

Evaluasi Kualitas Objek 3D Berbasis AI Menggunakan Hyperscore pada Subset Benchmark

Nurul Humaera B^{1*}, Farida², Amhar Akram Haidar³

^{1,2,3}Sistem Informasi, Institut Teknologi dan Bisnis Bina Adinata, Bulukumba, Indonesia

Email: ^{1*}humairah777@gmail.com, ²faridahvaryd4@gmail.com, ³amharakram90@gmail.com

Email Penulis Korespondensi: ¹humairah777@gmail.com

Abstrak: Akselerasi teknologi AI generatif teks-ke-3D (*Text-to-3D*) menjadi salah satu pendekatan untuk meningkatkan efisiensi dalam pembuatan aset 3D. Evaluasi menggunakan full Benchmark membutuhkan sumber daya yang relatif tinggi akibat biaya, kuota API, dan kebutuhan komputasi, sehingga diperlukan pendekatan subset Benchmark yang lebih efisien. Penelitian ini bertujuan untuk mengevaluasi kualitas hasil generasi objek 3D menggunakan skema *subset evaluation* pada Benchmark Hyperscore. Penelitian ini adalah analisis empiris awal dengan ruang lingkup terbatas pada 24 prompt dan dihasilkan menggunakan platform Meshy.AI. Generasi objek 3D dilakukan menggunakan platform Meshy.AI, kemudian dievaluasi menggunakan empat dimensi Hyperscore, yaitu *Alignment*, *Geometry*, *Texture*, dan *Overall Quality*. Hasil evaluasi menunjukkan adanya variasi kualitas antarkategori *prompt*. Kategori *Basic* memperoleh skor *Overall Quality* tertinggi sebesar 7,171, sedangkan kategori *Action* memperoleh skor terendah sebesar 5,513. Kategori *Spatial* juga menunjukkan variasi hasil yang relatif besar pada dimensi *Geometry*. Temuan ini menunjukkan bahwa skema subset evaluation berpotensi menjadi alternatif evaluasi yang lebih efisien untuk penelitian Text-to-3D tanpa menghilangkan kemampuan untuk menggambarkan variasi kualitas antarkategori. Namun, penelitian ini belum membandingkan hasil *subset* secara langsung dengan keseluruhan 160 *prompt*, sehingga kesetaraan *subset* terhadap Benchmark penuh belum dapat disimpulkan. Hasil penelitian ini dapat menjadi dasar untuk kajian lebih lanjut mengenai strategi pemilihan *subset* benchmark dalam evaluasi generasi *Text-to-3D*.

Kata kunci: *Text-to-3D*; *Generative AI*; *Hyperscore*; *3D Object Generation*; *Subset evaluation*

Abstract: The rapid advancement of generative artificial intelligence for Text-to-3D has emerged as a promising approach to improving the efficiency of 3D asset creation. However, evaluation using the full Hyperscore Benchmark requires substantial computational resources, API quotas, and operational costs, creating the need for a more efficient subset-based evaluation approach. This study aims to evaluate the quality of AI-generated 3D objects using a subset evaluation scheme based on the Hyperscore Benchmark. A total of 24 prompts was selected from 160 benchmark prompts while preserving the representation of all eight prompt categories. The 3D objects were generated using the Meshy.AI platform and evaluated across four Hyperscore dimensions: Alignment, Geometry, Texture, and Overall Quality. The results indicate variations in generation quality across prompt categories. The Basic category achieved the highest Overall Quality score of 7.171, whereas the Action category obtained the lowest score of 5.513. The Spatial category also exhibited relatively high variation in the Geometry dimension. These findings suggest that the subset evaluation scheme has the potential to serve as a more efficient evaluation approach for Text-to-3D research while preserving the ability to capture quality variations across prompt categories. Nevertheless, this study does not directly compare the 24-prompt subset with the complete set of 160 prompts; therefore, the equivalence of the subset to the full benchmark cannot yet be established. The findings provide a foundation for future research on subset selection strategies in Text-to-3D generation evaluation.

Keywords: Text-to-3D; Generative AI; Hyperscore; 3D Object Generation; Subset evaluation

1. PENDAHULUAN

Kemunculan AI Generatif atau GenAI dalam satu dekade terakhir merevolusi berbagai lini industri dan organisasi melalui otomatisasi serta pembuatan konten media orisinal[1][2]. Khususnya pada periode 2023-sekarang [3][4], akselerasi teknologi GenAI berbasis pemrosesan teks-ke-3D (*Text-to-3D*) berpotensi meningkatkan efisiensi pembuatan aset 3D[5][6] pada bidang gim, simulasi robotika, hingga *Augmented/Virtual Reality* (AR/VR) yang sebelumnya membutuhkan alokasi waktu dan tenaga kerja yang besar, serta diikuti oleh perkembangan NeRF (*Neural Radiance Field*)[7] sebagai salah satu fondasi penting bagi kemajuan generasi objek 3D modern. Selain pendekatan berbasis teks, perkembangan model generatif juga mengarah pada rekonstruksi objek 3D dari citra tunggal melalui pendekatan *Zero-1-to-3*[8], sehingga menunjukkan semakin beragamnya strategi generasi aset 3D berbasis AI.

Berbagai penelitian telah mengembangkan metode untuk meningkatkan kualitas dan efisiensi dalam proses generasi objek 3D. Pendekatan seperti Point-E[9] dan Shap-E[10] memperkenalkan representasi objek 3D dari instruksi tekstual, sedangkan DreamFusion[2], Magic3D[3], Fantasia3D[11], DreamGaussian[12], hingga Hunyuan3D[4] menunjukkan perkembangan lebih lanjut melalui pemanfaatan model difusi dan Gaussian Splatting. Perkembangan tersebut menunjukkan perkembangan dalam menghasilkan representasi objek 3D dari instruksi maupun informasi visual, sehingga dapat menghasilkan aset 3D dengan kualitas yang semakin baik.

Kendati demikian, efisiensi dan kualitas objek 3D yang dihasilkan oleh model generator sangat bergantung pada karakteristik instruksi tekstual (*prompt*) yang diberikan[13] karena proses pemetaan informasi bahasa ke representasi visual memanfaatkan model vision language yang dilatih dari pasangan data teks dan citra berskala besar[14]. Deskripsi



prompt yang ambigu atau kurang terperinci kerap menghasilkan bias penilaian, di mana model sering kali dinilai berkinerja buruk karena keterbatasan pemahaman bahasa, padahal kualitas fisik seperti geometri dan tekstur objek yang dihasilkan secara visual tergolong baik.

Benchmark Hyperscore[15] telah dikembangkan sebagai pendekatan evaluasi kualitas generasi objek *Text-to-3D*[16][17] dengan instrumen yang lebih terukur [18][19]. Pendekatan evaluasi kualitas ini diperlukan untuk mengurangi subjektivitas dan bias evaluasi manusia terhadap aspek estetika visual[20]. Namun, tantangan praktis muncul ketika proses evaluasi dilakukan menggunakan keseluruhan Benchmark Hyperscore (*full Benchmark*) sebanyak 160 *prompt*. Evaluasi secara menyeluruh membutuhkan alokasi sumber daya perangkat keras yang besar, waktu rendering yang lama, serta biaya operasional yang tinggi, terutama pada model *Text-to-3D* komersial[21] yang membatasi akses ke *Application Programming Interface* (API) melalui skema kuota harian atau langganan berbayar. Kendala finansial dan keterbatasan infrastruktur tersebut sering kali membatasi keterjangkauan proses evaluasi, khususnya bagi peneliti dengan sumber daya yang terbatas.

