Portfolio

Lutao Yan

Personal Homepage: http://lutaoyan.github.io

985 **Bachelor** **Data Science and Big Data Technology**

Keywords:

- Dataset
- MLLM (Multimodal Large Language Model)
- Data and Visual Analysis
- Computer Vision
- Artificial Intelligence
- Prompt and Fine-tuning on LLM

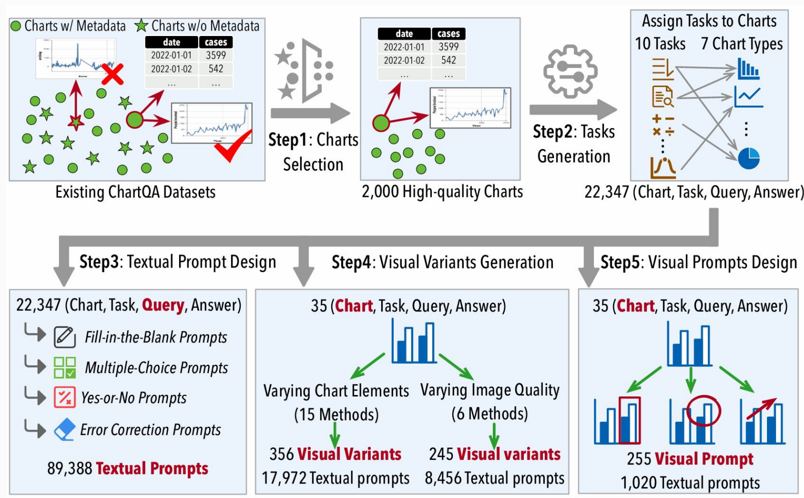
Note: 🔗 **means link access to other site**

Chart-Insight: A Large-Scale Dataset for Visual Analytics

Contribution & Honor

- Collaborated with HKUST & RUC
- A Recommendation Letter From Prof. LUO
- First Dataset focus on Low-Level Visual Task
- Available Metadata(eg. tables,pics,QA pairs)
- Abundant Visual and Textual Variants
- Support Investigating Performance of MLLMs

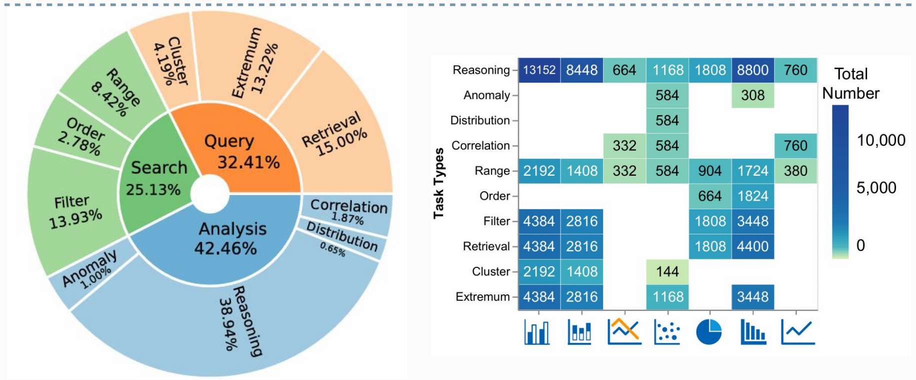
Pipeline for Dataset Construction



Highlight

- 5 key steps to construct Chart-Insight
- 2K high-quality charts
- Extensive and measurable difficulties
- Innovative 4 forms of textual prompt design
- Compared with Existing Datasets, available metadata facilitate future research
- Average 44.5 questions / chart, deep excavation of chart

Overview of Chart-Insights

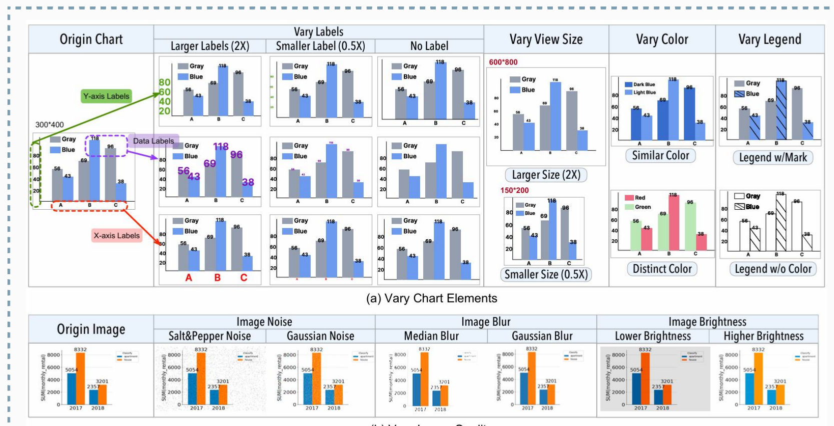


Low-level Tasks	Analysis Tasks (42.46%)				Search Tasks (25.13%)			Query Tasks (32.41%)		
	Reasoning	Anomaly	Distribution	Correlation	Determine Range	Order	Filter	Retrieve Value	Find Extremum	Find Cluster

