
Разработка алгоритма аугментации обучающих данных для систем визуального контроля поверхностных дефектов композиционных материалов с использованием технологии преобразования текста в изображение

П.А. Королев, П.А. Ухов
МАИ (национальный исследовательский университет), Москва

Повышение метрик определения дефектов моделями машинного обучения за счет обогащения обучающего набора данных с учетом зон возможного размещения дефектов.

Development of an algorithm for augmenting training data for systems for visual inspection of surface defects in composite materials using text-to-image conversion technology

P.A. Korolev, P.A. Ukhov
MAI, Moscow, Russia

Improving the metrics for detecting defects by machine learning models by enriching the training data set, taking into account the zones of possible location of defects.

1. Введение

На сегодняшний день системы с искусственным интеллектом на базе компьютерного зрения широко применяют в промышленности [1]. Данные системы используют как для контроля внешних дефектов композиционных материалов, так и для обнаружения дефектов структуры композиционного материала. Например, для полимерных композиционных материалов (ПКМ) возможно выявление дефектов структуры на базе импедансного метода и алгоритмов машинного обучения для анализа сигналов датчика. Кроме того, для поточного производства композиционных материалов возможна реализация автоматизированных методов визуального контроля ПКМ на наличие поверхностных трещин, раковин и вмятин, растрескивание связующего, нахлест слоев или витков армирующих материалов, наличие складок армирующего материала, крупных царапин и других видов повреждений.

Машинное обучение позволяет получить эффективные и высокоточные алгоритмы для контроля только при наличии большого объема исходных данных, которые собирают в виде датасетов соответствующей структуры. При этом важным является наличие большого количества примеров дефектов того или иного класса в датасете и равномерность распределения различных дефектов по классам. Нередко этого сложно достигнуть ввиду малого количества исходных данных, а при использовании методов компьютерного зрения для визуального контроля влияют такие факторы, как

освещение и параметры камеры, использованной для получения изображений дефектов.

Для получения эффективных алгоритмов машинного обучения необходимо выполнить сбор данных в виде датасета, его балансировку по видам дефектов, а нередко предварительную обработку изображений, разметку в зависимости от решаемой задачи (сегментация или детекция дефектов) и аугментацию. Только после выполнения данных шагов становится возможным выбирать архитектуры моделей машинного обучения и анализировать их метрики на собранном датасете. На реальном производстве ПКМ обнаружение дефектов является довольно редким событием, что не позволяет сформировать сбалансированный датасет для машинного обучения в кратчайшие сроки. Кроме того, частота появления дефектов различна, что может существенно влиять на процесс обучения. Например, для ПКМ нарушение сплошности в виде трещин и крупных воздушных или газовых раковин являются более частыми, чем нахлест слоев или растрескивание связующего, что затрудняет сбор сбалансированного датасета.

Для увеличения объема данных в датасете прибегают к обогащению данных на основе имеющихся [2], такой прием называется аугментацией данных. Стандартная аугментация данных для решения задач дефектоскопии на основе машинного зрения реализуется путем изменения изображений через различные операции, такие как поворот, масштабирование, обрезка области интереса изображения, добавление шумов, повороты

изображений и т.д. Такой метод аугментации хорошо работает при наличии достаточного количества дефектов на изображениях, а также при одинаковом распределении классов дефектов в обучающем наборе данных. Однако для разработки алгоритмов машинного обучения в условиях, где дефекты встречаются довольно редко, такие технологии неэффективны и не позволяют получить высокие метрики работы алгоритмов и обобщающую способность модели, достаточные для использования решения в системах технической диагностики ответственных изделий, например элементов планера летательного аппарата или гондолы авиационного двигателя.

Чтобы решить проблему распределения классов дефектов на изображениях ПКМ, в данной работе предлагается использовать более сложный метод аугментации обучающих данных, основанный на генерации изображений по текстовому описанию [3]. Указанная технология основана на использовании специального класса алгоритмов машинного обучения – диффузионных моделей. Указанные модели были предложены для генерации изображений. К ним можно отнести DALLE 2, Midjourney, Imagen.

