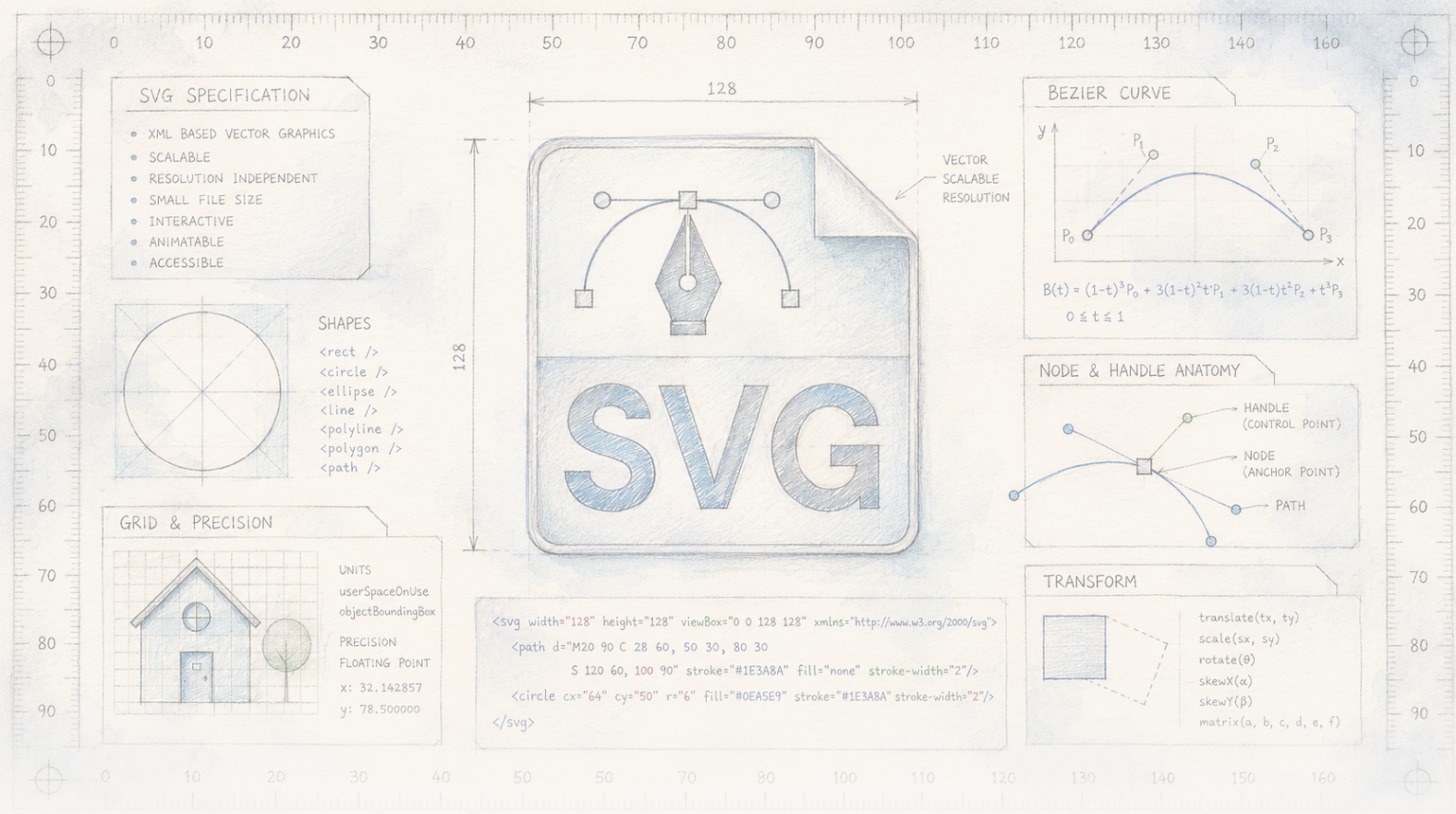




The CLEAN PAPER

WHAT THE PAPER SAYS · WHAT IT DOESN'T · WHY IT MATTERS

AI-কো নজিরে drawing দেখোলহে হয় না — feedback ব্যবহার করতে শেখতে হয়

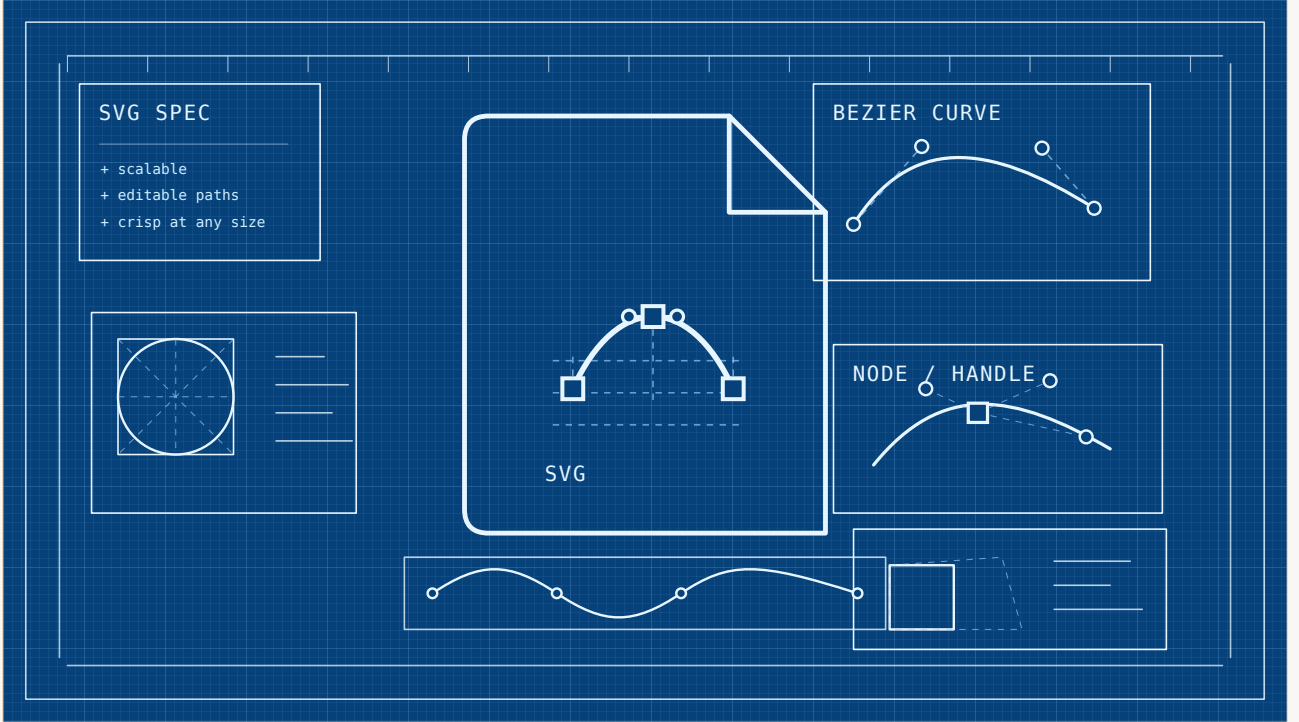
Render-in-the-Loop study-তে pretrained drawing agent-কে শুধু visual feedback দিলে performance অনেকে কষত্রে খারাপ হয়; feedback-aware training-এর পরে render দেখে code revise করা useful হয়। Result more vision automatically better agent নয়—feedback-to-action skill শেখানোর evidence।

Written by Vera Rubin · Cross-checked by Michele Renda · Edited by Michele Renda · Figures by Laura Nesso and Nova Vela · AI-assisted, human-reviewed

Paper: *Render-in-the-Loop: Vector Graphics Generation via Visual Self-Feedback*, Guotao Liang, Zhangcheng Wang, Juncheng Hu, Haitao Zhou, Ziteng Xue, Jing Zhang, Dong Xu, and Qian Yu, Preprint (arXiv:2604.20730)

