

Super-résolution par apprentissage profond pour les jeux vidéo

Projet TSI 2026–27

Encadrant : Michaël Clément – michael.clement@enseirb-matmeca.fr

Mots-clés : Apprentissage profond, super-résolution, temps réel, jeux vidéo

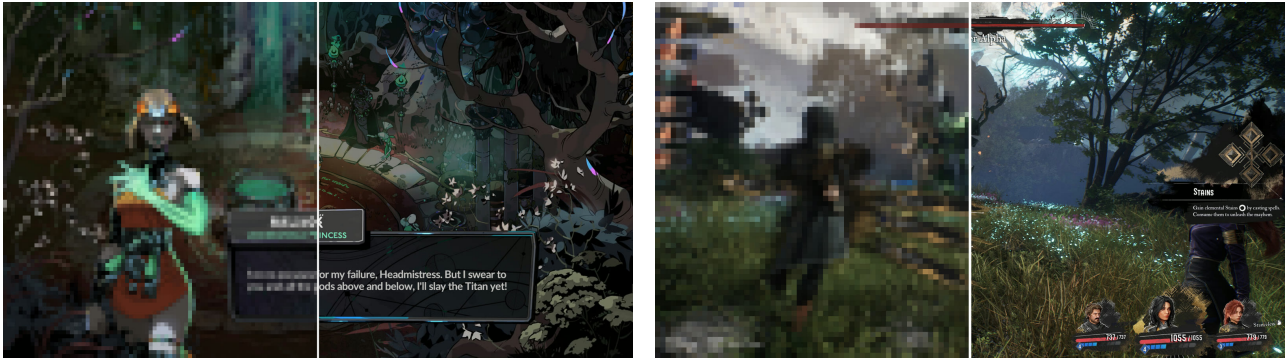


FIGURE 1 – Exemples de super-résolution $\times 16$ sur des captures de jeux vidéo : image très basse résolution (à gauche) et image haute résolution d’origine (à droite).

Contexte

Le rendu d’un jeu vidéo en 4K et à haute fréquence d’images est très coûteux, même pour les cartes graphiques récentes. Pour alléger ce coût, il est possible de rendre le jeu en basse résolution puis de l’agrandir à l’aide d’un réseau de neurones (Figure 1), ce qui permet de gagner en fluidité avec une perte de qualité visuelle limitée. Popularisée par le DLSS de NVIDIA, cette approche est aujourd’hui proposée par de nombreux constructeurs (AMD FSR 4, Intel XeSS, Sony PSSR). Elle s’appuie sur la super-résolution d’images et de vidéos par apprentissage profond [1, 2, 3]. Les méthodes les plus avancées exploitent aussi l’information temporelle, en combinant plusieurs images successives et en tenant compte du mouvement dans la scène [4, 5]. La principale difficulté reste la contrainte de temps réel : tout le traitement doit tenir en quelques millisecondes par image.

Objectifs du projet

L’objectif de ce projet est de concevoir un « DLSS maison » : une chaîne complète de super-résolution par apprentissage profond pour des jeux vidéo, des données d’entraînement jusqu’à l’utilisation en temps réel. Le projet s’articulera autour des axes suivants :

- **État de l’art.** Étude bibliographique des méthodes de super-résolution d’images et de vidéos par apprentissage profond, en particulier des architectures légères adaptées au temps réel.
- **Données et entraînement.** Constitution de bases d’apprentissage à partir de jeux de données publics et de captures de jeux vidéo. Entraînement et *fine-tuning* de différents modèles, et comparaison de leurs performances (PSNR, SSIM, temps par image).
- **Optimisation.** Réduction du temps d’exécution des modèles : quantification, élagage de poids, export ONNX, TensorRT, etc.
- **Amélioration de la qualité.** Exploitation de plusieurs images successives et/ou du mouvement pour améliorer la qualité et la stabilité temporelle. Ouverture possible vers le transfert de style.
- **Intégration dans un moteur de jeu.** Développement d’un petit jeu (Godot, Unity) ou reprise d’un jeu open source pour embarquer le modèle directement dans le moteur, avec accès à la profondeur et aux vecteurs de mouvement.

Références

- [1] W. Shi et al. « Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network ». *CVPR*, 2016.
- [2] X. Wang et al. « Real-ESRGAN: Training real-world blind super-resolution with pure synthetic data ». *ICCVW*, 2021.
- [3] C. Wan et al. « Swift parameter-free attention network for efficient super-resolution ». *CVPRW*, 2024.
- [4] K. C. K. Chan et al. « BasicVSR++: Improving video super-resolution with enhanced propagation and alignment ». *CVPR*, 2022.
- [5] L. Xiao et al. « Neural supersampling for real-time rendering ». *ACM TOG*, 2020.