

PENGEMBANGAN ALGORITMA *DENOISING* DENGAN KONSEP *DEEP BACK-PROJECTION*

IMPROVEMENT OF *DENOISING* ALGORITHM BASED ON *DEEP BACK-PROJECTION*

Muhammad Haris¹, Eni Heni Hermaliani², Sri Rahayu³, Dwiza Riana⁴

^{1,2,3,4} Universitas Nusa Mandiri

¹muhammad.uhs@nusamandiri.ac.id, ²enie_h@nusamandiri.ac.id,
³sriarahayu.rry@nusamandiri.ac.id, ⁴dwiza@nusamandiri.ac.id

Abstrak

Penelitian ini mengangkat salah satu permasalahan umum pada citra digital, terutama pada *smartphone*, yaitu *noise* yang disebabkan oleh kecilnya *aperture* dan ukuran *sensor* pada kamera digital. Untuk menyelesaikan permasalahan tersebut, peneliti mengembangkan metode *denoising* untuk mentransformasi citra kotor menjadi citra bersih. Terinspirasi dari konsep *deep back-projection*, penelitian ini melakukan modifikasi terhadap implementasi asli *deep back-projection*, menggunakan *down-projection layer* sebagai tahapan untuk menghilangkan *noise* dengan intuisi menggunakan bahwa proses *down-sampling* mampu menghilangkan *noise* pada citra masukan. Lalu citra kembali diperbesar sesuai ukuran aslinya. Pada eksperimen yang dilakukan, penelitian ini membuktikan bahwa teknik yang diajukan mampu memperoleh hasil terbaik dibandingkan metode lainnya pada *dataset* SIDD. Hasil eksperimen memperlihatkan peningkatan kualitas sebesar 6 dB dibandingkan model lainnya.

Kata kunci: *denoising, back-projection, citra kotor, citra bersih*

Abstract

This research discusses one of the general problems in digital image, especially the one produced by smartphones, which is produced by small aperture and sensor on the smartphone. To solve this problem, the researcher develops denoising methods that can transform noisy images to clean images. Inspired by deep back-projection networks, this research modifies the original implementation of deep back-projection. We use the down-projection layer as a stage that can intuitively remove the noise by down-sampling procedure, then bring back the image to original size using an up-sampling layer. In the experimental result, we show that the proposal is superior compared to other methods on SIDD dataset. The proposal has better quality, shown by the increase of 6 dB compared to other models.

Keywords: *denoising, back-projection, noisy image, clean image*

1. PENDAHULUAN

Dengan banyaknya jumlah *smartphone* terjual tiap tahun, kemampuan atau teknologi kamera pada *smartphone* semakin diharapkan menyamai kemampuan kamera *Digital Single Lens Reflex* (DSLR). Keunggulan dan kemudahan penggunaan merupakan salah satu penyebab *smartphone* semakin berkembang dan semakin banyak peminat. Namun, salah satu permasalahan yang dihadapi citra digital pada *smartphone* diantaranya adalah *denoising*. *Denoising* diakibatkan kecilnya ukuran *sensor* dan lensa pada *smartphone* dibandingkan DSLR.

Noise pada citra digital bukan hanya menyebabkan citra digital menjadi kurang baik secara estetika. Namun, berdasarkan riset yang dilakukan pada penelitian sebelumnya, citra kotor seperti pada Gambar 1, menyulitkan algoritma *computer vision*, seperti *object detection*, *segmentation*, untuk berfungsi dengan baik, atau bisa juga mengurangi performa dari model tersebut. Intuisi ini

dibuktikan pada penelitian sebelumnya [1]. Haris et al [1]., menunjukkan citra resolusi rendah dapat mengurangi akurasi deteksi objek.



