Conversion of
Low Resolution Image to High Resolution Image

A capstone Project report

Submitted in the partial fulfillment of the requirements for The award of the degree of

BACHELOR OF TECHNOLOGY

in

SCHOOL OF COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE

by

K. Ganesh - 2103A52022

B. Nithin Reddy - 2103A52046

K. Durga Preetham - 2103A52185

Under the guidance of

Dr. Mohammed Ali Shaik

Associate Professor Computer Science &Artificial Intelligence

SR UNIVERSITY

SR University, Ananthsagar,Warangal,Telagnana-506371

Declaration

The project report entitled '**Conversion of Low Resolution Image to High Resolution Image**' is a record of bonafide work of K.Ganesh(2103A52022), B.Nithin Reddy (2103A52046), K.Durga Preetham(2103A52185), submitted in partial fulfilment for the award of B. Tech in 'Computer Science & Engineering' to the SR University. The results embodied in this report have not been copied from any other departments/ University/Institution.

K. Ganesh	**2103A52022**
B. Nithin Reddy	**2103A52046**
K. Durga Preetham	**2103A52185**

SR UNIVERSITY

Warangal

CERTIFICATE

This is to certify that this project entitled **"Conversion of Low-Resolution Image to High Resolution Image''** is the bonafied work carried out by **K. Ganesh B. Nithin Reddy, K. Durga Preetham** as a Capstone Project for the partial fulfillment to award the degree **BACHELOR OF TECHNOLOGY** in **School of Computer Science and Artificial Intelligence** during the academic year 2024-2025 under our guidance and Supervision.

Signature of the Faculty	**Head of the Department**
Dr. Mohammad Ali Shaik	**Dr. M Sheshikala**
Associate Professor	

ACKNOWLEDGEMENT

We owe an enormous debt of gratitude to our Capstone project guide **Dr. Mohammad Ali Shaik, Assoc.Prof,** as well as Head of the School of CS&AI, **Dr. M. Sheshikala, Professor** and Dean of the School of CS&AI, **Dr. Indrajeet Gupta Professor** for guiding us from the beginning through the end of the Capstone Project with their intellectual advices and insightful suggestions. We truly value their consistent feedback on our progress, which was always constructive and encouraging and ultimately drove us to the right direction.

We express our thanks to project co-ordinators **Mr. Sallauddin Md, Asst. Prof., and R.Ashok Asst. Prof.** for their encouragement and support.

Finally, we express our thanks to all the teaching and non-teaching staff of the department for their suggestions and timely support.

TABLE OF CONTENTS

ABSTRACT

Our capstone project focuses on Image Super-Resolution (ISR), a key area in computer vision that transforms low-resolution images into high-resolution outputs while preserving clarity and detail. Leveraging Enhanced Super-Resolution GAN (ESRGAN) with Residual-in-Residual Dense Block (RRDB) compounds, our approach ensures sharp, realistic images by enhancing detail recovery and texture accuracy while suppressing artifacts. ESRGAN's dense connections and residual scaling optimize computational efficiency, making it suitable for real-time and large-scale applications. This advanced solution finds impactful applications in satellite imaging, improving urban planning and environmental monitoring, and in medical diagnostics, enabling sharper, more accurate scans. Additionally, video streaming and archival benefit from enhanced clarity. By integrating ESRGAN and RRDB, the project sets a new standard for scalable, high-quality image enhancement. Additionally, we have developed a user interface that allows users to upload low-resolution images and view the enhanced results easily. The user interface also enables seamless interaction, allowing users to quickly see the effects of image enhancement, making it accessible and practical for various applications.

