BAB IV
HASIL DAN PEMBAHASAN

Bab ini menyajikan pembahasan mengenai hasil evaluasi dari pelatihan dan pengujian model *encoder-decoder* yang digunakan dalam proses kompresi gambar. Pembahasn mencakup analisis peforma model saat pelatihan, pengaruh kompleksitas dan jenis *loss function* terhadap hasil kompresi. Seluruh hasil yang disajikan bertujuan untuk menjawab untuk mencapai tujuan penelitian.

4.1 Evaluasi Pelatihan Model

Pengamatan dilakukan terhadap nilai loss dari 20 model encoder-decoder pada tiga titik *epoch*, yaitu *epoch* ke-1, ke-50, dan ke-100 untuk mengevaluasi peforma pelatihan. Tabel 4.1 menyajikan hasil evaluasi nilai *training loss* dari masing-masing model pada ketiga titik tersebut. Nilai *loss* yang lebih rendah menunjukkan performa pelatihan yang lebih baik.

Tabel 4.1 Hasil pengukuran *training loss*

Epoch	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
1	0.01958	0.10378	0.46098	0.01262	0.10546	0.01880	0.10000	0.45259	0.01312	0.10038
50	0.00484	0.04638	0.23127	0.00459	0.04607	0.00422	0.04290	0.19935	0.00420	0.04261
100	0.00464	0.04508	0.22631	0.00450	0.04487	0.00403	0.04165	0.19494	0.00411	0.04141

Epoch	Model 11	Model 12	Model 13	Model 14	Model 15	Model 16	Model 17	Model 18	Model 19	Model 20
1	0.01872	0.10150	0.45757	0.01255	0.09948	0.01580	0.09131	0.38925	0.01198	0.09293
50	0.00373	0.04092	0.17946	0.00400	0.04090	0.00460	0.04445	0.23356	0.00460	0.04499
100	0.00354	0.03967	0.17548	0.00390	0.03970	0.00447	0.04342	0.23164	0.00455	0.04394

Berdasarkan tabel 4.1, didapatkan bahwa hasil evaluasi menunjukkan seluruh model mengalami penurunan nilai *training loss* seiring bertambahnya jumlah *epoch*., yang mengindikasikan proses pelatihan berlangsung dengan konvergen dan stabil.

4.2 Evaluasi Pengujian Model

Penelitian ini juga melakukan evaluasi terhadap performa model dalam merekontruksi gambar. Namun, lebih fokus terhadap seberapa efektif kompresi tambahan yang diterapkan pada dalam mengurangi kembali ukuran *latent space*

Muhammad Iqbal Fauzan, 2025
OPTIMALISASI LATENT SPACE PADA CNN ENCODER-DECODER UNTUK KOMPRESI DAN EFISIENSI TRANSMISI GAMBAR PADA KENDARAAN LISTRIK OTONOM
Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

setelah dilakukan *encoding*. Evaluasi dilakukan dengan menggunakan 50 gambar terkait kondisi jalan dan lingkungan disekitarnya dengan ukuran 80x80 piksel. Keberagaman ini diterapkan untuk mengevaluasi model dalam menangkap detail penting dan menghasilkan gambar rekonstruksi serta memberikan peniliain menyeluruh terhadap efektifitas dalam konteks dunia nyata dengan keadaan lingkungan yang kompleks. Adapun gambar yang digunakan untuk pengujian ditampilkan pada tabel 4.3.

