Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209

Benchmark Interpolasi Otomatis terhadap Ground Truth Frame-by-Frame pada Prinsip Exaggeration dalam Animasi 2D

Deni Sinaga*✉, Ema Utami✉
Magister Teknik Informatika, Universitas Amikom Yogyakarta, Indonesia

Abstrak: Interpolasi otomatis semakin sering dipakai untuk mempercepat produksi animasi 2D, tetapi kemampuannya mereproduksi prinsip *exaggeration*, seperti *smear frame* dan *squash and stretch*, belum diukur secara objektif terhadap gambar tangan. Penelitian ini mengukur penurunan kinerja rekonstruksi interpolasi otomatis pada frame berdeformasi tinggi dibandingkan seluruh frame, menggunakan *ground truth frame-by-frame* dari satu adegan animasi 2D berisi 192 frame yang digambar dengan Toon Boom Harmony 25.2 dan diekspor sebagai PNG berkanal alfa dengan latar abu-abu. Metode yang digunakan adalah *benchmark hold-out* dengan interval frame kunci 2, 4, dan 8. Frame yang dibuang direkonstruksi oleh lima metode (*hold*, *cross-dissolve*, *optical flow* DIS, RIFE v4.25, dan FILM), lalu dinilai dengan SSIM pada area karakter, F-score kontur, IoU siluet, indeks *ghosting*, dan galat rasio aspek. Frame berdeformasi tinggi ditandai dengan indeks deformasi siluet yang tidak terpengaruh translasi. Hasilnya, F-score kontur median pada frame berdeformasi tinggi (0.176–0.431) jauh lebih rendah daripada pada seluruh frame (0.313–0.721). RIFE menghasilkan *ghosting* terendah (selisih rerata terhadap *cross-dissolve* hingga −1.058; $p < 0.001$), tetapi tidak ada model yang meningkatkan kecocokan kontur secara konsisten dibanding *cross-dissolve*. *Ghosting* rendah pada RIFE tidak selalu disertai kecocokan bentuk karena sebagian frame kehilangan sebagian siluet. Pemisahan *layer* tidak memperbaiki hasil, dan latar belakang terutama memengaruhi SSIM. Temuan ini menunjukkan bahwa interpolasi yang diuji belum menggantikan gambar tangan pada *smear frame*, serta menyediakan kerangka evaluasi yang dapat direplikasi untuk alur kerja hibrida.
Kata Kunci: Animasi 2D; *Benchmark*; *Exaggeration*; Interpolasi Otomatis; *Optical Flow*;

Abstract: Automated interpolation is increasingly used to speed up 2D animation production, yet its ability to reproduce the exaggeration principle, such as smear frames and squash and stretch, has not been measured objectively against hand-drawn drawings. This study quantifies the decline in reconstruction performance of automated interpolation on highly deformed frames compared with all frames, using hand-drawn frame-by-frame ground truth from a single 192-frame 2D animation scene produced in Toon Boom Harmony 25.2 and exported as alpha-channel PNG on a gray background. We apply a hold-out benchmark with keyframe gaps of 2, 4, and 8. The removed frames are reconstructed by five methods (hold, cross-dissolve, DIS optical flow, RIFE v4.25, and FILM) and assessed with region-of-interest SSIM, contour F-score, silhouette IoU, a ghosting index, and aspect-ratio error. Highly deformed frames are flagged with a translation-invariant silhouette deformation index. The median contour F-score on highly deformed frames (0.176–0.431) is far lower than on all frames (0.313–0.721). RIFE yields

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209

the lowest ghosting (mean difference versus cross-dissolve down to −1.058; $p < 0.001$), but no model consistently improves contour agreement over cross-dissolve. Low ghosting for RIFE is not always accompanied by shape agreement because some frames lose part of the silhouette. Separating layers does not improve the results, and the background mainly affects SSIM. These findings indicate that the tested interpolation methods do not yet replace hand-drawn work on smear frames, and offer a replicable evaluation framework for hybrid workflows.
Keywords: 2D Animation; Benchmark; Exaggeration; Automated Interpolation; Optical Flow;

1. PENDAHULUAN

Animasi 2D merupakan media komunikasi visual yang banyak dimanfaatkan untuk promosi, pendidikan, dan penyampaian narasi budaya [1]. Tinjauan terkini menunjukkan bahwa teknik dan alat animasi terus berkembang [2]. Pada konten promosi berdurasi pendek, dampak emosional harus tercapai dalam hitungan detik, sehingga kualitas gerak pada adegan krusial menentukan efektivitas pesan. Kondisi ini membuat pertanyaan tentang kapan produksi dapat diotomasi dan kapan gambar tangan tidak tergantikan menjadi relevan, baik secara praktis maupun ilmiah.

Daya tarik animasi 2D bertumpu pada prinsip-prinsip gerak klasik, antara lain pelebaran dan pemipihan bentuk (*squash and stretch*) serta *exaggeration* [3], [4]. Pada gerakan berkecepatan tinggi, animator menggunakan *smear frame*, yaitu gambar yang bentuknya sengaja didistorsi selama satu atau dua frame untuk menyampaikan kecepatan dan menggantikan gambar antara yang akurat secara anatomis; teknik ini bahkan direproduksi pada animasi berbasis 3D [5], [6]. Penerapan prinsip animasi yang konsisten juga terbukti penting pada film pendek animasi [7]. Namun, produksi *frame-by-frame* sangat padat karya karena setiap gambar kunci, *breakdown*, dan gambar antara dibuat satu per satu. Alternatif berbasis kerangka (*rigging*) memang mengurangi kerumitan dan menjaga konsistensi pose [8], tetapi bentuk yang kaku membatasi deformasi ekstrem seperti pada *smear frame*.

Upaya otomasi telah menyentuh beberapa tahap produksi. Even dkk. mengusulkan penyisipan gambar antara untuk animasi 2D kasar yang memuat gerak non-linear [9], sedangkan Barroso dkk. mengusulkan penyisipan otomatis untuk animasi lukis berbasis sapuan [10]. Otomasi juga merambah perambatan warna antarframe [11] dan dekomposisi video menjadi lapisan yang konsisten [12]. Interpolasi frame video (*video frame interpolation*, VFI) sendiri telah dikaji luas, baik berbasis aliran optik, kernel, maupun fase [13]. Perkembangannya mencakup estimasi aliran dua arah dengan peta visibilitas [14], penalaran kedalaman untuk oklusi [15], aliran yang disesuaikan tugas [16], *forward warping* dengan *softmax splatting* [17], arsitektur berbasis *Transformer* [18], serta korelasi semua pasangan piksel [19] yang bertumpu pada estimator aliran optik seperti PWC-Net [20] dan RAFT [21]. Model modern seperti RIFE [22] dan FILM [23] menyasar video alami dengan gerak besar.

Pada animasi, interpolasi menghadapi tantangan yang berbeda. Narita dkk. memperlihatkan bahwa gambar garis tanpa gradien kecerahan memerlukan transformasi jarak agar aliran optik dapat dipakai [24]. AnimeInterp mendefinisikan masalah interpolasi video animasi secara formal dan menunjukkan bahwa minimnya tekstur dan gerak non-linear yang besar menyulitkan metode untuk video alami [25]. Chen dan Zwicker menambahkan bahwa gerak non-linear yang berlebihan juga menyulitkan penyusunan data latih [26]. Pendekatan generatif [27] dan pendekatan geometris untuk gambar garis [28], [29] kemudian dikembangkan, sementara pada gerak 3D, *in-betweening* berbasis pembelajaran mendalam dipelajari secara luas [30], [31], [32].

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209

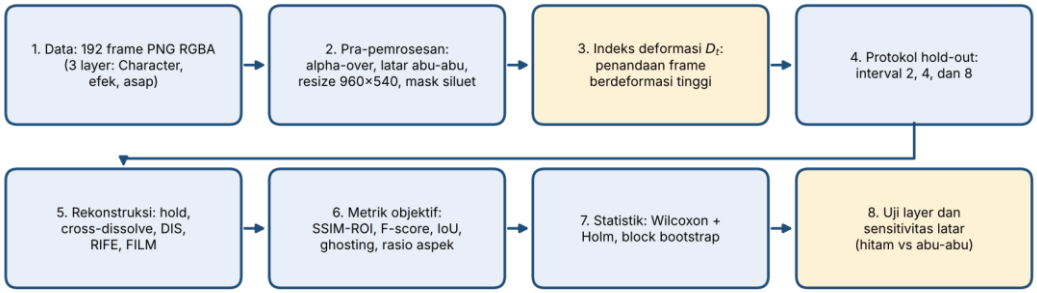
Meskipun demikian, kinerja pada studi-studi tersebut umumnya dilaporkan pada himpunan data publik atau kategori gerak, bukan pada jenis frame *exaggeration* dari produksi nyata. Studi pada jurnal yang sama telah mengukur *trade-off* otomasi prinsip animasi secara kuantitatif pada animasi 3D [33], tetapi belum menyentuh animasi 2D gambar tangan. Berdasarkan literatur yang ditinjau, belum ditemukan analisis terstratifikasi yang membandingkan interpolasi otomatis dengan *ground truth* gambar tangan penuh dari satu adegan promosi, lengkap dengan ukuran beban kerja berbasis jumlah gambar. Selain itu, adegan produksi lazimnya disusun dalam beberapa *layer*, misalnya karakter dan efek, tetapi belum jelas apakah memisahkan *layer* sebelum interpolasi memperbaiki rekonstruksi, dan apakah warna latar yang dipakai saat komposisi memengaruhi skor.

Penelitian ini menjawab kesenjangan tersebut dengan *benchmark hold-out* pada satu adegan animasi 2D yang seluruh frame-nya digambar tangan menggunakan Toon Boom Harmony 25.2. Frame dibuang secara periodik, direkonstruksi oleh lima metode (*hold*, *cross-dissolve*, *optical flow* DIS, RIFE, dan FILM), lalu dibandingkan dengan gambar asli. Kebaruan penelitian ini terletak pada gabungan empat hal: *ground truth* gambar tangan penuh sehingga tidak diperlukan penilai manusia, analisis terstratifikasi berdasarkan indeks deformasi siluet yang tidak terpengaruh translasi untuk memisahkan frame *exaggeration* dari frame biasa, skema pengujian per *layer* yang memanfaatkan ekspor *layer* berkanal alfa, dan analisis sensitivitas terhadap warna latar. Tujuan penelitian adalah mengukur seberapa besar kinerja rekonstruksi menurun pada frame berdeformasi tinggi, apakah pemisahan *layer* mengubah hasil tersebut, dan seberapa kuat kesimpulan bergantung pada latar komposisi. Seluruh tahapan otomatis sehingga dapat direplikasi.

2. METODE PENELITIAN

2.1. Tahapan Penelitian

Penelitian ini berupa eksperimen komputasional kuantitatif dengan pendekatan *hold-out*. Seluruh tahapan, dari pengolahan data hingga analisis statistik, dilakukan secara otomatis sehingga dapat diulang dengan hasil yang sama. Alur penelitian ditunjukkan pada Gambar 1. Penelitian ini tidak melibatkan responden manusia; seluruh pengukuran berasal dari perhitungan terhadap citra.