Gambar 1. Contoh perbandingan citra bersih dan citra kotor

Pada penelitian awal terkait *denoising*, banyak riset menggunakan *noise* buatan atau sintesis, seperti *gaussian noise*. Hal ini membuat metode *denoising* menjadi tidak aplikatif pada permasalahan nyata. Namun, kemajuan teknologi dan AI, memacu para periset untuk dapat menghasilkan model AI yang lebih aplikatif. *Dataset SIDD* [2] menyajikan banyak tipe citra kotor dari 5 tipe kamera pada *smartphone* dan menyediakan *ground truth image* yang diambil dari kamera DSLR. Dataset ini terdiri dari ~30,000 citra digital baik tipe sRGB maupun Raw. Pada penelitian ini kami fokus menggunakan dataset SIDD sebagai data untuk melatih dan mengevaluasi model yang dibuat.

Penelitian terkait *denoising* sudah banyak dilakukan oleh beberapa peneliti, seperti [3, 4, 5, 6, 7, 8]. Namun, pada awal penelitiannya, banyak periset bergantung pada *noise* sintesis dalam membuat dan mengevaluasi model *denoising*, seperti menggunakan *white gaussian noise* (AWGN) [8, 9]. Tren kenaikan performa *deep learning* juga searah dengan perbaikan cara mengevaluasi model *denoising*. Penelitian *denoising* semakin terdorong menggunakan data asli [2, 10]. Pada era *deep learning*, topik *denoising* juga menjadi salah satu topik yang banyak diteliti hingga sekarang [8]. Selain itu, permasalahan *denoising* umum juga dibahas pada pemrosesan sinyal [11].

Adapun kontribusi paper ini terbagi menjadi 2 yaitu sebagai berikut:

- **Deep back-projection network untuk denoising.** Penulis menunjukkan perubahan urutan pada *up-* dan *down-projection* modul dapat memodifikasi fungsi dari *deep back-projection network* untuk penyelesaian masalah *denoising*. *Down-projection module* dapat digunakan sebagai *module denoising* dengan cara memperkecil ukuran dari fitur citra, lalu *up-*

projection digunakan sebagai *module upsampling* agar fitur tersebut kembali ke resolusi semula.

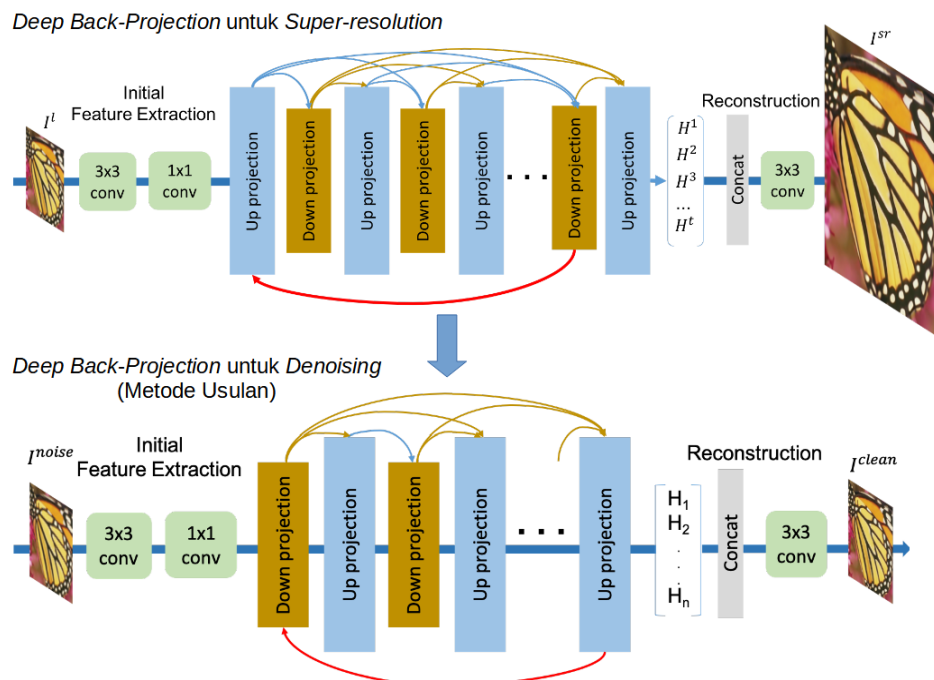
- **Eksperimen yang menyeluruh pada data non artifisial.** Pada hasil eksperimen, penulis menunjukkan performa model yang diajukan menggunakan data publik dengan komprehensif, serta membandingkan performanya dengan 7 metode lain. Model usulan sukses melampaui kualitas dari model lain dengan selisih sampai 6 dB.