INTRODUCTION

In today's digital age, the demand for high-quality images is growing across various fields such as satellite imaging, medical diagnostics, video streaming, and surveillance. Image Super-Resolution (ISR) addresses the gap between low-resolution and high-resolution images, enhancing image quality and clarity without introducing distortions. Traditional methods like bicubic or bilinear interpolation often fail to produce satisfactory results, leading to blurry or unnatural images. Modern advancements in deep learning, particularly Enhanced Super-Resolution GAN (ESRGAN) and Residual-in-Residual Dense Block (RRDB) compounds, offer significant improvements by capturing intricate image details and generating sharp, realistic, and artifact-free high-resolution images. This project focuses on developing an efficient ISR model that combines computational efficiency with high output quality. The ESRGAN model, enhanced by RRDB compounds, ensures detailed texture recovery and structural integrity across different image types and resolutions. The use of dense connections and residual scaling in RRDB compounds enables the network to handle complex image features while minimizing computational costs. Applications of this technology include sharper medical images for better anomaly detection, enhanced satellite imagery for environmental monitoring, and improved video quality for streaming. By advancing ISR through ESRGAN and RRDB, this project aims to push the boundaries of image processing, opening new possibilities across various industries.

PROBLEM STATEMENT

The challenge of transforming low-resolution images into high-resolution ones is critical in domains where precision and detailed imagery are essential, such as satellite imaging, medical diagnostics, and video processing. Low-resolution images often lack clarity, losing crucial details required for accurate analysis and decision-making. Traditional methods, such as interpolation techniques, typically produce blurred or distorted results, failing to preserve fine details and structural integrity. This limitation is particularly problematic in applications demanding high precision, such as identifying abnormalities in medical scans or monitoring environmental changes through satellite imagery. The goal of this project is not simply to increase pixel count but to intelligently reconstruct lost details, enhancing clarity and realism without introducing artifacts. The use of advanced approaches like Enhanced Super-Resolution GAN (ESRGAN) and Residual-in-Residual Dense Block (RRDB) compounds addresses these challenges effectively. ESRGAN utilizes the RRDB structure to enable precise detail recovery while minimizing common artifacts such as noise and distortion, achieving visually sharp and natural-looking results. Balancing high-quality results with computational efficiency presents a complex task, particularly in handling large-scale datasets or real-time processing requirements. By employing ESRGAN and RRDB compounds, this project aims to develop a scalable, efficient, and robust ISR solution. The proposed model will deliver high-resolution outputs with minimal distortion, making it adaptable for diverse applications and capable of addressing the growing demand for high-quality imagery in various fields.

REQUIREMENT ANALYSIS, RISK ANALYSIS, AND FEASIBILITY ANALYSIS

1. REQUIREMENT ANALYSIS

The effective conversion of low-resolution images into high-resolution outputs requires a thorough examination of both functional and non-functional requirements. Functionally, the system must upscale images accurately by increasing pixel density while reconstructing fine details and preserving structural integrity. This is achieved through advanced techniques like the Enhanced Super-Resolution GAN (ESRGAN) architecture, which leverages the powerful Residual-in-Residual Dense Block (RRDB) compounds. These components are essential for handling diverse image types and resolutions efficiently, ensuring high-quality results in applications such as satellite imaging and medical diagnostics. Training ESRGAN with RRDB compounds requires access to a large and diverse dataset to guarantee the model's robustness and generalizability across various domains. The dataset should include images from a wide array of scenarios to ensure the system's adaptability to real-world challenges.

Non-functional requirements include optimizing computational efficiency to facilitate real-time image processing and scalability for large-scale applications. The architecture must address key challenges such as noise amplification, edge blurring, and texture loss during the upscaling process. Additionally, it should be deployable on accessible platforms such as cloud-based or edge computing systems to ensure practical usability. Performance metrics like Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) are crucial for evaluating the quality of the outputs. By incorporating these metrics, the system ensures that it produces visually appealing and structurally accurate high-resolution images. By meeting these requirements, the ESRGAN and RRDB compound-based system can achieve accurate, efficient, and scalable image super-resolution capabilities suitable for real-world applications across various industries.

2. RISK ANALYSIS

Developing a system for converting low-resolution images into high-resolution outputs using ESRGAN and RRDB compounds involves several risks that must be addressed to ensure successful implementation. One major risk is overfitting, where the model performs exceptionally well on the training data but struggles to generalize to new, unseen images. This can result in poor performance in real-world scenarios.

