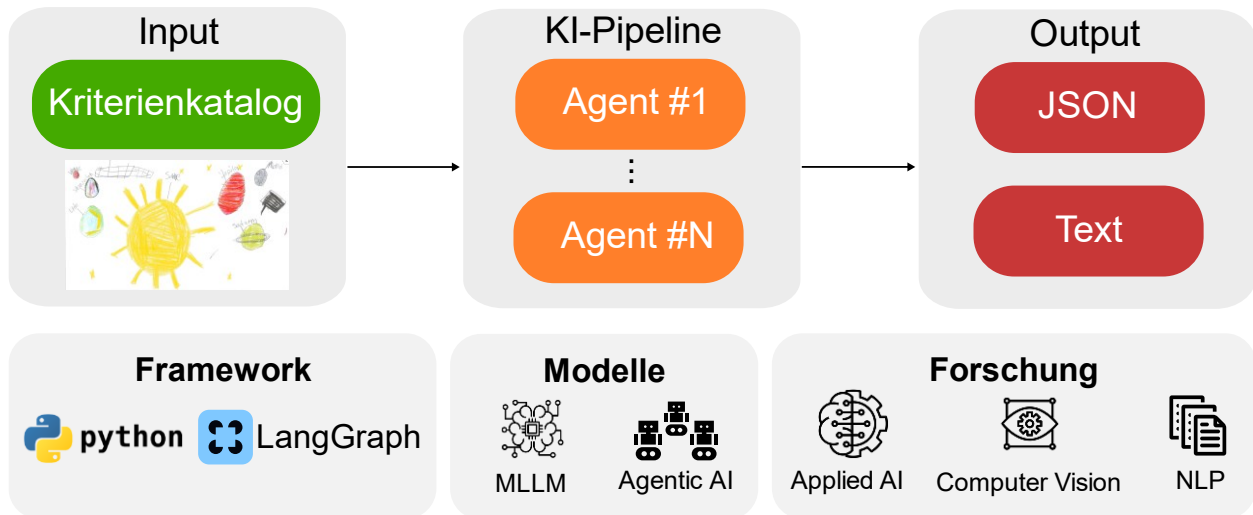


Ausschreibung Bachelor-Thesis / Master-Thesis

Entwicklung einer MLLM-basierten Pipeline zur automatisierten semantischen Analyse von Schülerzeichnungen



Ausgangslage

Im Rahmen eines Forschungsprojekts der AG Didaktik des Sachunterrichts zu Vorstellungen vom „Aufbau des Sonnensystems“ in unterschiedlichen Bildungsphasen entsteht ein kontinuierlich wachsender Bilddatensatz. Dieser umfasst handgezeichnete Darstellungen von Vorschulkindern, Schüler*innen und Studierenden. Die Auswertung der visuellen Konzepte in den Bildern erfolgt bislang manuell anhand eines qualitativen Kriterienkatalogs. Aufgrund des hohen Aufwands und der steigenden Anzahl an Bildern ist dieses Verfahren jedoch nur eingeschränkt skalierbar. Ziel ist es daher, den Prozess mithilfe einer automatisierten, KI-gestützten Bildanalyse- und NLP-Pipeline effizient und reproduzierbar zu gestalten.

Problemstellung

Multimodale Large Language Models (MLLMs) ermöglichen die simultane Verarbeitung von Text- und Bildmodalitäten und bieten mittels Zero- oder Few-Shot-Prompting fortgeschrittene Möglichkeiten zur semantischen Bildanalyse. Die Herausforderung besteht darin, eine robuste End-to-End-Pipeline zu entwickeln, die komplexe visuelle Konzepte in Kinderzeichnungen kriteriengeleitet extrahiert und strukturiert validiert. Hierzu soll die Eignung moderner MLLM-Agentenarchitekturen für die Orchestrierung des Analyseprozesses sowie für eine Fehlerkorrektur und Konsistenzprüfung untersucht und eine entsprechende Implementierung (z. B. mit Frameworks wie LangGraph oder AutoGen) vorgenommen werden.