Highlight

- 10 basic analytic tasks across 7 widely-used chart types
- Distribution on vast fine-graine task vs. chart, eg. Bar, Line, Scatter, Pie
- 10 low-level tasks into 3 categories, eg. Analysis, Search, and Query
- 89,388 quartets (chart, task,question,answer)
- In-depth evaluation on impact of basic chart element

Visual Element on Charts

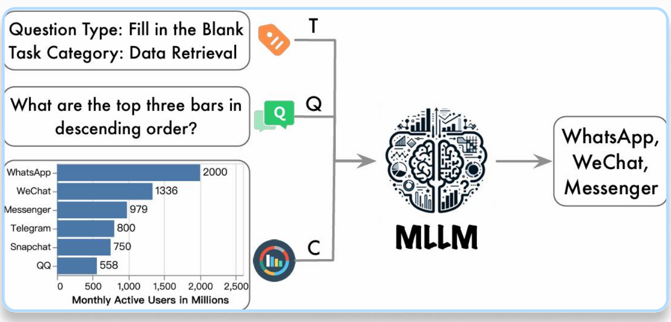


Chain-of-Charts: A Novel Method to Improve Performance of MLLM

Contribution & Honor

- Summited to IEEE VIS (Flagship Conference)
- Co-first Author of Paper
- Improve MLLM Performance by 24% in the field of Visual Analysis
- Transferable like Chain-of-Thought
- Can be combined with Visual Prompts