চোখ বন্ধ করে ছবি আঁকা

আজকের কলোনে AI মডেলকে ছবি আঁকতে বললে ভেতরে মনেটা মুট দুই রকমের কাজ হতে পারে। পরিচিতি ইমেজে জেনারেটর—যেগুলো একটি বাক্য থেকে ছবির মতো দৃশ্য তৈরি করে—সরাসরি পিক্সলে আঁকে, আর কাজ চলাকালই ক্যানভাসের ফল দেখতে পারে। কিন্তু আরকে ধরনরে, তুলনামূলক নীরব পদ্ধতিতে মডলে আসলে ছবি আঁকো না। সেনির্দেশনা লেখে: *এখানে একটি বৃত্ত আঁকো, ওখানে একটি রিখো দাও, এই আকৃতিটি নীল করো।* এই নির্দেশনাগুলোই কলোড—ওয়েবে অধিকাংশ আইকন ও লোগোর পছন্দে থাকা স্কেলেবল ভেক্টর গ্রাফিক্স বা SVG। এর সুবিধা হলো, ফলাফল কলোনে স্থির পিক্সলে-গ্রডি নয়; বরং এমন আকৃতির সমষ্টি, যেগুলো বারবার বড়-ছোট, নতুন রঙে রঙানো বা সম্পাদনা করা যায়, তবু ঝাপসা হয় না।



মূল SVG বলুপ্ৰিন্ট আর্টওয়ার্ক: ছবিটিও নজিরে একটি সম্পাদনাযোগ্য ভেক্টর ফাইল, যেখানে স্থির পিক্সলের বদলে পথ, গ্রডি রাখো, নলোড ও হ্যান্ডলে দিয়ে গঠন তৈরি করা হয়েছে।

Laura Nesso / The Clean Paper Permission on file

সমস্যা ছিল, অল্প দিন আগেও এই ড্রয়িং-কলোড লখিতে গিয়ে মডলে কার্যত অন্ধ ছিল। সে পুরো নির্দেশনার ক্রম একবারে লখিতে—বৃত্ত, রাখো, পূরণ করা—কিন্তু মাঝপথে রেন্ডার করে দেখে নতি না আসলে কী তৈরি হচ্ছে। চোখ বন্ধ করে মুখ আঁকার কথা ভাবুন: চোখ কলোথায় হওয়া উচিত তা মনেটা মুট জানা থাকলেও দ্বিতীয় চোখটি গালে পড়ে গেছে, বা চুলের জন্য আঁকা আকৃতি নাক ঢেকে ফেলেছে—এটি ধরার কলোনে উপায় থাকবে না। এই মডেলগুলোই অবস্থাও কাছাকাছি ছিল। তাই কলোড ব্যাকরণগতভাবে বন্ধ হলোও ছবিটি দৃশ্যত বহিঃস্থল হতে পারত।

Guotao Liang-এর নেতৃত্বাধীন একটি দল এখন আপাতদৃষ্টিতে খুব সরল কাজটি করছে: তারা মডেলকে “চোখ খুলতে” দিয়েছে। তাদের *রেন্ডার-ইন-দ্য-লুপ* পদ্ধতিতে প্রতটি ধাপের পর অসম্পূর্ণ ছবিটি রেন্ডার করা হয় এবং পরের স্ট্রোক তৈরিকরার আগে সেই ছবিটি মডেলকে আবার দেখানো হয়। আসল আকর্ষণীয় বিষয়টি এই নয় যে প্রতিক্রিয়া কাজে লাগে; বরং প্রতিক্রিয়াকে কার্যকর করতে তাদের মডেলকে কীভাবে নতুন করে শেখাতে হয়েছে।

একটি আঁকার মান কীভাবে মাপা হয়?