Vorgehensweise und Erwartete Ergebnisse

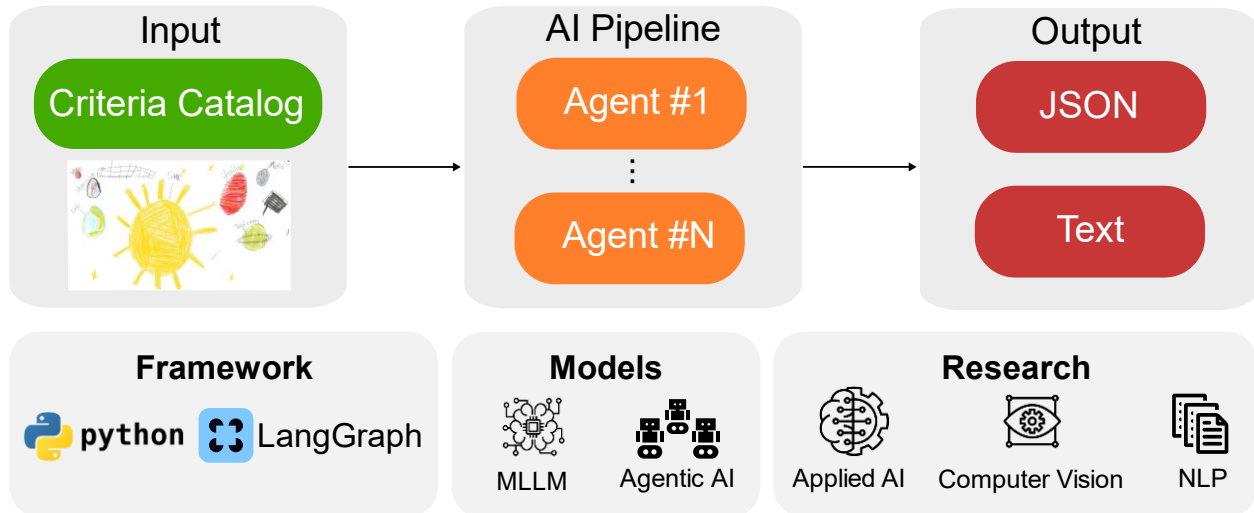
Zunächst führen Sie eine systematische Literaturrecherche zu MLLMs, Agenten, KI-gestützten Bildanalyseverfahren und für diesen Anwendungsfall geeigneten Evaluationsmetriken durch. Darauf aufbauend konzipieren und implementieren Sie eine MLLM- und agentenbasierte Pipeline zur kriteriengeleiteten visuellen Analyse der Zeichnungen. Abschließend evaluieren Sie die Performance und Robustheit Ihres Ansatzes durch einen quantitativen und qualitativen Abgleich der automatisiert erzeugten Analyseergebnisse mit den vorliegenden manuell erstellten Analysedaten unter Nutzung des vorgegebenen Kriterienkatalogs.

Ansprechpartner

Dr. Stephan Tilgner | E-Mail: tilgner@uni-wuppertal.de

Bachelor's thesis / Master's thesis

Development of an MLLM-based Pipeline for the Automated Semantic Analysis of Student Drawings



Initial Situation

A continuously growing image dataset is being compiled as part of a research project conducted by the Research Group for the Didactics of Primary Science and Social Studies, focusing on conceptions of the “structure of the solar system” across different educational stages. The dataset comprises hand-drawn representations produced by preschool children, school pupils, and university students. To date, the visual conceptions represented in these images have been analyzed manually using a qualitative coding framework. However, due to the considerable effort involved and the increasing number of images, this approach is only scalable to a limited extent. The aim is therefore to make the analysis process more efficient and reproducible by means of an automated, AI image analysis and NLP pipeline.

Problem Definition

Multimodal large language models (MLLMs) can process text and images simultaneously and offer advanced capabilities for semantic image analysis through zero-shot or few-shot prompting. However, developing a robust end-to-end pipeline that can extract complex visual concepts from children’s drawings in a criteria-driven manner and validate it in a structured way remains challenging. In order to achieve this objective, the feasibility of employing contemporary MLLM agent architectures for orchestrating the analysis process, error correction, and consistency checking should be investigated and implemented (e.g., via frameworks such as LangGraph or AutoGen).

Methods and Expected Results

First, you will conduct a systematic literature review of MLLMs, AI agent systems, AI-based image analysis methods, and evaluation metrics suitable for this use case. Then, you will design and implement an MLLM- and agent-based pipeline for criteria-driven visual analysis of students’ drawings. Finally, you will evaluate the performance and robustness of your approach by comparing the automated results with existing manually generated analysis data quantitatively and qualitatively while considering the criteria catalog.

Contact Person

Dr. Stephan Tilgner | **E-Mail:** tilgner@uni-wuppertal.de