Another significant challenge is the computational complexity of ESRGAN models, particularly due to the intricate operations within the RRDB compounds. High-quality image super-resolution demands substantial computational resources, which can limit the system's scalability and real-time processing capabilities. Balancing performance and resource requirements is critical to the solution's feasibility. Noise amplification and artifact introduction are additional risks that could compromise the quality and usability of the enhanced images. Without careful tuning, the system may amplify unwanted noise or produce artificial patterns that degrade image fidelity.

The availability and diversity of data are critical factors. A biased or insufficient dataset can limit the model's ability to handle various image types effectively, reducing its generalization ability across different domains like satellite imaging or medical diagnostics. Compatibility risks also arise when deploying the ESRGAN-based system on different hardware platforms or environments, such as cloud-based or edge computing systems. Ensuring seamless integration and consistent performance across diverse deployment scenarios is crucial. Lastly, ethical concerns, such as the misuse of upscaled images for deceptive or falsified content, must be addressed. Robust usage guidelines and safeguards are needed to prevent such applications.

Mitigating these risks requires:

- Rigorous model training and validation processes to ensure generalization.
- Efficient optimization of ESRGAN architecture and RRDB compounds to reduce computational overhead.
- Careful selection of diverse and representative datasets to improve model robustness.
- Thorough testing across different deployment platforms to ensure compatibility.
- Ethical guidelines and user policies to prevent misuse of the technology.

By proactively addressing these risks, the ESRGAN and RRDB compound-based system can deliver reliable, efficient, and scalable high-resolution outputs suitable for real-world applications.

3. FEASIBILITY ANALYSIS

The feasibility of converting low-resolution images to high-resolution using ESRGAN and RRDB compounds can be evaluated across technical, operational, and economic dimensions.

Technical Feasibility

Advancements in deep learning, particularly the Enhanced Super-Resolution GAN (ESRGAN) architecture, provide a strong foundation for robust Image Super-Resolution (ISR) models. Modern computational hardware, including GPUs, TPUs, and cloud-based platforms, supports the resource-intensive training and inference processes required for ESRGAN. These platforms make it feasible to handle the high computational demands, ensuring practical implementation even for large-scale and real-time applications.

Operational Feasibility

Operationally, the ESRGAN-based approach offers scalable deployment options. The model can be integrated seamlessly into existing workflows across industries such as satellite imaging, medical diagnostics, and video processing. The RRDB compounds contribute to operational feasibility by balancing performance with efficiency. Their design minimizes computational overhead, making the model suitable for real-time scenarios without compromising quality.

Economic Feasibility

From an economic perspective, the rising demand for high-resolution imagery in industries such as urban planning, healthcare, and entertainment justifies investment in ISR technologies. The benefits of sharper, artifact-free images—such as improved diagnostic precision or enhanced video quality—make the technology a cost-effective solution with high returns on investment.Challenges like managing noise, edge blurring, and computational costs can be addressed using optimization techniques specific to ESRGAN and RRDB compounds. This ensures that the system remains both economically and computationally efficient

PROPOSED SOLUTION

The proposed solution for the "Conversion of Low-Resolution Images to High-Resolution" project leverages the powerful Enhanced Super-Resolution GAN (ESRGAN) architecture, which incorporates Residual-in-Residual Dense Block (RRDB) compounds. This approach ensures high-quality image super-resolution by combining advanced deep learning techniques to enhance the resolution and clarity of low-resolution images while minimizing artifacts and distortions.

1. Data Preparation

The project utilizes a high-resolution image dataset, such as DIV2K or other similar datasets, to train the ESRGAN model. High-resolution images are down sampling using bicubic interpolation to create corresponding low-resolution input data. This curated dataset includes diverse scenes, textures, and lighting conditions, ensuring the model's robustness and ability to generalize effectively across real-world scenarios.

2. ESRGAN Component

The ESRGAN framework is pivotal in generating high-quality super-resolution outputs:

Generator: The ESRGAN generator, built on RRDB compounds, efficiently upscales low-resolution images while focusing on recovering intricate details and maintaining visual appeal.