Gambar 1. Alur penelitian dari data hingga analisis statistik dan sensitivitas latar.

2.2. Dataset dan *Ground Truth*

Dataset berupa satu adegan animasi 2D yang digambar sepenuhnya dengan tangan oleh penulis pertama menggunakan Toon Boom Harmony 25.2 Premium. Adegan ini merupakan bagian dari film animasi promosi berjudul *Information Technology*

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209

Undergraduate Program Profile [34] (Gambar 2) dan menampilkan karakter yang memegang tablet dengan ekspresi panik, dua rangkaian gambar ekor yang dicambukkan dan menggantikan karakter selama beberapa frame, kemudian karakter yang terpental dan jatuh disertai efek asap. Tampilan proyek pada perangkat lunak tersebut ditunjukkan pada Gambar 3. Adegan diekspor sebagai urutan PNG RGBA berkanal alfa beresolusi 3556 × 2000 piksel untuk tiga *layer*, yaitu karakter (*Character*), efek (*effect*), dan asap (*smoke*), serta frame komposit. Karya ini adalah ciptaan orisinal penulis pertama sehingga tidak ada pembatasan hak cipta atas penggunaannya sebagai data.



Gambar 2. Cuplikan layar hasil akhir film *Information Technology Undergraduate Program Profile* sebagai konteks produksi adegan yang diteliti.



Gambar 3. Antarmuka Toon Boom Harmony 25.2 Premium pada proyek adegan yang diteliti: jendela Camera, Node View, dan Timeline.

Pemilihan satu adegan dengan *ground truth* penuh didasarkan pada pertimbangan metodologis. *Benchmark hold-out* hanya valid bila setiap frame yang dibuang masih tersedia sebagai gambar tangan, sehingga galat rekonstruksi dapat dihitung tanpa penilai manusia. Adegan ini dipilih karena memuat tiga jenis gerak sekaligus: gerak halus karakter, penggantian gambar secara menyeluruh pada rangkaian ekor, dan gerak jatuh yang cepat disertai efek. Keragaman ini memungkinkan perbandingan antara frame yang mudah dan sulit direkonstruksi dalam satu data yang sama.

2.3. Pra-pemrosesan

Setiap frame RGBA diperkecil menjadi 960 × 540 piksel dengan interpolasi berbasis luas area yang dihitung pada kanal ternormalisasi alfa (*premultiplied*), lalu digabung dengan latar polos melalui operasi *alpha-over*. Latar utama berwarna abu-abu (nilai 128), sedangkan latar hitam dibuat dari sumber yang sama sebagai pembanding sensitivitas. *Layer* ditumpuk dari bawah ke atas dengan urutan asap, efek, dan karakter. Siluet

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209

diperoleh dari piksel yang selisihnya terhadap latar melebihi 8, diikuti penutupan morfologis dengan kernel 5 × 5 dan pengisian lubang. Frame dengan luas siluet kurang dari 500 piksel dianggap kosong dan dikeluarkan dari pengukuran, agar bayangan tepi tipis tidak dihitung sebagai objek.

2.4. Protokol *Hold-out*

Dari urutan 192 frame, frame kunci dipertahankan pada indeks 0, g, 2g, dan seterusnya, ditambah frame terakhir, dengan interval g ∈ {2, 4, 8}. Frame di antara dua frame kunci a dan b direkonstruksi pada posisi relatif s menurut Persamaan (1), kemudian dibandingkan dengan gambar tangan asli pada posisi yang sama.

$$s = \frac{t-a}{b-a} \tag{1}$$

Jumlah gambar kunci yang diperlukan pada tiap interval dibandingkan dengan jumlah gambar unik yang digambar pada versi manual untuk menghasilkan ukuran beban kerja pengganti (*proksi*). Ukuran ini bukan pengukuran jam kerja; ia hanya mencerminkan banyaknya gambar yang harus dibuat tangan.

2.5. Metode Rekonstruksi

Lima pembanding digunakan. Pertama, *hold*, yaitu menahan frame kunci sebelumnya tanpa interpolasi. Kedua, *cross-dissolve*, yaitu percampuran linier kedua frame kunci menurut Persamaan (2), yang meniru interpolasi paling sederhana tanpa estimasi gerak.

$$\hat{I}_t = (1-s)I_a + sI_b \tag{2}$$

Ketiga, *optical flow* berbasis pencarian invers padat (*dense inverse search*, DIS) dengan pengaturan sedang [35]. Aliran dua arah dihitung pada citra keabuan, diskalakan sesuai posisi s, lalu dipakai untuk menggeser kedua frame kunci sebelum dicampur seperti pada Persamaan (3), dengan W(·) adalah operator pergeseran citra dan F aliran optik antarframe kunci.

$$\hat{I}_t = (1-s)\cdot W(I_a, s\cdot F_{b\to a}) + s\cdot W(I_b, (1-s)\cdot F_{a\to b}) \tag{3}$$

Keempat, RIFE versi 4.25 [22], model pembelajaran mendalam yang mengestimasi aliran antara secara langsung, dijalankan dari implementasi Practical-RIFE dengan bobot pralatih tanpa pelatihan ulang. Kelima, FILM [23], model pralatih dari TensorFlow Hub yang ditujukan untuk gerak besar. Keduanya dipanggil pada posisi relatif s menurut Persamaan (1) dengan frame beresolusi analisis 960 × 540 piksel (dengan *padding* ke kelipatan 64 piksel untuk RIFE), sehingga hasilnya mencerminkan pemakaian bawaan (*out-of-the-box*) oleh animator. *Hold* dan *cross-dissolve* sengaja dipilih sebagai batas bawah yang transparan tanpa data latih, sedangkan model generatif ToonCrafter [27] tidak diuji dan dibahas sebagai arah pengembangan.

2.6. Metrik Evaluasi

Kualitas rekonstruksi dihitung pada area karakter (*region of interest*, ROI), yaitu kotak pembatas gabungan siluet prediksi dan siluet asli ditambah margin 16 piksel, agar latar yang luas tidak mendominasi skor. PSNR dihitung menurut Persamaan (4) dan SSIM mengikuti [36]. Pemilihan SSIM dan PSNR sebagai metrik pelengkap mengikuti praktik umum dengan memperhatikan bahwa keduanya berkaitan erat secara analitis [37].

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{255^2}{\mathrm{MSE}}\right) \tag{4}$$

Kecocokan kontur diukur dengan F-score atas tepi hasil deteksi Canny [38] pada ambang 30 dan 90, dengan toleransi jarak 2 piksel, seperti Persamaan (5), dengan P adalah presisi dan R adalah *recall* tepi prediksi terhadap tepi *ground truth*. Pendekatan F-score pada kontur dengan toleransi jarak mengikuti evaluasi batas objek [39], [40].

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : https://doi.org/10.58602/itsecs.v5i1.504
Page 187-209

$$F = \frac{2PR}{P+R} \quad (5)$$

Selain itu, IoU siluet dan galat luas relatif dihitung antara siluet prediksi dan siluet asli. Indeks *ghosting* G pada Persamaan (6) mengukur luas piksel prediksi yang berada di luar siluet *ground truth* yang dikembangkan 7 × 7 piksel, dinormalisasi oleh luas siluet asli; nilai lebih dari 1 berarti luas bayangan melebihi luas objek. Galat rasio aspek pada Persamaan (7) memakai AR = tinggi/lebar kotak pembatas siluet dan menjadi proksi *squash and stretch*.

$$G = \frac{|M_p \setminus \delta(M_g)|}{|M_g|} \quad (6)$$

$$E_{AR} = \frac{|AR_p - AR_g|}{AR_g} \quad (7)$$

Metrik berbasis piksel dan kontur belum tentu mencerminkan kualitas visual yang dirasakan penonton. Karena itu, pada data utama ditambahkan dua metrik perseptual berbasis fitur jaringan saraf, yaitu LPIPS dengan tulang punggung AlexNet [41] dan DISTS [42], yang dihitung pada ROI yang sama dengan SSIM-ROI. Nilai yang lebih rendah menandakan kemiripan perseptual yang lebih baik. Keduanya diperlakukan sebagai pelengkap, bukan pengganti penilaian manusia.

Frame berdeformasi tinggi ditentukan dengan indeks deformasi D_t pada Persamaan (8), yaitu satu dikurangi IoU antara siluet frame sebelumnya yang digeser agar pusat massanya berimpit dengan frame saat ini (operator T) dan siluet frame saat ini. Karena translasi dihilangkan, indeks ini menangkap perubahan bentuk, termasuk penggantian gambar secara menyeluruh pada *smear frame*. Frame dengan D_t pada kuartil ke-3 teratas disebut berdeformasi tinggi.

$$D_t = 1 - IoU(T(M_{t-1}), M_t) \quad (8)$$

Seluruh metrik dihitung hanya pada frame uji, yaitu frame yang tidak menjadi frame kunci dan memiliki siluet asli yang valid. Frame kunci tidak dievaluasi karena rekonstruksinya identik dengan gambar asli. Pemisahan ini mencegah skor terdongkrak oleh frame yang trivial. Frame berdeformasi tinggi diidentifikasi sekali dari urutan *ground truth* penuh, sehingga himpunan ini tidak bergantung pada interval atau metode yang diuji dan perbandingan antarinterval tetap adil.

2.7. Skema Pengujian per *Layer*

Selain pengujian pada frame komposit (skema gabungan langsung, seperti pada bagian sebelumnya), empat skema per *layer* diuji dengan lima metode dan tiga interval yang sama. Pada skema *Character*, hanya *layer Character* yang diinterpolasi dan hasilnya dibandingkan dengan *layer Character* asli. Skema efek berlaku sama untuk gabungan *layer* efek dan asap, dan skema asap untuk *layer* asap saja. Pada skema komposit per *layer*, ketiga *layer* diinterpolasi secara terpisah, lalu pada setiap piksel dipilih *layer* dengan selisih terbesar terhadap latar B menurut Persamaan (9), kemudian hasilnya dibandingkan dengan frame komposit asli.

$$\hat{\mathrm{I}}_t^{\mathrm{komp}} = \hat{\mathrm{I}}_t^{l*}, \;\; l* = \mathrm{argmax}_l |\hat{\mathrm{I}}_t^l - B| \quad (9)$$

Pada Persamaan (9), l adalah indeks *layer* (karakter, efek, atau asap) dan B adalah warna latar. Siluet, ROI, dan indeks deformasi dihitung per *layer* pada skema per *layer*, dengan ambang kuartil ketiga yang dihitung ulang pada tiap *layer*. Frame yang *layer*-nya kosong dikeluarkan, sehingga jumlah frame uji berbeda antarskema. Skema komposit per *layer* memakai himpunan frame uji yang sama dengan frame komposit sehingga dapat dibandingkan langsung dengan interpolasi pada gambar gabungan menggunakan uji berpasangan.