Penelitian mengenai evaluasi kualitas *Text-to-3D* terus berkembang. Salah satu di antaranya adalah T3Bench yang memanfaatkan evaluasi otomatis berbasis multi-view, sedangkan Hyperscore menawarkan penilaian multidimensi melalui delapan kategori *prompt*. Baik T3Bench maupun Hyperscore dirancang untuk mengevaluasi kualitas hasil generasi *Text-to-3D*. Namun, kedua benchmark tersebut masih menggunakan evaluasi berbasis *full benchmark* dan belum mengkaji strategi pemilihan subset sebagai alternatif evaluasi yang lebih efisien. Meskipun demikian, seluruh evaluasi tersebut masih sepenuhnya menggunakan benchmark sehingga membutuhkan jumlah *prompt*, waktu komputasi, dan pemanggilan API yang relatif besar. Hingga saat ini, kajian mengenai penggunaan skema subset evaluation sebagai alternatif evaluasi yang lebih efisien pada Benchmark Hyperscore masih belum banyak dilaporkan. Oleh karena itu, penelitian ini mengevaluasi apakah subset yang mewakili seluruh delapan kategori tetap mampu menggambarkan variasi kualitas hasil generasi objek 3D secara deskriptif. Berbeda dengan penelitian sebelumnya yang mengevaluasi Benchmark secara penuh, penelitian ini mengusulkan penggunaan subset sebanyak 24 *prompt* yang mewakili delapan kategori sebagai pendekatan evaluasi yang lebih efisien. Pendekatan subset ini mampu mereduksi *volume* pemanggilan API, konsumsi kredit komersial dan total estimasi waktu rendering sebesar 85% dibandingkan skema pengujian penuh (160 *prompt*). Berdasarkan kondisi tersebut, perlu dilakukan kajian lanjut terkait penggunaan *subset Benchmark* secara sistematis yang tetap mencakup variasi kategori *prompt* untuk mengevaluasi kualitas objek 3D komersial secara lebih efisien. Penelitian ini menggunakan Meshy.AI sebagai *platform* komersial *Text-to-3D* untuk menghasilkan objek 3D berdasarkan *prompt* yang telah dipilih dari Benchmark Hyperscore.

Masalah tersebut menjadi dasar penggunaan *subset evaluation* dalam penelitian ini, yang merupakan langkah taktis untuk tetap mempertahankan keterwakilan dari delapan kategori *prompt*. Penelitian ini bertujuan untuk mengevaluasi kualitas objek 3D berbasis AI dengan memanfaatkan *subset Benchmark* yang terdiri dari 24 *prompt* dari total 160 *prompt* pada Benchmark Hyperscore [15] yang diuji pada *platform* komersial berbasis *web*, yakni Meshy.AI. Melalui penerapan metode Hyperscore, penelitian ini menganalisis distribusi skor dan variasi kualitas hasil generasi objek 3D pada berbagai kategori *prompt* berdasarkan empat dimensi Hyperscore, yaitu *Alignment*, *Geometry*, *Texture* dan *Overall Quality*. Kontribusi penelitian ini terletak pada analisis empiris penggunaan subset Benchmark Hyperscore sebagai pendekatan evaluasi yang lebih efisien, dengan tetap mempertimbangkan variasi kualitas pada seluruh kategori *prompt*.

2. METODOLOGI PENELITIAN

Metode penelitian ini menggunakan pendekatan eksperimen kuantitatif untuk mengukur dan menganalisis kualitas objek 3D hasil generasi AI komersial *Text-to-3D*[2], [3], [4] (Meshy.AI)[21]. Pengukuran dilakukan secara terukur menggunakan metrik evaluasi[18] Hyperscore berdasarkan skema *subset evaluation* untuk memperoleh nilai *Alignment*, *Geometry*, *Texture* dan *Overall Quality* sebagai instrumen kualitas objek 3D yang dihasilkan.

2.1. Dataset dan Objek Penelitian

Penelitian ini menggunakan Benchmark Hyperscore sebagai sumber *prompt* untuk menghasilkan objek 3D. Benchmark tersebut terdiri atas 160 *prompt* yang terbagi menjadi delapan kategori, yaitu *Basic*, *Refined*, *Complex*, *Fantastical*, *Group*, *Spatial*, *Action* dan *Imaginative*, dengan masing-masing kategori terdiri atas 20 *prompt*. Lihat pada Tabel 1. Penyesuaian jumlah *prompt* dilakukan dengan mempertimbangkan keterbatasan kuota API pada *platform* AI komersial, biaya penggunaan layanan, serta kebutuhan komputasi selama proses evaluasi. Oleh karena itu, penelitian ini menerapkan pendekatan *subset evaluation* dengan menggunakan 24 *prompt*, yaitu tiga *prompt* pertama dari setiap kategori. Pemilihan tersebut dilakukan untuk mempertahankan keterwakilan seluruh kategori dalam *subset* yang digunakan.

Tabel 1. Pembagian Subset Prompt

No.	Kategori Prompt	Subset Prompt
1	Basic	A green pepper
2	Basic	A grey mouse
3	Basic	A striped zebra
4	Refined	A black umbrella with a curved handle
5	Refined	A bookmark with a tassel
6	Refined	A badge shaped like a shield
7	Complex	A black bat with luminous red eyes, outstretched wings, and a detailed claw with subtle ridges
8	Complex	A square table with intricate carvings along the legs, a scratched surface, and a missing corner
9	Complex	A bronze goblet encrusted with rubies, with swirling engravings of dragons coiling up its stem
10	Fantastical	A candleholder resembling a twisted, gnarled tree branch
11	Fantastical	A castle built entirely from interlocking giant seashells
12	Fantastical	A cactus blooming flowers made of glowing embers
13	Group	A blue umbrella and a silver key
14	Group	A bronze helmet and a golden crown
15	Group	A black camera and a metal tripod
16	Spatial	A white orchid in a ceramic vase
17	Spatial	A white candle in an iron holder
18	Spatial	A polished silver bell resting on a wooden stand
19	Action	A dog is barking at a cat with its tail wagging
20	Action	A dog is licking an ice cream cone with its tail wagging
21	Action	A penguin is waddling with a fish in its beak
22	Imaginative	A crowned koala is inspecting a mystical golden staff
23	Imaginative	A bear astronaut is planting a flag with a paw print
24	Imaginative	A fox magician is pulling a rabbit from a hat

Setiap *prompt* digunakan sebagai *prompt* input pada *platform* Meshy.AI untuk menghasilkan objek 3D. Objek hasil generasi kemudian digunakan sebagai objek evaluasi dengan metode Hyperscore. Seluruh hasil generasi objek disimpan dalam format *.obj* untuk selanjutnya diproses dan dievaluasi.

