Composite Inpainting

Addressing Shortcomings in GAN and Patch-Based Inpainting with a Hybrid Approach

Kai Zhu, Maria Alejandra Escalante
University of Toronto

Motivation

Deep learning-based inpainting models such as generative adversarial networks (GANs) and stable diffusion offer state of the art generation of realistic content in complex scenes using contextual surroundings. There exist, however, drawbacks for each method: GANs are difficult to train and often fail to capture textures effectively [1]; Stable diffusion is prone to hallucinations [2].

Left: A partial hand hallucinated by stable diffusion
Right: Distorted and blurry fence texture created by DeepFill GAN

Classical **patch-based methods** reportedly excel at replicating textural details such as architectural features and repeating patterns, but encounter difficulty when generating larger structurally consistent content [1][3]. These methods also have the advantage of operating on a single image without requiring lengthy training time and data sets.

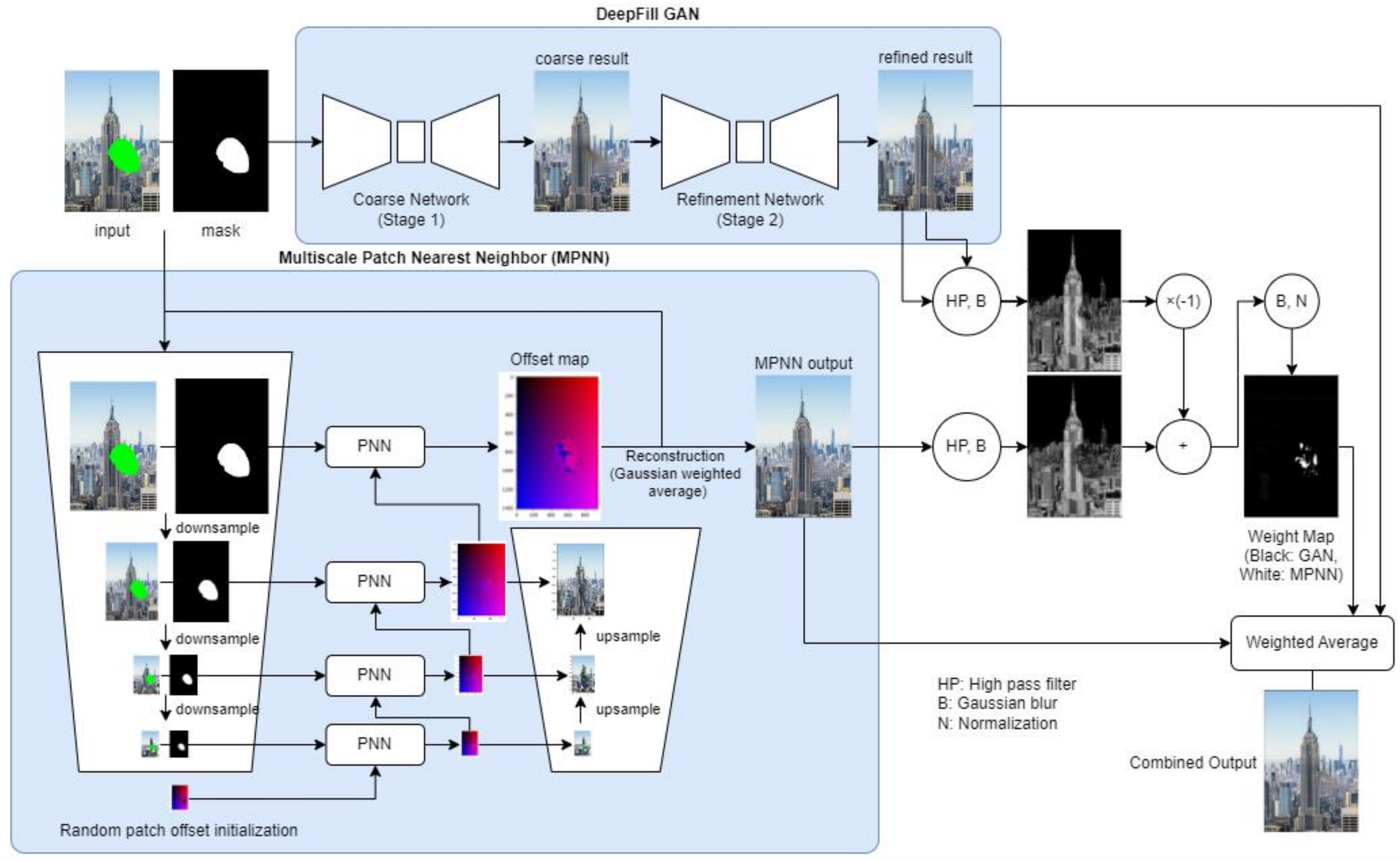
We propose a **hybrid approach** using GAN and patch-based methods in hopes of leveraging their strengths and minimizing their shortcomings to improve both structural and textural consistency.