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : https://doi.org/10.58602/itsecs.v5i1.504
Page 187-209

2.8. Analisis Sensitivitas terhadap Latar dan Data Awal

Karena warna latar komposisi dapat memengaruhi metrik berbasis piksel, seluruh *benchmark* diulang pada latar hitam dari sumber alfa yang sama, sehingga selisihnya hanya mencerminkan latar. Hasil juga dibandingkan dengan percobaan awal pada ekspor lama berlatar hitam tanpa kanal alfa. Perbandingan terakhir tidak sepenuhnya murni karena ekspor ulang dilakukan secara terpisah, sehingga dipakai hanya untuk menilai apakah urutan metode berubah, bukan untuk menghitung efek latar.

2.9. Analisis Statistik

Perbandingan antarmetode pada frame uji yang sama dilakukan dengan uji peringkat bertanda Wilcoxon, dengan koreksi Holm atas sepuluh pasangan metode untuk setiap kombinasi interval, himpunan frame, dan metrik. Ukuran efek dilaporkan sebagai korelasi biserial peringkat pada Persamaan (10), dengan W_+ dan W_- adalah jumlah peringkat selisih positif dan negatif.

$$r = \frac{w^+ - w^-}{w^+ + w^-} \tag{10}$$

Karena frame berurutan saling berkorelasi, asumsi kemandirian uji Wilcoxon tidak terpenuhi sepenuhnya sehingga nilai p diperlakukan sebagai indikator pendukung. Rentang kepercayaan 95% untuk rerata selisih dihitung dengan *block bootstrap* (blok 5 frame, 5000 resampel, benih acak 42) dan dijadikan dasar utama interpretasi. Statistik deskriptif dilaporkan sebagai median dan kuartil. Seluruh perhitungan dijalankan dengan Python 3 (OpenCV, scikit-image, NumPy [43], SciPy [44], dan pandas); RIFE dan FILM dijalankan pada GPU di Google Colab, sedangkan analisis statistik dijalankan pada CPU. Hasil kelima metode pada frame komposit berasal dari satu proses eksekusi yang sama. Skrip dan data hasil per frame disertakan sebagai berkas pendukung naskah.

2.10. Lingkungan Komputasi dan Reproduksibilitas

Agar eksperimen dapat direplikasi pada lingkungan komputasi lain, konfigurasi perangkat keras, versi pustaka, dan parameter pemrosesan dirangkum pada Tabel 1. Metode klasik (*hold*, *cross-dissolve*, dan DIS) bersifat deterministik, sedangkan RIFE dan FILM dijalankan pada GPU sehingga digit terakhir keluaran dapat berbeda antarperangkat keras atau versi pustaka; pada dua kali menjalankan RIFE dan FILM di lingkungan yang sama, selisih F-score per frame tidak melebihi 0.005. Batas waktu *keyframe* mengikuti indeks 0, g, 2g, dan seterusnya, ditambah frame terakhir bila belum termasuk, dan posisi relatif frame yang direkonstruksi dihitung sebagai s = (t − a)/(b − a) untuk frame kunci a dan b.

Tabel 1. Konfigurasi komputasi, versi pustaka, dan parameter pemrosesan.

Komponen	Konfigurasi
Perangkat keras	GPU NVIDIA Tesla T4 (15360 MiB; *driver* 580.82.07) pada Google Colab
Bahasa dan pustaka	Python 3.13.16; NumPy 2.4.6; SciPy 1.16.3; pandas 2.2.3; scikit-image 0.25.2; OpenCV 5.0.0.93
Kerangka model	PyTorch 2.14.1+cu130 (RIFE, LPIPS, DISTS); TensorFlow 2.21.0 dan TensorFlow Hub 0.16.1 (FILM)
Pra-pemrosesan	Perkecil ke 960 × 540 (INTER_AREA, kanal *premultiplied*); latar abu-abu bernilai 128 (hitam 0); siluet: selisih terhadap latar > 8, penutupan 5 × 5, pengisian lubang; frame kosong: luas < 500 piksel
Hold dan *cross-dissolve*	*Hold*: salinan frame kunci sebelumnya; *cross-dissolve*: $(1 - s)\cdot I_a + s\cdot I_b$
Optical flow (DIS)	OpenCV DISOpticalFlow, preset MEDIUM, pada citra abu-abu; aliran dua arah; *warping* bilinear dengan nilai batas sama dengan warna latar

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209

Komponen	Konfigurasi
RIFE v4.25	Repositori Practical-RIFE (hzwer); bobot v4.25; masukan *BGR→RGB* ternormalisasi 0–1; *timestep* s; bantalan replikasi ke kelipatan 64 piksel (960 × 576), dipangkas kembali ke 960 × 540; pengaturan lain bawaan; versi *commit*: weights: ['IFNet_HDv3.py', 'RIFE_HDv3.py', '__pycache__', 'flownet.pkl', 'refine.py']
FILM	Model TensorFlow Hub google/film/1 (cadangan: Kaggle); masukan x0, x1, dan *time* = s; bantalan kelipatan 64 piksel; pengaturan bawaan
Metrik	SSIM (scikit-image, jendela bawaan 7 × 7, data_range 255) pada citra abu-abu; ROI: kotak pembatas gabungan siluet prediksi dan asli + margin 16 piksel; Canny 30/90 dengan toleransi 2 piksel; dilasi 7 × 7 untuk *ghosting*; LPIPS AlexNet dan DISTS pada ROI
Statistik	Uji Wilcoxon berpasangan, koreksi Holm, *block bootstrap* (blok 5 frame, 5000 ulangan, *seed* 42); ambang deformasi tinggi: kuartil ketiga D_t (0.331; n = 183)
Ketersediaan skrip	[penulis: cantumkan tautan repositori atau pernyataan ketersediaan skrip dan data, bila ada]

Rincian versi tercatat pada berkas lingkungan.txt yang dihasilkan notebook penelitian.

3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil *benchmark* pada frame komposit berlatar abu-abu sebagai data utama, hasil uji statistik, hasil skema per *layer*, analisis sensitivitas latar, dan pembahasan yang menghubungkan temuan dengan penelitian terdahulu. Seluruh angka berasal dari satu adegan dan satu animator sehingga diperlakukan sebagai bukti awal, bukan estimasi populasi.

3.1. Hasil Penelitian

Karakteristik dataset dirangkum pada Tabel 2. Dari 192 frame komposit, 151 frame (78.6%) merupakan gambar unik dan 40 frame identik dengan frame sebelumnya, yang menandakan sebagian kecil gerak ditampilkan dengan penahanan gambar. *Layer Character*, gabungan efek dan asap, serta asap saja masing-masing memuat 128, 90, dan 67 gambar unik. Frame 1 dan frame 186 sampai 192 tidak memuat objek, sehingga 184 frame (frame 2 sampai 185) memenuhi ambang luas siluet dan dianalisis.

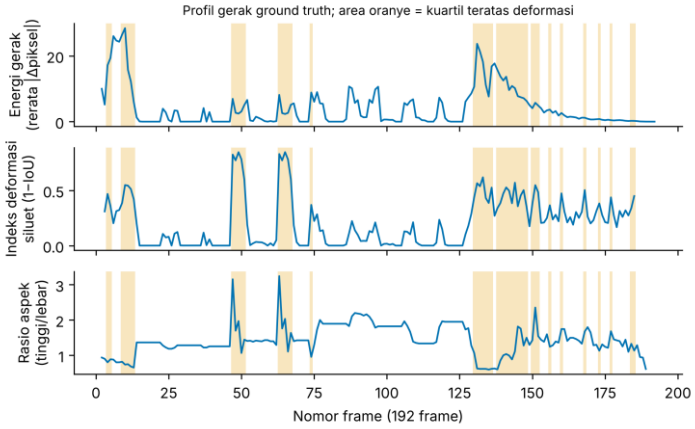
Tabel 2. Karakteristik dataset adegan animasi 2D.

Parameter	Nilai
Perangkat lunak produksi	Toon Boom Harmony 25.2 Premium
Jumlah frame (komposit)	192
Format ekspor / *layer*	PNG RGBA / 3 (karakter, efek, asap)
Resolusi ekspor / resolusi analisis	3556 × 2000 / 960 × 540 piksel
Frame unik (komposit)	151 (78.6%)
Frame identik dengan frame sebelumnya	40
Frame dianalisis (luas siluet ≥ 500 piksel)	184 (frame 2–185)

Gambar 4 menyajikan profil gerak *ground truth*. Energi gerak rerata antarframe sebesar 4.06 dengan nilai maksimum 28.43, sedangkan indeks deformasi memiliki median 0.139 dan kuartil ketiga 0.331 (n = 183). Dengan ambang kuartil ketiga tersebut, 46 frame ditandai sebagai berdeformasi tinggi. Frame-frame ini terkonsentrasi pada tiga segmen: pembukaan adegan, dua rangkaian gambar ekor yang menggantikan karakter selama lima frame (frame 47 sampai 51 dan 63 sampai 67) dengan rasio aspek siluet melonjak hingga sekitar 3.2, serta sekuens karakter terpental dan jatuh pada frame sekitar 130 sampai 185. Pola ini menunjukkan bahwa deformasi tinggi pada adegan ini sebagian besar berupa penggantian bentuk secara menyeluruh, bukan sekadar perubahan proporsi, sehingga merepresentasikan pemakaian *smear* dan gambar efek yang disengaja.

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209



Gambar 4. Profil gerak *ground truth*; area oranye menandai frame berdeformasi tinggi.

Konfigurasi tiga interval dirangkum sebagai berikut. Pada interval 2, 4, dan 8, jumlah gambar kunci masing-masing 97, 49, dan 25, sehingga *proksi* beban kerja berkurang 35.8%, 67.5%, dan 83.4% terhadap 151 gambar unik pada versi manual (1 − gambar kunci/151), sementara jumlah frame uji meningkat dari 92 menjadi 138 dan 161, dengan 24, 35, dan 40 di antaranya berdeformasi tinggi. Pada interval 4 dan 8, tiap gambar kunci mewakili empat sampai delapan frame.

Jumlah frame dilaporkan secara seragam pada Tabel 3 untuk setiap interval dan skema pengujian. Dari 192 frame, jumlah frame non-kunci yang dibuang adalah 95, 143, dan 167 pada interval 2, 4, dan 8. Frame dengan luas siluet < 500 piksel (frame 1 dan frame 186–192, yang berisi latar kosong sebelum dan sesudah adegan) dikeluarkan dari pengujian karena metrik berbasis siluet tidak terdefinisi pada gambar kosong. Frame tersebut tidak dibuang dari data; hanya frame non-kunci yang kosong yang tidak dinilai, yaitu 3, 5, dan 6 frame pada interval 2, 4, dan 8, sehingga frame uji menjadi 92, 138, dan 161. Frame berdeformasi tinggi didefinisikan dari kuartil ketiga indeks deformasi atas 183 frame yang memiliki frame sebelumnya valid, lalu dihitung pada himpunan frame uji tiap skema. Pada skema per *layer*, frame uji berkurang karena *layer* efek dan asap kosong pada sebagian frame.

Tabel 3. Rincian jumlah frame: frame non-kunci, frame uji, dan frame berdeformasi tinggi untuk tiap skema dan interval.