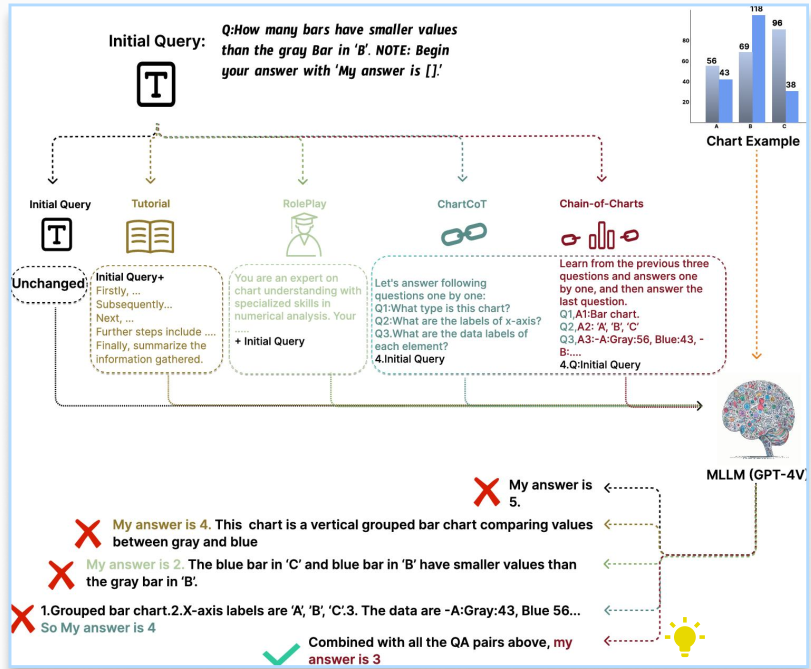
Evaluation Framework



Highlight

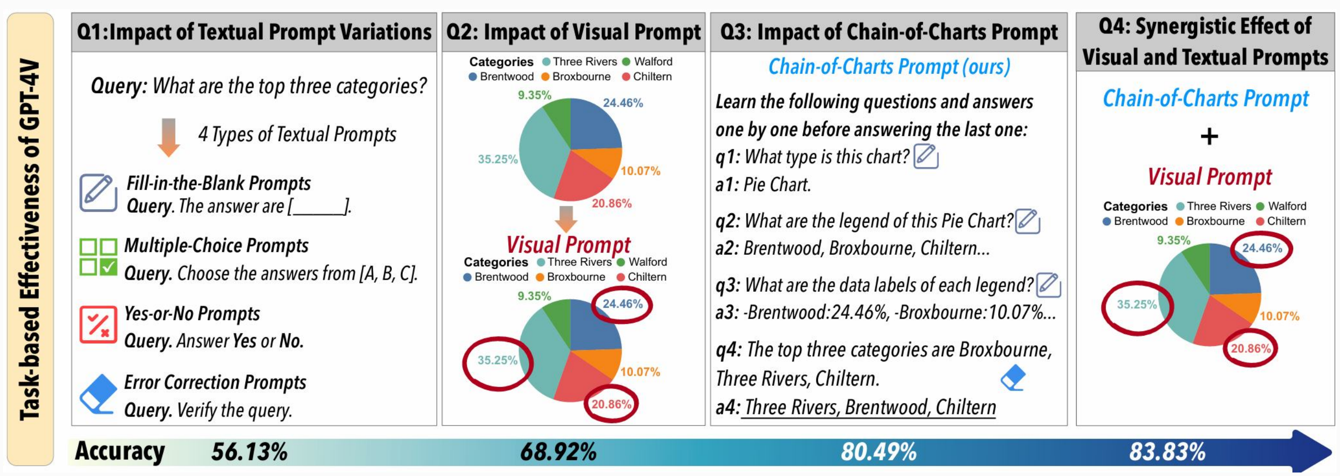
- A reasonable framework (T,Q,C)to evaluate MLLM
- With the bonus of Visual Prompt & Chain-of-Charts, the accuracy rate is increased from 56.13% to 83.83%!
- Shed light on the capabilities and limitations of MLLM
- Offer valuable insights for future research

Chain-of-Charts vs. Other Prompt Method



Highlight

- Chain-of-charts demonstrates its effectiveness and Interpretability
- Better than other common methods of enhancement (eg. Tutorial, Role-Play)
- Progressively guide the model towards a deeper understanding of charts
- Significantly improved GPT-4V's capabilities across 10 different tasks
- Developing visual prompts specifically is a promising research direction



3D Lipstick Effect: A Tool to Make Face Fancy

Feature & Honor

- Top 3% Course Design (3D Vision and AI)
- Based on Google MediaPipe AI Framework
- Core Algorithm: Face Mesh
- Extract Key Points to Make Effect on Face

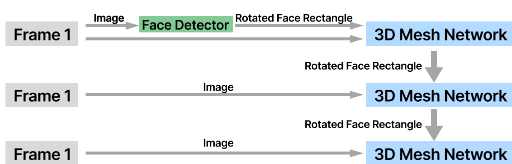
Background

In the current context of rapidly evolving **artificial intelligence (AI)** and **computer vision** technologies, my project - **3D lipstick effects** based on MediaPipe - is an innovative practice in this trend. Combining the **high-precision facial tracking** technology with **3D graphics rendering**, this technology has a wide range of practical applications, especially in the **e-commerce and social media**.

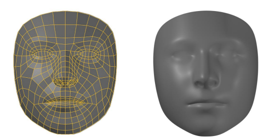
Principle & Demo

Theory

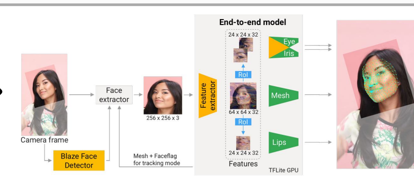
Reality



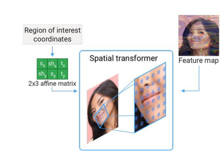
Face Detection with Rectangle



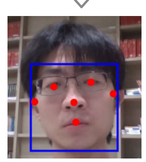
468 Key Point Detection & 3D Modeling



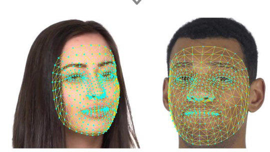
Attention Mesh to Predict Important Region



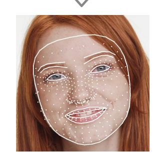
Lip Attention Mechanism



Dynamic Capture



Face Mesh Prediction



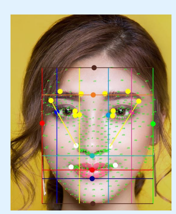
Key Contours: Eyes and Lip

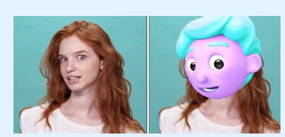
AR Makeup

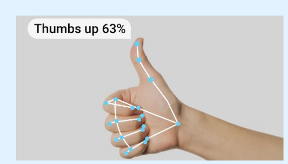
Details

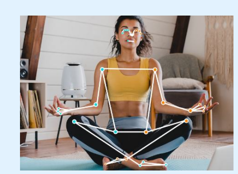
- Available Source Code of this programme
- Demo Video showing basic function
- More open sourced improved applications
- Complete Design Report

