

УДК 004.85

Автоматизация анализа поверхностей с использованием методов искусственного интеллекта

Касютин Илья Алексеевич

Спасенов Алексей Юрьевич (научный руководитель)

МГТУ им. Н.Э. Баумана, Москва, Россия

Рассмотрена задача автоматизации анализа поверхностей металлов с применением методов машинного обучения. Цель — классификация типов поверхностей и оценка шероховатости. Работа содержит в себе математическую постановку задачи классификации, изучение существующих подходов к оценке шероховатости металлов, описание процесса сбора обучающего набора данных, обучение и сравнение моделей компьютерного зрения в задаче классификации.

Ключевые слова: классификация, критерии оптимальности, нейронные сети, компьютерное зрение, технологическая обработка материалов

Введение. Проблема анализа профилограмм возникает в тех областях, где требуется высокоточный контроль качества выходных материалов. Она обусловлена сильной вариативностью входных данных в зависимости от конкретной предметной области. Наряду с классическими методами компьютерного зрения, в последнее время все чаще начинают применяться методы глубокого обучения для анализа одномерных и двумерных данных благодаря гибкости и универсальности архитектур. В частности, передовые архитектуры нейронных сетей достигли человеческого уровня качества во многих задачах детекции и классификации изображений, что говорит об их потенциале в схожих задачах более узких предметных областей.

Обзор существующих подходов. Многие статьи из сферы процессов обработки металлических поверхностей [1, 2], которые решают задачи контроля на производстве, например акцентируют внимание на детекции внештатных ситуаций [3] или определении ориентации заготовки [4], не касаются непосредственно анализа структуры материала. Существующие исследования по анализу материалов, решающие задачу контроля качества конечного продукта, направленные на решение задачи регрессии (например, определение диаметра, биения, радиуса скругления углов фрезы в работе [5]) имеют достаточно сложные экспериментальные установки и не используют модели компьютерного зрения для анализа профилей материалов.

Непосредственно задачи классификации до сих пор пытались решать методами классического компьютерного зрения (например, предобработка изображений и выделение признаков с помощью фильтра Габора) в комбинации с достаточно простыми моделями машинного обучения, не предназначенных для работы с двумерными изображениями [6].

Если решать задачу анализа профилограмм в контексте классической задачи поиска аномалий, как сделали авторы статьи [7], то понадобится большой объем материалов с различными видами дефектов, для которых требуется ручная разметка областей, на которых эти дефекты выявлены. Например, в упомянутой статье одним из наборов данных для обучения является GC10-DET. Этот набор данных был собран из реальных образцов, полученных на производстве. Он содержит порядка 2300 изображений с 10 различными классами дефектов: пробитие материала, линия сварного шва, серповидный разрыв, пятно воды, масляное пятно, шлейф из царапин, контрастные примеси, сколы, складки и царапины.

В работе предлагается решение задачи классификации с помощью различных моделей компьютерного зрения сверточной и трансформерной архитектур.

Математическая постановка задачи классификации. Задача классификации формулируется следующим образом. Пусть $X = \{x_1, x_2, \dots, x_n\}$ — множество объектов, где n — количество объектов в выборке. Также дано $Y = \{1, \dots, K\}$ — множество классов, где K — количество классов. Цель заключается в определении для каждого x_i класса y_i , принадлежащего одному из K заранее заданных классов.

Математически задача сводится к построению алгоритма (классификатора), который отображает каждый объект x_i в один из классов Y .

Для построения алгоритма используется параметрическое семейство алгоритмов $A = \{a(x, \theta) | \theta \in \Theta\}$, где $a(x, \theta): X \rightarrow Y$ — алгоритм с параметрами θ , которые определяют его поведение; Θ — пространство параметров.

Метод обучения классификатора μ строит алгоритм $a = \mu(X^l)$, используя обучающую выборку $X^l = \{(x_1, y_1), (x_2, y_2), \dots, (x_\ell, y_\ell)\}$.

