

# Roadmap PSVO-2024

---

Concept	Description
Traitement d'image	Prétraitement de l'image pour la rendre exploitable. Inclut le redimensionnement, le filtrage, la conversion en niveaux de gris, la suppression de bruit, etc.
Segmentation	Division de l'image en régions distinctes pour identifier et isoler les objets ou les zones d'intérêt.
Détection de caractéristiques	Extraction de points d'intérêt comme les contours, les coins, ou les textures. Utilise des algorithmes comme SIFT, SURF, et Canny.
Classification d'image	Assignation d'une étiquette ou catégorie à une image entière. Par exemple, classer une image comme étant un chat, un chien, etc.
Détection d'objets	Localisation et identification d'objets spécifiques dans une image avec des boîtes englobantes (par exemple, avec YOLO, SSD, ou R-CNN).
Suivi d'objets (Object Tracking)	Suivi du mouvement et de la position d'un objet à travers plusieurs images ou une vidéo.
Reconstruction 3D	Création de modèles 3D à partir d'images 2D, souvent via des stéréoscopies ou des caméras à profondeur.
Reconnaissance faciale	Identification et vérification de visages dans une image ou une vidéo, utilisée dans des systèmes de sécurité, biométrie, etc.
Segmentation sémantique	Attribution d'une étiquette à chaque pixel de l'image, permettant de créer

	des masques précis pour chaque objet.
Segmentation d'instances	Différenciation de chaque instance d'un objet dans une image, même s'il y a plusieurs objets du même type.
Reconnaissance de texte (OCR)	Détection et conversion du texte dans une image en texte numérique exploitable, utile pour la numérisation de documents.
Réseaux génératifs adverses (GANs)	Utilisation des GANs pour générer de nouvelles images réalistes, transformer des images, ou améliorer la qualité d'image (super-résolution).
Super-résolution	Amélioration de la qualité et des détails d'une image en augmentant sa résolution avec des algorithmes d'apprentissage profond.
Reconnaissance d'action	Analyse des mouvements dans les vidéos pour identifier des actions humaines, utilisée pour la surveillance, le sport, etc.
Vision basée sur les événements	Capture d'images en fonction des changements dans une scène, permettant une analyse plus rapide et plus efficace en termes d'énergie.
Vision multicapteur	Fusion de données provenant de plusieurs capteurs (caméra, lidar, radar) pour obtenir une compréhension plus complète de l'environnement, utilisé notamment dans les voitures autonomes.
Perception 3D en temps réel	Modélisation en 3D d'environnements en temps réel pour les applications robotiques ou drones, permettant la navigation autonome.
Vision augmentée et mixte	Combinaison d'images réelles avec des informations virtuelles pour

enrichir la perception d'une scène, utilisée dans la réalité augmentée (AR).

Détection d'anomalies

Identification d'objets ou de comportements anormaux dans des images ou vidéos, crucial pour la sécurité ou la surveillance.

Apprentissage multimodal

Combinaison de données provenant de différentes modalités (image, texte, audio) pour obtenir une meilleure compréhension des scènes ou des actions.