Tabel 4.3 Gambar Data Test untuk pengujian

Gambar input berukuran 80x80										
	Gambar 1	Gambar 2	Gambar 3	Gambar 4	Gambar 5	Gambar 6	Gambar 7	Gambar 8	Gambar 9	Gambar 10
	Gambar 11	Gambar 12	Gambar 13	Gambar 14	Gambar 15	Gambar 16	Gambar 17	Gambar 18	Gambar 19	Gambar 20
	Gambar 21	Gambar 22	Gambar 23	Gambar 24	Gambar 25	Gambar 26	Gambar 27	Gambar 28	Gambar 29	Gambar 30
	Gambar 31	Gambar 32	Gambar 33	Gambar 34	Gambar 35	Gambar 36	Gambar 37	Gambar 38	Gambar 39	Gambar 40
	Gambar 41	Gambar 42	Gambar 43	Gambar 44	Gambar 45	Gambar 46	Gambar 47	Gambar 48	Gambar 49	Gambar 50

Setelah proses pelatihan dilakukan, seluruh model diuji untuk mengukur performa model dalam merekonstruksi gambar serta efisiensi kompresi pada *latent space*. Evaluasi dilakukan dengan tiga skema pengujian, yaitu tanpa kompresi tambahan, dengan kompresi menggunakan Blosc, dan dengan Zlib. Metrik yang digunakan meliputi nilai PSNR sebagai indikator kualitas hasil rekonstruksi, ukuran *latent* dalam satuan *byte*, serta estimasi waktu pengiriman data *latent* melalui

jaringan LoRa berdasarkan referensi bitrate sebesar 37,5 kbps. Pengujian ini bertujuan untuk menilai seberapa jauh representasi laten hasil encoding dapat dikompresi lebih lanjut dengan meminimalisir dampak dari kompresi, serta mengidentifikasi model-model yang paling optimal untuk diterapkan pada sistem transmisi data dengan keterbatasan bandwidth. Rangkuman hasil pengujian seluruh model ditampilkan pada Tabel 4.4.

Tabel 4.4 Hasil pengujian model

Nomor Model	Kompresi Tambahan	Rerata SSIM	Rerata PSNR (dB)	PSNR Tertinggi (dB)	PSNR Terendah (dB)	Rerata Ukuran Latent (bytes)	Estimasi Waktu Pengiriman (detik)
1	-	0.46	29.49	31.22	28.39	640	0.14
	Blosc	0.46	29.49	31.22	28.39	528	0.11
	Zlib	0.46	29.49	31.22	28.39	523	0.11
2	-	0.47	29.44	31.20	28.31	640	0.14
	Blosc	0.47	29.44	31.20	28.31	528	0.11
	Zlib	0.47	29.44	31.20	28.31	523	0.11
3	-	0.55	28.61	29.89	27.88	640	0.14
	Blosc	0.55	28.61	29.97	27.88	528	0.11
	Zlib	0.55	28.61	29.96	27.88	523	0.11
4	-	0.46	29.14	30.52	28.34	640	0.14
	Blosc	0.46	29.14	30.52	28.34	528	0.11
	Zlib	0.46	29.14	30.52	28.33	523	0.11
5	-	0.47	29.66	31.51	28.46	640	0.14
	Blosc	0.47	29.66	31.51	28.46	528	0.11
	Zlib	0.47	29.66	31.51	28.46	523	0.11
6	-	0.50	29.59	31.27	28.47	896	0.19
	Blosc	0.50	29.59	31.27	28.47	784	0.17
	Zlib	0.50	29.59	31.27	28.47	770.84	0.16
7	-	0.51	29.48	30.95	28.44	896	0.19
	Blosc	0.51	29.48	30.95	28.44	784	0.17
	Zlib	0.51	29.48	30.95	28.44	771.32	0.16
8	-	0.59	28.61	29.96	27.83	896	0.19
	Blosc	0.59	28.61	29.96	27.83	784	0.17
	Zlib	0.59	28.61	29.96	27.83	772.14	0.16
9	-	0.50	29.10	29.75	28.29	896	0.19
	Blosc	0.50	29.10	29.75	28.29	784	0.17
	Zlib	0.50	29.10	29.75	28.29	769.44	0.16
10	-	0.50	29.57	31.10	28.44	896	0.19
	Blosc	0.50	29.57	31.10	28.44	784	0.17
	Zlib	0.50	29.57	31.10	28.44	773.44	0.16
11	-	0.53	29.54	30.36	28.58	1152	0.25
	Blosc	0.53	29.54	30.36	28.58	1040	0.22
	Zlib	0.53	29.54	30.36	28.58	1009.92	0.22
12	-	0.53	29.57	30.59	28.26	1152	0.25
	Blosc	0.53	29.57	30.59	28.26	1040	0.22
	Zlib	0.53	29.56	30.59	28.26	1009.08	0.22