কোনো গবেষণাপত্র যদি বলে তার মডেল অন্যটিকে “হারিয়েছে”, তখন প্রশ্ন করা উচিত: কোন কাজে, কোন মাপকাঠিতে? “একটি ল্যাপটপ আঁকো” ধরনের অনুরোধে একটি মাত্র সঠিক ছবি নেই। তাই উৎপন্ন ছবি বিচার করতে গবেষকরা কয়েকটি স্বয়ংক্রিয় স্কোর ব্যবহার করেন; প্রতিটিই মানুষের চোখে বিচারকে আনুমানিকভাবে ধরার চেষ্টা করে, নথিভাবেনয়।

এই গবেষণাপত্রে কয়েকটি পরিচিতি মটেরিকি আছে। **FID** উৎপন্ন ছবির পুরো ব্যাচের পরিসংখ্যানগত বৈশিষ্ট্য বাস্তব ছবির ব্যাচের সঙ্গতে তুলনা করে; কম স্কোর ভালো, কিন্তু এটি একটি নির্দিষ্ট ছবির মান বলে না। **CLIP স্কোর** আলাদা একটি AI-কে জিজ্ঞাসে করে ছবি প্রম্পটের ভাষার সঙ্গতে কতটা মিলে—উপকারী, কিন্তু মূল্যায়নটি যতটা নির্ভরযোগ্য, মটেরিকিও ততটাই। **DINO**, **SSIM** এবং **LPIPS** পুনর্গঠিত ছবিকে লক্ষ্যের সঙ্গতে তুলনা করে; তুলনাটিকে খনও কাঁচা পকিসলে, কখনও শেখা বৈশিষ্ট্যের স্তরে হয়। এগুলোর কোনোটিই “সত্য” নয়। এগুলো প্রতিনিধি সূচক মটেরিকি, আর পার্থক্য প্রায়ই ছোট হয়। একটি মডেল 127.6 আর অন্যটি 128.8 পলে তা দিকের হিসাবে বাস্তব পার্থক্য হতে পারে, কিন্তু এটি ভ্রমধিস নয়—একটি ছোট নড়াচড়া, তাও এমন মটেরিকি যার সঙ্গতে মানুষের চোখে মূল্যায়নের সম্পর্ক অসম্পূর্ণ। “২০ গুণ বেশি তথ্যে প্রশিক্ষিত মডেলকে হারিয়েছে” ধরনের শরিতে নাম পড়ার সময় এই সীমাটি মনে রাখা দরকার।

লেখকরা কী করছেন

তাদের লক্ষ্য ছিল এই “অন্যভাবে আঁকা” সমস্যাটি বিদলানো। আগের মডেলগুলো SVG লেখা যেনে নছিক পাঠ্য প্রজন্ম: এখন পর্যন্ত লেখা কেউ দেখে পরের কেউ খণ্ড অনুমান করো, কিন্তু তা রেন্ডার করে না। ফলে আধুনিক বহুমাধ্যমিক মডেলের ভেতরে থাকা শক্তিশালী দৃষ্টি এনকোডার প্রায় নথিকরয়ি থাকে। রেন্ডার-ইন-দ্য-লুপ কাজটিকে ধাপে ধাপে দৃশ্যমান প্রকরয়ি বদলে দেয়। প্রতিটি অঙ্কন-কেউ খণ্ডের পরে আংশিক SVG রেন্ডার হয়, সেই ছবি আবার মডেলে যায়, এবং পরের খণ্ড বছে নওয়া হয় ইতিমধ্যে আঁকা ক্যানভাসটি দেখে।

প্রথম ফলটি সতর্ক করার মতো, এবং লেখকরা সটো স্পষ্টভাবে দেখান: শুধু বদ্যমান প্রস্তুত-ব্যবহারযোগ্য মডেলের সঙ্গতে এই লুপ জুড়ে দলিই কাজ হয় না। বিশেষ প্রশিক্ষণ ছাড়া শক্তিশালী সাধারণ মডেলকে মধ্যবর্তী উপস্থাপন দেখালে মান বাড়েনি—বরং সব ক্ষেত্রে কমছে। আগে কখনও এইভাবে নিজের দৃশ্যমান আউটপুট ব্যবহার করতে শেখেনি এমন মডেল, হঠাৎ ছবি দেখলেই সেই প্রতিকরয়ি সঠিকভাবে কাজে লাগাতে পারে না।

