

E DIN EN 18288:2026-06 (E)

Erscheinungsdatum: 2026-05-15

Artificial Intelligence - Taxonomy of AI tasks in computer vision; English version prEN 18288:2026

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.....	6
5.1 General.....	6
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.....	7
6.2.3 Object classification.....	8
6.2.4 Object detection	8
6.2.5 Semantic segmentation	9
6.2.6 Instance segmentation.....	9
6.2.7 Panoptic segmentation.....	10
6.2.8 Image registration.....	10
6.2.9 Image similarity	11
6.2.10 Image classification	12
6.2.11 Face analysis	12
6.2.12 Saliency detection	12
6.2.13 Scene Graph Generation (SGG).....	13
6.2.14 Instance identification.....	13
6.2.15 Feature extraction.....	14
6.2.16 Geometric feature detection.....	14
6.2.17 Image stitching.....	15
6.2.18 Image retrieval.....	15
6.3 Spatial analysis.....	16
6.3.1 General.....	16
6.3.2 Depth estimation	16
6.3.3 Object / 3D reconstruction.....	17
6.3.4 Object pose estimation	18
6.3.5 Camera pose estimation.....	19
6.3.6 Simultaneous Localization and Mapping (SLAM).....	19
6.4 Temporal analysis.....	20
6.4.1 General.....	20
6.4.2 Multiple Object Tracking (MOT).....	20
6.4.3 Action recognition.....	20
6.4.4 Event detection	21

6.4.5	Object re-identification (Re-ID).....	21
6.4.6	Visual odometry.....	22
6.4.7	Change detection.....	22
6.4.8	Deepfake detection.....	23
6.4.9	Video summarization.....	23
6.4.10	Optical flow estimation	23
6.4.11	Motion detection	24
6.4.12	Single object tracking.....	24
6.5	Tasks producing image as an output	25
6.5.1	General	25
6.5.2	Image generation	25
6.5.3	Image denoising.....	25
6.5.4	Super-resolution.....	26
6.5.5	Image inpainting.....	26
6.5.6	Image-to-image translation	26
6.5.7	Style transfer	27
6.5.8	Object swapping	27
6.5.9	Novel view synthesis	28
6.6	Tasks producing video as an output	28
6.6.1	General	28
6.6.2	Video enhancement.....	28
6.6.3	Video generation	29
6.6.4	Video stabilization.....	29
6.7	Multi-modal tasks	30
6.7.1	General	30
6.7.2	Visual question answering (VQA)	30
6.7.3	Image captioning.....	30
6.7.4	Visual grounding	31
6.7.5	RGB-depth fusion	31
	Bibliography.....	32

Figures

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.....	7
Figure 2	— Object detection of a dog and a person, here both objects are localized using a bounding box, the colour represents the different classes these objects are categorized into.	9
Figure 3	— Human pose estimation is often conducted on multiple rigid body based model assumption	18