2. DENOISING DEEP BACK-PROJECTION

Denoising merupakan salah satu *inverse problem* yang populer di domain citra digital. Selain *denoising*, ada banyak *inverse problem* lain yg populer seperti *super-resolution*, *image enhancement*, dan lain-lain. Pada problem *super-resolution*, *Deep Back-Projection Network* (DBPN) [12, 13] diperkenalkan oleh Haris et al. DBPN berhasil menjadi model terbaik pada kompetisi *Super-resolution* di NTIRE2018 [14].

DBPN mengusulkan modul *iterative up-projection* dan *down-projection* yang digunakan untuk memperkuat relasi dua arah antara fitur resolusi rendah dan fitur resolusi tinggi. Selain itu, DBPN juga memperkenalkan modul *back-projection* [15] pada *deep learning* yang digunakan untuk meminimalisir *error feedback* untuk merekonstruksi citra yang lebih baik.

Terinspirasi dari model DBPN, penulis memanfaatkan modul *iterative down-projection* dan *up-projection* untuk menyelesaikan problem *denoising*. Transformasi yang dilakukan tergambar pada Gambar 2. Model usulan tetap mempertahankan komponen yang digunakan pada desain asli *deep back-projection network*.



Gambar 2. Metode usulan dengan melakukan perubahan pada *deep back-projection* [12, 13] untuk

Denoising. Perubahan dilakukan dengan mengganti urutan projection module pada *deep back-projection network*

Secara umum, komponen yang digunakan untuk model *denoising* sama dengan desain asli *deep back-projection network*. Pada model usulan, hanya keluaran fitur dari *up-projection network* yang digabungkan untuk merekonstruksi *clean image*. Namun, perbedaan terletak pada tahapan awal module *back-projection*. Pada *super-resolution*, tahapan pertama merupakan *up-projection* yang digunakan untuk memperbesar resolusi dari fitur citra tersebut. Sedangkan pada *denoising*, tahapan pertama merupakan *down-projection* yang digunakan untuk memperkecil resolusi dari fitur yg mengharapkan *noise* akan diperkecil melalui operator *down-sampling*.

3. PEMBAHASAN

Pada metode yang kami ajukan, *projection unit* menggunakan 2 tipe *filter* dengan ukuran 8×8 *striding* 4 dan *padding* 2 untuk dataset sRGB serta *filter* dengan ukuran 12×12 *striding* 8 dan *padding* 2 untuk dataset Raw. Inisialisasi masing-masing *layer* menggunakan standar deviasi dengan perhitungan pada [16]. Seluruh *convolutional* and *deconvolutional layer* dipasangkan dengan *parametric rectified linear units* (PreLUs).

Dataset SIDD [2] digunakan untuk melatih model yang diusulkan. Dataset ini terdiri dari ~30000 citra kotor yang diambil pada 10 kondisi *lighting* yang berbeda menggunakan 5 jenis *smartphone*. Dataset ini memiliki 2 format citra yaitu sRGB dan RAW. Masing-masing citra memiliki ukuran 256×256 .