Skema	Non-kunci 2	Non-kunci 4	Non-kunci 8	Uji / deform. 2	Uji / deform. 4	Uji / deform. 8
Gabungan langsung / komposit	95	143	167	92 / 24	138 / 35	161 / 40
Character	95	143	167	90 / 24	135 / 36	157 / 40
Efek + asap	95	143	167	40 / 12	62 / 16	73 / 18
Asap	95	143	167	30 / 7	45 / 9	53 / 12

Sel kanan: frame uji / frame berdeformasi tinggi. Non-kunci: seluruh frame yang dibuang pada interval tersebut (95, 143, dan 167). Persentase frame berdeformasi tinggi terhadap frame uji pada skema gabungan: 26.1%, 25.4%, dan 24.8%.

3.2. Hasil Pengujian pada Seluruh Frame

Pada seluruh frame uji, SSIM-ROI median berada pada 0.732 sampai 0.775 pada interval 2 dan 0.655 sampai 0.758 pada interval 8 (Tabel 4), dengan *cross-dissolve* tertinggi pada ketiga interval (0.775, 0.749, dan 0.758). F-score kontur membedakan

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209

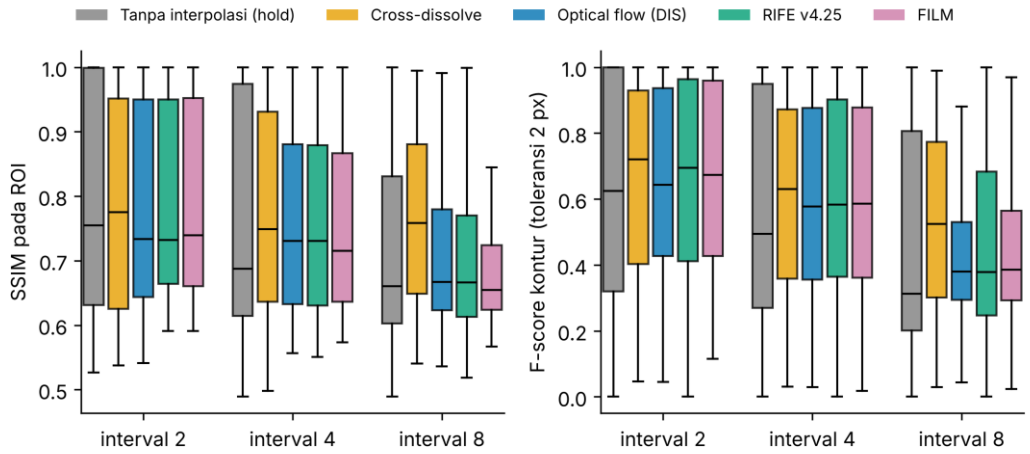
metode lebih jelas, dan *cross-dissolve* kembali tertinggi pada ketiga interval (0.721, 0.631, dan 0.524), sedangkan model pembelajaran mendalam tidak melampauinya: RIFE mencapai 0.694, 0.583, dan 0.378, FILM 0.674, 0.586, dan 0.386, dan DIS 0.643, 0.577, dan 0.379. *Hold* terendah pada ketiga interval (0.625, 0.494, dan 0.313). Indeks *ghosting cross-dissolve* naik tajam seiring interval (0.111, 0.186, dan 0.322), sedangkan RIFE hanya 0.019, 0.022, dan 0.043, terendah di antara kelima metode pada ketiga interval. *Hold* memiliki *ghosting* rendah (0.043 sampai 0.071) karena tidak mencampur frame, tetapi IoU dan F-score-nya rendah.

Tabel 4. Median metrik pada seluruh frame uji (SSIM-ROI, F-score kontur, IoU, dan indeks *ghosting*).

Interval	Metode	SSIM-ROI	F-score	IoU	*Ghosting*
2	*Hold*	0.755	0.625	0.808	0.043
	Cross-dissolve	0.775	0.721	0.837	0.111
	Optical flow (DIS)	0.734	0.643	0.849	0.052
	RIFE v4.25	0.732	0.694	0.865	0.019
	FILM	0.739	0.674	0.863	0.032
4	*Hold*	0.687	0.494	0.741	0.069
	Cross-dissolve	0.749	0.631	0.769	0.186
	Optical flow (DIS)	0.730	0.577	0.792	0.082
	RIFE v4.25	0.730	0.583	0.842	0.022
	FILM	0.715	0.586	0.790	0.043
8	*Hold*	0.661	0.313	0.618	0.071
	Cross-dissolve	0.758	0.524	0.715	0.322
	Optical flow (DIS)	0.667	0.379	0.632	0.188
	RIFE v4.25	0.666	0.378	0.662	0.043
	FILM	0.655	0.386	0.613	0.118

n frame uji: 92 (interval 2), 138 (interval 4), 161 (interval 8). Sebaran lengkap pada Gambar 5.

Sebaran yang lebar pada Gambar 5 menunjukkan bahwa nilai agregat mencampur frame statis yang mudah direkonstruksi dengan frame dinamis yang sulit. Hal ini mendorong analisis terstratifikasi berikut.



Gambar 5. Sebaran SSIM-ROI dan F-score kontur pada seluruh frame uji untuk lima metode dan tiga interval.

3.3. Hasil Pengujian pada Frame Berdeformasi Tinggi

Kinerja kelima metode turun tajam pada frame berdeformasi tinggi. F-score median hanya 0.176 sampai 0.431 (Tabel 5), dibandingkan 0.313 sampai 0.721 pada seluruh

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

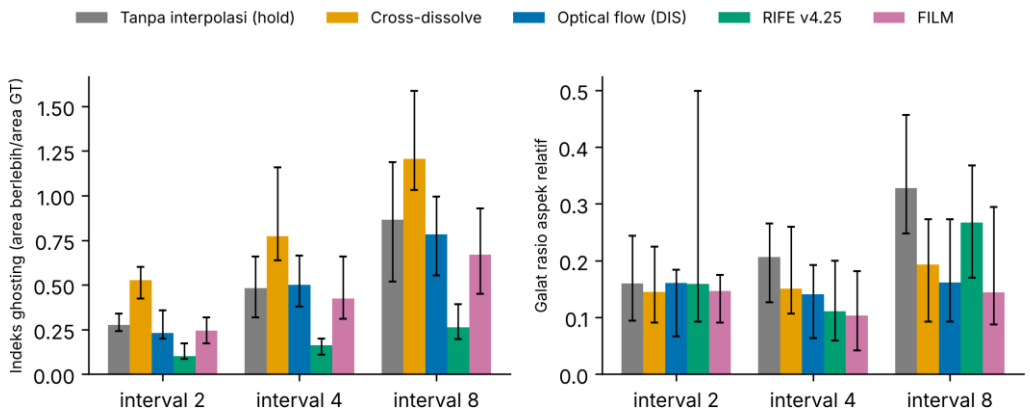
Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209

frame (Tabel 4). Pada interval 4, penurunan relatif F-score adalah 58.2% untuk *cross-dissolve* (0.631 menjadi 0.264), 48.0% untuk RIFE (0.583 menjadi 0.303), 48.6% untuk FILM (0.586 menjadi 0.301), 53.4% untuk DIS (0.577 menjadi 0.269), dan 51.2% untuk *hold* (0.494 menjadi 0.241). Indeks *ghosting cross-dissolve* mencapai 0.527, 0.773, dan 1.207 pada interval 2, 4, dan 8, sehingga pada interval 8 luas bayangan melebihi luas siluet asli. RIFE menekan *ghosting* menjadi 0.100, 0.163, dan 0.262 (Gambar 6a), sedangkan DIS dan FILM berada pada 0.230 sampai 0.782. Galat luas siluet RIFE juga terendah (0.080, 0.157, dan 0.172; *cross-dissolve* 0.537, 0.715, dan 0.967). Namun galat rasio aspek tidak membaik secara konsisten (Gambar 6b): pada interval 8, RIFE (0.267) mendekati *hold* (0.328) dan lebih besar daripada *cross-dissolve* (0.193), sedangkan FILM terendah (0.144). Pada IoU siluet, RIFE tertinggi pada interval 2 dan 4 (0.688 dan 0.559) tetapi turun menjadi 0.293 pada interval 8, di bawah FILM (0.376).

Tabel 5. Median metrik pada frame berdeformasi tinggi.

Interval	Metode	SSIM-ROI	F-score	IoU	*Ghosting*
2	*Hold*	0.716	0.313	0.493	0.275
	Cross-dissolve	0.701	0.372	0.534	0.527
	Optical flow (DIS)	0.731	0.416	0.647	0.230
	RIFE v4.25	0.725	0.425	0.688	0.100
	FILM	0.730	0.431	0.662	0.244
4	*Hold*	0.681	0.241	0.308	0.483
	Cross-dissolve	0.724	0.264	0.455	0.773
	Optical flow (DIS)	0.727	0.269	0.491	0.500
	RIFE v4.25	0.731	0.303	0.559	0.163
	FILM	0.708	0.301	0.517	0.425
8	*Hold*	0.683	0.176	0.166	0.864
	Cross-dissolve	0.756	0.251	0.311	1.207
	Optical flow (DIS)	0.766	0.246	0.342	0.782
	RIFE v4.25	0.731	0.217	0.293	0.262
	FILM	0.691	0.256	0.376	0.669

n frame uji berdeformasi tinggi: 24 (interval 2), 35 (interval 4), 40 (interval 8).



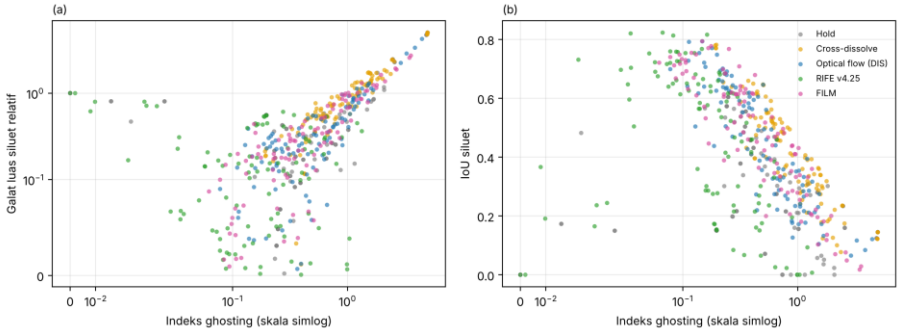
Gambar 6. Median indeks *ghosting* (a) dan galat rasio aspek relatif (b) pada frame berdeformasi tinggi, dengan rentang kepercayaan 95% *bootstrap*.

Ghosting rendah pada RIFE perlu dibaca bersama ukuran bentuk. Pada frame berdeformasi tinggi, *ghosting* tidak berkorelasi dengan galat luas siluet pada RIFE (Spearman $\rho = -0.04$), padahal pada *cross-dissolve*, DIS, dan FILM keduanya sangat berkaitan ($\rho = 0.94$, 0.78, dan 0.87). Pada metode yang menumpuk dua gambar, *ghosting* tinggi adalah akibat langsung dari luas siluet yang berlebih, sehingga indeks tersebut

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209

mengikuti galat luas. RIFE tidak menumpuk gambar, sehingga *ghosting*-nya rendah meskipun bentuk tidak cocok. Dari 99 frame berdeformasi tinggi, 31 frame RIFE memiliki *ghosting* < 0.1, tetapi 7 di antaranya memiliki IoU < 0.5 (Gambar 7). Dua frame pada interval 8 memiliki galat luas ≥ 0.95, yaitu prediksi dengan luas siluet jauh berbeda dari gambar asli. Dengan demikian, *ghosting* rendah menandakan tidak adanya bayangan ganda, bukan kesesuaian bentuk; sebagian frame RIFE kehilangan bagian siluet atau menghasilkan bentuk yang rapi tetapi salah. Hal ini menjelaskan mengapa penurunan *ghosting* terbesar (RIFE − *cross-dissolve* = −1.058 pada interval 8) tidak diikuti kenaikan F-score yang sepadan.