Related Work

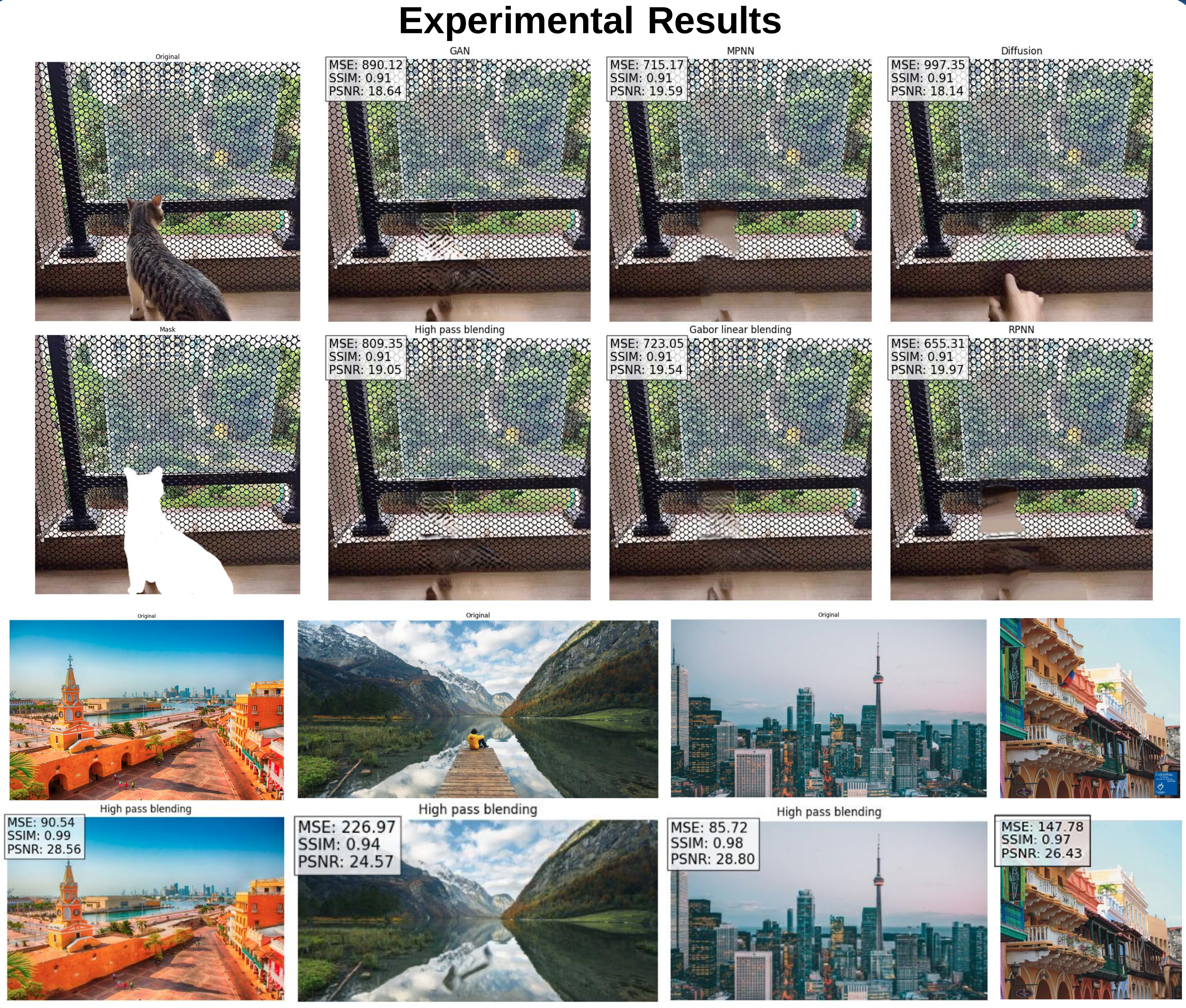
- **Free-Form Image Inpainting with Gated Convolution** presents a novel system using gated convolutions and SN-PatchGAN, enhancing inpainting quality and color consistency over prior methods [4].
- **Generative Patch Nearest-Neighbor (GPNN)** is an efficient, high quality patch-based single-image generation method. GPNN adopts SinGAN's multi-scale architecture while replacing the generator and discriminator with patch nearest-neighbor modules [3].
- **Latent diffusion models (LDM)** introduced by Rombach et al. enhances efficiency and quality in high-resolution image synthesis and inpainting [5]. The runwayml/stable-diffusion-inpainting model used in our comparison is based on this model.
- **Gabor filters** are recognized as a prominent method in texture classification applicable in textural analysis for inpainting. Bianconi et al [6] examine the impact of various Gabor filter parameters on texture discrimination. Their research indicates that while increasing frequencies and orientations has limited effect, smoothing parameters significantly enhance classification performance.