2.2. Variabel dan Lingkungan Eksperimen

Tabel 2. Konfigurasi Eksperimen

Komponen	Konfigurasi
Platform Generasi	Meshy.AI
Jenis Generasi	Text-to-3D

Dataset	Benchmark Hyperscore
Jumlah <i>prompt</i>	24 dari 160 <i>prompt</i>
Kategori	8 Kategori (3 <i>prompt</i> /kategori)
Format Luaran	.obj
Parameter Generasi	Pengaturan <i>default</i> Meshy.AI
Lingkungan Evaluasi	Google Colab
Bahasa Pemrograman	Python
Metrik Evaluasi	Hyperscore (<i>Alignment, Geometry, Texture, dan Overall Quality</i>)

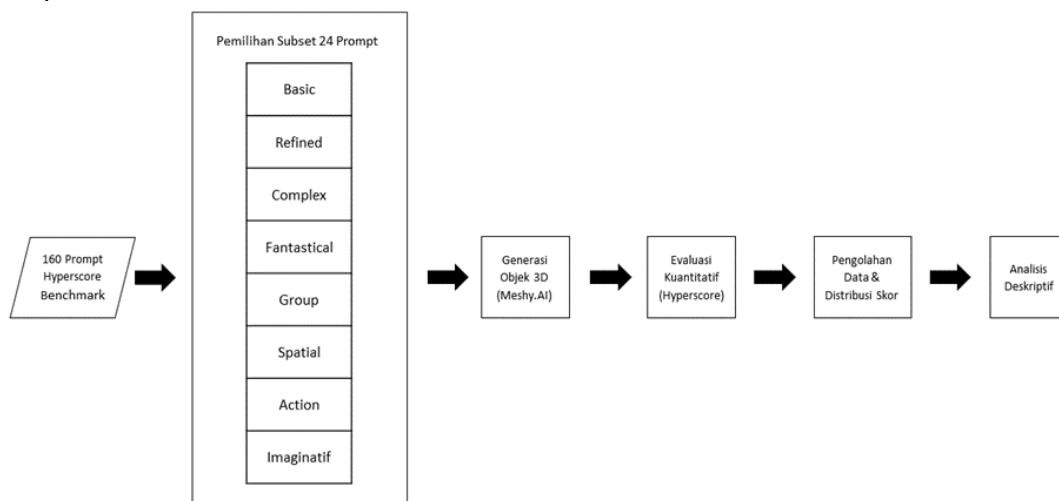
Penelitian ini mengontrol tiga jenis variabel utama. Variabel bebas dalam penelitian ini adalah *prompt* yang dipilih dari *subset* Benchmark Hyperscore yang mewakili delapan kategori *prompt*. Variabel terikat berupa skor evaluasi objek 3D berdasarkan metrik *Alignment, Geometry, Texture, dan Overall Quality*. Sementara itu, variabel kontrol dijaga dengan menyeragamkan generator AI menggunakan *platform* Meshy.AI, mengatur parameter generasi pada opsi default, serta menyeragamkan format luaran aset 3D dengan ekstensi berkas *.obj*.

Konfigurasi perangkat lunak, perangkat keras, serta parameter eksperimen yang digunakan dalam penelitian ini disajikan pada Tabel 2. Eksperimen dijalankan menggunakan kombinasi perangkat lunak dan perangkat keras pendukung. Perangkat lunak yang digunakan terdiri atas *platform* AI generatif Meshy.AI sebagai generator objek, *framework* evaluasi Hyperscore yang dieksekusi menggunakan bahasa pemrograman Python pada lingkungan Google Colab, serta peramban Google Chrome untuk mengakses *platform cloud*. Perangkat keras yang digunakan yaitu komputer jinjing Lenovo dengan memori RAM 5 GB yang digunakan untuk mengakses *platform cloud* serta mendukung proses evaluasi dan pengolahan data hasil eksperimen.

Proses generasi objek 3D dilakukan menggunakan *platform* Meshy.AI dengan mode Text-to-3D dan pengaturan parameter default. Semua *prompt* dijalankan secara independen hingga menghasilkan satu model 3D dalam format *.obj*. Selanjutnya, berkas keluaran kemudian dievaluasi menggunakan *framework* Hyperscore pada lingkungan Google Colab dengan memanfaatkan implementasi Python yang sama untuk seluruh sampel, sehingga setiap objek diproses menggunakan prosedur evaluasi yang konsisten.

2.3. Tahapan Penelitian dan Analisis Data

Alur pelaksanaan eksperimen dilakukan secara sistematis melalui enam tahapan utama seperti yang diilustrasikan pada Gambar 1.



Gambar 1. Alur Metodologi Penelitian

Tahap pertama diawali dengan mempersiapkan dataset Benchmark Hyperscore yang terdiri dari 160 *prompt*. Tahap kedua adalah seleksi *subset* dengan memilih tiga *prompt* pertama secara acak sistematis dari masing-masing delapan kategori, untuk memastikan variasi linguistic dan tingkat kesulitan *prompt* terwakili secara adil sehingga didapatkan total 24 *prompt* sebagai *subset* penelitian. Tahap ketiga adalah generasi objek 3D menggunakan Meshy.AI berdasarkan 24 *prompt* yang telah dipilih. Penelitian ini menggunakan mekanisme generasi bawaan Meshy.AI tanpa pengaturan seed secara manual, sehingga variasi stokastik yang berasal dari proses generasi model tidak dikendalikan secara khusus. Proses generasi menghasilkan output berupa berkas dengan ekstensi *.obj*. Tahap keempat yaitu

mengevaluasi objek 3D menggunakan metode Hyperscore pada Google Colab untuk memperoleh skor kualitas objek 3D. Setiap berkas .obj dievaluasi menggunakan prosedur yang sama sehingga seluruh sampel memperoleh skor pada empat dimensi, yaitu Alignment, Geometry, Texture, dan Overall Quality.

Tahap kelima adalah analisis data dengan menghitung ukuran pemusatan dan penyebaran data yang mencakup nilai rerata, nilai minimum, nilai maksimum, serta deviasi standar pada setiap submetrik evaluasi. Data tersebut kemudian dianalisis secara deskriptif kuantitatif untuk menginterpretasikan perbedaan skor kualitas objek 3D pada setiap kategori *prompt*. Tahap enam dirumuskan dalam bentuk interpretasi hasil untuk menganalisis perbedaan objek 3D berdasarkan kategori *prompt*.

2.4. Metrik Evaluasi

Evaluasi kualitas objek 3D dilakukan menggunakan empat dimensi penilaian dalam Hyperscore, yaitu *Alignment*, *Geometry*, *Texture*, dan *Overall Quality*. *Alignment* digunakan untuk menilai kesesuaian objek yang dihasilkan dengan informasi yang terdapat pada *prompt*. *Geometry* digunakan untuk menilai kualitas struktur dan bentuk objek 3D. *Texture* digunakan untuk menilai kualitas tampilan permukaan dan tekstur objek. Sementara itu, *Overall Quality* digunakan untuk menggambarkan kualitas objek 3D secara keseluruhan berdasarkan hasil evaluasi. Skor dari setiap dimensi digunakan untuk membandingkan kualitas hasil generasi antarkategori *prompt*.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Generasi Objek 3D

Generasi objek 3D dilakukan menggunakan 24 *prompt* pilihan (*subset*) dari Benchmark Hyperscore. Setiap *prompt* menghasilkan satu objek 3D dalam format .obj, menggunakan konfigurasi default di Meshy.AI. Hasil generasi menunjukkan adanya variasi kualitas antarobjek berdasarkan kompleksitas dan karakteristik yang didefinisikan dalam *prompt*. Objek yang dibentuk melalui deskripsi entitas tunggal cenderung memiliki keutuhan struktur yang lebih baik dibandingkan dengan objek yang dihasilkan dari *prompt* yang memerlukan relasi spasial maupun interaksi yang kompleks. Gambar 2 menunjukkan beberapa contoh hasil generasi objek 3D dari *prompt* yang digunakan dalam penelitian ini.