Gambar 7. Indeks *ghosting* terhadap IoU siluet pada frame berdeformasi tinggi (semua interval) untuk lima metode.

Gambar 8 memperlihatkan contoh kualitatif pada interval 4. Pada frame 47 dan 63, gambar asli hanya berisi ekor yang menggantikan karakter. *Hold* menampilkan karakter utuh dari frame kunci terdekat, *cross-dissolve* menampilkan tumpang tindih transparan antara ekor dan tubuh karakter, dan DIS menghasilkan tekstur robek. RIFE dan FILM menampilkan ekor dengan bayangan karakter yang lebih lemah, tetapi tampaknya karena keluarannya condong ke frame kunci terdekat. Pada frame 50, ekor berubah menjadi bentuk menyebar yang tidak ada pada kedua frame kunci sehingga tidak ada metode yang dapat merekonstruksinya. Ini adalah ciri frame *smear*: gambar diganti, bukan digeser.



Gambar 8. Contoh kualitatif pada interval 4 untuk frame 63, 47, dan 50: *ground truth*, *hold*, *cross-dissolve*, *optical flow* DIS, RIFE, dan FILM.

3.4. Uji Statistik

Tabel 6 merangkum uji berpasangan F-score dan *ghosting* pada frame berdeformasi tinggi (hasil SSIM-ROI, IoU, dan galat lain tersedia pada berkas pendukung). Semua metode berbasis gerak menurunkan *ghosting* dibanding *cross-dissolve* pada ketiga interval (p terkoreksi < 0.001, rentang kepercayaan tidak mencakup nol). Penurunan terbesar terjadi pada RIFE, dengan selisih rerata −0.407, −0.932, dan −1.058, dan RIFE juga lebih

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209

rendah daripada DIS pada ketiga interval (−0.186, −0.525, dan −0.551). Pada F-score, tidak ada metode yang unggul atas *cross-dissolve* pada frame berdeformasi tinggi: seluruh rentang kepercayaan pada interval 4 mencakup nol, dan RIFE pada interval 8 justru lebih rendah (−0.023; p terkoreksi 0.035). Terhadap *hold*, DIS dan FILM memperbaiki F-score secara konsisten (selisih +0.064 sampai +0.092, p terkoreksi ≤ 0.005), demikian pula RIFE (+0.065 sampai +0.070, p terkoreksi ≤ 0.030). Pada seluruh frame interval 8, *cross-dissolve* juga unggul atas ketiga interpolator pada F-score (selisih −0.105 untuk DIS, −0.072 untuk RIFE, dan −0.101 untuk FILM; p terkoreksi ≤ 0.005).

Tabel 6. Uji berpasangan pada frame berdeformasi tinggi: selisih rerata (Δ) dengan rentang kepercayaan 95% *block bootstrap*.

Interval	Metrik	DIS − Dissolve	RIFE − Dissolve	FILM − Dissolve	RIFE − DIS
2	F-score	+0.031 [0.003, 0.058]	+0.013 [−0.038, 0.044]	+0.029 [−0.013, 0.051]	−0.018 [−0.045, −0.007]
	Ghosting	−0.221*** [−0.294, −0.141]	−0.407*** [−0.483, −0.341]	−0.243*** [−0.296, −0.165]	−0.186*** [−0.284, −0.115]
4	F-score	−0.002 [−0.025, 0.011]	+0.005 [−0.044, 0.036]	+0.003 [−0.027, 0.017]	+0.006 [−0.033, 0.044]
	Ghosting	−0.407*** [−0.583, −0.269]	−0.932*** [−1.411, −0.575]	−0.455*** [−0.688, −0.263]	−0.525*** [−0.883, −0.290]
8	F-score	−0.003 [−0.040, 0.020]	−0.023* [−0.057, 0.003]	+0.004 [−0.035, 0.022]	−0.021 [−0.051, 0.004]
	Ghosting	−0.507*** [−0.710, −0.317]	−1.058*** [−1.420, −0.724]	−0.532*** [−0.740, −0.276]	−0.551*** [−0.793, −0.327]

Tanda bintang menunjukkan p terkoreksi Holm: * < 0.05, ** < 0.01, *** < 0.001; tanpa tanda berarti tidak signifikan. Δ > 0 berarti metode pertama lebih tinggi.

3.5. Validasi Perseptual Otomatis

Tabel 7 melaporkan LPIPS dan DISTS pada ROI untuk seluruh frame uji dan frame berdeformasi tinggi; nilai yang lebih rendah lebih baik. Metrik ini mengukur jarak fitur jaringan saraf sehingga lebih dekat dengan penilaian visual daripada SSIM, tetapi tetap merupakan estimasi, bukan penilaian penonton.

Tabel 7. Median LPIPS dan DISTS pada ROI (nilai lebih rendah lebih baik).

Interval	Metode	LPIPS (semua)	DISTS (semua)	LPIPS (deform.)	DISTS (deform.)
2	*Hold*	0.269	0.148	0.447	0.187
	Cross-dissolve	0.239	0.171	0.425	0.209
	Optical flow (DIS)	0.214	0.191	0.344	0.228
	RIFE v4.25	0.172	0.152	0.309	0.203
	FILM	0.193	0.180	0.317	0.266
4	*Hold*	0.375	0.182	0.538	0.243
	Cross-dissolve	0.300	0.187	0.486	0.231
	Optical flow (DIS)	0.280	0.204	0.465	0.247
	RIFE v4.25	0.246	0.174	0.405	0.221
	FILM	0.282	0.244	0.432	0.308
8	*Hold*	0.531	0.211	0.612	0.269
	Cross-dissolve	0.405	0.203	0.529	0.260
	Optical flow (DIS)	0.436	0.231	0.503	0.270
	RIFE v4.25	0.398	0.204	0.530	0.264
	FILM	0.440	0.297	0.566	0.326

Data utama (latar abu-abu, gabungan langsung).

Hasilnya sejalan dengan analisis bentuk. LPIPS sangat berkaitan dengan F-score kontur (Spearman ρ = −0.94 pada 1.955 pengamatan), sedangkan DISTS berkaitan lebih lemah

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

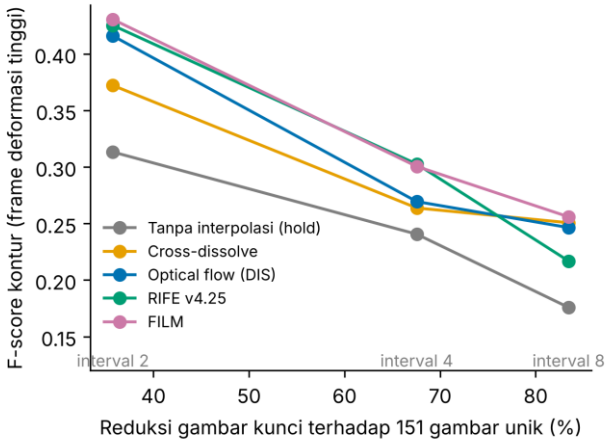
Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : https://doi.org/10.58602/itsecs.v5i1.504
Page 187-209

(ρ = −0.82). RIFE memiliki median LPIPS terendah pada seluruh frame di ketiga interval, dan pada frame berdeformasi tinggi di interval 2 dan 4 (0.309 dan 0.405); pada interval 8 selisihnya mengecil, dan DIS (0.503) lebih rendah daripada RIFE (0.530). Uji Wilcoxon berpasangan pada frame berdeformasi tinggi (eksploratori, tanpa koreksi Holm) menunjukkan LPIPS RIFE lebih rendah daripada *cross-dissolve* pada interval 2 dan 4 (selisih rerata −0.066 dan −0.051; $p \leq 0.006$) dan lebih rendah daripada *hold* pada ketiga interval (−0.086 sampai −0.109; $p \leq 0.001$), tetapi tidak berbeda nyata dari DIS dan FILM, kecuali FILM pada interval 4 ($p = 0.019$). DISTS lebih beragam: median terendah pada seluruh frame dicapai *hold* di interval 2, RIFE di interval 4, dan *cross-dissolve* di interval 8 dengan selisih sangat tipis terhadap RIFE. Pada frame berdeformasi tinggi, RIFE tidak berbeda jelas dari *cross-dissolve* dan *hold* menurut DISTS, tetapi lebih baik daripada FILM yang memiliki DISTS terburuk (selisih −0.048 sampai −0.077; $p < 0.001$). Dengan demikian, keunggulan RIFE pada *ghosting* hanya sebagian terkonfirmasi secara perseptual: terlihat pada LPIPS di interval kecil, tetapi tidak pada DISTS maupun pada interval 8.

Perlu ditegaskan bahwa kesesuaian kontur, siluet, dan metrik perseptual otomatis tidak sama dengan kualitas gerak yang dirasakan penonton. Gerak yang tampak halus saat diputar dapat berasal dari frame yang bentuknya berbeda dari gambar asli, dan sebaliknya frame yang cocok secara kontur dapat terasa kaku dalam urutan. Pada animasi, penyimpangan dari gambar tangan kadang justru bernilai artistik. Karena itu hasil penelitian ini dibaca sebagai ukuran *kedekatan terhadap gambar tangan*, bukan ukuran kualitas gerak. Uji persepsi oleh penonton pada urutan frame yang diputar menjadi pengembangan lanjutan yang diperlukan.

3.6. *Trade-off* Beban Kerja dan Kualitas

Gambar 9 memetakan F-score pada frame berdeformasi tinggi terhadap reduksi gambar kunci. Kenaikan reduksi dari 35.8% ke 67.5% dan 83.4% menurunkan F-score median RIFE dari 0.425 menjadi 0.303 dan 0.217, DIS dari 0.416 menjadi 0.269 dan 0.246, FILM dari 0.431 menjadi 0.301 dan 0.256, serta *cross-dissolve* dari 0.372 menjadi 0.264 dan 0.251. Pada reduksi 67.5% dan 83.4%, selisih antarmetode mengecil dan tidak konsisten (Tabel 6), sehingga pemilihan model yang lebih kompleks tidak memberi keuntungan kontur pada interval lebar. Dengan 97, 49, dan 25 gambar kunci, jumlah gambar yang harus digambar tangan turun dari 151 menjadi 25 pada interval 8, tetapi penghematan itu dibayar dengan penurunan F-score median pada frame berdeformasi tinggi sebesar 33% sampai 49% pada keempat metode non-*hold* dibanding interval 2.



Gambar 9. *Trade-off* antara reduksi gambar kunci (proksi beban kerja) dan F-score kontur pada frame berdeformasi tinggi.