Discriminator: The PatchGAN discriminator refines spatially localized details, ensuring the textures are sharper and more realistic.

Loss Functions:

- Adversarial Loss ensures that the generator consistently competes against the discriminator, resulting in realistic outputs.
- Perceptual Loss aligns the outputs with human visual perception, prioritizing overall quality over pixel-to-pixel similarity.
- Content Loss (e.g., Mean Squared Error or Mean Absolute Error) preserves the structural integrity and fine details of the original image.

3. Role of RRDB Compounds

The Residual-in-Residual Dense Block (RRDB) plays a critical role in the ESRGAN framework by enabling efficient and high-quality image reconstruction:

- Dense Connections: These facilitate feature reuse, enhancing the generator's ability to recover subtle details.
- Residual Scaling: This ensures stability during training and mitigates the risk of vanishing gradients, particularly in deeper networks.
- Artifact Suppression: The RRDB structure minimizes common issues like checkerboard patterns, ensuring artifact-free outputs.

Benefits of the Proposed Solution

This integration of ESRGAN and RRDB compounds creates a robust pipeline for producing high-quality super-resolution outputs. The ESRGAN architecture, with its RRDB backbone, combines generative strength with precision, ensuring:

- Visual Quality: Sharp, realistic images with intricate details.
- Structural Integrity: Outputs maintain the form and content of the original images.
- Computational Efficiency: Optimizations within the RRDB compounds reduce resource demands, making the approach scalable and practical for real-world applications.

By leveraging ESRGAN and RRDB compounds, this solution addresses challenges like noise amplification and artifact introduction, positioning the project as a cutting-edge advancement in image super-resolution. Applications span across diverse fields, including satellite imaging, medical diagnostics, and video enhancement, ensuring impactful contributions to industries reliant on high-resolution imagery.