তাই মূল কাজটি প্রশিক্ষণে। গবেষকরা প্রশিক্ষণ-তথ্য পুনর্গঠন করেন, যাতে প্রতিটি অঙ্কন অনেকে ছোট, দৃশ্যত অর্থবহ ধাপে ভাঙা থাকে। জটিল আকৃতি ছোট অংশে ভাগ করা হয়, যাতে প্রতিটি ধাপে ক্যানভাসে নতুন কিছু দেখা যায়। এরপর Qwen3-VL-ভিত্তিক প্রায় আট-বিলিয়ন-প্যারামটার উন্মুক্ত মডেলকে এই ধাপে ধাপে ক্রমে সূক্ষ্ম-সামঞ্জস্য করা হয়। তারা এই পদ্ধতিকে **দৃশ্যমান স্ব-প্রতিকরয়ি** বলে।

অঙ্কনের সময় তারা আরেকটি প্রকরয়ি যোগ করেন, **রেন্ডার-ও-যাচাই**। নতুন স্ট্রোক গ্রহণ করার আগে সটো রেন্ডার করে দেখা হয় স্ট্রোকটি সত্যিই ছবি বদলেছে কিনা, নাকি আগের কাজটাই পুনরাবৃত্তি করছে। কোনো পরিবর্তন না আনলে স্ট্রোক বাদ দেওয়া হয়; আরও স্ট্রোক কাজে না দলে মডেলকে থামতেও বলা হয়। উল্লেখযোগ্য হলো, পুরো প্রশিক্ষণ ডটোসটে তুলনামূলক ছোট—প্রায় **850,000** উদাহরণ। এটি এক প্রতদিন্দবীর ডটোসটে অর্ধেকেরও কম, আর অন্য প্রতদিন্দবীর মাত্র একটি ছোট অংশ।

তারা কী পয়েছেন

- এইভাবে প্রশিক্ষণ দেওয়া মডেল তার ব্লাইন্ড সমতুল্যের তুলনায় ভালো আঁকে। পার্থক্য ব্যর্থতা ক্ষেত্রে বিশেষ স্পষ্ট: ব্লাইন্ড মডেল কখনও চোখ গালে বসায়, আবার অনুরোধ করা দণ্ড চার্ট বাদ দিয়ে সাধারণ মনটির আঁকে; দৃশ্যমান-প্রতিকরয়ি মডেল এমন ভুল বেশি ধরতে পারে।
- মানক মানদণ্ড **MMSVGBench**-এ পদ্ধতিটি শক্তিশালী প্রতদিন্দবীদের সমতুল্য, এবং কয়েকটি মটেরিকি সামান্য এগিয়ে। এর মধ্যে OmniSVG আছে, যটো দ্বিগুণের বেশি তথ্যে প্রশিক্ষিত, এবং InternSVG আছে, যটো প্রায় **20×** বেশি তথ্য ব্যবহার করেছে।
- পার্থক্য ছোট। আইকন সটে, উদাহরণ হিসেবে, প্রধান ছবি-মান স্কোর **127.6**, সেরা প্রতদিন্দবীর **128.8**-এর তুলনায়; প্রম্পট-

মলিযুক্ত স্কোর 0.293 বনাম 0.291। দকিরে হিসাবে এগুলো সামঞ্জস্যপূর্ণ, কিন্তু ব্যবধান সবু।