Technology Application









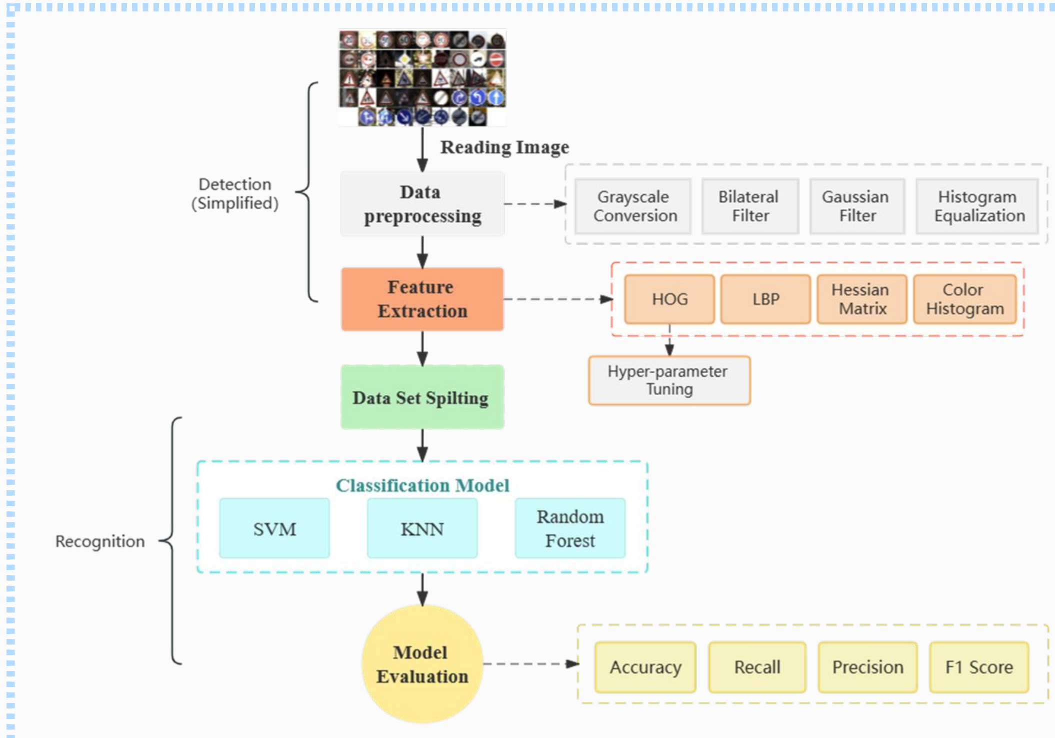
- *Age prediction*
- *Puppeteering*
- *Gesture recognition*
- *Posture detection*

Traffic-sign Detection and Recognition

Feature & Honor

- School of Computing Programme in NUS
- Distinction (Top) Assessment Certificate
- Traffic Sign dataset-based Deep learning
- Analysis Below 3% Error Rates

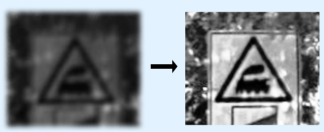
Overview of Project



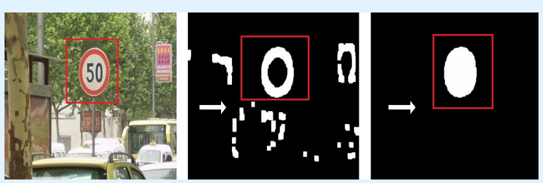
Highlight

- Sort large-scale dataset into 31,367 training images, 7,842 validation images, and 12,630 testing images

- 4 Images Pre-processing Combination Method, Effectively Improve Image Quality



- Feature extraction algorithm (Especially HOG) to filter out the misdetected non-traffic sign areas Precise Positioning Step by Step



- Multi-Evaluation to Find Best Model Combination

Model	Accuracy	Precision	Recall	F1 Score
HoG + KNN	0.939	0.912	0.907	0.909
Hessian + KNN	0.932	0.933	0.926	0.929
HoG + Random Forest	0.991	0.996	0.983	0.989
Hessian + random forest	0.977	0.989	0.961	0.975
LBP + SVM (linear kernel)	0.928	0.969	0.901	0.934
HoG + SVM (poly kernel)	0.986	0.991	0.981	0.986

Gallery

Defense Scene