Pelatihan model menggunakan augmentasi untuk meningkatkan generalisasi pada model sehingga menghasilkan performa yang lebih baik. Pelatihan menggunakan *batch size* = 8 dan *patch size* 256×256 . *Learning rate* yang digunakan pada saat inisialisasi adalah $1e-4$ untuk semua layer, lalu berkurang dengan skala $1/10$ di tiap 10^4 iterasi dengan total 2×10^4 iterasi. Untuk optimisasi, *network* menggunakan Adam dengan *momentum* = 0.9. Seluruh eksperimen menggunakan PyTorch 1.0 pada NVIDIA TITAN X GPUs, dan Python 3.5.

Model yang digunakan mengkombinasikan *dense connection*, *recurrent network*, dan *residual learning* seperti yang direkomendasikan pada [13]. Model juga menggunakan $n_0 = 256$ dan $n_R = 64$, $t = 7$, dengan iterasi = 3. Pada layer rekonstruksi, model menggunakan layer conv(3,3). Untuk meningkatkan performa model, kami juga mengimplementasikan *pseudo-ensemble*. *Pseudo-ensemble* menggunakan *horizontal flipping*, *vertical flipping*, dan *transpose*. Penggunaan *pseudo-ensemble* terbukti mampu meningkatkan performa model hingga ~0.2dB.

Tabel 1. Hasil performa model usulan dibandingkan dengan model lainnya pada dataset SIDD. Angka merah menunjukan nilai tertinggi dan angka biru menunjukan nilai kedua tertinggi. Nilai tertinggi pada PSNR dan SSIM menunjukkan performa lebih baik.

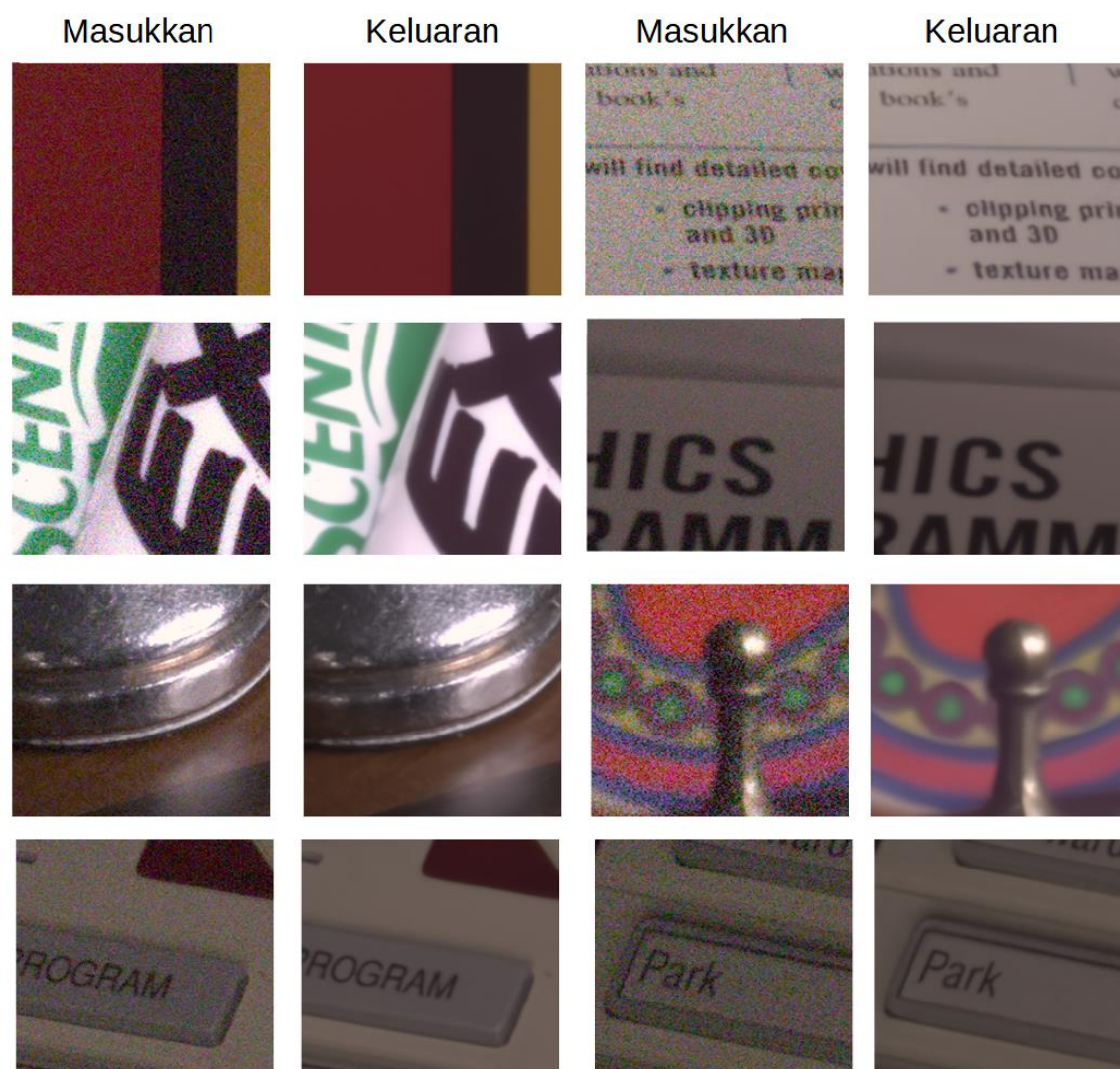
	Dataset	Kinerja Usulan	BM3D [3]	KSVD-DCT [4]	KSVD-G [5]	LPG-PCA [6]	WNNM [7]	EPLL [8]	DnCNN [9]
PSNR	Raw	51.68	45.52	42.70	42.50	42.79	44.85	40.73	43.30
	sRGB	39.48	25.65	27.51	27.19	24.49	25.78	27.11	23.66
SSIM	Raw	0.996	0.980	0.970	0.969	0.974	0.975	0.935	0.965