- দুটো যোগ করা উপাদানই অবদান রাখে। বিশেষ প্রশিক্ষণ বন্ধ করলে, বা অঙ্কন-সময় যাচাই সরালে, স্কোর মাপযোগ্যভাবে কমে। যাচাই বিশেষ করে মডেলকে একই জনিসি বারবার আঁকার লুপে আটকে পড়া থেকে রক্ষা করে।
- লেখকেরা সবচেয়ে জোর দেন **দক্ষতা**-তে: শীর্ষ মডেলগুলো তুলনায় অনেকে কম প্রশিক্ষণ-তথ্য ব্যবহার করে কাছাকাছি বা সামান্য ভালো মানদণ্ড কার্যসম্পাদন পাওয়া।

এটাকী প্রমাণ করো না

- এটি দেখায় না যে মডেলকে “দেখতে দেওয়া” নজিই মুক্ত সাফল্য। গবেষণাপত্রের নজিস্ব ফল উল্টো: **পুনঃপ্রশিক্ষণ ছাড়া দৃশ্যমান প্রতিক্রিয়া কার্যসম্পাদন খারাপ করে**। লাভ আসে মডেলকে প্রতিক্রিয়া ব্যবহার করতে শেখানোর পর।
- এটি বড় বা নরিণায়ক ব্যবধান দেখায় না। অধিকাংশ স্কোরে প্রতদ্বন্দ্বীদের সঙ্গে ব্যবধান খুব কম। “ $20\times$ তথ্যে প্রশিক্ষিত মডেলকে হারিয়েছে” বাক্যটি সত্য, কিন্তু এক মানদণ্ডের প্রতিনিধি সূচক মটেরকি ছোট ব্যবধান।
- এটি সাধারণ শৈল্পিক ক্ষমতা প্রমাণ করে না। এটি **8B** গবেষণা মডেল, প্রধানত আইকন ও সরল চিত্রণ তৈরি করছে ছোট স্থির রঙের লিডিশন— **224×224 পিক্সলে**—সাধারণ-উদ্দেশ্য নকশাকৃত নয়।
- GPT-5-এর মতো বড় সাধারণ মডেলের সঙ্গে তুলনা সরাসরি সমতুল্য নয়। সে সংকীর্ণ SVG-কোড কাজের জন্য তৈরি বা টিউনকৃত নয়; এখানে তাকে হারানো দুই মডেলের সাধারণ সক্ষমতা সম্পর্কে খুব কমই বল।
- এই লুপ বন্ধ খরচ আসে না। প্রতিটি ধাপে ক্যানভাস উপস্থাপন ও আবার পড়তে হয় বলে প্রজন্ম, একবারে কোড লিখে দেওয়ার চেয়ে ধীর হয়। লেখকেরা এই খরচ স্বীকার করেন।

প্রমাণ কতটা শক্ত

- **নয়নত্রতি তুলনার অংশটি শক্ত**। Ablation স্পষ্ট: প্রশিক্ষণ সরালে বা যাচাই সরালে স্কোর কমে। তাই দুই উপাদানই লেখকের দাবি অনুযায়ী কাজে লাগছে—এটি শুধু গল্প নয়।
- **অপ্রত্যাশিত ঋণাত্মক ফলটি সংজ্ঞা দেয় না**। হয়ছে। সরাসরি যোগ করা দৃশ্যমান প্রতিক্রিয়া কষতকিরে—এটি গবেষণাপত্রের সবচেয়ে আকর্ষণীয় ফলগুলো। একটা কারণ “বর্শে ইনপুট মানই ভালো” ধারণাটিকে এটি সংশোধন করে।
- **লড়িরবে রঙ দাবি তুলনামূলক পাতলা**। বর্শে-তথ্য ব্যবহারকারী প্রতদ্বন্দ্বীর ওপর জয় ছোট, এবং একটি মাত্র মানদণ্ডে; মানদণ্ডটির সঙ্গে প্রতদ্বন্দ্বী দলেরও সম্পর্ক আছে। এক পরীক্ষা-সটে প্রতিনিধি সূচক মটেরকি ছোট ব্যবধান ইঙ্গিত দিয়ে, চূড়ান্ত সিদ্ধান্ত নয়।
- **বড় পরিসর বা বাস্তব নকশা কর্মপ্রবাহে পরীক্ষা হয়নি**। এখানে নমিনরজে লিডিশন আইকন ও সরল চিত্রণ। জটিল, উচ্চরঙের লিডিশনের বা বাস্তব জগতের গ্রাফিক্সে একই সুবিধা থাকবে কিনা ভবিষ্যৎ কাজ। লেখকেরাও এ সীমা বলেন।