(a) A green pepper



(b) A grey mouse



(c) A striped zebra



(d) A polished silver bell resting on a wooden stand



(e) A white orchid in a ceramic vase



(f) A white candle in an iron holder



(g) A penguin is waddling with a fish in its beak



(h) A dog is barking at a cat with its tail wagging



(i) A dog is licking an ice cream cone with its tail wagging

Gambar 2. Sampel Visual Hasil Generasi Objek 3D Menggunakan Meshy.AI pada Berbagai Kategori *Prompt*

Berdasarkan pengamatan visual pada Gambar 2, terlihat perbedaan karakteristik objek 3D yang dihasilkan dari beberapa kategori *prompt*. Objek tunggal sederhana pada kategori *Basic* (Gambar 2(a), 2(b), dan 2(c)) menunjukkan pembentukan geometri yang utuh[19], topologi permukaannya rapi, dan pemetaan tekstur sesuai dengan deskripsi warna maupun material pada *prompt*. Kondisi tersebut menunjukkan bahwa *prompt* dengan struktur deskripsi yang relatif sederhana dapat menghasilkan representasi objek yang lebih mudah dikenali secara visual

Berbeda dengan kategori *Basic*, objek pada kategori *Spatial* (Gambar 2(d), 2(e), dan 2(f)) menunjukkan adanya ketidaktepatan penempatan atribut. Hubungan spasial antareleman pada beberapa hasil mengalami distorsi atau penggabungan geometri yang tidak wajar[15]. Kondisi tersebut menunjukkan bahwa generator tidak hanya perlu mengenali objek yang disebutkan dalam *prompt*, tetapi juga perlu mempertahankan posisi dan hubungan antareleman yang dijelaskan secara tekstual. Objek yang berhasil dibuat dengan generator secara individual belum tentu diikuti oleh keberhasilan dalam mempertahankan hubungan spasial antarobjek.

Perbedaan yang lebih terlihat terdapat pada objek dinamis dalam kategori *Action* (Gambar 2(g), 2(h), dan 2(i)). Pada hasil tersebut terlihat degradasi ketajaman tekstur dan hilangnya detail pada area lipatan maupun ekstremitas tubuh. Karakteristik tersebut menunjukkan bahwa representasi pose atau aktivitas yang kompleks dapat menjadi tantangan bagi generator ketika informasi tekstual harus diterjemahkan menjadi struktur 3D. Generator perlu mempertahankan identitas objek sekaligus merepresentasikan tindakan dan konfigurasi tubuh yang dijelaskan dalam *prompt*. Kondisi ini dapat menyebabkan sebagian informasi dalam instruksi tidak direpresentasikan secara optimal pada objek akhir.

Hasil pengamatan visual tersebut memberikan indikasi awal bahwa kompleksitas *prompt* berkaitan dengan variasi kualitas hasil generasi. Namun, pengamatan visual saja belum cukup untuk digunakan sebagai dasar perbandingan antarkategori. Oleh karena itu, hasil generasi selanjutnya dievaluasi secara kuantitatif menggunakan metode Hyperscore untuk memperoleh pengukuran pada beberapa dimensi kualitas.

3.2 Hasil Evaluasi Hyperscore

Hasil evaluasi menggunakan Hyperscore menunjukkan adanya variasi skor pada empat dimensi penilaian, yaitu *Alignment*, *Geometry*, *Texture*, dan *Overall Quality*. Keempat dimensi tersebut digunakan untuk menilai keselarasan instruksi dengan objek yang dihasilkan, kualitas struktur geometri, kualitas tekstur permukaan, serta kualitas objek secara keseluruhan. Variasi kualitas antarkategori tersebut sejalan dengan penelitian sebelumnya yang menunjukkan bahwa objek 3D hasil generasi AI dapat mengalami perbedaan kualitas dan berbagai jenis distorsi yang dapat diamati melalui evaluasi kualitas secara sistematis[22]. Ringkasan nilai evaluasi untuk masing-masing kategori *prompt* disajikan pada Tabel 3.

Tabel 3. Rata-rata Skor Hyperscore Berdasarkan Kategori

No.	Kategori Prompt	Alignment	Geometry	Texture	Overall
1	<i>Basic</i>	7.883	6.850	6.932	7.171
2	<i>Refined</i>	7.119	6.561	6.633	6.799
3	<i>Complex</i>	6.490	6.771	6.155	6.448
4	<i>Fantastical</i>	7.175	6.801	6.401	6.815
5	<i>Group</i>	6.591	6.278	6.337	6.258
6	<i>Spatial</i>	6.554	5.976	6.118	6.152
7	<i>Action</i>	6.071	5.643	5.439	5.513
8	<i>Imaginative</i>	6.261	5.965	5.738	5.819

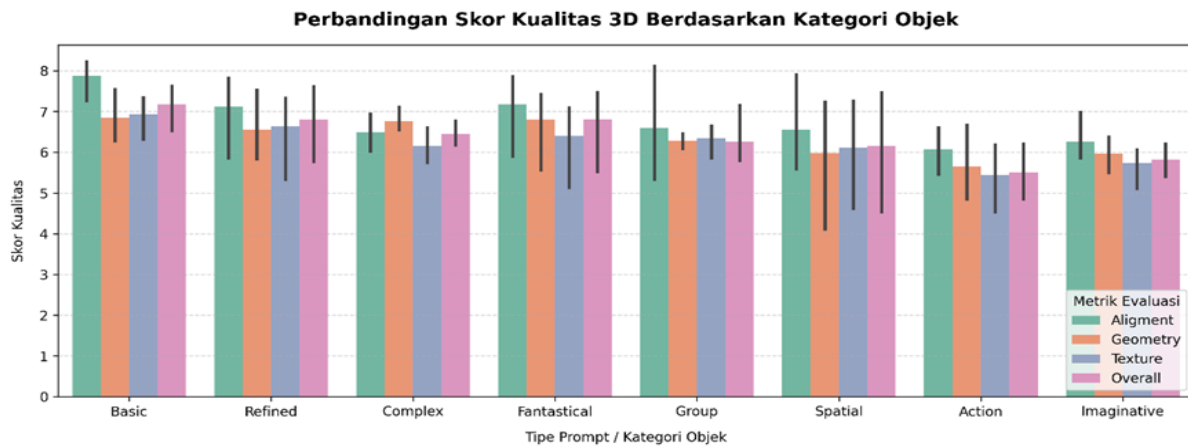
Berdasarkan nilai *Overall Quality*, kategori *Basic* memperoleh rata-rata skor tertinggi sebesar 7,171, sedangkan kategori *Action* memperoleh rata-rata skor terendah sebesar 5,513. Perbedaan tersebut menunjukkan bahwa kualitas hasil generasi tidak seragam di seluruh kategori *prompt*. Kategori dengan deskripsi objek yang relatif sederhana cenderung memperoleh skor lebih tinggi, sedangkan kategori yang membutuhkan pemahaman terhadap aktivitas atau hubungan yang lebih kompleks menunjukkan skor yang relatif lebih rendah.

Pola tersebut juga terlihat pada dimensi evaluasi lainnya. Skor *Alignment* tertinggi dicapai oleh kategori *Basic* dengan rata-rata skor mendekati 7,883, sedangkan skor terendah berada pada kategori *Action* dengan nilai sekitar 6,071. Pada dimensi *Geometry*, kategori *Basic* memperoleh skor sekitar 6,850 dan kategori *Complex* sekitar 6,771, sedangkan kategori *Action* berada pada kisaran 5,643. Pada dimensi *Texture*, kategori *Basic* kembali menunjukkan skor tertinggi sekitar 6,932, sementara kategori *Action* memperoleh skor terendah sekitar 5,439. Sementara itu, pada dimensi *Overall Quality*, kategori *Basic* memperoleh skor sekitar 7,171, diikuti oleh kategori *Refined* dan *Fantastical* yang berada pada kisaran 6,799 dan 6,815, sedangkan kategori *Action* berada pada kisaran 5,513.