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209

3.7. Hasil Pengujian per *Layer*

Skema per *layer* diuji pada tiga interval yang sama. Pada interval 4, jumlah frame uji adalah 138 untuk skema komposit, 135 untuk *layer Character*, 62 untuk gabungan efek dan asap, dan 45 untuk asap saja, dengan frame berdeformasi tinggi sebanyak 35, 36, 16, dan 9 (pada interval 2 dan 8: 24 dan 40 untuk *Character*, 12 dan 18 untuk efek, serta 7 dan 12 untuk asap). Tabel 8 dan Gambar 10 merangkum hasil interval 4; hasil ketiga interval disertakan pada berkas pendukung.

Tabel 8. Median F-score kontur, indeks *ghosting*, dan galat luas pada interval 4 untuk lima skema pengujian.

Skema	Metode	F-score (semua)	*Ghosting* (semua)	F-score (deform.)	*Ghosting* (deform.)	Galat luas (deform.)
Gabungan langsung	*Hold*	0.494	0.069	0.241	0.483	0.377
	Cross-dissolve	0.631	0.186	0.264	0.773	0.715
	Optical flow (DIS)	0.577	0.082	0.269	0.500	0.304
	RIFE v4.25	0.583	0.022	0.303	0.163	0.157
	FILM	0.586	0.043	0.301	0.425	0.347
Komposit per *layer*	*Hold*	0.475	0.069	0.241	0.483	0.377
	Cross-dissolve	0.600	0.186	0.263	0.773	0.715
	Optical flow (DIS)	0.559	0.083	0.259	0.546	0.334
	RIFE v4.25	0.577	0.015	0.289	0.160	0.158
	FILM	0.552	0.043	0.299	0.425	0.284
Character	*Hold*	0.522	0.045	0.252	0.463	0.267
	Cross-dissolve	0.638	0.110	0.340	0.829	0.837
	Optical flow (DIS)	0.667	0.055	0.332	0.329	0.293
	RIFE v4.25	0.645	0.008	0.385	0.136	0.121
	FILM	0.636	0.027	0.349	0.324	0.199
Efek + asap	*Hold*	0.155	0.173	0.078	0.385	0.513
	Cross-dissolve	0.236	0.460	0.121	1.027	0.501
	Optical flow (DIS)	0.263	0.197	0.150	0.715	0.324
	RIFE v4.25	0.222	0.069	0.147	0.074	0.567
	FILM	0.280	0.139	0.163	0.342	0.279
Asap	*Hold*	0.133	0.235	0.067	0.322	0.819
	Cross-dissolve	0.219	0.486	0.100	0.560	0.552
	Optical flow (DIS)	0.247	0.195	0.089	0.553	0.549
	RIFE v4.25	0.198	0.069	0.078	0.033	0.632
	FILM	0.258	0.125	0.119	0.273	0.288

n frame uji (semua / berdeformasi tinggi): gabungan langsung 138 / 35, komposit 138 / 35, *Character* 135 / 36, efek + asap 62 / 16, asap 45 / 9.

Pada *layer Character*, pola yang muncul mirip dengan gambar gabungan. Pada frame berdeformasi tinggi, F-score kontur median berada pada 0.252 sampai 0.385 untuk kelima metode, dengan *hold* terendah dan RIFE tertinggi (0.385), sedangkan RIFE juga menghasilkan *ghosting* terendah (0.136 dibanding 0.829 untuk *cross-dissolve*). Pada seluruh frame, DIS memiliki F-score median tertinggi pada *layer* ini (0.667 dibanding 0.638 untuk *cross-dissolve*), perbedaan kecil yang tidak diuji secara statistik karena *ground truth*-nya berbeda dari frame komposit.

Menginterpolasi *layer* secara terpisah lalu menggabungkannya tidak memperbaiki rekonstruksi dibanding interpolasi pada gambar gabungan. Pada frame berdeformasi tinggi interval 4, F-score median skema komposit per *layer* identik atau lebih rendah hingga 0.014 dari skema gabungan langsung (*hold* 0.241 pada keduanya, FILM 0.299 dan 0.301, RIFE 0.289 dan 0.303, dan DIS 0.259 dan 0.269). Uji Wilcoxon berpasangan atas 45 kombinasi (lima metode, tiga metrik, tiga interval) pada frame berdeformasi tinggi hanya

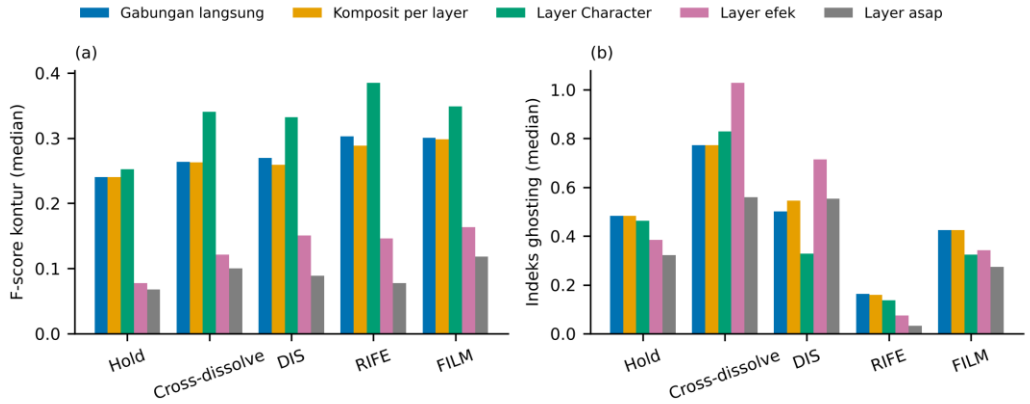
Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209

menghasilkan tujuh selisih signifikan setelah koreksi Holm. Selisih F-score signifikan justru negatif (*cross-dissolve* pada interval 2 dan 4 serta DIS pada interval 8, antara −0.019 dan −0.028), *ghosting* DIS pada interval 8 naik 0.095, dan tiga selisih SSIM-ROI berukuran 0.004 sampai 0.005 yang dapat diabaikan secara praktis.

Layer efek lebih sulit direkonstruksi dan datanya lebih sedikit. Pada frame berdeformasi tinggi, F-score median gabungan efek dan asap hanya 0.078 sampai 0.163 (n = 16 pada interval 4), dan *ghosting cross-dissolve* mencapai 1.027, artinya luas bayangan melebihi luas efek. RIFE kembali menghasilkan *ghosting* terendah (0.074), tetapi galat luas siluetnya justru tertinggi (0.567 dibanding 0.279 sampai 0.513 pada keempat metode lain), yang mengindikasikan bahwa *ghosting* rendah pada *layer* efek dapat dicapai dengan menghilangkan sebagian efek, bukan dengan merekonstruksinya. Pada *layer* asap saja pola yang sama muncul (F-score 0.067 sampai 0.119; galat luas RIFE 0.632 dengan *ghosting* 0.033). Karena jumlah frame berdeformasi tinggi pada kedua *layer* ini hanya 9 sampai 18, hasilnya bersifat indikatif.

Ukuran sampel memengaruhi keterpercayaan perbandingan antar-*layer*. Rentang kepercayaan 95% *block bootstrap* untuk median F-score RIFE pada frame berdeformasi tinggi interval 4 berlebar 0.204 pada skema komposit (n = 35), 0.376 pada *Character* (n = 36), 0.561 pada efek dan asap (n = 16), dan 0.047 pada asap (n = 9); lebar untuk *cross-dissolve* berturut-turut 0.145, 0.297, 0.609, dan 0.079. Rentang pada *layer* efek dan asap mencakup hampir seluruh selang nilai F-score (0.023 sampai 0.584 untuk RIFE), sehingga urutan metode pada *layer* tersebut tidak dapat dipastikan. Rentang asap yang sempit dengan n = 9 mencerminkan nilai yang rendah dan seragam, bukan presisi yang tinggi. Karena itu, hasil per *layer* dibaca sebagai indikasi arah, dan kesimpulan yang lebih tegas hanya dapat ditarik pada skema komposit dan *Character*. Ketiadaan keunggulan RIFE pada *layer* efek dan asap konsisten dengan penjelasan di atas: bentuk efek berubah drastis antar-frame, sehingga RIFE menghasilkan *ghosting* rendah tetapi galat luas tinggi (0.567), dan jumlah frame yang kecil mencegah perbedaan tersebut diuji secara andal.



Gambar 10. Median F-score kontur (a) dan indeks *ghosting* (b) pada frame berdeformasi tinggi interval 4 untuk skema gabungan langsung, komposit per *layer*, *layer Character*, efek, dan asap.

3.8. Sensitivitas terhadap Latar dan Data Awal

Pengulangan *benchmark* pada latar hitam dari sumber alfa yang sama memperlihatkan bahwa latar terutama memengaruhi SSIM-ROI. Pada seluruh frame, SSIM-ROI pada latar abu-abu lebih tinggi 0.025 sampai 0.117 daripada pada latar hitam, dan selisih itu signifikan setelah koreksi Holm pada 14 dari 15 kombinasi metode dan interval (kecuali RIFE pada interval 2). Pengaruh pada F-score jauh lebih kecil: selisih rerata antarlatar berada pada −0.015 sampai +0.011, walaupun sebagian kombinasi tetap signifikan (10

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : https://doi.org/10.58602/itsecs.v5i1.504
Page 187-209

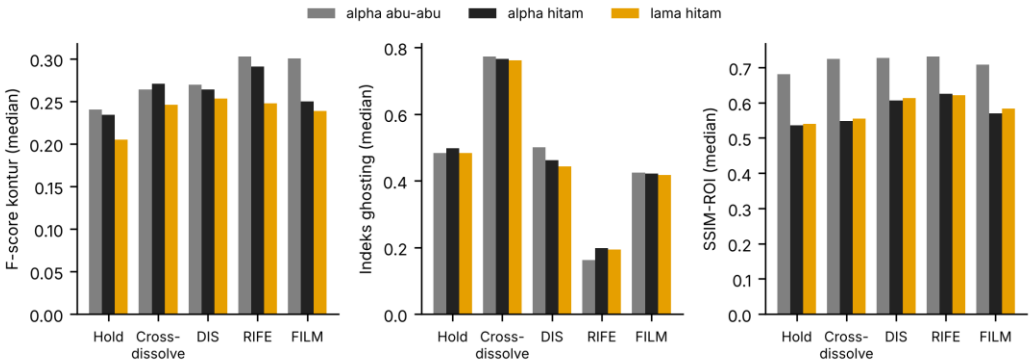
dari 15) karena selisih yang konsisten pada frame berurutan. Urutan metode tidak berubah pada interval 4 (Tabel 9): *cross-dissolve* tertinggi (pada data awal setara dengan RIFE) dan *hold* terendah pada F-score di ketiga data (abu-abu, hitam, dan data awal), sedangkan posisi kedua bergantian antara RIFE dan FILM. Dengan demikian kesimpulan tentang kecocokan kontur tidak bergantung pada warna latar, tetapi nilai SSIM-ROI antarpenelitian dengan latar berbeda tidak dapat dibandingkan langsung.

Tabel 9. Median SSIM-ROI dan F-score kontur pada interval 4 (seluruh frame) untuk tiga data: alfa abu-abu (utama), alfa hitam, dan ekspor awal berlatar hitam.