Methods

- We implemented a multi-scale patch nearest neighbor (MPNN) method based on the architecture of GPNN[3] using the patchmatch algorithm.
  - Low scale → generate structure
  - High scale → generate texture
- Outputs of GAN and MPNN were composited using high-pass filter weighed blending to maximize texture density.
- Gabor filtering and refined patch nearest neighbor (RPNN, using coarse GAN output as PNN input) were also explored as alternative compositing methods.



Experimental Results



Mean Statistics for Single and Hybrid Methods (* denotes hybrid methods)

	GAN	MPNN	Diffusion	*High pass	*Gabor	*RPNN
MSE	890.1	715.2	997.3	809.3	723.1	655.3
SSIM	0.908	0.912	0.907	0.909	0.911	0.914
PSNR	18.6	19.6	18.1	19.0	19.5	20.0

References

[1] Omar Elharrouss et al. "Image Inpainting: A Review". In: *Neural Processing Letters* 51.2 (Dec. 2019), pp. 2007–2028. doi: 10.1007/s11063-019-10163-0.
[2] M. El Helou and S. Süsstrunk, "BIGPrior: Toward Decoupling Learned Prior Hallucination and Data Fidelity in Image Restoration," in *IEEE Transactions on Image Processing*, vol. 31, pp. 1628-1640, 2022, doi: 10.1109/TIP.2022.3143006.
[3] Granot, N., Feinstein, B., Shocher, A., Bagon, S., & Irani, M. (2022). Drop the GAN: In defense of patches nearest neighbors as single image generative models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 13460-13469).
[4] Yu, Jiahui, et al. "Free-form image inpainting with gated convolution." Proceedings of the IEEE/CVF international conference on computer vision. 2019.
[5] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-Resolution Image Synthesis With Latent Diffusion Models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 10684-10695).
[6] Bianconi, Francesco, and Antonio Fernández. "Evaluation of the effects of Gabor filter parameters on texture classification." *Pattern recognition* 40.12 (2007): 3325-3335.