Secara keseluruhan, pola tersebut menunjukkan konsistensi antara pengamatan visual dan hasil evaluasi kuantitatif. Kategori *Basic* yang secara visual menunjukkan bentuk objek yang relatif utuh juga memperoleh skor tinggi pada beberapa dimensi evaluasi. Sebaliknya, kategori *Action* yang menunjukkan degradasi tekstur dan detail objek juga memperoleh skor yang relatif rendah. Kesesuaian antara hasil pengamatan visual dan skor Hyperscore menunjukkan bahwa variasi kualitas yang terlihat pada objek hasil generasi juga tercermin dalam pengukuran kuantitatif.

3.3 Analisis Berdasarkan Karakteristik Prompt

Generasi objek 3D dilakukan menggunakan 24 *prompt* pilihan (*subset*) dari Benchmark Hyperscore. Setiap *prompt* menghasilkan satu objek 3D dalam format *.obj*. menggunakan konfigurasi default di Meshy.AI.



Gambar 3. Distribusi Skor Evaluasi Kuantitatif *Hyperscore* Berdasarkan Kategori *Prompt*. *Error bar* menunjukkan deviasi standar (*standard deviation/SD*)

Tabel 4. Statistik deskriptif skor Overall Quality berdasarkan kategori prompt

Kategori	Mean $\pm$ SD	Minimum	Maksimum
<i>Basic</i>	7,171 $\pm$ 0,572	6,530	7,629
<i>Refined</i>	6,799 $\pm$ 0,942	5,765	7,619
<i>Complex</i>	6,444 $\pm$ 0,294	6,181	6,768
<i>Fantastical</i>	6,607 $\pm$ 1,032	5,528	7,518
<i>Group</i>	6,434 $\pm$ 0,712	5,799	7,158
<i>Spatial</i>	6,050 $\pm$ 1,574	4,539	7,469
<i>Action</i>	5,513 $\pm$ 0,693	4,856	6,201
<i>Imaginative</i>	5,898 $\pm$ 0,435	5,406	6,206

Pengukuran menggunakan metrik Hyperscore mencakup empat dimensi penilaian, yaitu *Alignment*, *Geometry*, *Texture*, dan *Overall Quality*. Keempat dimensi tersebut digunakan sebagai dasar analisis kualitas objek 3D pada setiap kategori *prompt*. Pendekatan evaluasi multidimensi tersebut sejalan dengan perkembangan penelitian evaluasi generasi 3D yang menekankan penilaian yang fine-grained dan interpretable terhadap karakteristik aset 3D, sehingga kualitas objek tidak hanya direpresentasikan melalui satu skor keseluruhan[23].

Berdasarkan agregasi data pada Gambar 3 dan Tabel 4, performa generasi objek 3D menunjukkan variasi skor pada seluruh kategori *prompt*. Kategori *Basic* memperoleh rata-rata Overall Quality tertinggi sebesar $7,171 \pm 0,572$, sedangkan kategori *Action* memiliki rata-rata terendah sebesar $5,513 \pm 0,693$. Kategori *Spatial* menunjukkan deviasi standar terbesar ($SD = 1,574$). Hal ini menunjukkan bahwa variasi kualitas hasil generasi lebih tinggi dibandingkan dengan kategori lainnya. Gambar 3 menampilkan distribusi skor evaluasi Hyperscore berdasarkan kategori *prompt*. Error bar pada grafik merepresentasikan deviasi standar (SD), sehingga menunjukkan variasi kualitas objek 3D hasil generasi *Text-to-3D* [2], [3], [24] pada setiap kategori. Perbedaan skor menunjukkan adanya variasi kualitas hasil generasi objek 3D antarkategori. Skor *Alignment* tertinggi dicapai oleh kategori *Basic* dengan rata-rata skor mendekati 7,883, sedangkan skor terendah berada pada kategori *Action* sebesar 6,071. Skor *Geometry* tertinggi didominasi oleh kategori *Basic* sebesar 6,850 dan *Complex* sebesar 6,771, sementara skor terendah berada pada kategori *Action* sebesar



5,643. Skor *Texture* tertinggi diperoleh pada kategori *Basic* sebesar 6,932, sedangkan skor terendah tercatat pada kategori *Action* sebesar 5,439. Adapun skor Overall tertinggi dipimpin oleh kategori *Basic* sebesar 7,171, diikuti oleh *Refined* dan *Fantastical* sebesar 6,799 dan 6,815, sedangkan skor terendah dicapai oleh kategori *Action* sebesar 5,513.

Perbedaan skor antarkategori menunjukkan bahwa karakteristik *prompt* menunjukkan pola tertentu yang selajalan dengan kualitas objek 3D yang dihasilkan. Kategori *Basic* menunjukkan performa yang relatif lebih baik karena *prompt* pada kategori ini cenderung mendeskripsikan objek tunggal dengan karakteristik yang lebih sederhana. Pada kondisi tersebut, proses pemetaan informasi tekstual dari *prompt* menjadi representasi bentuk dan tampilan objek 3D relatif sederhana dibandingkan dengan instruksi yang melibatkan beberapa objek, hubungan spasial, atau aktivitas.

Temuan pada kategori *Basic* terlihat konsisten pada keempat dimensi evaluasi. Skor *Alignment* yang tinggi menunjukkan bahwa informasi utama dalam *prompt* relatif berhasil direpresentasikan dalam objek yang dihasilkan. Skor *Geometry* dan *Texture* yang relatif tinggi menunjukkan bahwa kategori dengan instruksi yang lebih sederhana cenderung memiliki kualitas geometri dan tekstur yang lebih baik dalam eksperimen ini. Hal ini menunjukkan bahwa kategori *Basic* dapat dipandang sebagai kategori yang menghasilkan kondisi evaluasi yang relatif lebih stabil dalam eksperimen ini.

Sebaliknya, kategori *Action* menunjukkan performa yang relatif lebih rendah. *Prompt* pada kategori ini mengandung informasi mengenai aktivitas atau gerakan objek, sehingga generator perlu mempertahankan hubungan antara objek, posisi, dan tindakan yang dijelaskan dalam teks. Informasi tersebut memiliki tingkat kompleksitas yang lebih tinggi dibandingkan dengan deskripsi objek tunggal. Ketika generator harus merepresentasikan beberapa informasi tersebut secara bersamaan, terdapat kemungkinan bahwa sebagian detail semantik tidak dipertahankan sepenuhnya dalam objek 3D.