Metode	SSIM abu-abu	SSIM hitam	SSIM awal	F abu-abu	F hitam	F awal
Hold	0.687	0.628	0.641	0.494	0.489	0.421
Cross-dissolve	0.749	0.619	0.626	0.631	0.603	0.555
Optical flow (DIS)	0.730	0.652	0.660	0.577	0.573	0.523
RIFE v4.25	0.730	0.674	0.699	0.583	0.595	0.555
FILM	0.715	0.623	0.635	0.586	0.567	0.528

n = 138 frame uji pada ketiga data. Data awal adalah ekspor lama berlatar hitam tanpa kanal alfa.

Dibanding percobaan awal pada ekspor lama, F-score pada data utama lebih tinggi 0.008 sampai 0.075 (rerata 0.046 pada 15 kombinasi). Perbedaan ini tidak dapat dikaitkan seluruhnya dengan latar karena ekspor ulang dilakukan terpisah, termasuk kemungkinan perbedaan tepi dan penghalusan kanal alfa; yang dapat disimpulkan hanyalah bahwa urutan metode dan temuan utama bertahan pada kedua ekspor. Hasil ini juga menunjukkan bahwa laporan skor mutlak perlu menyebut cara ekspor dan warna latar secara eksplisit.



Gambar 11. Sensitivitas SSIM-ROI (a) dan F-score kontur (b) terhadap warna latar dan data awal pada interval 4.

3.9. Analisis Perbandingan

Temuan utama penelitian ini adalah bahwa penurunan kinerja interpolasi pada frame berdeformasi tinggi bukan sekadar akibat gerak yang lebih cepat, melainkan akibat sifat frame itu sendiri. Pada frame *smear*, bentuk yang digambar tidak berada pada lintasan antara dua frame kunci terdekat; interpolasi, yang menurut definisinya mencari korespondensi, tidak memiliki dasar untuk merekonstruksinya. Hal ini terlihat dari konsistensi hasil: F-score median pada frame berdeformasi tinggi berada di kisaran 0.18 sampai 0.43 untuk kelima metode dan semua interval, termasuk dua model pembelajaran mendalam. Pola ini sejalan dengan pengamatan bahwa gerak non-linear yang berlebihan dan gerak besar menyulitkan interpolasi animasi [25], [26], [27], dan dengan temuan bahwa pencocokan korespondensi pada gambar garis memerlukan penanganan khusus [24].

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209

Keunggulan setiap metode bersifat spesifik. RIFE paling efektif menekan *ghosting* dan galat luas siluet, tetapi galat rasio aspeknya pada interval 8 mendekati *hold*, dan contoh kualitatif memperlihatkan keluarannya cenderung mengikuti frame kunci terdekat (Gambar 8). Artinya, *ghosting* yang rendah tidak selalu berarti bentuk yang benar. FILM mirip DIS pada *ghosting*, tetapi menghasilkan karakter kabur pada frame ekor. Tidak ada model yang meningkatkan kecocokan kontur secara konsisten dibanding *cross-dissolve*, yang mendukung argumen bahwa pencocokan piksel untuk animasi bergaya garis dan blok warna bermasalah [28] dan bahwa model untuk video alami [22], [23], [19] tidak otomatis cocok untuk animasi [25]. Hasil ini juga mengingatkan bahwa metrik berbasis piksel seperti SSIM dan PSNR hanya sebagian menangkap kualitas persepsi [41], sehingga perbandingan SSIM-ROI yang hampir setara pada seluruh frame tidak berarti hasilnya setara bagi penonton.

Kegagalan pada frame *smear* tidak berarti metode yang diuji buruk secara umum. Pada seluruh frame, interpolator masih mengungguli *hold* secara konsisten pada interval kecil dan menengah, dan pada segmen penahanan gambar metode *hold* sudah memadai. Kesulitan terkonsentrasi pada sekitar seperempat frame yang paling berpengaruh terhadap kesan gerak, sehingga keputusan otomasi sebaiknya dibuat per frame atau per segmen, bukan per adegan. Pada interval 8, *cross-dissolve* bahkan mengungguli ketiga interpolator pada F-score seluruh frame; ini menandakan bahwa percampuran sederhana mempertahankan kontur lebih baik meskipun menghasilkan bayangan yang lebih luas.

Hasil per *layer* memperkuat interpretasi ini. Karena pemisahan *layer* tidak meningkatkan F-score, dan sebagian kombinasi justru menurun, kesulitan pada frame *smear* tampaknya bukan akibat tumpang tindih objek dalam satu gambar, melainkan akibat penggantian bentuk itu sendiri; interpolasi per *layer* tidak memberikan jalan pintas. Pada *layer* efek dan asap, tingginya galat luas RIFE menunjukkan bahwa ukuran *ghosting* perlu dibaca bersama ukuran bentuk, tidak berdiri sendiri. Analisis sensitivitas menambahkan bahwa kesimpulan ini tahan terhadap perubahan latar, sedangkan SSIM-ROI mutlak tidak.

Penelitian ini melengkapi studi interpolasi animasi yang umumnya melaporkan metrik agregat pada himpunan data publik [25], [26] dengan analisis terstratifikasi berdasarkan deformasi siluet pada gambar tangan produksi nyata. Berbeda dari [33], yang mengukur *trade-off* efisiensi dan kedekatan terhadap kurva referensi pada animasi 3D berbasis F-Curve, penelitian ini mengukur *trade-off* yang analog pada animasi 2D gambar tangan, dengan ukuran beban kerja berbasis jumlah gambar. Perbandingan dengan model generatif seperti [27], yang menyasar interpolasi kartun dengan gerak non-linear dan oklusi, menjadi pengembangan yang paling penting.

Hasil ini mengarah pada alur kerja hibrida: frame *smear* dan frame efek digambar sebagai gambar kunci, sedangkan segmen gerak halus dapat diisi interpolasi pada interval kecil. Indeks deformasi D_t dapat dipakai sebagai daftar periksa awal untuk menandai frame yang harus digambar tangan sebelum interpolasi dijalankan. Karena reduksi gambar kunci meningkat dari 35.8% menjadi 67.5% pada interval 4 dengan penurunan F-score seluruh frame yang lebih moderat dibanding interval 8, interval 4 tampak sebagai titik kompromi yang lebih aman pada adegan ini, meskipun kesimpulan ini harus dikonfirmasi pada adegan lain. Kualitas yang dinilai di sini bersifat geometris; penerimaan estetika oleh penonton harus diuji terpisah.

3.10. Keterbatasan

Penelitian ini memiliki beberapa keterbatasan. Pertama, data berasal dari satu adegan dan satu animator sehingga generalisasi ke gaya, durasi, dan tingkat keahlian lain belum terbukti. Kedua, RIFE dan FILM diuji tanpa pelatihan ulang pada gaya gambar penulis; model yang dilatih atau disesuaikan khusus untuk animasi [25], [27] dapat menghasilkan hasil berbeda. Ketiga, tidak ada evaluasi persepsi manusia; LPIPS [41] dan DISTS [42]

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209

hanya memperkirakan kesamaan visual terhadap gambar tangan dan belum menggantikan uji responden pada urutan frame yang diputar, yang disarankan pada penelitian lanjutan. Keempat, siluet diperoleh dengan ambang selisih terhadap latar sehingga piksel yang sangat mirip latar dapat terlewat, dan analisis dilakukan pada resolusi 960 × 540. Kelima, jumlah frame berdeformasi tinggi per interval relatif kecil (24 sampai 40) dan frame berurutan saling berkorelasi, sehingga nilai p bersifat pendukung. Keenam, hasil per *layer*, terutama *layer* efek dan asap, bertumpu pada 7 sampai 18 frame berdeformasi tinggi dengan rentang kepercayaan yang lebar sehingga hanya indikatif. Ketujuh, perbandingan dengan ekspor awal tidak murni karena ekspor dilakukan terpisah. Kedelapan, ukuran beban kerja adalah *proksi* berbasis jumlah gambar, bukan pengukuran waktu.

4. KESIMPULAN

Penelitian ini mengukur secara objektif bagaimana interpolasi otomatis, termasuk model pembelajaran mendalam, berperforma terhadap gambar tangan pada satu adegan animasi 2D. Kinerja rekonstruksi menurun tajam pada frame berdeformasi tinggi: F-score kontur median hanya 0.176 sampai 0.431 dibanding 0.313 sampai 0.721 pada seluruh frame. RIFE paling efektif menekan *ghosting*, tetapi tidak ada metode yang meningkatkan kecocokan kontur secara konsisten dibanding *cross-dissolve*, dan kelima metode pada dasarnya tidak mampu merekonstruksi frame *smear* yang menggantikan gambar secara menyeluruh. Memisahkan *layer* sebelum interpolasi tidak memberi perbaikan, *layer* efek dan asap lebih sulit dengan bukti yang masih indikatif, dan warna latar terutama memengaruhi SSIM tanpa mengubah urutan metode. Dengan demikian, hasil ini mendukung penggunaan interpolasi hanya pada segmen gerak halus dengan interval kecil, sementara frame *exaggeration* tetap perlu digambar tangan. Pada adegan ini, interval 4 tampak sebagai kompromi yang lebih aman bagi segmen non-*smear*. *Ghosting* rendah pada RIFE tidak menjamin kecocokan bentuk, dan kesesuaian kontur atau metrik perseptual otomatis tidak sama dengan kualitas gerak yang dirasakan penonton.

Kelebihan penelitian ini adalah protokol yang dapat direplikasi sepenuhnya, tidak bergantung pada responden, dan memberikan analisis terstratifikasi berdasarkan deformasi siluet, skema per *layer*, serta sensitivitas latar. Keterbatasannya adalah cakupan satu adegan, model tanpa pelatihan ulang, dan tidak adanya uji persepsi manusia. Pengembangan selanjutnya mencakup model generatif, penambahan adegan dan animator, metrik persepsi, serta survei responden.

PERNYATAAN KONTRIBUSI PENULIS (CRediT)

Deni Sinaga: Konseptualisasi, Metodologi, Perangkat lunak, Investigasi, Analisis formal, Kurasi data, Visualisasi, Penulisan – draf awal. **Ema Utami:** Supervisi, Validasi, Penulisan – peninjauan dan penyuntingan.

PERNYATAAN KONFLIK KEPENTINGAN

Para penulis menyatakan bahwa tidak terdapat kepentingan finansial atau hubungan pribadi yang diketahui yang dapat memengaruhi hasil penelitian yang dilaporkan dalam artikel ini.

PERNYATAAN PENDANAAN

Penulis tidak menerima dukungan finansial untuk penelitian, kepenulisan, dan/atau publikasi artikel ini.