Nomor Model	Kompresi Tambahan	Rerata SSIM	Rerata PSNR (dB)	PSNR Tertinggi (dB)	PSNR Terendah (dB)	Rerata Ukuran Latent (bytes)	Estimasi Waktu Pengiriman (detik)
	-	0.62	28.81	29.97	28.02	1152	0.26
13	Blosc	0.62	28.81	29.97	28.02	1040	0.22
	Zlib	0.62	28.81	29.97	28.02	1009.4	0.22
	-	0.54	29.33	30.78	28.41	1152	0.25
14	Blosc	0.54	29.33	30.77	28.41	1038.36	0.22
	Zlib	0.54	29.33	30.77	28.41	1004.02	0.21
	-	0.53	29.73	31.52	28.65	1152	0.25
15	Blosc	0.53	29.73	31.52	28.65	1040	0.22
	Zlib	0.53	29.73	31.52	28.65	1009.1	0.22
	-	0.55	30.12	32.06	28.73	928	0.20
16	Blosc	0.55	30.12	32.07	28.73	806.16	0.17
	Zlib	0.55	30.12	32.06	28.73	803.28	0.17
	-	0.57	30.14	31.88	28.85	928	0.20
17	Blosc	0.57	30.14	31.88	28.85	809.84	0.17
	Zlib	0.57	30.14	31.88	28.85	805.24	0.17
	-	0.64	29.09	31.16	27.89	928	0.20
`18	Blosc	0.64	29.09	31.16	27.89	810.72	0.17
	Zlib	0.64	29.09	31.16	27.89	802.84	0.17
	-	0.55	29.52	31.66	28.58	928	0.20
19	Blosc	0.55	29.52	31.66	28.58	804.86	0.17
	Zlib	0.55	29.52	31.66	28.88	801.68	0.17
	-	0.55	29.66	30.88	28.77	928	0.20
20	Blosc	0.55	29.66	30.88	28.77	808.8	0.17
	Zlib	0.55	29.66	20.88	28.77	804.44	0.17

Tabel 4.4 menyajikan hasil pengujian dari 20 model encoder-decoder yang dievaluasi dalam tiga skema berbeda: tanpa kompresi tambahan (*baseline*), dengan kompresi menggunakan Blosc, dan dengan kompresi menggunakan Zlib. Parameter yang diukur meliputi rerata SSIM dan PSNR, nilai PSNR tertinggi dan terendah, ukuran *latent space* dalam byte setelah proses encoding, serta estimasi waktu pengiriman berdasarkan ukuran tersebut. Hasil pada tabel terlihat angka rerata SSIM berada di rentang 0.46-0.64. Ini menunjukkan adanya degradasi hasil gambar rekonstruksi setelah melewati proses encoding-decoding terhadap gambar input awal. Kompresi tambahan secara konsisten mengurangi ukuran *latent* antara 100 hingga 140 byte pada sebagian besar model. Selain itu, waktu pengiriman juga mengalami penurunan yang sepadan, tanpa memengaruhi nilai SSIM dan PSNR secara signifikan. Perubahan nilai SSIM dan PSNR antara ketiga skema pengujian umumnya berada dalam rentang desimal kecil, menunjukkan bahwa proses kompresi tambahan tidak memberikan degradasi kualitas rekonstruksi yang berarti.