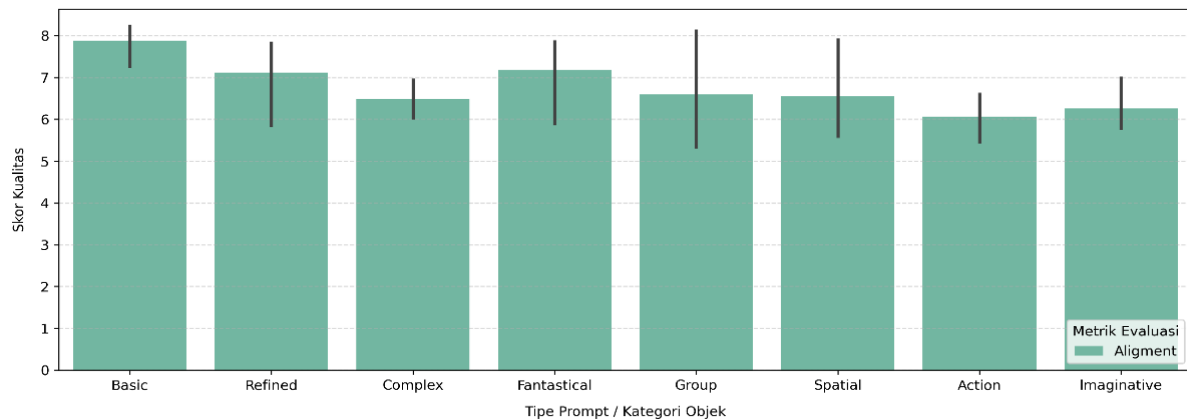
Pada kategori *Action*, penurunan terjadi pada dimensi *Geometry* dan *Texture*. Skor *Geometry* sebesar 5,643 menunjukkan bahwa kompleksitas aktivitas berkaitan dengan kesulitan dalam mempertahankan struktur dan bentuk objek. Skor sebesar 5,439 pada dimensi *Texture* menunjukkan adanya penurunan kualitas tampilan permukaan. Berdasarkan pengamatan visual, penurunan tersebut terjadi pada ekstremitas, lipatan, dan bagian objek yang terlibat dalam pose dinamis. Adapun kesesuaian makna *prompt*, representasi bentuk, dan tampilan objek merupakan permasalahan yang terjadi pada kategori *Action* yang secara visual terlihat pada Gambar 2.

Kategori *Spatial* memberikan temuan yang berbeda. Meskipun skor rata-ratanya pada dimensi *Geometry* berada pada kisaran 5,976, kategori ini memiliki rentang variasi hasil yang relatif lebar, yaitu sekitar 4,539 hingga 7,469 pada Tabel 4. Rentang tersebut menunjukkan bahwa hasil generasi pada kategori *Spatial* tidak selalu konsisten. Pada beberapa *prompt*, generator dapat menghasilkan struktur spasial yang relatif baik, tetapi pada *prompt* lainnya dapat terjadi distorsi atau penggabungan geometri. Variasi tersebut penting karena menunjukkan bahwa nilai rata-rata saja belum sepenuhnya menggambarkan karakteristik hasil generasi pada kategori dengan hubungan spasial yang kompleks.

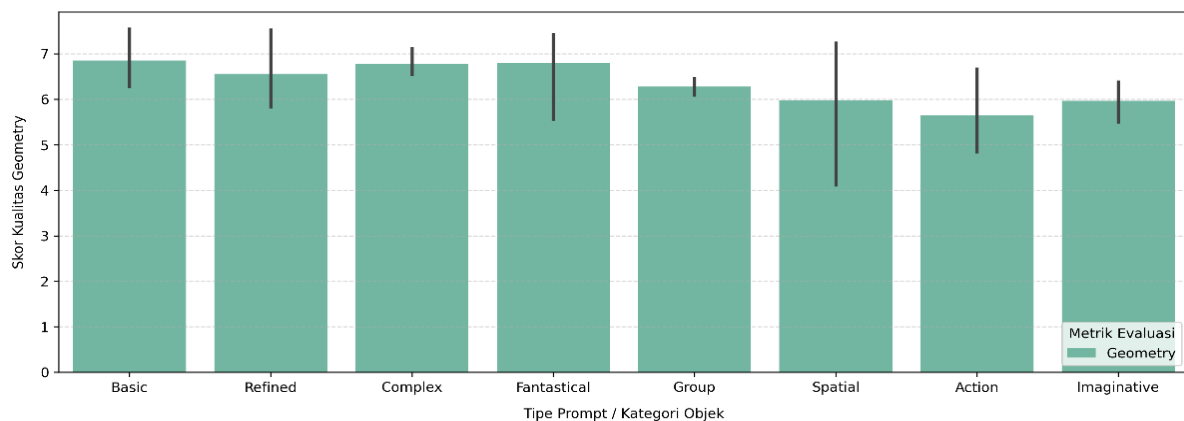
Temuan pada kategori *Spatial* juga menunjukkan pentingnya mempertimbangkan penyebaran data dalam evaluasi objek 3D. Dua kategori dapat memiliki nilai rata-rata yang berdekatan, tetapi tingkat variasi hasilnya dapat berbeda. Oleh karena itu, penggunaan ukuran pemusatan dan penyebaran seperti rerata (mean), nilai minimum, nilai maksimum, dan deviasi standar memberikan informasi yang lebih lengkap dibandingkan dengan hanya menggunakan satu nilai rata-rata. Dalam konteks penelitian ini, ukuran tersebut membantu menunjukkan apakah hasil generasi pada suatu kategori cenderung konsisten atau memiliki variasi kualitas yang besar.

Perbedaan skor antarkategori pada penelitian ini dianalisis menggunakan pendekatan statistik deskriptif melalui ukuran pemusatan (mean) dan ukuran penyebaran (minimum, maksimum, serta deviasi standar). Oleh karena itu, hasil penelitian menggambarkan kecenderungan variasi kualitas generasi objek 3D pada setiap kategori *prompt* dan belum digunakan untuk menyimpulkan adanya perbedaan yang signifikan secara statistik.

