

Mis à jour le 03/02/2026

S'inscrire

Formation Moondream AI

2 jours (14 heures)

Présentation

Moondream AI est un modèle de vision-langage léger pour décrire des images, extraire des informations et répondre à des questions visuelles. Vous apprendrez à l'intégrer rapidement pour automatiser des cas d'usage comme le contrôle qualité, l'indexation d'images, ou l'assistance à la saisie via OCR et compréhension de scène.

Cette formation vise une prise en main opérationnelle de Moondream AI : chargement du modèle, inférence sur images, prompts efficaces et structuration des sorties (JSON, champs métiers). Les exercices s'appuient sur des scénarios concrets (photos produits, documents, captures d'écran) et des métriques simples pour valider la qualité.

L'approche est centrée sur des ateliers et démos reproductibles : pipeline de prétraitement, gestion des performances (CPU/GPU), et intégration dans une API. Les livrables incluent un notebook/mini-projet, un script d'inférence, et une checklist de bonnes pratiques (prompts, latence, erreurs).

Objectifs

- Installer et configurer Moondream AI dans un environnement Python.
- Réaliser des inférences image?texte et des questions/réponses visuelles.
- Concevoir des prompts robustes et des sorties structurées exploitables.
- Optimiser latence et consommation mémoire selon CPU/GPU.
- Exposer le modèle via une API simple et tester en conditions réelles.

Public visé

- Développeurs Python souhaitant intégrer de la vision IA.
- Data scientists/ML engineers orientés prototypage rapide.
- Ingénieurs backend voulant industrialiser une brique VLM.
- Chefs de projet technique évaluant des cas d'usage vision.

Pré-requis

- Bonnes bases en Python (packages, environnements, scripts).
- Notions de machine learning et d'inférence de modèles.
- Manipulation de fichiers image et formats (JPEG/PNG) et JSON.
- Connaissances de base en API HTTP (requêtes, payloads).

Pré-requis techniques

- Ordinateur avec 16 Go de RAM recommandés (8 Go minimum).
- Windows (idéalement WSL2), macOS ou Linux.
- Python 3.10+, gestion d'environnements (venv/conda) et pip.
- Option GPU : NVIDIA avec pilotes et stack compatible (selon framework), sinon exécution CPU.
- Éditeur de code (VS Code ou équivalent) et un terminal.

Programme de formation Moondream AI

[Jour 1 - Matin]

Découvrir Moondream AI et préparer l'environnement

- Comprendre le positionnement : VLM (Vision-Language Model) et cas d'usage (caption, VQA, extraction visuelle)
- Choisir le mode d'exécution : local CPU/GPU, contraintes mémoire, latence et coût
- Installer et valider l'environnement Python (venv, dépendances, versions) et un premier script d'inférence
- Structurer un projet : gestion des entrées (images), sorties (JSON), logs et configuration
- Atelier pratique : Installer Moondream, exécuter une inférence sur 10 images et exporter les résultats en JSON.

[Jour 1 - Après-midi]

Prompts vision & extraction d'informations fiables

- Concevoir des prompts orientés tâches : description, question/réponse, extraction structurée (champs attendus)
- Réduire l'ambiguïté : consignes, format de sortie, contraintes (unités, listes fermées, priorité aux éléments visibles)
- Gérer les cas difficiles : faible luminosité, texte petit, objets multiples, arrière-plan bruité
- Mettre en place des garde-fous : validation de schéma, règles métier, détection d'incertitude et reprise manuelle
- Atelier pratique : Construire un extracteur (JSON) pour analyser des photos de produits (nom, couleur, présence de logo, état).

[Jour 2 - Matin]

Industrialiser un pipeline de traitement d'images

- Prétraitement : redimensionnement, recadrage, normalisation, gestion des formats et orientation EXIF
- Traitement par lots : files d'attente, parallélisme, reprise sur erreur et idempotence
- Mesurer la qualité : jeux de test, critères (exactitude, complétude), échantillonnage et revue humaine
- Optimiser les performances : cache, choix de résolution, limitation du contexte et gestion des timeouts
- Atelier pratique : Créer un pipeline CLI qui ingère un dossier, traite en batch et produit un rapport de qualité (CSV + erreurs).

[Jour 2 - Après-midi]

Exposer Moondream en service et sécuriser l'usage

- Créer une API simple (endpoint analyse image) avec gestion des uploads et réponses JSON stables
- Bonnes pratiques de robustesse : limites de taille, quotas, validation MIME, gestion des erreurs et observabilité
- Déploiement : packaging, variables d'environnement, configuration, logs et stratégie de montée en charge
- Conformité et risques : données sensibles, anonymisation, rétention, traçabilité et politiques d'accès
- Atelier pratique : Déployer un microservice d'analyse d'images avec endpoint /analyze et tests de charge basiques.

Sociétés concernées

Cette formation s'adresse à la fois aux particuliers ainsi qu'aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée ou bien à acquérir des connaissances métiers spécifiques ou des méthodes modernes.

Positionnement à l'entrée en formation

Le positionnement à l'entrée en formation respecte les critères qualité Qualiopi. Dès son inscription définitive, l'apprenant reçoit un questionnaire d'auto-évaluation nous permettant d'apprécier son niveau estimé sur différents types de technologies, ses attentes et objectifs personnels quant à la formation à venir, dans les limites imposées par le format sélectionné. Ce questionnaire nous permet également d'anticiper certaines difficultés de connexion ou de sécurité interne en entreprise (intraentreprise ou classe virtuelle) qui pourraient être problématiques pour le suivi et le bon déroulement de la session de formation.

Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format numérique à tous les participants.

Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de réflexions, et de travail en groupe.

Validation

À la fin de la session, un questionnaire à choix multiples permet de vérifier l'acquisition correcte des compétences.

Sanction

Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.