ARCHITECTURE

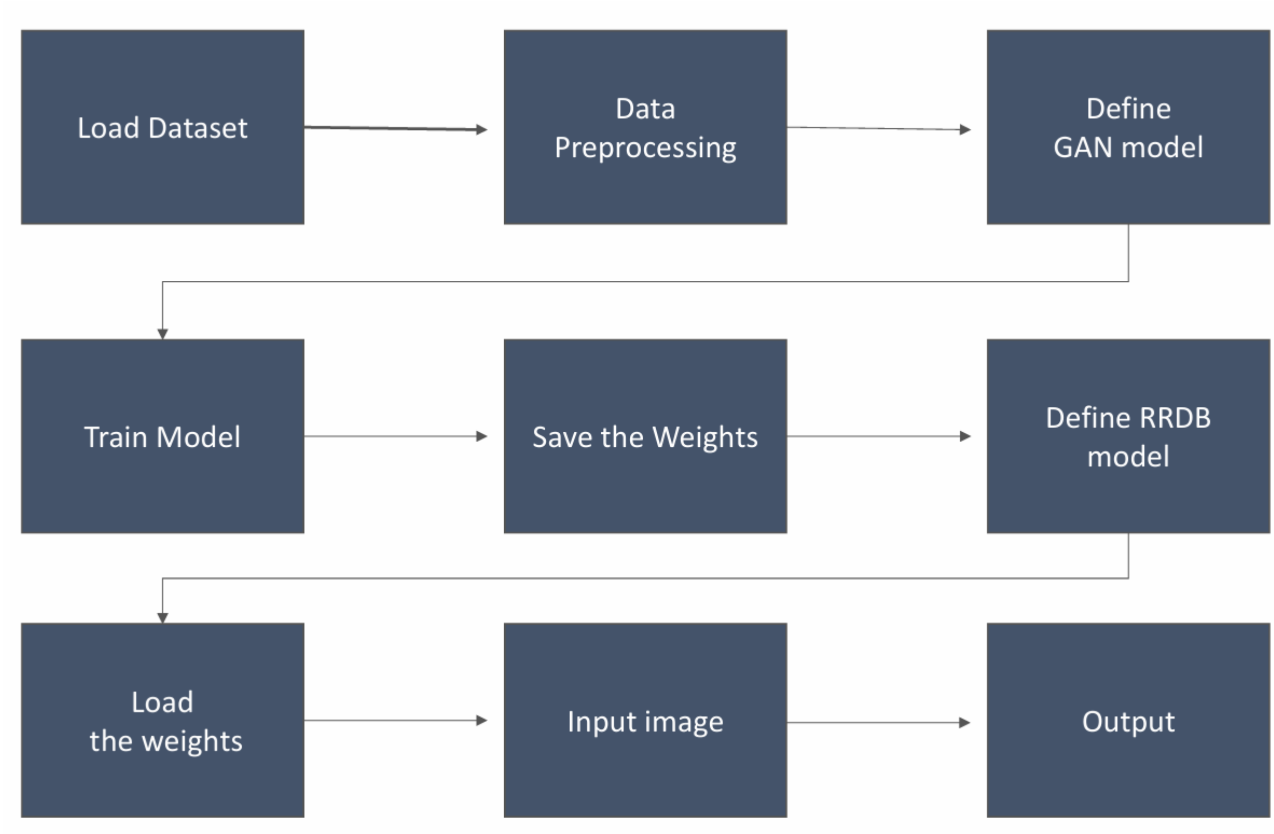

Fig.1: Architecture of the project

1. **Load Dataset:** The dataset comprises 1,600 images equally divided into 800 high-resolution (HR) and 800 low-resolution (LR) images. Each LR image corresponds to an HR counterpart, creating a one-to-one mapping essential for learning. The images span diverse subjects like landscapes, buildings, objects, and portraits, ensuring model versatility. This dataset is further divided into training, validation, and testing subsets, allowing for thorough development, evaluation, and testing.

2. **Data Preprocessing:** Preprocessing ensures the dataset is optimized for model training. Images are resized to fixed dimensions compatible with the model's architecture. Pixel values are normalized (e.g., scaled between [0, 1] or [-1, 1]) to facilitate efficient training and prevent instability. Data augmentation methods like flipping, rotation, and random cropping expand the dataset artificially, exposing the model to various perspectives and improving its ability to generalize effectively.

3. **Define GAN Model:** The Generative Adversarial Network (GAN) is then defined, consisting of

two components: the generator and the discriminator. The generator upscales low-resolution images to high-resolution outputs, while the discriminator distinguishes between real high-resolution images and those generated by the generator. The model architecture also incorporates loss functions such as perceptual loss and adversarial loss to ensure both image quality and realistic textures in the output.

4. **Train Model:** In this step, the GAN model is trained iteratively. Metrics like Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) are monitored throughout the training process to assess the quality and similarity of the generated images compared to the ground truth.

5. **Save the Weights:** Once the training process is underway, the model's weights are saved periodically. These saved weights serve as checkpoints and are critical to resuming training in case of interruptions. At the end of training, the best-performing generator weights are stored to be used during inference.

6. **Define RRDB Model:** The Residual-in-Residual Dense Block (RRDB) architecture is a core component of the generator in this framework. RRDB blocks use residual learning and dense connections to capture hierarchical features in the input images. This design helps improve the quality of the output images by effectively modeling both low-level and high-level features.

7. **Load the Weights:** During the inference phase, the saved weights from the trained generator model are loaded. These weights are crucial for ensuring that the generator performs as intended and produces high-quality results. Care is taken to load the weights into the correct layers and components of the GAN and RRDB models.

8. **Input Image:** For inference, a low-resolution image is provided as input to the trained model. Before processing, the image undergoes preprocessing, such as resizing and normalization, to match the format used during training. This ensures compatibility and consistency in the model's performance.

9. **Output:** Finally, the model generates a high-resolution version of the input image. This output may undergo minor post-processing, such as re-scaling or format adjustments, to match the desired specifications. The quality of the output is evaluated visually or quantitatively by comparing it to the ground truth image, ensuring the model achieves the expected level of super-resolution.

