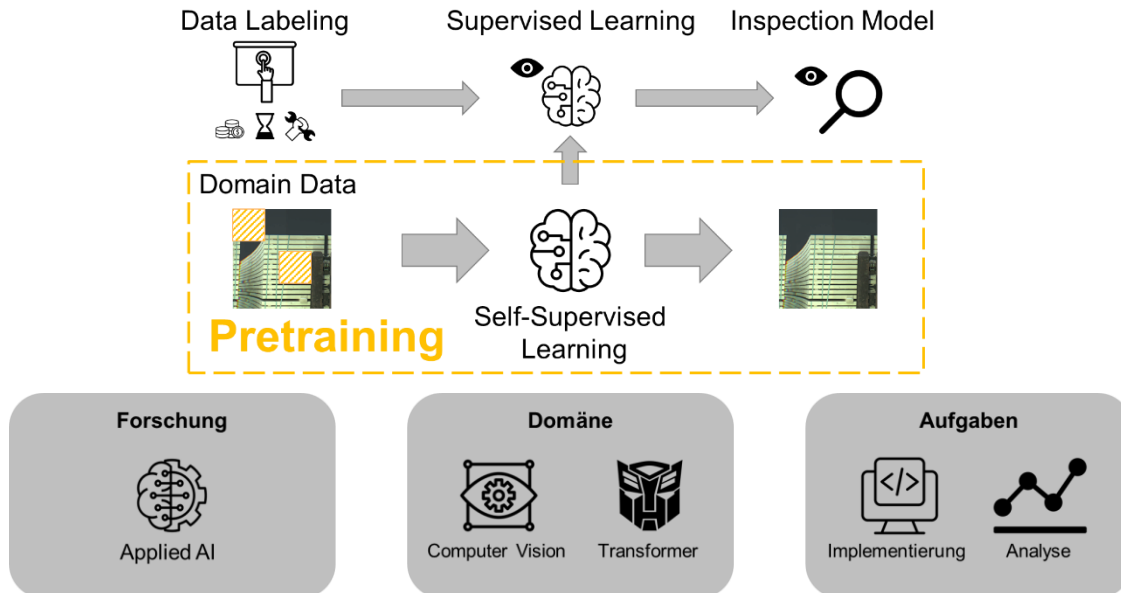


----- English version below -----

Masterarbeit Domain specific pretraining für die automatisierte visuelle Inspektion



Ausgangslage

In industriellen Anwendungsfällen der automatisierten visuellen Inspektion (AVI) erfolgt das Training von Computer Vision Modellen fast ausschließlich mittels Supervised Learning. Um die begrenzte Datenlage in industriellen Anwendungen kompensieren zu können, wird in den meisten Anwendungen zusätzlich auf Transfer Learning gesetzt. Hierbei wird ein auf einem großen Benchmark-Datensatz vortrainiertes Modell mit einem kleineren, Datensatz, der spezifisch für den vorliegenden Anwendungsfall ist, gefinetuned.

Problemstellung

Jüngst wurden in den Bereichen Natural Language Processing und Computer Vision neue Bestmarks in etablierten Forschungsbenchmarks erreicht, die auf einer Kombination von semi- oder self-Supervised Learning Verfahren und der Nutzung von Transformer Modellen basiert. Dies hat zum einen den Vorteil, dass keine Kosten für das Labeling von Daten entstehen, und zum anderen, dass die trainierten Modelle eine höhere Generalisierungsfähigkeit aufweisen. Die Herausforderung besteht nun darin diese Vorteile in Anwendungsfälle der AVI, mit ihrer begrenzten Verfügbarkeit an Trainingsdaten, zu übertragen.