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209

PERNYATAAN PENGGUNAAN AI GENERATIF DAN TEKNOLOGI BERBANTUAN AI DALAM PROSES PENULISAN

Gagasan, perumusan masalah, rancangan penelitian, pembuatan seluruh data animasi gambar tangan, pemilihan metrik dan protokol pengujian, pelaksanaan benchmark, analisis, interpretasi hasil, serta penarikan kesimpulan merupakan karya penulis. Penulis hanya memanfaatkan Claude (Anthropic) secara terbatas sebagai alat bantu pendukung, yaitu untuk membantu penyuntingan bahasa, penataan naskah sesuai templat jurnal, dan penyusunan skrip pendukung berdasarkan rancangan serta arahan penulis. Seluruh keluaran alat tersebut telah ditinjau, diuji, dan diverifikasi oleh penulis, yang bertanggung jawab penuh atas keaslian, ketepatan, dan isi naskah ini.

REFERENCES

[1] M. Chaerudin, A. Hariawan, and E. Purnawati, "Optimizing frame techniques in 2-dimensional video models in a learning perspective: Case study: Movie clip entitled "Perjalanan Jenderal Soedirman"," Journal of Multimedia Trend and Technology, vol. 3, no. 1, Apr. 2024, doi: 10.35671/jmtt.v3i1.45.

[2] X. Wang and W. Zhong, "Evolution and innovations in animation: A comprehensive review and future directions," Concurrency and Computation: Practice and Experience, vol. 36, no. 2, 2024, doi: 10.1002/cpe.7904.

[3] F. Thomas and O. Johnston, The Illusion of Life: Disney Animation. New York, NY, USA: Hyperion, 1981.

[4] J. Lasseter, "Principles of traditional animation applied to 3D computer animation," ACM SIGGRAPH Computer Graphics, vol. 21, no. 4, pp. 35–44, Jul. 1987, doi: 10.1145/37402.37407.

[5] S.-T. Lin and M.-T. Chi, "3D-to-2D animation smear effect technique based on Japanese hand-drawn animation style," in SIGGRAPH Asia 2024 Posters, 2024, doi: 10.1145/3681756.3697891.

[6] J. Basset, P. Bénard, and P. Barla, "SMEAR: Stylized motion exaggeration with art-direction," in ACM SIGGRAPH 2024 Conference Papers, 2024.

[7] S. A. Zega, G. J. P. Tambun, W. Adhitya, and I. N. Hamdi, "Penerapan prinsip-prinsip animasi pada film pendek animasi "Nohoax"," Journal of Applied Multimedia and Networking, vol. 6, no. 1, pp. 1–11, 2022, doi: 10.30871/jamn.v6i1.4179.

[8] A. Rahmi, M. F. A. Nasrullah, S. Prasetyaningsih, S. Indriyani, U. K. Nasution, H. N. Rizky, and D. Leviana, "Rig implementation using Duik Bassel plugin in 2D animation "Terjebak Hoaks"," in Proceedings of the 5th International Conference on Applied Engineering (ICAE 2022), 2023, doi: 10.4108/eai.5-10-2022.2327862.

[9] M. Even, P. Bénard, and P. Barla, "Non-linear rough 2D animation using transient embeddings," Computer Graphics Forum, vol. 42, no. 2, pp. 411–425, 2023, doi: 10.1111/cgf.14771.

[10] N. Barroso, A. Fondevilla, and D. Vanderhaeghe, "Automatic inbetweening for stroke-based painterly animation," Computer Graphics Forum, vol. 44, no. 1, Art. no. e15201, 2025, doi: 10.1111/cgf.15201.

[11] H. H. Gowtham, A. D. Parakkat, and M.-P. Cani, "PropColor: Interactive color propagation for 2D animations," in Vision, Modeling, and Visualization (VMV 2024), 2024, doi: 10.2312/vmv.20241210.

[12] Y. Kasten, D. Ofri, O. Wang, and T. Dekel, "Layered neural atlases for consistent video editing," ACM Transactions on Graphics, vol. 40, no. 6, pp. 1–12, 2021, doi: 10.1145/3478513.3480546.

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209

[13] J. Dong, K. Ota, and M. Dong, "Video frame interpolation: A comprehensive survey," ACM Transactions on Multimedia Computing, Communications, and Applications, 2023, doi: 10.1145/3556544.
[14] H. Jiang, D. Sun, V. Jampani, M.-H. Yang, E. Learned-Miller, and J. Kautz, "Super SloMo: High quality estimation of multiple intermediate frames for video interpolation," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 9000–9008, 9008.
[15] W. Bao, W.-S. Lai, C. Ma, X. Zhang, Z. Gao, and M.-H. Yang, "Depth-aware video frame interpolation," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3703–3712, 3712.
[16] T. Xue, B. Chen, J. Wu, D. Wei, and W. T. Freeman, "Video enhancement with task-oriented flow," International Journal of Computer Vision, vol. 127, no. 8, pp. 1106–1125, 2019, doi: 10.1007/s11263-018-01144-2.
[17] S. Niklaus and F. Liu, "Softmax splatting for video frame interpolation," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5437–5446, 5446, doi: 10.1109/CVPR42600.2020.00548.
[18] L. Lu, R. Wu, H. Lin, J. Lu, and J. Jia, "Video frame interpolation with Transformer," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3522–3532, 3532, doi: 10.1109/CVPR52688.2022.00352.
[19] Z. Li, Z.-L. Zhu, L.-H. Han, Q. Hou, C.-L. Guo, and M.-M. Cheng, "AMT: All-pairs multi-field transforms for efficient frame interpolation," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023, doi: 10.1109/CVPR52729.2023.00945.
[20] D. Sun, X. Yang, M.-Y. Liu, and J. Kautz, "PWC-Net: CNNs for optical flow using pyramid, warping, and cost volume," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 8934–8943, 8943.
[21] Z. Teed and J. Deng, "RAFT: Recurrent all-pairs field transforms for optical flow," in Computer Vision – ECCV 2020, pp. 402–419, 2020, doi: 10.1007/978-3-030-58536-5_24.
[22] Z. Huang, T. Zhang, W. Heng, B. Shi, and S. Zhou, "Real-time intermediate flow estimation for video frame interpolation," in Computer Vision – ECCV 2022, vol. 13674, pp. 624–642, 1367, doi: 10.1007/978-3-031-19781-9_36.
[23] F. Reda, J. Kontkanen, E. Tabellion, D. Sun, C. Pantofaru, and B. Curless, "FILM: Frame interpolation for large motion," in Computer Vision – ECCV 2022, vol. 13667, pp. 250–266, 1366, doi: 10.1007/978-3-031-20071-7_15.
[24] R. Narita, K. Hirakawa, and K. Aizawa, "Optical flow based line drawing frame interpolation using distance transform to support inbetweenings," in 2019 IEEE International Conference on Image Processing (ICIP), pp. 4200–4204, 4204, doi: 10.1109/ICIP.2019.8803506.
[25] L. Siyao, S. Zhao, W. Yu, W. Sun, D. Metaxas, C. C. Loy, and Z. Liu, "Deep animation video interpolation in the wild," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 6587–6595, 6595, doi: 10.1109/CVPR46437.2021.00652.
[26] S. Chen and M. Zwicker, "Improving the perceptual quality of 2D animation interpolation," in Computer Vision – ECCV 2022, vol. 13677, 2022, doi: 10.1007/978-3-031-19790-1_17.
[27] J. Xing, H. Liu, M. Xia, Y. Zhang, X. Wang, Y. Shan, and T.-T. Wong, "ToonCrafter: Generative cartoon interpolation," ACM Transactions on Graphics, vol. 43, no. 6, pp. 1–11, 2024, doi: 10.1145/3687761.
[28] L. Siyao, T. Gu, W. Xiao, H. Ding, Z. Liu, and C. C. Loy, "Deep geometrized cartoon line inbetweening," in Proceedings of the IEEE/CVF International Conference on

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209

Computer Vision (ICCV), pp. 7291–7300, 7300, doi: 10.1109/ICCV51070.2023.00670.
[29] H. Mo, C. Gao, and R. Wang, "Joint stroke tracing and correspondence for 2D animation," ACM Transactions on Graphics, vol. 43, no. 3, pp. 1–17, 2024, doi: 10.1145/3649890.
[30] C. Mo, K. Hu, C. Long, and Z. Wang, "Continuous intermediate token learning with implicit motion manifold for keyframe based motion interpolation," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023.
[31] B. N. Oreshkin, A. Valkanas, F. G. Harvey, L.-S. Ménard, F. Bocquelet, and M. J. Coates, "Motion in-betweening via deep Δ-interpolator," IEEE Transactions on Visualization and Computer Graphics, vol. 30, no. 8, pp. 5693–5704, 2024, doi: 10.1109/TVCG.2023.3309107.
[32] S. Starke, Y. Zhao, F. Zinno, and T. Komura, "Neural animation layering for synthesizing martial arts movements," ACM Transactions on Graphics, vol. 40, no. 4, Art. no. 92, 2021, doi: 10.1145/3450626.3459881.
[33] A. Z. Rahman, M. F. Filza, and B. K. Umri, "Otomatisasi efek follow-through dan overlapping action pada animasi 3D berbasis offset keyframe F-Curve prosedural dan simulasi mass-spring-damper menggunakan Blender," Jurnal Informatika, vol. 15, no. 1, pp. 29–42, 2026, doi: 10.55340/jiu.v15i1.2678.
[34] amikomjogja, "Information Technology Undergraduate Program Profile," YouTube, Feb. 1, 2023. [Online Video]. Available: https://www.youtube.com/watch?v=hoL7tuBQtm8 (accessed Nov. 21, 2026).
[35] T. Kroeger, R. Timofte, D. Dai, and L. Van Gool, "Fast optical flow using dense inverse search," in Computer Vision – ECCV 2016, pp. 471–488, 2016, doi: 10.1007/978-3-319-46493-0_29.
[36] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600–612, 2004, doi: 10.1109/TIP.2003.819861.
[37] A. Horé and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in 2010 20th International Conference on Pattern Recognition (ICPR), pp. 2366–2369, 2369, doi: 10.1109/ICPR.2010.579.
[38] J. Canny, "A computational approach to edge detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986, doi: 10.1109/TPAMI.1986.4767851.
[39] D. R. Martin, C. C. Fowlkes, and J. Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 5, pp. 530–549, May. 2004.
[40] F. Perazzi, J. Pont-Tuset, B. McWilliams, L. Van Gool, M. Gross, and A. Sorkine-Hornung, "A benchmark dataset and evaluation methodology for video object segmentation," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 724–732, 2016, doi: 10.1109/CVPR.2016.85.
[41] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 586–595, 2018.
[42] K. Ding, K. Ma, S. Wang, and E. P. Simoncelli, "Image quality assessment: Unifying structure and texture similarity," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 5, pp. 2567–2581, 2022, doi: 10.1109/TPAMI.2020.3045810.
[43] C. R. Harris et al., "Array programming with NumPy," Nature, vol. 585, no. 7825, pp. 357–362, Sep. 2020, doi: 10.1038/s41586-020-2649-2.

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 5 Number 1, January 2027
DOI : **https://doi.org/10.58602/itsecs.v5i1.504**
Page 187-209

[44] P. Virtanen et al., "SciPy 1.0: Fundamental algorithms for scientific computing in Python," Nature Methods, vol. 17, no. 3, pp. 261–272, 2020, doi: 10.1038/s41592-019-0686-2.

Deni Sinaga: Penulis Korespondensi
Copyright © 2027, Deni Sinaga and Ema Utami.