IJIS: Vol.1, Issue 10, November 2025, Page: 21-25 ISSN: 3049-3277

International Journal of Integrative Studies (IJIS)

Journal homepage:www.ijis.co.in

# Tile-Based Image Upscaling Using AI

**Dr. Padmaja Pulicherla*[1], Kaushik Pendem[2]**

*Head of Department,Hyderabad Institute of Technology and Management,Medchal,Telangana UG Student,Hyderabad Institute of Technology and Management,Medchal,Telangana

## Abstract

**This project addresses the growing demand for high-quality visual content by developing an AI-driven image upscaling solution using diffusion models. The system enhances low-resolution images through a VRAM-efficient, tile-based processing approach that enables super-resolution on standard GPU hardware.**

**The project leverages Stable Diffusion x4 upscaling technology combined with intelligent tiling strategies to overcome memory constraints typically associated with processing large images. By dividing images into manageable tiles and applying diffusion-based enhancement techniques, the system achieves significant detail improvement while maintaining computational efficiency.**

**Keywords**: *AI-powered upscaling using state-of-the-art diffusion models, Tile-based processing for VRAM optimization, Compatibility with standard GPU configurations, Enhanced accessibility to super-resolution technology with minimal hardware requirements*

*The prototype demonstrates effective image quality enhancement, making advanced AI upscaling techniques more accessible to users without high-end computing resources.*

**DOI: 10.63856/ejd96466**

## 1. Introduction

The Tile-Based Image Upscaling using AI Diffusion Prototype is an innovative solution designed to address the increasing demand for high-quality visual content in digital media. This project implements an artificial intelligence-driven approach to enhance low-resolution images using advanced diffusion models. The system employs a tile-based processing strategy that optimizes VRAM usage, making sophisticated image upscaling accessible on standard GPU hardware. The project leverages Stable Diffusion x4 upscaling technology, which uses probabilistic diffusion processes to reconstruct high-resolution details from low-quality source images. By breaking down large images into smaller, manageable tiles, the system overcomes traditional memory limitations while maintaining visual coherence across tile boundaries. This approach democratizes access to professional-grade image enhancement tools, enabling users with modest hardware configurations to achieve results previously requiring high-end computing resources.

## Literature Survey

The introduction of Convolutional Neural Networks (CNNs) revolutionized image processing tasks. Dong et al. (2014) pioneered the application of deep learning to super-resolution with SRCNN (Super-Resolution Convolutional Neural Network), demonstrating that learned features significantly outperform hand-crafted methods. This work established that neural networks could learn the complex mapping between low and high-resolution image spaces.

Kim et al. (2016) advanced the field with VDSR (Very Deep Super-Resolution), implementing deeper architectures with residual learning. Their research showed that increasing network depth with proper training techniques leads to substantial quality improvements. Ledig et al. (2017) introduced SRGAN (Super-Resolution Generative Adversarial Network), which employed adversarial training to generate perceptually convincing high-resolution images, marking a shift toward generating realistic textures rather than merely minimizing pixel-wise errors.

**Problem Definition**

In many practical scenarios, users work with low-resolution images due to constraints like limited storage, slow internet speeds, or legacy hardware. Traditional image upscaling algorithms, such as bicubic and bilinear interpolation, often produce blurry results and fail to recover genuine details, especially when processing high-resolution outputs. Additionally, deep learning-based upscaling methods may require significant hardware resources, which are inaccessible to most users with standard GPUs.

The challenge is to develop a solution that enhances the quality and detail of low-resolution images using AI-driven diffusion techniques, while maintaining computational efficiency and accessibility. The proposed system must overcome VRAM limitations through tile-based processing, using pretrained diffusion models (like Stable Diffusion x4) to deliver high-quality, visually detailed images without the need for expensive hardware or lengthy processing times.

**Framework Overview**

**Ingestion**

The system accepts a low-resolution input image and analyses its size. For large images, it automatically segments them into smaller, overlapping tiles to ensure each tile fits within standard GPU memory limits, enabling efficient upscaling without heavy VRAM usage.