Vorgehensweise und Erwartete Ergebnisse

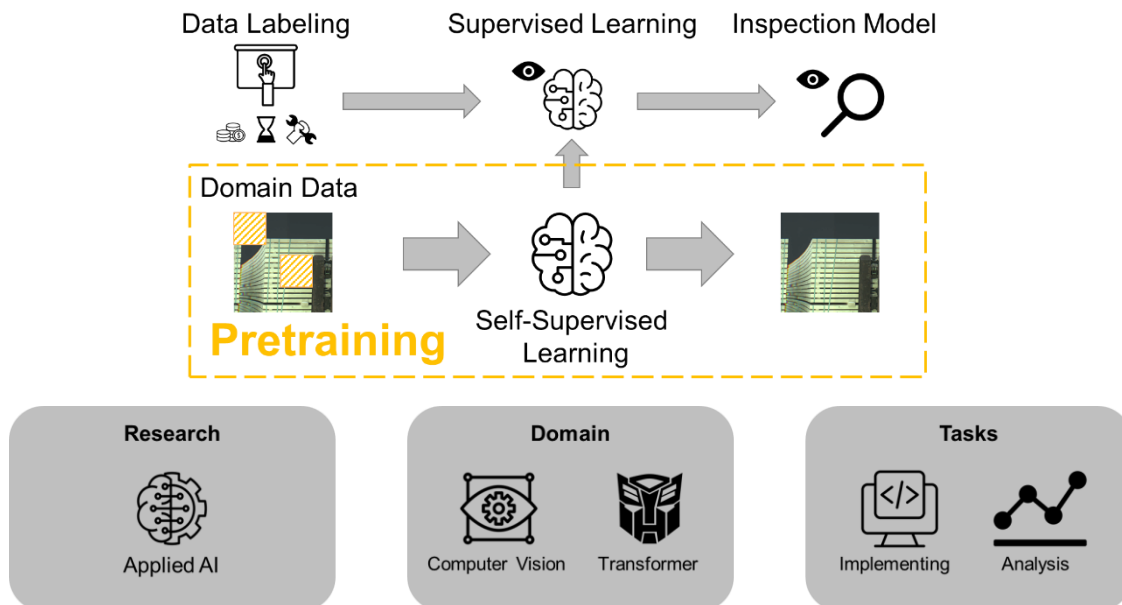
In dieser Arbeit soll untersucht werden, welchen Einfluss die Verwendung von domänenspezifischen Daten anstelle der bisher üblichen Benchmarkdaten auf das Pretraining von Vision Transformer Modellen mit self-supervised Methoden hat. Die Qualität der vortrainierten Modelle wird in einem zweiten Schritt durch supervised finetuning auf AVI Aufgaben bewertet. Mögliche Untersuchungsparameter könnten in diesem Kontext die Größe der verwendeten Pretraining und Finetuning Datensätze, die Zusammenstellung der Pretraining Datensätze, die verwendete Pretraining Methodik, die Wahl des Vision Transformer Modells, die Wahl der Finetuning Aufgabe und noch viele weitere sein.

Ansprechpartner

Nils Hütten | E-Mail: nhuetten@uni-wuppertal.de

Master thesis

Domain specific pretraining for automated visual inspection



Initial Situation

In industrial applications of automated visual inspection (AVI), computer vision models are trained almost exclusively using supervised learning. In order to compensate for the limited data available in industrial applications, transfer learning is also used in most applications. Here, a model pre-trained on a large benchmark data set is fine-tuned with a smaller data set that is specific to the application at hand.

Problem Definition

Recently, new records have been achieved in established research benchmarks in the fields of natural language processing (NLP) and computer vision (CV), which are based on a combination of semi- or self-supervised learning methods and the use of transformer models. On the one hand, this has the advantage that costs for labelling the data can be saved and, on the other hand, that the trained models have a higher generalisation capability. The challenge now is to transfer these advantages to AVI applications with their limited availability of training data.

Procedure and Expected Results

This thesis investigates the influence of using domain-specific data instead of the usual benchmark data on the pre-training of vision transformer models with self-supervised methods. In a second step, the quality of the pre-trained models will be evaluated by supervised finetuning on AVI tasks. Possible investigation parameters in this context could be the size of the pretraining and finetuning data sets used, the composition of the pretraining data sets, the pretraining methodology used, the choice of vision transformer model, the choice of finetuning task and many more.

Contact Person

Nils Hütten| **E-Mail:** nhuetten@uni-wuppertal.de