Secara umum, nilai rerata PSNR antar model berada pada kisaran 28,6 dB hingga 30,1 dB. Model dengan hasil *latent space* lebih besar (misalnya 1152 byte) tidak selalu menghasilkan PSNR yang lebih tinggi dibanding model yang lebih ringan, yang menunjukkan bahwa ukuran *latent* bukan satu-satunya faktor penentu kualitas rekonstruksi. Beberapa model yang menggunakan *loss function* SSIM (seperti Model 3 dan 8) mencatatkan PSNR yang cenderung lebih rendah dibandingkan model lainnya, meskipun nilai PSNR tertinggi mereka masih berada di atas 29 dB. Hal ini menunjukkan bahwa fungsi *loss* yang lebih perseptual dapat berdampak pada metrik numerik seperti PSNR.

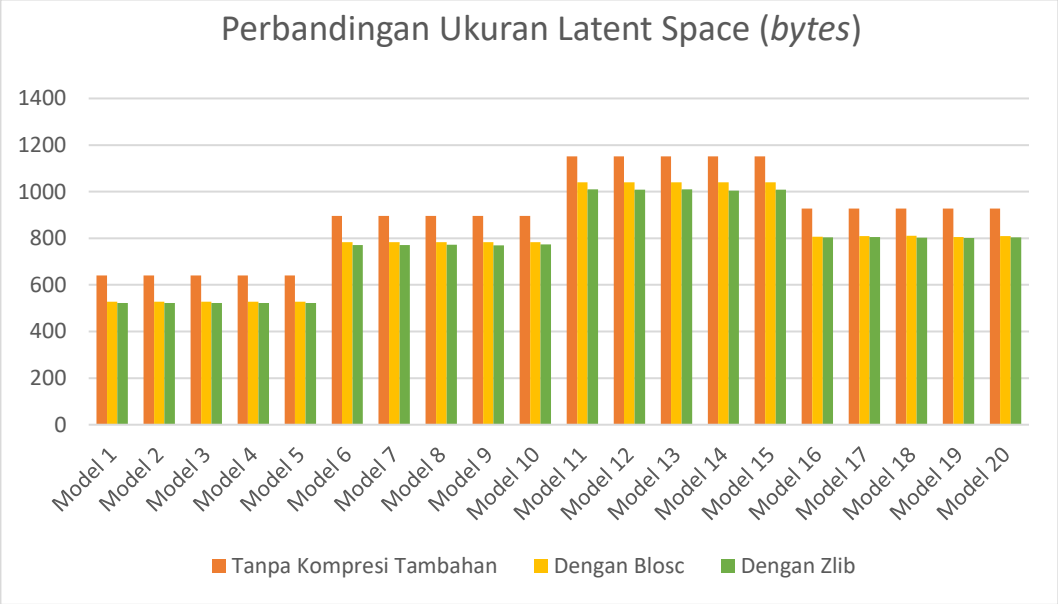
Ketika diterapkan kompresi tambahan menggunakan Blosc dan Zlib, ukuran *latent* pada setiap model mengalami penurunan yang signifikan. Namun, terdapat perbedaan yang mencolok antara kelompok model dengan *latent space* kecil (seperti model 1–5 dan 6–10) dan kelompok dengan *latent* besar (seperti model 11–15). Pada kelompok kecil, hasil kompresi antar model cenderung seragam; misalnya pada model 1–5, ukuran hasil kompresi dengan Zlib berkisar sangat sempit antara 523 hingga 524 byte. Hal ini menunjukkan dimensi dan kedalaman fitur terbatas, variasi nilai antar *channel* dalam *latent space* juga cenderung homogen. Akibatnya, algoritma kompresi bekerja dengan efisien yang serupa antar model.