	sRGB	0.971	0.685	0.780	0.771	0.681	0.809	0.870	0.583
--	------	--------------	-------	-------	-------	-------	-------	--------------	-------

Untuk mengkonfirmasi kemampuan network yang dihasilkan, kita bandingkan performa model dengan beberapa paper terkait. Perbandingan menggunakan tujuh model denoising lainnya yaitu: BM3D [3], KSVD-DCT [4], KSVD-G [5], LPG-PCA [6], WNNM [7], EPLL [8], DnCNN [9]. Perbandingan ini menggunakan dataset SIDD pada 2 format citra yaitu sRGB dan Raw.

Tabel 1 menunjukkan hasil komparasi pada seluruh model *denoising* yang dievaluasi menggunakan dataset SIDD. Pada dataset SIDD, terdapat 2 jenis dataset yaitu sRGB dan Raw. Pada kedua dataset tersebut, model usulan memiliki performa terbaik dibandingkan seluruh model lain. Pada dataset Raw, model usulan lebih tinggi sebanyak ~6 dB, sedangkan untuk metrik SSIM, model usulan lebih tinggi 0.016. Pada dataset sRGB, model usulan lebih tinggi sebanyak ~12 dB, sedangkan untuk metrik SSIM, model usulan lebih tinggi 0.101.

Performa terbaik kedua pada dataset Raw yaitu BM3D [3]. Model BM3D mampu mengalahkan DnCNN [9] yang merupakan model deep learning. Hal ini membuktikan bahwa pelatihan menggunakan data sintetis berpotensi mengurangi performa model jika dievaluasi menggunakan data asli.



Gambar 3. Evaluasi kualitatif pada model usulan. Noise citra merupakan masukan dari model, sedangkan keluaran dari model usulan menghasilkan image yang berhasil dibersihkan dari noise.