**Featurization**

Each image tile is preprocessed—resized and normalized—then passed through the encoder of a diffusion-based AI model. The encoder extracts high-level image features such as edges, textures, and semantic patterns, preparing the tiles for detailed enhancement during upscaling.

**C. Credibility Features**

The framework includes mechanisms to ensure output reliability, including tracking the specific pretrained diffusion model version used, logging processing parameters, and verifying that each tile has been consistently upscaled. This guarantees reproducibility and trustworthy results.

**D. Visual Features**

The system uses the diffusion AI model to reconstruct and enhance details in every tile, restoring sharpness and realistic textures. Overlapping regions are blended to form a seamless final image. The quality of upscaling is validated using objective image metrics like SSIM, PSNR, and LPIPS, ensuring visually appealing and accurate results, as seen in Fig 1.1 and 1.2

Fig 1.1

Fig 1.2

**E. Network Features**

The model architecture utilizes advanced neural network components—such as transformer blocks and attention mechanisms—to efficiently capture local and global image dependencies. These layers enhance the model's ability to reconstruct context-aware details and maintain coherence across tiles.

**F. Modelling**

The encoded features are enhanced using a pretrained diffusion model. Hyperparameters (like noise steps and tile overlap) are optimized to balance efficiency and visual quality, leveraging modern neural network design to recover lost details with high fidelity

**G. Decisioning**

Automated quality checks are performed using objective metrics (SSIM, PSNR, LPIPS) to evaluate upscaled tiles. Only outputs passing set thresholds are assembled, with failed tiles flagged for reprocessing.

**Dataset And Preprocessing**

A benchmark set of low-resolution images was curated from public image enhancement datasets and real-world samples. All images were assessed for variety in content, resolution, and typical degradation (such as blur and compression artifacts). Before upscaling, every image was standardized: (a) Resized to input resolution (b) Colour channels normalized.

A. Splits and Metrics

Upscaled outputs were evaluated using quantitative metrics: (a) SSIM (Structural Similarity Index): Assesses perceptual quality. (b) PSNR (Peak Signal-to-Noise Ratio): Measures signal fidelity. (c) LPIPS (Learned Perceptual Image Patch Similarity): Evaluates similarity using deep features, ensuring realism and preserving details.

**METHODS**

**A. Image Classifier**

A deep learning model is applied to the tiles to distinguish genuine image content from artifacts and noise, ensuring that upscaling is performed only on authentic regions

**B. Metadata Classifier**

Auxiliary metadata—including image origin, source device, and file attributes—is analysed using a lightweight classifier. This step supports credibility checks and model selection for optimal enhancement pathways.

**C. Visual Classifier**

The system performs post-upscaling visual inspections.Automated checks compare the final output against gold-standard references, quantifying visual improvements to maintain the fidelity of details.

**D. Graph Outlier Detection**

| Metric | Purpose | Value |
|---|---|---|
| SSIM | Perceptual similarity assessment | 0.82–0.98 |
| PSNR | Signal fidelity measurement | 24–32 dB |
| LPIPS | Deep feature-based quality benchmark | 0.09–0.22 |

**Table 1.1**

### E. Fusion and Calibration

A logistic fusion layer combines normalized subsystem outputs. We apply temperature scaling and isotonic regression on the validation set.

## Experiments

### A. Baselines

The proposed method is benchmarked against traditional and advanced upscaling techniques to assess performance improvements.

| Method | Description |
|---|---|
| Bicubic Interpolation | Classic interpolation |
| SRGAN | Super-Resolution GAN |
| ESRGAN | Enhanced SRGAN |
| Stable Diffusion (Ours) | Tile-based upscaling |

**Table 2.1**

### B. Ablations

To understand the impact of different components, several ablation studies were performed:

- Comparison with and without tile-based processing to assess VRAM efficiency and continuity.
- Experiments disabling visual verification and graph outlier detection to evaluate their contribution to visual quality and artifact reduction.
- Adjustment of overlap size and diffusion steps to study effects on seamlessness and detail recovery.