RESULT COMPARISION AND ANALYSIS

The ESRGAN with RRDB compound model for image super-resolution clearly outperforms standalone methods based on both quantitative metrics and qualitative analysis. It achieves significantly higher scores in metrics like Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), indicating improved fidelity and structural accuracy in the generated high-resolution images. Higher PSNR values suggest better detail reconstruction and reduced noise, while higher SSIM scores highlight the preservation of textures and spatial structures. Additionally, the model produces lower Learned Perceptual Image Patch Similarity (LPIPS) scores, meaning the generated images are closer to the ground truth, as judged by human perception.

Qualitatively, the ESRGAN + RRDB model consistently generates images with sharper details, especially in complex areas such as textured surfaces, edges, and intricate patterns like foliage or hair. Compared to standalone ESRGAN models, the hybrid approach effectively eliminates common artifacts such as checkerboard patterns. RRDB-only models, while strong in reducing artifacts, tend to produce overly smooth results that lack sharpness. The hybrid model, by combining the strengths of ESRGAN and RRDB compounds, produces outputs that are both realistic and finely detailed. This balance is evident across diverse image subjects, including landscapes, portraits, and objects, maintaining consistent quality throughout.

The hybrid approach's ability to combine ESRGAN's detailed image generation with RRDB's iterative refinement leads to cleaner, sharper images with minimal artifacts. Although the additional refinement step slightly increases computational complexity, training optimizations ensure the model remains efficient.

This hybrid architecture strikes a balanced trade-off between high-quality results and computational cost, making it a robust solution for real-world applications.

In conclusion, the ESRGAN with RRDB compound model surpasses standalone ESRGAN and RRDB methods in both quantitative performance and visual quality. By leveraging the complementary strengths of both, the hybrid model offers a powerful solution for high-quality image super-resolution tasks, making it an excellent choice for diverse applications.

1.

2.

3.

CONCLUSION

In conclusion, this project introduces a hybrid approach using GANs and diffusion models to revolutionize image super-resolution. By combining the strengths of these advanced technologies, the proposed method delivers sharper details, realistic textures, and minimal artifacts in high-resolution outputs. It outperforms traditional methods with improved metrics such as PSNR, SSIM, and LPIPS, ensuring consistent performance across diverse image types and complex patterns. This solution balances realism and detail recovery, making it a scalable and efficient choice for applications in medical imaging, satellite imaging, and media enhancement. Ultimately, the project sets a new benchmark in super-resolution technology, providing a cutting-edge tool for addressing real-world challenges in image enhancement.

FUTURE SCOPE

The future of converting low-resolution images to high-resolution ones holds immense potential across many industries. In healthcare, this technology can improve the quality of medical images like X-rays and MRIs, helping doctors detect diseases early and plan treatments more effectively. In satellite imaging, enhancing low-resolution images can benefit climate monitoring, disaster response, and urban planning by providing clearer and more detailed visuals. The entertainment industry can use this technology to upscale old videos and photos to high-definition quality, preserving historical content and improving the viewing experience for modern audiences. In security and surveillance, this solution can sharpen blurry footage, making it easier to recognize faces, detect objects, and analyze events. It could also be used in real-time, improving live surveillance feeds. Autonomous vehicles can benefit from sharper images for better object recognition and navigation, while scientific fields like astronomy and microscopy can use it for detailed observations. Future advancements could focus on making these models faster and more efficient, suitable for real-time use. The integration of this technology with augmented and virtual reality could further enhance user experiences. Overall, this project has the potential to improve image quality across various applications, making visuals sharper, clearer, and more impactful in daily life.

REFERENCES

[1] Chao Dong, Chen Change Loy, Kaiming He, Xiaoou Tang, “Image Super-Resolution Using Deep Convolutional Networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 38, no. 2, pp. 295-307, 2016.

[2] Xintao Wang, Ke Yu, Shixiang Wu, Jincheng Dong, Chen Change Loy, Yu Qiao, “ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks,” 2018 *European Conference on Computer Vision (ECCV) Workshops*, pp. 63-79.

