



Problem

Industrielle Qualitätskontroll-Workflows erfordern spezialisierte klassische Computer Vision-Lösungen zur Defekterkennung, was traditionell sowohl Domänenexpertise als auch fortgeschrittene Programmierkenntnisse voraussetzt. Die Erstellung filterbasierter Bildverarbeitungs-Workflows ist zeitaufwändig und teuer, da Computer Vision-Ingenieurinnen das Wissen der Domänenexpertinnen in funktionsfähigen Code übersetzen müssen.

Viele Unternehmen, insbesondere kleinere, haben Schwierigkeiten bei der Implementierung effektiver Computer Vision-Workflows aufgrund dieser Expertise-Lücke und des Mangels an visuellen Bearbeitungsmöglichkeiten. Der traditionelle Entwicklungsprozess ist ineffizient und erfordert mehrere Iterationen zwischen Domänenexpertinnen, die industrielle Defekte verstehen, und Ingenieurinnen, die klassische Computer Vision-Algorithmen implementieren können. Selbst wenn Lösungen entwickelt werden, erfordert deren Modifikation und Anpassung die Rückkehr zu Programmierungsspezialist*innen, was zu anhaltenden Engpässen führt.

Lösung

Wir entwickeln ein multimodales Large Language Model (MLLM), das in visuelle Programmierplattformen integriert ist und klassische, filterbasierte Computer Vision-Workflows aus Beispielbildern und natürlichsprachlichen Beschreibungen generieren kann, um sie dann visuell bearbeitbar zu machen. Dieses Modell kombiniert die Fähigkeiten von Vision-Language-Modellen (VLM) und erklärbarer Code-Generierung, um die Lücke zwischen visuellem Verständnis und interaktiver Workflow-Implementierung zu schließen.

Unter Nutzung von Few-Shot-Learning-Prinzipien können Nutzerinnen wenige Beispielbilder (wie Defektmuster) zusammen mit natürlichsprachlichen Beschreibungen als Anforderungen bereitstellen, und das Modell generiert entsprechende klassische Computer Vision-Workflows. Diese Workflows werden dann in intuitive visuelle Programmierungsschnittstellen integriert, wo Nutzerinnen Komponenten per Drag-and-Drop verwenden, Parameter über grafische Steuerelemente anpassen und Verarbeitungsschritte ohne Programmierkenntnisse modifizieren können.

Impact

Diese Lösung ermöglicht es Domänenexpertinnen, ihr Fachwissen direkt in reproduzierbare und visuell anpassbare Computer Vision-Workflows zu übertragen, ohne umfangreiche Programmierkenntnisse zu benötigen. Die visuellen Bearbeitungsmöglichkeiten erlauben Echtzeitanpassungen und iterative Verbesserungen und eliminieren den Kommunikationsengpass zwischen Domänenexpertinnen und Programmierungsspezialist:innen.

Sie beschleunigt sowohl die anfängliche Entwicklung als auch die laufende Anpassung industrieller Computer Vision, indem visuelle Beispiele und textuelle Anforderungen direkt in bearbeitbare, erklärbare Workflows übersetzt werden. Dies reduziert Entwicklungszeit und -kosten erheblich und ermöglicht kontinuierliche Anpassungen, wodurch Unternehmen Computer Vision-Lösungen für Qualitätskontrollanwendungen schnell prototypisieren und eigenständig warten können.

Partners

1. Scope Sorting GmbH, Anton-Huber-Straße 20, 73430 Aalen
2. Mountain IT-Services & Security Vadim Mitin, Anton-Huber-Straße 20, 73430 Aalen.

Project participants, KI-Werkstatt:

Prof. Dr. Tim Dahmen
Ziwei He



Kofinanziert von der Europäischen Union



Baden-Württemberg



Problem

Industrial quality control workflows require specialized classical computer vision solutions for defect detection, which traditionally demands both domain expertise and advanced programming skills. Creating filter-based image processing workflows is time-consuming and expensive, as it requires computer vision engineers to translate domain experts' knowledge into working code.

Many companies, especially smaller ones, struggle to implement effective computer vision workflows due to this expertise gap and the lack of visual editing capabilities. The traditional development process is inefficient, requiring multiple iterations between domain experts who understand the industrial defects and engineers who can implement the classical computer vision algorithms. Even when solutions are developed, modifying and adapting them requires going back to programming specialists, creating ongoing bottlenecks.

Solution

We create a multimodal Large Language Model (MLLM) integrated with visual programming platforms that can generate classical, filter-based computer vision workflows from sample images and natural language descriptions, then make them visually editable. This model combines the capabilities of vision-language models (VLM) and explainable code generation to bridge the gap between visual understanding and interactive workflow implementation.

Leveraging few-shot learning principles, users can provide a few example images (like defect patterns) along with natural language descriptions as requirements, and the model generates appropriate classical computer vision workflows. These workflows are then integrated into intuitive visual programming interfaces where users can drag-and-drop components, adjust parameters through graphical controls, and modify processing steps without programming knowledge.

Impact

This solution enables domain experts to directly transfer their expertise to reproducible and visually customizable computer vision workflows without requiring extensive programming knowledge. The visual editing capabilities allow for real-time adjustments and iterative improvements, eliminating the communication bottleneck between domain experts and programming specialists..

It accelerates both initial development and ongoing adaptation of industrial computer vision by translating visual examples and textual requirements directly into editable, explainable workflows. This significantly reduces development time and costs while enabling continuous customization, allowing companies to rapidly prototype and independently maintain their computer vision solutions for quality control applications.

Partners

1. Scope Sorting GmbH, Anton-Huber-Straße 20, 73430 Aalen
2. Mountain IT-Services & Security Vadim Mitin, Anton-Huber-Straße 20, 73430 Aalen.

Project participants, KI-Werkstatt:

Prof. Dr. Tim Dahmen
Ziwei He



Kofinanziert von der
Europäischen Union



Baden-Württemberg