Sebaliknya, pada kelompok model dengan ukuran *latent* besar, seperti model 11–15, hasil kompresi menunjukkan variasi ukuran yang lebih tinggi meskipun dimensi *latent* awal identik. Misalnya, pada model 11 hingga 15, ukuran hasil kompresi Zlib berkisar antara 976 hingga 1040 byte. Perbedaan ini menunjukkan bahwa meskipun ukuran *latent space* secara arsitektural sama (1152 byte), isi atau distribusi nilai di dalamnya memiliki tingkat entropi yang berbeda-beda antar model. Hal ini dapat disebabkan oleh jenis *loss function* yang digunakan saat pelatihan, karena fungsi *loss* memengaruhi bagaimana fitur dikodekan dalam *latent space*. Model dengan fungsi *loss* berbasis struktur seperti SSIM atau kombinasi dengan Gradient kemungkinan menghasilkan representasi yang lebih bervariasi atau kurang terstruktur secara statistik, sehingga tingkat kompresibilitasnya menjadi lebih rendah. Sebaliknya, model dengan *loss* seperti

MSE cenderung menghasilkan distribusi nilai yang lebih terkonsentrasi, sehingga lebih mudah dikompresi.

Kelompok model 16 hingga 20, yang memiliki konfigurasi paling sederhana tanpa lapisan konvolusi 2×2 pada segmen pengembangan, ukuran *latent space* awal tercatat konsisten sebesar 928 *byte*. Meskipun secara dimensi lebih kecil dari kelompok 11–15, hasil kompresi pada kelompok ini menunjukkan variasi yang moderat, tidak sehomogen kelompok model kecil, namun juga tidak setinggi variasi pada kelompok besar. Dalam skenario kompresi dengan Zlib, ukuran hasil kompresi berkisar antara 780 hingga 813 byte, sementara dengan Blosc berkisar antara 784 hingga 818 byte. Perbedaan ini menunjukkan bahwa meskipun struktur arsitektur antar model dalam kelompok ini identik, distribusi nilai dalam *latent space* tetap dipengaruhi oleh perbedaan fungsi *loss* yang digunakan. Karena model-model ini tidak memiliki lapisan pengolahan fitur paling dalam (2×2), kompleksitas fitur yang dihasilkan kemungkinan berada pada tingkat menengah, sehingga menghasilkan entropi yang tidak seragam, tetapi juga tidak ekstrem. Dengan demikian, kelompok model ini menempati posisi transisi antara model-model berukuran *latent* kecil yang sangat homogen dan model besar yang sangat bervariasi dalam hasil kompresinya.

Selanjutnya, grafik 4.2 berikut menyajikan perbandingan ukuran *latent space* yang dihasilkan oleh 20 model encoder-decoder dalam bentuk *latent space* dan *latent space* terkompresi, yaitu tanpa kompresi tambahan, dengan kompresi menggunakan Blosc, dan dengan Zlib. Masing-masing model diuji terhadap 50 gambar uji, dan nilai yang ditampilkan merupakan rata-rata ukuran file *latent* hasil encoding dalam satuan byte. Grafik ini bertujuan untuk menggambarkan efisiensi representasi *latent* yang dihasilkan oleh setiap model serta dampak penerapan kompresi tambahan terhadap ukuran akhir data sebelum transmisi.



Gambar 4.2 Grafik perbandingan ukuran *latent space*

Grafik yang disajikan memperlihatkan perbandingan ukuran rata-rata *latent space* yang dihasilkan oleh 20 model encoder-decoder: tanpa kompresi tambahan, dengan kompresi Blosc, dan dengan kompresi Zlib. Data ini memberikan gambaran mengenai efisiensi representasi data hasil kompresi awal model serta efektivitas tambahan kompresi entropik terhadap pengurangan beban data sebelum proses transmisi. Jika dibandingkan dengan ukuran rata-rata citra asli yang berada pada kisaran 10819,74 byte, ukuran *latent* hasil encoding pada semua model sudah menunjukkan pengurangan yang sangat signifikan bahkan sebelum kompresi tambahan dilakukan, yaitu berkisar antara 640 hingga 1152 byte, atau sekitar 6% hingga 11% dari ukuran gambar asli.