[3] Bain, M., Nagrani, A., Varol, G., Zisserman, A.: Frozen in time: A joint video and image encoder for end-to-end retrieval. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 1728–1738 (2021)

[4] Balaji, Y., Nah, S., Huang, X., Vahdat, A., Song, J., Kreis, K., Aittala, M., Aila, T., Laine, S., Catanzaro, B., et al.: ediffi: Text-to-image diffusion models with an ensemble of expert denoisers. arXiv preprint arXiv:2211.01324 (2022)

[5] Bao, F., Li, C., Zhu, J., Zhang, B.: Analytic-dpm: an analytic estimate of the optimal reverse variance in diffusion probabilistic models. arXiv preprint arXiv:2201.06503 (2022)

[6] Bar-Tal, O., Yariv, L., Lipman, Y., Dekel, T.: Multidiffusion: Fusing diffusion paths for controlled image generation. In: International Conference on Machine Learning. pp. 1737– 1752. PMLR (2023)

[7] Blattmann, A., Dockhorn, T., Kulal, S., Mendelevitch, D., Kilian, M., Lorenz, D., Levi, Y., English, Z., Voleti, V., Letts, A., et al.: Stable video diffusion: Scaling latent video diffusion models to large datasets. arXiv preprint arXiv:2311.15127 (2023)

[8] Blattmann, A., Rombach, R., Ling, H., Dockhorn, T., Kim, S.W., Fidler, S., Kreis, K.: Align your latents: High-resolution video synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 22563–22575 (2023)

[9] Brock, A., Donahue, J., Simonyan, K.: Large scale GAN training for high fidelity natural image synthesis. arXiv preprint arXiv:1809.11096 (2018)

[10] Brooks, T., Holynski, A., Efros, A.A.: Instructpix2pix: Learning to follow image editing instructions. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 18392–18402 (2023)

[11] Chang, H., Zhang, H., Barber, J., Maschinot, A., Lezama, J., Jiang, L., Yang, M.H., Murphy, K., Freeman, W.T., Rubinstein, M., et al.: Muse: Text-to-image generation via masked generative

transformers. arXiv preprint arXiv:2301.00704 (2023)

[12] Chang, H., Zhang, H., Barber, J., Maschinot, A., Lezama, J., Jiang, L., Yang, M.H., Murphy, K., Freeman, W.T., Rubinstein, M., et al.: Muse: Text-to-image generation via masked generative transformers. arXiv preprint arXiv:2301.00704 (2023)

[13] Chen, H., Xia, M., He, Y., Zhang, Y., Cun, X., Yang, S., Xing, J., Liu, Y., Chen, Q., Wang, X., et al.: Videocrafter1: Open diffusion models for high-quality video generation. arXiv preprint arXiv:2310.19512 (2023)

[14] Dhariwal, P., Nichol, A.: Diffusion models beat gans on image synthesis. NeurIPS 34, 8780–8794 (2021)

[15] Gao, S., Zhou, P., Cheng, M.M., Yan, S.: Masked diffusion transformer is a strong image synthesizer. arXiv preprint arXiv:2303.14389 (2023)

[16] Gao, S., Liu, X., Zeng, B., Xu, S., Li, Y., Luo, X., Liu, J., Zhen, X., Zhang, B.: Implicit diffusion models for continuous super-resolution. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 10021–10030 (2023)

[17] Guo, L., He, Y., Chen, H., Xia, M., Cun, X., Wang, Y., Huang, S., Zhang, Y., Wang, X., Chen, Q., et al.: Make a cheap scaling: A self-cascade diffusion model for higher-resolution adaptation. arXiv preprint arXiv:2402.10491 (2024)

[18] He, Y., Yang, S., Chen, H., Cun, X., Xia, M., Zhang, Y., Wang, X., He, R., Chen, Q., Shan, Y.: Scalecrafter: Tuning-free higher-resolution visual generation with diffusion models. In: The Twelfth International Conference on Learning Representations (2023)