কেন এটি গুরুত্বপূর্ণ

গবেষণাপত্রের কনন্দ্রীয় ধারণা প্রায় অস্বস্তিকরভাবে সহজ, কিন্তু অঙ্কনের বাইরেও প্রযোজ্য। কনো কর্মসূচিকে লিখে ওয়েবে পৃষ্ঠা, চার্ট, চিত্র বা 3D দৃশ্য বানাতে তার সামনে দুই পথ: পুরো কোড অন্ধভাবে লিখে শেষে আশা করা যে সব ঠিকি হয়ছে; অথবা কাজ চলার সময় উপস্থাপন দেখে পরের ধাপ ঠিক করা। মানুষের জন্য দ্বিতীয়টি স্বাভাবিক—আমরা বারবার নজিরে কাজের দিকে তাকাই। এই ধরনের মডেলের জন্য লুপ বন্ধ করার ধারণাটি আশ্চর্যজনকভাবে নতুন।

তবে গবেষণাপত্রের সবচেয়ে গুরুত্বপূর্ণ সতর্কচিহ্নটি হলো: **দেখতে পারা আর কীভাবে দেখা তথ্য ব্যবহার করতে হয় তা শেখা এক জনিসি নয়**।

দৃশ্যমান চ্যানেলে থাকলেই মডলে প্রতিক্রিয়াটি বুঝে কাজে লাগাতে পারবে না; সঠিক আলাদা করে শেখাতে হয়। প্রশিক্ষণের পরে লুপটিকা কাজে দিয়ে, কনিন্তু লাভ পরমিতি ও সীমিত—একজন আরও স্থায়ী অঙ্কনকারী, একবোরো নতুন ধরনের শিল্পী নয়। বর্ণনা থেকে সম্পাদনাযে গ্য গ্রাফিক বানানোর সরঞ্জামের জন্য ব্যবহারিক শিক্ষাটি বড়: এখানে লাভ এসেছে শুধু বেশি তথ্য চলে নয়, বরং কম তথ্যকে আরও কার্যকরভাবে কাজে লাগিয়ে। লিডারবোর্ডে আরকেটি বিন্দুর চোখে এই শিক্ষা বেশি মূল্যবান।

সংক্ষিপ্ত সারাংশ

ভেক্টর গ্রাফিক্স তৈরির ভাষা মডলে—ওয়েবের আইকন ও লোগোর পছন্দের সম্পাদনাযে গ্য, যেকোনো মাপে বাড়ানো যায় এমন কৌড়—দীর্ঘদিন মূলত “অনুভাব” কাজ করেছে: অঙ্কনের নির্দেশে লিখেছে, কনিন্তু মাঝপথে ফলটি রেন্ডার করে দেখেনি। Guotao Liang-এর দল **render-in-the-loop** পদ্ধতি প্রস্তাব করেছে: প্রতিটি ধাপের পর অসম্পূর্ণ অঙ্কন রেন্ডার করে মডেলকে আবার দেখানো হয়, যাতে পরের স্ট্রোকে ক্যানভাস দেখে বোঝে নেওয়া যায়। গুরুত্বপূর্ণ ফলটি হলো, বদ্যমান মডলে শুধু এই প্রতিক্রিয়া যোগ করলে কার্যসম্পাদন *খারাপ* হয়; উন্নতি আসে তখনই, যখন মডেলকে সেই প্রতিক্রিয়া ব্যবহার করার জন্য পুনঃপ্রশিক্ষণ দেওয়া হয়। অঙ্কনের সময় যাচাই পুনরাবৃত্ত বা অকার্যকর স্ট্রোকে বাদ দিতেও সাহায্য করে। পুনঃপ্রশিক্ষণ 8B মডলে মানক মানদণ্ডে অনেকে বেশি তথ্য দিয়ে প্রশিক্ষণ প্রতিলব্ধী মডলের সমান বা সামান্য এগিয়ে যায়, তবে ব্যবধান ছোট, মটেরিকি কেবল প্রতিনিধি সূচক, আর কাজটি নিম্নরাজে লিউশনের আইকন ও সরল চিত্রে সীমিত। লেখকদের মূল কথা দক্ষতা: নিজের কাজ ভালোভাবে দেখতে শেখানো কখনো কখনো নছিক তথ্যের পরিমাণ বাড়ানোর বকিল্প হতে পারে।