В августе 2022 года компанией Stability AI была выпущена нейросеть Stable Diffusion, которая предназначена для генерации изображений по текстовому запросу. Stable Diffusion – это алгоритм, состоящий из множества компонентов и моделей. Это не единая монолитная модель. В основе решения – компонент понимания текста, компонент создания информации об изображении и декодер изображения.

Для кодирования текста используется модель ClipText, получающая на вход текст и преобразующая его в 77 векторов эмбедингов токенов, каждый в 768 измерениях. Далее эта информация подается на вход сети UNet+Scheduler для постепенной обработки (диффузии) информации в пространстве информации (скрытом пространстве) – массив 4, 64, 64, который передается декодеру автокодировщика, рисующего готовое изображение при помощи массива обработанной информации. Результатом работы алгоритма является готовое изображение размера (3, 512, 512), что соответствует каналам цвета (красный/зеленый/синий), ширине и высоте.

Указанный алгоритм позволяет создавать новые изображения на основе текстового описания, которое задается человеком. Кроме этого, модель можно дообучить на собственных данных, что является ее существенным достоинством. Таким образом, на основе текстового описания можно сгенерировать новое изображение с дефектом, которое может быть добавлено в обучающий набор данных, после чего

применить стандартные методы аугментации, что в совокупности повысит валидационные метрики модели на распознавание объектов на изображениях.

Технология подготовки данных выглядит следующим образом. На изображениях ПКМ без дефектов эксперт предварительно указывает места, в которых, по его мнению, могут возникать те или иные поверхностные дефекты. Шаблоны с местами возможного расположения дефектов размечаются, к ним добавляется текстовое описание дефектов ПКМ. После этого на подготовленных данных выполняется так называемый тюнинг модели Stable Diffusion и получается алгоритм для генерации указанных дефектов.

Далее подготовленный на экспертных данных алгоритм позволяет сгенерировать произвольное количество дефектов и получить сбалансированный по классам датасет достаточного объема для обучения модели обнаружения дефектов на изображениях.

2. Заключение

Предложенный метод аугментации был протестирован на датасете дефектов поверхностей изделий из дерева [4] и датасете дефектов наружной поверхности металлических труб. Благодаря аугментации с генерацией изображений по текстовому описанию удалось добиться равномерного распределения классов дефектов в обучающем наборе данных и повысить валидационные метрики моделей обнаружения объектов на изображениях, в первую очередь mAP.

В перспективе указанные метод может быть использован для обнаружения методами компьютерного зрения поверхностных дефектов ПКМ на поточном производстве, что позволит существенно сократить трудоемкость контрольных операций и производить дефектовку изделий из ПКМ в автоматизированном режиме.

Список использованных источников

- [1] Ramentol E., Olsson T., Barua S. Machine Learning Models for Industrial Applications // AI and Learning Systems. Industrial Applications and Future Directions. 2021. Chap. 7. 19 p. URL: https://www.researchgate.net/publication/349408537_Machine_Learning_Models_for_Industrial_Applications.
- [2] Mikołajczyk A., Grochowski M. Data augmentation for improving deep learning in image classification problem // 2018 International Interdisciplinary PhD Workshop (IIPhDW). IEEE, 2018. P. 117–123. URL:

-
- https://www.researchgate.net/publication/325920702_Data_augmentation_for_improving_deep_learning_in_image_classification_problem.
- [3] High-resolution image synthesis with latent diffusion models / R. Rombach, A. Blattmann, D. Lorenz, P. Esser, B. Ommer // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, Louisiana 19–24 June 2022 : proceedings. 2022. P. 10684–10695. DOI 10.48550/arXiv.2112.10752.
- [4] Kodytek P., Bodzas A., Bilik P. A large-scale image dataset of wood surface defects for automated vision-based quality control processes [version 2; peer review: 2 approved] // F1000Research. 2022. 10:581. 16 p. DOI 10.12688/f1000research.52903.2.