Ketika diterapkan kompresi tambahan menggunakan Blosc dan Zlib, terjadi penurunan ukuran *latent* lebih lanjut, rata-rata sebesar 15% hingga 20% dari ukuran awalnya. Zlib umumnya memberikan hasil kompresi yang lebih besar (lebih kecil ukurannya) dibandingkan Blosc, walaupun perbedaannya tidak besar. Sebagai contoh, Model 1 memiliki *latent size* baseline sebesar 640 byte, yang dikompresi menjadi 528 byte dengan Blosc dan 523 byte dengan Zlib. Model-model dengan struktur *bottleneck* yang lebih besar, seperti Model 11–15 (1152 byte baseline), juga menunjukkan reduksi yang cukup signifikan, menjadi sekitar 960–1040 byte

setelah dikompresi. Ini menunjukkan bahwa efisiensi kompresi juga tergantung pada seberapa terstruktur dan berulang pola data dalam *latent space* yang dihasilkan oleh model.

Sementara itu, pada kelompok model 16 hingga 20, yang memiliki, ukuran *latent space* awal tercatat konsisten sebesar 928 byte. Hasil kompresi pada kelompok ini menunjukkan variasi yang moderat, tidak sehomogen kelompok model kecil, namun juga tidak setinggi variasi pada kelompok besar. Dalam skenario kompresi dengan Zlib, ukuran hasil kompresi berkisar antara 780 hingga 813 byte, sementara dengan Blosc berkisar antara 784 hingga 818 byte. Perbedaan ini menunjukkan bahwa meskipun struktur arsitektur antar model dalam kelompok ini identik, distribusi nilai dalam *latent space* tetap dipengaruhi oleh perbedaan fungsi *loss* yang digunakan. Karena model-model ini tidak memiliki lapisan pengolahan fitur paling dalam (2×2), kompleksitas fitur yang dihasilkan kemungkinan berada pada tingkat menengah, sehingga menghasilkan entropi yang tidak seragam, tetapi juga tidak ekstrem. Dengan demikian, kelompok model ini menempati posisi transisi antara model-model berukuran *latent* kecil yang sangat homogen dan model besar yang sangat bervariasi dalam hasil kompresinya.

Dari estimasi waktu pengiriman terlihat bahwa perbedaan ukuran *latent* antar model dapat mempengaruhi waktu pengiriman secara signifikan. Ketika kompresi tambahan diterapkan, waktu pengiriman dapat berkurang hingga puluhan milidetik. Misalnya, ukuran 640 byte yang dikompresi menjadi 523 byte (Zlib) akan menghasilkan waktu pengiriman sekitar 111,6 ms, lebih singkat dibandingkan *baseline*-nya. Oleh karena itu, kompresi tambahan secara nyata memberikan kontribusi dalam efisiensi waktu transmisi, yang sangat krusial dalam sistem komunikasi berbitrate rendah seperti LoRa, khususnya jika pengiriman dilakukan secara real-time atau dalam jumlah besar secara berkala.

Berdasarkan data pada tabel, tidak hanya kualitas visual atau numerik seperti PSNR yang menjadi pertimbangan dalam memilih model, tetapi juga aspek praktis seperti ukuran *latent* dan estimasi waktu transmisi. Dalam konteks ini, Model 1 menonjol sebagai model dengan ukuran *latent* terkecil secara konsisten dalam ketiga skema pengujian, yaitu masing-masing 640 byte (baseline), 528 byte

(Blosc), dan 523 byte (Zlib), yang berimplikasi pada waktu pengiriman tercepat di antara semua model. Namun, ketika mempertimbangkan keseimbangan antara efisiensi ukuran dan kualitas rekonstruksi, Model 16 dan Model 17 tetap menunjukkan performa yang patut diperhitungkan. Meskipun memiliki ukuran *latent* yang sedikit lebih besar (832 byte), keduanya menghasilkan nilai rerata PSNR tertinggi di antara semua model, serta masih mempertahankan estimasi waktu pengiriman yang relatif singkat, yaitu sekitar 177,5 ms (baseline) dan sekitar 140–145 ms (Zlib).