as a specific problem to solve that involves the interpretation or transformation of visual data (e.g. [images \(3.1.2\)](#), videos) using computational methods.

5.3 Categories of computer vision tasks

Computer vision tasks are categorized according to their functional goals which can encompass a set of subtasks and functionalities. These categories provide a shared taxonomy across domains and applications and are explained in [Clause 6](#).

EXAMPLE Image analysis encompasses object localization and object classification.

5.4 Task variants and versions

Variants of a task arise when modifications are made to the input modalities, output formats, or operational conditions while maintaining the same functional goal. For example, the localization task for 3D object is an extension of the localization task for 2D objects, producing 3D bounding boxes instead of 2D ones.

5.5 Role of context on computer vision tasks

Context influences both the interpretation and the difficulty of a computer vision task. Accounting for context ensures evaluation of vision tasks in operational environments is reproducible.

EXAMPLE In object detection tasks for autonomous driving, the context (e.g. weather, lighting, speed) strongly influences both the visual appearance of objects and the reliability of detections. For instance, detecting pedestrians at night in heavy rain at high speed introduces motion blur and reflections that degrade model accuracy compared to clear daylight conditions.

6 Core computer vision tasks

6.1 General

This clause introduces the fundamental computer vision tasks that serve as building blocks for more complex computer vision systems. It defines their properties and input-output structures.

6.2 Image analysis

6.2.1 General

Image analysis encompasses tasks aimed at extracting meaningful information from visual data. It forms the basis for understanding visual scenes by:

- localizing objects
- classifying objects
- detecting objects
- segmenting objects
- making links between visual elements in an image (e.g. generating scene graph, stitching images)
- extracting features

6.2.2 Object localization

Object localization is a computer vision task that deals with localization of objects in a digital image which involves a geometric description of the spatial extent of the object. The most frequent geometric

description is the minimum bounding rectangle, i.e. the bounding box that envelops the object fully, see [Figure 1](#). However, other descriptors, such as polygons can be used [\[3\]](#).

Bounding boxes around objects are typically described by four of values values:

- either two corners, e.g. upper-left and bottom-right with x and y coordinates, i.e. [upper-left x, upper-left y, bottom-right x, bottom-right y]
- one corner with x, y coordinates, height, width, [upper-left x, upper-left y, height, width]
- object centre x, y coordinates, object height, object width, i.e. [centre x, centre y, height, width]



Figure 1 — In object localization tasks, most often the minimal bounding rectangle, i.e. the bounding box, describes where an object is located in an image, here around the person.

EXAMPLE 1 The localization of a face in the front-facing camera of mobile phones, for the purpose of unlocking the screenlock.

EXAMPLE 2 The localization of cars in satellite imagery for the purpose of traffic density estimation.

NOTE 1 Often localization coordinates and height/width are not described in absolute pixel values but fractions of the image width and height

NOTE 2 By convention the $(x,y) = (0,0)$ coordinate system origin is placed at the top left of the image, with x values increasing to the right and y values increasing towards the bottom. However, this convention is not universal and might not hold everywhere.

6.2.3 Object classification

Object classification is a computer vision task focused on identifying and categorizing objects in a digital image into specific groups, categories, or classes. The assigned classification of a given object can follow several several structures but are commonly Binary, Multi-class, Multi-label classification.

- Input of the object classification task:
 - image or video sequence containing the object targeted for classification.