### C. Implementation Details

- All experiments conducted on a workstation with an NVIDIA RTX 4060 GPU (8GB VRAM) with Google T4 GPU
- Models implemented in Python using PyTorch and Hugging Face Diffusers libraries.
- Tile overlap set to 32 pixels for blending, and upscaling factor fixed at 4×.
- Evaluation performed on a set of public benchmark datasets, results averaged across splits for objectivity.

## Interpretability And Explanations

Feature visualizations show what the model focuses on during upscaling.Saliency maps highlight important regions for enhancement.Metrics (SSIM, PSNR, LPIPS) diagnose tile quality.Outlier detection explains flagged tiles, supporting reliable results.

## Ethical Considerations

Ensuring all image data is sourced from public, non-restricted datasets to respect privacy and copyright.The upscaling system avoids altering content or adding misleading details, focusing only on clarity and enhancement.Transparent reporting of model limitations and potential biases is maintained throughout the project.Results are used strictly for legitimate purposes such as restoration, research, and accessibility.

## Limitations

The approach may produce minor visible seams in very large images, despite tile blending.Upscaling quality depends on the diversity and quality of training data; results may vary for unusual image types.Hardware constraints limit real-time processing for extremely high-resolution inputs.Model outputs may occasionally favour texture realism over exact reconstruction of original content.

**Deployment Blueprint**

The solution can be packaged as a Python-based command-line tool or REST API for integration into imaging workflows.Requires a workstation with an NVIDIA GPU (minimum 8GB VRAM) and Python (preferably 3.8+).Dependencies include PyTorch, Hugging Face Diffusers, and supporting image-processing libraries.Batch processing and automated logging are supported for scalable deployment in production environments.Documentation and configuration files guide users through installation, GPU setup, and pipeline customization.

**Conclusion**

The proposed tile-based AI upscaling framework efficiently enhances image detail while remaining accessible on standard hardware. By combining advanced diffusion models, automated tile handling, and transparent quality checks, the system consistently produces high-fidelity results. This approach supports practical applications in restoration, research, and digital archiving, demonstrating reliable performance and scalable deployment.

**Acknowledgments**

I gratefully acknowledge the mentorship of Dr. Padmaja Pulicherla, whose expert advice and encouragement were instrumental in shaping this project. I appreciate the guidance and feedback from the Computer Science and Engineering faculty and my peers at Hyderabad Institute of Technology and Management. Thanks to the open-source developers and researchers for making advanced AI tools and datasets publicly available, enabling meaningful innovation and learning throughout this work.

**Reproducibility Checklist**

All code and data sources are documented and accessible.Hyperparameters and evaluation metrics are clearly specified.Instructions provided for repeating experiments on standard hardware.

**Appendix B Policy Configuration**

Tile size: 512×512 pixels; overlap: 32 pixels.Minimum SSIM: 0.80 for tile acceptance.Default model: Stable Diffusion x4.Output format: PNG (default), JPEG optional.

**References**

1. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., &Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 10684-10695.
2. Saharia, C., Ho, J., Chan, W., Salimans, T., Fleet, D. J., & Norouzi, M. (2022). Image super-resolution via iterative refinement. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(4), 4713-4726.
3. Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., & Loy, C. C. (2018). ESRGAN: Enhanced super-resolution generative adversarial networks. Proceedings of the European Conference on Computer Vision (ECCV) Workshops.
4. NVIDIA Corporation. (2023). NVIDIA DLSS (Deep Learning Super Sampling). https://www.nvidia.com/en-us/geforce/technologies/dlss/
5. Intel Corporation. (2023). Intel XeSS (Xe Super Sampling). https://www.intel.com/content/www/us/en/products/docs/discrete-gpus/arc/technology/xess.html
6. AMD Technologies. (2023). AMD FidelityFX Super Resolution(FSR). https://www.amd.com/en/technologies/fidelityfx-super-resolution
7. Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric. Proceedings of the IEEE Conference on Computer Vision and Pattern Recogniti on, 586-595.