Gambar 3 menunjukkan keluaran dari model usulan. Model usulan berhasil menghilangkan *noise* yang ada pada citra masukan. Pada citra yang berisi *text*, terlihat bahwa *text* menjadi lebih terbaca dengan penghilangan *noise* pada gambar. Hal yang sama pada gambar lain, objek yang ada pada gambar semakin jelas dan tajam, sehingga meningkatkan estetika gambar tersebut.

Penulis juga menghitung komputasi yang dibutuhkan oleh metode usulan untuk memproses sebuah gambar berukuran 256 x 256 dengan format sRGB. Metode usulan membutuhkan waktu sekitar 0.25 detik untuk memproses masing-masing gambar.

4. KESIMPULAN

Penelitian ini menunjukkan penggunaan *deep back-projection* untuk *denoising*. Dengan perubahan yang minimal yaitu mengubah urutan dari *up-* dan *down-projection unit*, penelitian ini menunjukkan performa yang sangat baik. *Down-projection unit* digunakan untuk menghilangkan

noise pada citra digital pada *domain* fitur, sedangkan *up-projection unit* digunakan untuk mengembalikan resolusi fitur pada resolusi semula. Model usulan mampu memiliki nilai PSNR dan SSIM paling tinggi dibandingkan model lainnya, dengan selisih dibandingkan model terbaik kedua mencapai 6 dB pada dataset Raw dan 11 dB pada dataset sRGB. Hal ini membuktikan bahwa model usulan berhasil mengatasi permasalahan *denoising* pada dataset SIDD.

DAFTAR PUSTAKA

- [1] Haris, M., Shakhnarovich, G., & Ukita, N. (2021, December). Task-driven super resolution: Object detection in low-resolution images. In *International Conference on Neural Information Processing* (pp. 387-395). Springer, Cham.
- [2] Abdelhamed, A., Lin, S., & Brown, M. S. (2018). A high-quality denoising dataset for smartphone cameras. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1692-1700).
- [3] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3D transform-domain collaborative filtering. *IEEE TIP*, 16(8):2080–2095, 2007
- [4] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE TIP*, 15(12):3736–3745, 2006.
- [5] L. Zhang, W. Dong, D. Zhang, and G. Shi. Two-stage image denoising by principal component analysis with local pixel grouping. *Pattern Recognition*, 43(4):1531–1549, 2010.
- [6] S. Gu, L. Zhang, W. Zuo, and X. Feng. Weighted nuclear norm minimization with application to image denoising. In *CVPR*, 2014.
- [7] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. In *ICCV*, 2011.
- [8] K. Zhang et al. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE TIP*, 2017.
- [9] A. Buades, B. Coll, and J. Morel. A non-local algorithm for image denoising. In *CVPR*, 2005.
- [10] T. Plotz and S. Roth. Benchmarking denoising algorithms with real photographs. In *CVPR*, 2017.
- [11] Setiowati, S., Sitompul, MEH. 2020. Denoising of Electrocardiogram (ECG) Signal Using Fast Fourier Transform Method in Drowsiness Detection System. *JETT*. Vol. 7 No. 1. 789-796.
- [12] Haris, M., Shakhnarovich, G., & Ukita, N. (2018). Deep back-projection networks for super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1664-1673).
- [13] Haris, M., Shakhnarovich, G., & Ukita, N. (2020). Deep back-projectinetworks for single image super-resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(12), 4323-4337.
- [14] R. Timofte, S. Gu, J. Wu, L. Van Gool, L. Zhang, M.-H. Yang, M. Haris, G. Shakhnarovich, N. Ukita et al., “Ntire 2018 challenge on single image super-resolution: Methods and results,” in *Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2018 IEEE Conference on, 2018.
- [15] M. Irani and S. Peleg, “Improving resolution by image registration,” *CVGIP: Graphical models and image processing*, vol. 53, no. 3, pp. 231–239, 1991.
- [16] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” *arXiv preprint arXiv:1512.03385*, 2015.