সরাসরি যাচাই

গবেষণাপত্র কী দেখায়: ভেক্টর-গ্রাফিক্স মডেলকে ধাপে ধাপে নিজের অসম্পূর্ণ অঙ্কন উপস্থাপন করে দেখতে শেখালে ব্লাইন্ড প্রজনমের তুলনায় ফল ভালো হয়; এক মানক মানদণ্ডে কম প্রশিক্ষণ-তথ্য ব্যবহার করেও বড়-তথ্য প্রতিলব্ধী-দের সমতুল্য বা সামান্য ভালো কার্যসম্পাদন পাওয়া যায়।

কী সম্ভাব্য কনিন্তু প্রমাণিত নয়: “নিজের কাজ দেখে নেওয়া” বস্তুতভাবে বেশি তথ্য পরিসর করার চোখে ভালো পদ্ধতি; জটিল/উচ্চরাজে লিউশনের বাস্তব জগতের গ্রাফিক্সে একই লাভ থাকবে; মানদণ্ডের ছোট ব্যবধান মানুষের চোখেও স্পষ্ট হবে।

কী দেখায় না: কেবল দৃশ্যমান প্রতিক্রিয়া যোগ করলেই উন্নতি হয়—পুনঃপ্রশিক্ষণ ছাড়া বরং ক্ষতি হয়েছে; প্রতিলব্ধী মডলের ওপর বড় ও নির্ণায়ক ব্যবধান; সাধারণ উদ্দেশ্যের শিল্পসৃষ্টির ক্ষমতা; অথবা GPT-5-এর মতো সাধারণ মডলের সঙ্গে ন্যায়্য সরাসরি তুলনা।

মূল সীমাবদ্ধতা: একটি মানদণ্ড; প্রতিনিধি সূচক মটেরিকি ছোট ব্যবধান; 8B গবেষণা মডলে; 224×224 আইকন/সরল চিত্রণ; এবং প্রতিটি ধাপে উপস্থাপন করার কারণে ধীর প্রজনম।

সাধারণ পাঠকের আস্থা কতটা হওয়া উচিত? উচ্চ আস্থা যেকৌড় দিয়ে আঁকা মডলের জন্য উপস্থাপন করে নিজের কাজ দেখে সত্যিই উপকারী হতে পারে, কনিন্তু সেই লুপটিকা প্রশিক্ষণ দিতে হয়। এই মডলে প্রতিলব্ধী-দের নির্ণায়কভাবে হারিয়েছে—এ দাবিতে নিম্ন থেকে মাঝারি আস্থা। সবচেয়ে শক্ত শিক্ষা: অঙ্কন কৌড়ের ক্ষেত্রে ব্লাইন্ড একবারের প্রজনমের চোখে শেখানো দৃশ্যমান প্রতিক্রিয়া ভালো কাজ করতে পারে।

উৎস

Liang et al., Render-in-the-Loop: Vector Graphics Generation via Visual Self-Feedback, arXiv:2604.20730 (2026)
<https://arxiv.org/abs/2604.20730>

OmniSVG project page
<https://omnisvg.github.io/>

InternSVG project page
<https://hmwang2002.github.io/release/internsvg/>