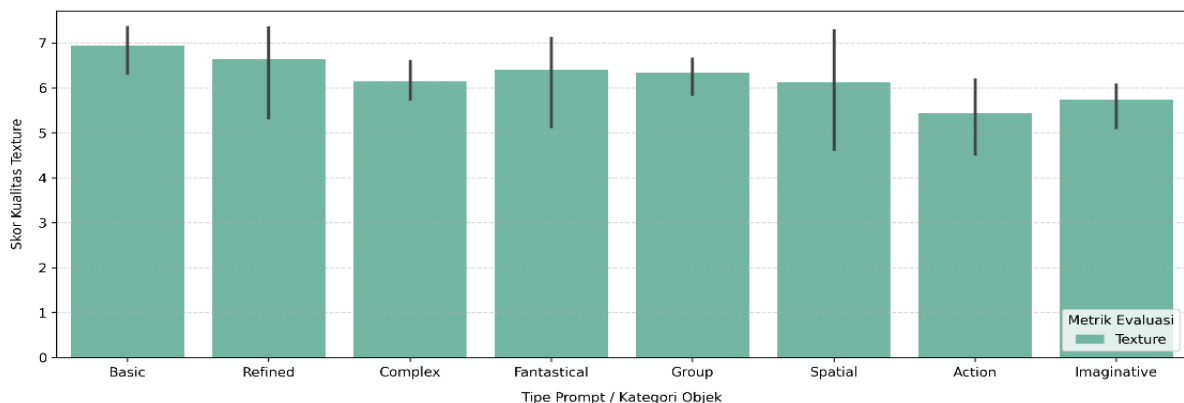
3.4 Analisis Dimensi Hyperscore



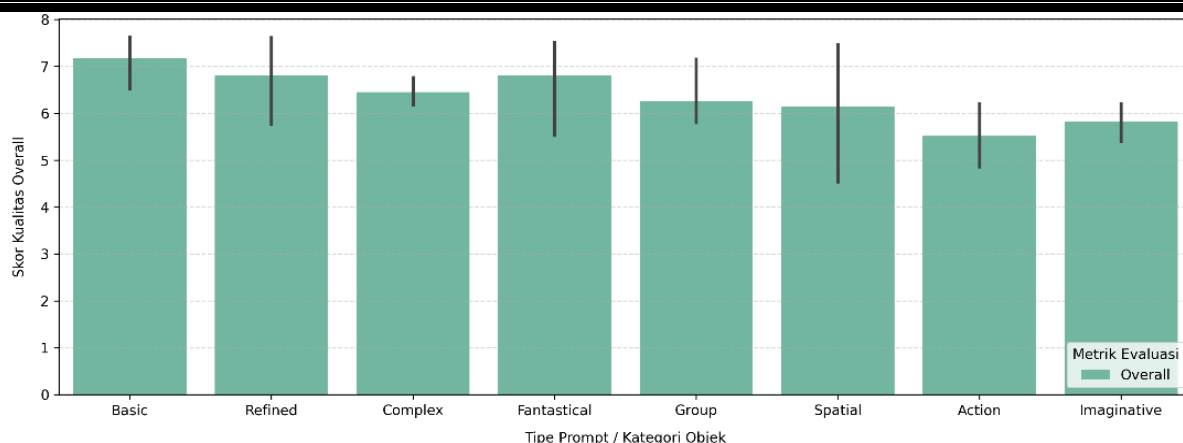
(a) Skor *Alignment* Per Kategori



(b) Skor *Geometry* Per Kategori



(c) Skor *Texture* Per Kategori



(d) Skor Overall Per Kategori

Gambar 4. Distribusi Skor Evaluasi Kuantitatif *Hyperscore* Per Dimensi: (a) *Alignment*, (b) *Geometry*, (c) *Texture*, dan (d) *Overall Quality*

Pada dimensi *Alignment*, kategori *Basic* menunjukkan tingkat kepatuhan model yang paling dominan (Gambar 4(a)) dengan skor rata-rata tertinggi sebesar 7,883. Hal ini menunjukkan bahwa *prompt* yang mendeskripsikan objek tunggal menghasilkan tingkat keselarasan yang lebih tinggi dibandingkan dengan kategori lainnya[13]. Sebaliknya, pada kategori *Action* dan *Imaginative*, nilai *Alignment* mengalami penurunan hingga kisaran 6,071–6,261. Fenomena ini mengindikasikan adanya penurunan kemampuan model dalam mempertahankan kesesuaian antara *prompt* dan objek yang dihasilkan ketika instruksi menjadi lebih kompleks.

Pada dimensi *Geometry* (Gambar 4(b)), kategori *Basic*, *Complex*, dan *Fantastical* secara konsisten mempertahankan skor di kisaran 6,771–6,850. Hal menarik terlihat pada kategori *Spatial*; meskipun nilai rata-ratanya berada pada angka 6,050, kategori ini memiliki error bar atau rentang deviasi paling panjang (Tabel 4), berkisar antara 4,539 hingga 7,469. Rentang yang lebar ini menunjukkan adanya variasi hasil generasi yang lebih besar dibandingkan dengan kategori lainnya. Generator terkadang berhasil membentuk geometri spasial dengan sangat baik, namun di saat lain mengalami kegagalan struktur yang cukup parah[15][19].

Grafik dimensi *Texture* (Gambar 4(c)) mengonfirmasi bahwa kualitas sintesis warna permukaan berbanding lurus dengan kejelasan sampel. Kategori *Basic* dengan skor 6,932 dan *Refined* pada skor 6,633 menghasilkan visualisasi permukaan yang paling halus. Sebaliknya, kategori *Action* mencatatkan skor *Texture* terendah sebesar 5,439. Degradasi tekstur pada kategori *Action* menunjukkan bahwa kualitas visual cenderung menurun ketika *prompt* memuat deskripsi aksi yang lebih kompleks[18][19]. Keterbatasan ini menyebabkan *texture map* secara simetris sulit dipetakan dengan akurat pada geometri objek yang sedang memunculkan pose dinamis atau terdistorsi, sebagaimana diidentifikasi dalam pendekatan neural stylization berbasis objek 3D[25]. Secara kumulatif pada metrik Overall (Gambar 4(d)), terjadi penurunan kinerja generator secara gradual dari kategori *Basic* dengan skor 7,171 menuju *Action* dengan skor 5,513. Temuan ini menunjukkan bahwa peningkatan kompleksitas deskripsi pada *prompt* cenderung diikuti oleh penurunan kualitas objek 3D yang dihasilkan.

Penelitian ini memiliki beberapa keterbatasan yang perlu dipertimbangkan dalam menginterpretasikan hasil penelitian. Evaluasi dibatasi pada 24 *prompt* dari total 160 *prompt* Benchmark *Hyperscore* dan hanya menggunakan satu platform *Text-to-3D* komersial, yaitu Meshy.AI. Proses generasi menggunakan mekanisme bawaan Meshy.AI tanpa pengaturan seed secara manual, sehingga variasi stokastik dari model generatif belum dapat dikendalikan sepenuhnya. Selain itu, evaluasi kualitas sepenuhnya menggunakan skema penilaian otomatis *Hyperscore*, sehingga penelitian ini belum membandingkan hasilnya dengan penilaian manusia. Penelitian ini juga belum melakukan validasi silang menggunakan human evaluation maupun evaluator alternatif sehingga interpretasi kualitas masih mengacu pada metrik otomatis *Hyperscore*. Keterbatasan tersebut menyebabkan hasil penelitian lebih tepat dipahami sebagai analisis empiris terhadap penggunaan subset Benchmark pada satu lingkungan eksperimen, bukan sebagai generalisasi terhadap seluruh model *Text-to-3D*.

4. KESIMPULAN

Penelitian ini mengevaluasi kualitas generasi objek 3D berbasis AI menggunakan subset Benchmark *Hyperscore* sebagai analisis empiris awal dalam ruang lingkup terbatas yang terdiri atas 24 *prompt* yang dipilih dari total 160 *prompt*

dengan mempertahankan keterwakilan delapan kategori Benchmark Hyperscore dan dihasilkan menggunakan *platform* Meshy.AI. Hasil evaluasi menggunakan empat dimensi Hyperscore, yaitu *Alignment*, *Geometry*, *Texture*, dan *Overall Quality*, menunjukkan bahwa kualitas hasil generasi objek 3D bervariasi pada setiap kategori *prompt*.

Kategori *Basic* memperoleh performa relatif paling baik dengan nilai *Overall Quality* rata-rata sebesar 7,171, sedangkan kategori *Action* memperoleh nilai terendah sebesar 5,513. Hasil tersebut menunjukkan bahwa *prompt* yang mendeskripsikan objek secara sederhana cenderung menghasilkan objek dengan kesesuaian, struktur geometri, dan kualitas tekstur yang lebih baik dibandingkan dengan *prompt* yang melibatkan aktivitas atau hubungan yang lebih kompleks. Kategori *Spatial* juga menunjukkan variasi hasil yang relatif besar pada dimensi *Geometry*, yang mengindikasikan bahwa hubungan spasial antareleman berkaitan dengan variasi konsistensi hasil generasi objek 3D.

Penggunaan subset 24 *prompt* secara deskriptif menunjukkan potensi sebagai pendekatan evaluasi yang lebih efisien untuk menggambarkan variasi kualitas antarkategori pada Benchmark Hyperscore. Meskipun demikian, penelitian ini belum membandingkan hasil subset secara langsung dengan keseluruhan 160 *prompt* dalam Benchmark Hyperscore. Oleh karena itu, temuan yang diperoleh masih bersifat deskriptif dan belum dapat digunakan untuk menyatakan bahwa subset memiliki tingkat kesetaraan dengan full Benchmark. Penelitian selanjutnya dapat melakukan validasi *subset* melalui perbandingan langsung dengan Benchmark penuh, menggunakan strategi pemilihan *prompt* yang berbeda, serta melibatkan lebih dari satu *platform Text-to-3D* untuk menguji konsistensi hasil evaluasi.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada Institut Teknologi dan Bisnis Bina Adinata atas dukungan yang diberikan selama pelaksanaan penelitian ini. Penulis juga mengucapkan terima kasih kepada pihak-pihak yang telah memberikan dukungan, masukan, dan bantuan selama proses pelaksanaan penelitian hingga penyusunan artikel ini. Dukungan tersebut turut membantu kelancaran proses eksperimen, analisis data, dan penyelesaian penelitian.

