

Defect.AI 2

Data Efficient AI for Visual Defect Detection in Media and Manufacturing

Programm / Ausschreibung	Kooperationsstrukturen, Kooperationsstrukturen, Bridge Ausschreibung 2023	Status	laufend
Projektstart	01.10.2023	Projektende	31.03.2026
Zeitraum	2023 - 2026	Projektlaufzeit	30 Monate
Keywords	AI; manufacturing; media; defect; inspection		

Projektbeschreibung

Die Erkennung von Defekten mit Hilfe von Computer-Vision-Methoden hat sich in den letzten Jahren in verschiedenen Industriezweigen, z. B. in der Medien- und Fertigungsindustrie, zunehmend durchgesetzt. Dazu gehören zum Beispiel die Erkennung von Staub und Kratzern in Filmmaterial, Blocking Störungen in übertragenen oder digitalisierten Video und Oberflächen- und strukturelle Fehler in RGB und Hyperspektralbildern produzierter Güter. Die meisten aktuell eingesetzten Systeme verwenden hochentwickelte Algorithmen, in die das Fachwissen von Experten eingeflossen ist.

Zwar haben Deep-Learning-Ansätze in vielen Bildverarbeitungsanwendungen erhebliche Fortschritte gebracht, doch die Spezifität vieler Detektionsaufgaben und der Mangel an Trainingsdaten (sowohl wegen der Kosten für die Erstellung von Datensätzen als auch wegen des seltenen Auftretens vieler relevanter Fehlertypen) haben den Einsatz dieser Methoden bisher erschwert. Die Anwendbarkeit von Deep-Learning-Methoden in diesen Detektionsaufgaben würde es ermöglichen, sie an spezifische Fehlertypen anzupassen, die bei bestimmten Anwendern auftreten, und langfristig die Entwicklungs- und Wartungskosten zu senken.

Defect.AI zielt darauf ab, Ansätze für dateneffizientes Training von visuellen Detektionsmethoden zu erforschen. Der Ansatz wird auf aktuelle Fortschritte im maschinellen Lernen aufbauen, wie z.B. Foundation Models, die selbstüberwacht auf großen unannotierten Datensätzen vortrainiert werden können und dann mit Hilfe von Transfer-Learning an eine spezifische Aufgabe angepasst werden. Ein weiterer wichtiger Aspekt ist die Erfassung des Expertenwissens, das in den vorhandenen, manuell gebauten Detektoren steckt. Knowledge Distillation wird eingesetzt, um dieses spezifische Wissen zu lernen. Dies ist skalierbar und ohne zusätzliche manuelle Annotation möglich. Zur Überbrückung größerer Domänenlücken (Domain Adaptation) werden generative Ansätze wie GANs und Diffusionsmodelle untersucht, um z.B. Style Transfer zwischen Domänen wie RGB und Infrarot durchzuführen.

Die Verfügbarkeit von Deep-Learning-basierten Defekterkennungsansätzen, die mit kleinen Datensätzen an neue Varianten von Defekten angepasst werden können, eröffnet mittelfristig die Perspektive, dass Benutzer die auf ihren Systemen eingesetzten Modelle effizient anpassen können. Defect.AI hat das Ziel, die im Projekt entwickelten Ansätze in drei Use Cases zu validieren: Qualitätskontrolle in der Fertigung/Industrie, Fehlererkennung in der Filmrestaurierung und Qualitätskontrolle bei der Digitalisierung von visuellen Medien.

Abstract

Detection of impairments and defects using computer vision methods has been increasingly adopted in recent years in different industry sectors, for example in media and manufacturing. This includes for example the detection of dust and scratches in film material, block and playback distortions due to data loss in digitised or digitally transmitted video, surface or structural impairments of manufactured products in RGB and hyperspectral images, and many more. Most deployed systems use highly sophisticated algorithms incorporating the specialised know-how of domain experts.

While deep-learning based approaches have brought significant progress in many computer vision applications, the specific nature of many defect detection tasks and the lack of training data (both due to the costs of creating datasets, but also due to the rare occurrence of many relevant types of defects) has hindered the adoption of these methods so far. Making deep-learning based methods applicable to these defect detection tasks would enable adjusting them to specific types of defects occurring at specific users, as well as reduce development and maintenance costs in the long run.

Defect.AI aims to research approaches for data-efficient training of visual defect detection methods. The approach will make use of recent advances in machine learning such as foundation models, which can be pre-trained in a self-supervised manner on large unannotated data sets and are then adjusted to a specific task using transfer learning. Another important aspect is to capture the expert knowledge embedded in existing hand-crafted detectors. Knowledge distillation will be used in order to learn this specific knowledge, which can be done at scale without requiring manual annotation. In order to bridge more significant domain gaps (domain adaptation), generative approaches such as GANs and diffusion models will be studied to perform e.g. style transfer between domains such as RGB and infrared.

The availability of deep-learning based defect detection approaches that can be adjusted to new variants of defects with small data sets opens a mid-term perspective of enabling users to adjust the models deployed on their systems efficiently. Defect.AI aims to validate the approaches developed in the project in three use cases: quality control for manufacturing/industrial inspection, defect detection for film restoration and quality control for visual media digitisation.

Endberichtkurzfassung

The researched foundation model methods combine vision transformer foundation models, a multi input-image approach, knowledge distillation from a legacy detector, mass synthetisation of high-quality training data and fine-tuning based on real-world samples to create a versatile foundation for specialized defect detection applications. The research results show that transformer-based segmentation models combined with high-quality synthetic data pretraining and targeted real data fine-tuning can effectively detect and localize various types of defects with very few training samples. The multi input-image architecture based on a Swin Transformer backbone leverages diverse sensor modalities while remaining adaptable to real-world constraints. Backbone optimization using Swin-B delivered significant gains in inference speed and memory efficiency without compromising accuracy.

For knowledge distillation and domain adaptation we extracted knowledge embedded in a specialised, hand-crafted legacy detector to generate pseudo ground truth data and then used this data for the training of a distilled model. Qualitative analysis by experts revealed that the model successfully replicated the legacy detector's strengths but also inherited and amplified specific vulnerabilities, as there are significant false detections in areas of complex object motion or difficult scene content. To tackle these limitations without extensive manual labelling, we developed a procedural defect synthesis framework. This framework extracts single-frame defects from a small number of real-world sample images and injects them with varying defect properities (luminance, transparency...) into a large set of high-resolution video and film sequences with difficult content (motion, reflections...) and varying realistic noise. To replicate the complexity of real film, we modelled the

varying transparency levels and edge-blending characteristics of dust, hair and other single frame defects. This approach allows the model to encounter the defects against a vast diversity of background content, simulating the unpredictable nature of historical film archives while providing the ground-truth precision necessary for robust foundation model training. The adoption of highly realistic synthetic data as a foundation for our training process proved to be very effective in mitigating the issues associated with the legacy detector's ground truth data.

The quantitative and qualitative validation of the Defect.AI detection approaches and models reveal that highest-quality detection of single frame defects (e.g. dust, dirt, random scratches, hair...) and robustness against false detections (caused by moving objects, reflections...) is possible, even with a low number of real defect samples. By this the Defect.AI approaches and methods are an excellent base for future project result exploitation.

Further information can be found at the public project website at <https://www.joanneum.at/digital/en/projects/defect-ai/>.

Projektkoordinator

- JOANNEUM RESEARCH Forschungsgesellschaft mbH

Projektpartner

- HS-ART Digital Service GmbH