

Évaluation de modèles vidéo à poids ouverts en inférence locale : application au service communication de l'ENSMA

Localisation : Site du Futuroscope, France

Date de début : mai 2027 (flexible)

Durée : 2 mois maximum

Encadrants : Mickaël BARON baron@ensma.fr)

Mots clés : intelligence artificielle, génération vidéo, modèles à poids ouverts, inférence locale, benchmark, ComfyUI, consommation électrique.

Contexte du stage

Les progrès de l'intelligence artificielle permettent aujourd'hui de produire ou de transformer différents types de contenus : textes, images, sons et vidéos. Dans le domaine audiovisuel, des outils comme ComfyUI¹ facilitent la création de workflows visuels : les différentes étapes de production peuvent être assemblées sous forme de blocs, configurées et réutilisées.

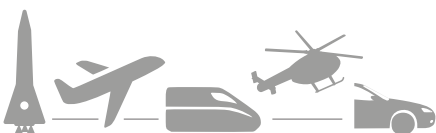
Pour le service communication de l'ISAE-ENSMA, ces outils pourraient ouvrir de nouvelles possibilités de création vidéo et contribuer à accélérer la production de contenus. Leur utilisation soulève toutefois plusieurs questions : comment obtenir des résultats cohérents avec l'identité visuelle de l'École ? Comment rendre les workflows suffisamment simples et réutilisables ? Et quelles possibilités offrent les modèles à poids ouverts lorsqu'ils sont exécutés localement, sans dépendre uniquement de services en ligne ?

Objectif du stage

L'objectif principal du stage est de concevoir et de mettre en œuvre un benchmark permettant d'évaluer et de comparer des modèles d'intelligence artificielle à poids ouverts, exécutés localement, pour des tâches de création vidéo. Le benchmark prendra appui sur un cas d'usage concret : la production de contenus pour le service communication de l'ISAE-ENSMA, au moyen de workflows ComfyUI et de templates respectant l'identité visuelle de l'École.

Le travail portera principalement sur la définition des scénarios et des critères d'évaluation, puis sur l'expérimentation des modèles. Les résultats seront comparés selon la qualité des vidéos produites, les temps de génération, l'utilisation des ressources matérielles, la consommation électrique et la facilité d'intégration dans les workflows. Le ou la stagiaire formulera ainsi des recommandations sur les modèles et les configurations les plus adaptés. Du matériel équipé de cartes graphiques sera mis à sa disposition pour mener ces expérimentations.

1. <https://comfy.org/>



Missions principales

Les missions principales sont décrites ci-dessous.

- Analyser les besoins du service communication de l'ISAE-ENSMA et définir des scénarios représentatifs de création vidéo.
- Concevoir le benchmark : choisir les modèles à poids ouverts à tester, établir un protocole expérimental reproductible et définir les critères de comparaison.
- Mettre en place les expérimentations en local, notamment à l'aide de ComfyUI, sur le matériel équipé de cartes graphiques mis à disposition.
- Évaluer et comparer les modèles selon la qualité des vidéos produites, les temps de génération, l'utilisation des ressources matérielles et la consommation électrique.
- Créer un cas d'usage pour l'ISAE-ENSMA, en élaborant des workflows et des templates ComfyUI réutilisables, cohérents avec l'identité visuelle de l'École.
- Analyser les résultats et formuler des recommandations sur les modèles, les configurations et les conditions les plus adaptés à la production vidéo locale.
- Documenter le benchmark et les workflows afin de faciliter leur réutilisation et la poursuite des expérimentations.

Profil du candidat

Le candidat doit être étudiant en licence 3 dans le domaine de l'informatique.

La curiosité, l'autonomie, la rigueur et l'envie de travailler sur un projet associant logiciel et matériel sont également recherchées.

Documents à fournir

Toute candidature incomplète ne pourra pas être étudiée.

- Curriculum vitæ;
- les relevés de notes permettant d'apprécier le parcours et le potentiel du candidat;
- tout autre document utile pour appuyer la candidature (par exemple : présentation de projets, lien vers un dépôt GitHub ou portfolio).