PERNYATAAN PENGGUNAAN KECERDASAN BUATAN (AI)

Pada penulisan dan penyusunan artikel ini, penulis menggunakan *platform* kecerdasan buatan (GenAI) Meshy.AI sebagai instrument eksperimen yang menghasilkan objek 3D. Penulis juga menggunakan ChatGPT selama penyusunan naskah yang bertujuan untuk memperbaiki tata bahasa. Selanjutnya, penulis meninjau, mengedit dan memvalidasi kembali seluruh isi artikel secara penuh serta bertanggungjawab atas orisinalitas dan konten yang ada dalam artikel ini.

REFERENCES

- [1] R. Rombach, A. Blattmann, and D. Lorenz, "High-Resolution Image Synthesis with Latent Diffusion Models," *IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, pp. 10674–10685, 2022, doi: 10.1109/cvpr52688.2022.01042.
- [2] B. Poole, A. Jain, J. T. Barron, and B. Mildenhall, "DREAMFUSION: TEXT-TO-3D USING 2D DIFFUSION," in *International Conference on Learning Representations (ICLR)*, 2023, p. 2209.14988, [Online]. Available: <https://openreview.net/forum?id=FjNys5c7VyY>.
- [3] C. Lin *et al.*, "Magic3D: High-Resolution Text-to-3D Content Creation," *IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2023, doi: 10.48550/arXiv.2211.10440.
- [4] T. Hunyuan, "Hunyuan3D 1.0: A Unified Framework for Text-to-3D and Image-to-3D Generation," *arXiv*, 2025, doi: 10.48550/arXiv.2411.02293.
- [5] A. T. Prakash, S. Shaji, and S. Raghunath, "A SURVEY ON TEXT TO 3D MODEL GENERATOR," *IJRTI2501071 Int. J. Res. Trends Innov.*, vol. 10, no. 1, pp. 585–589, 2025, [Online]. Available: <https://www.ijrti.org/papers/IJRTI2501071.pdf>.
- [6] R. Ziatdinov and R. Nabiyev, "Generating Digital Models Using Text-to-3D and Image-to-3D Prompts: Critical Case Study," *Proc. - 2025 Int. Russ. Smart Ind. Conf. SmartIndustryCon 2025*, pp. 719–724, 2025, doi: 10.1109/SmartIndustryCon65166.2025.10986113.
- [7] B. Mildenhall, J. T. Barron, R. Ramamoorthi, and R. Ng, "NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis," *Commun. ACM*, vol. 65, no. 1, pp. 99–106, 2021, doi: 10.1145/3503250.
- [8] R. Liu, R. Wu, B. Van Hoorick, P. Tokmakov, S. Zakharov, and C. Vondrick, "Zero-1-to-3: Zero-shot One Image to 3D Object," *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 9264–9275, 2023, doi: 10.1109/ICCV51070.2023.00853.
- [9] A. Nichol, H. Jun, P. Dhariwal, P. Mishkin, and M. Chen, "Point-E: A System for Generating 3D Point Clouds from Complex Prompts," *arXiv*, 2022, doi: 10.48550/arXiv.2212.08751.
- [10] H. Jun and A. Nichol, "Shap-E: Generating Conditional 3D Implicit Functions," *IEEE/CVF Conf. Comput. Vis.*



- Pattern Recognit.*, 2023, doi: 10.48550/arxiv.2305.02463.
- [11] R. Chen, Y. Chen, N. Jiao, and K. Jia, "Fantasia3D: Disentangling Geometry and Appearance for High-quality Text-to-3D Content Creation," *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 22189–22199, 2023, doi: 10.1109/ICCV51070.2023.02033.
- [12] J. Tang, J. Ren, H. Zhou, Z. Liu, and Z. Gang, "DREAMGAUSSIAN: GENERATIVE GAUSSIAN SPLATTING FOR EFFICIENT 3D CONTENT CREATION," in *The International Conference on Learning Representations (ICLR)*, 2024, pp. 1–18, doi: 10.48550/arXiv.2309.16653.
- [13] B. Pecher, M. Spiegel, R. Belanec, and J. Cegin, "Revisiting Prompt Sensitivity in Large Language Models for Text," *ArXiv, abs*, p. 2602.04297, 2026, doi: 10.48550/arXiv.2602.04297.
- [14] A. Radford *et al.*, "Learning Transferable Visual Models From Natural Language Supervision," *Int. Conf. Mach. Learn. PMLR*, 2021, doi: 10.48550/arxiv.2103.00020.
- [15] Y. Zhang, B. Cui, Q. Yang, Z. Li, and Y. Xu, "Benchmarking and Learning Multi-Dimensional Quality Evaluator for Text-To-3D Generation," *Int. Conf. Comput. Vis.*, pp. 18563–18574, 2025, doi: 10.1109/ICCV51701.2025.01725.
- [16] S. Maiti, L. Agapito, and F. Kokkinos, "Gen3DEval: Using vLLMs for Automatic Evaluation of Generated 3D Objects," *IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, pp. 18552–18562, 2025, doi: 10.1109/cvpr52734.2025.01729.
- [17] Y. Zhang *et al.*, "3DGen-Bench: Comprehensive Benchmark Suite for 3D Generative Models," *IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2025, doi: 10.48550/arXiv.2503.21745.
- [18] X. Fei, A. Swaminathan, and P. Favaro, "A Quantitative Evaluation of Score Distillation Sampling Based Text-to-3D," *arXiv*, 2024, doi: 10.48550/arxiv.2402.18780.
- [19] Y. He *et al.*, "T3Bench: Benchmarking Current Progress in Text-to-3D Generation," 2024, doi: 10.48550/arXiv.2310.02977.
- [20] T. Wu *et al.*, "GPT-4V(ision) is a Human-Aligned Evaluator for Text-to-3D Generation," *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2024, doi: 10.48550/arXiv.2401.04092.
- [21] M. Sawant, "Advancements in AI for Text-to-3D Model Generation : A Comparative Study of Meshy.ai and LumaLabs.ai Genie," *Int. J. Adv. Comput. Eng. Commun. Technol.*, vol. 14, no. 1, pp. 485–489, 2025, doi: 10.65521/ijacect.v14i1.563.
- [22] Y. Zhou *et al.*, "3DGCQA: A Quality Assessment Database for 3D AI-Generated Contents," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, no. c, 2025, doi: 10.1109/ICASSP49660.2025.10889921.
- [23] S. Duggal *et al.*, "Eval3D: Interpretable and Fine-Grained Evaluation for 3D Generation," *IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, pp. 13326–13336, 2025, doi: 10.48550/arXiv.2504.18509.
- [24] B. Kerbl *et al.*, "3D Gaussian Splatting for Real-Time Radiance Field Rendering," *ACM Trans. Graph.*, vol. 42, no. 4, pp. 1–14, 2023, doi: 10.1145/3592433.
- [25] O. Michel, R. Bar-On, R. Liu, S. Benaim, and R. Hanocka, "Text2Mesh : Text-Driven Neural Stylization for Meshes," *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, pp. 13492–13502, 2022, doi: 10.1109/CVPR52688.2022.01313.