Процесс обучения задается отображением

$$\mu: X^l \rightarrow a,$$

где μ строит алгоритм a , который затем используется для решения задачи классификации.

Математическая постановка задачи оптимизации. Обучение моделей машинного обучения производится с целью максимизации качества работы на предоставленном наборе данных, или минимизации ошибки на нем. Ошибка в задачах машинного обучения представлена эмпирическим риском Q , который является функционалом качества алгоритма a на обучающей выборке Y^l :

$$Q(a, Y^l) = \frac{1}{l} \sum_{i=1}^l L(a(x_i), c_i),$$

где $L(a(x), c)$ — функция потерь, которая отражает величину ошибки алгоритма $a \in A$ на объекте $x \in X$.

Для задачи классификации обычно выбирается вид функции потерь, называемый кросс-энтропией

$$CE = -\sum_{i=1}^K c_i \log \hat{c}_i,$$

где c_i — метка класса (1, если объект x принадлежит классу c_i , и 0 в противном случае); $\hat{c}_i$ — предсказание алгоритма классификации.

Далее алгоритм машинного обучения подбирается с целью минимизации Q :

$$a = \mu(Y^l, \theta) = \operatorname{argmin}_{\theta \in \Theta} \sum_{i=1}^l CE_i.$$

Описание экспериментальных данных. Во всех экспериментах использовался цифровой микроскоп G5 10MP 1-1000X HD с максимальным увеличением в 1000 раз. Для сбора данных выполнялась съемка видеоряда с последующим извлечением самых четких фотографий. Обучение и тестирование работы моделей компьютерного зрения выполняется на единых образцах, но в разных его областях. Для избежания утечки данных разбиение на обучающую и тестовую выборки выполняется с учетом временной компоненты, как указано на рисунке 1. Параметры t_i и t_j равны соответственно 0,7 и 0,8. Такие значения позволяют взять достаточно данных, с одной стороны, для обучения, а с другой, — для получения достоверных результатов на отложенной выборке и исключения пересечения этих множеств.

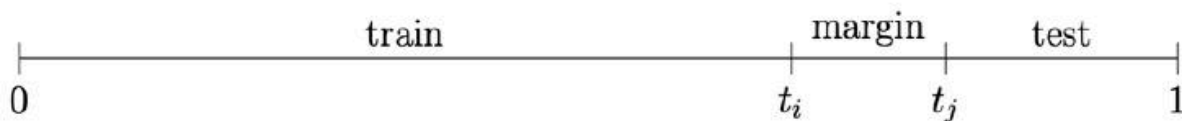


Рис. 1. Метод разбиения набора данных на обучающую и тестовую выборки

Таким образом было собрано два набора данных для анализа в разных задачах:

1) 27 классов поверхностей чугуна, представленных различными техническими процессами с классами чистоты поверхности, определенные в соответствии с ГОСТ 2789–59 [2] (5157 изображений) (табл. 1). Примеры собранных изображений представлены на рис. 1;

2) три класса поверхностей алюминия (2609 изображений): торцевое фрезерование; ленточное пиление; плоское шлифование. Примеры собранных изображений представлены на рис. 2.

Обучение моделей компьютерного зрения на полученных наборах данных. В ходе эксперимента использовалось несколько моделей компьютерного зрения, а именно:

- ResNet18;
- MobileNetV3Small;
- ResNeXt-50 (32x4d);
- EfficientNetV2-S;
- ViT-B/16;
- Swin-B.

Данные модели были выбраны из-за их широкого распространения в реальных задачах, связанных с компьютерным зрением. В частности, первые четыре модели относятся к классу сверточных нейронных сетей, которые отличаются высокой скоростью работы, а последние две — к классу трансформеров, которые обеспечивают лучше качество в большинстве задач компьютерного зрения, при этом более требовательны к программному обеспечению.

Таблица 1

Типы поверхностей чугуна, присутствующие в наборе данных

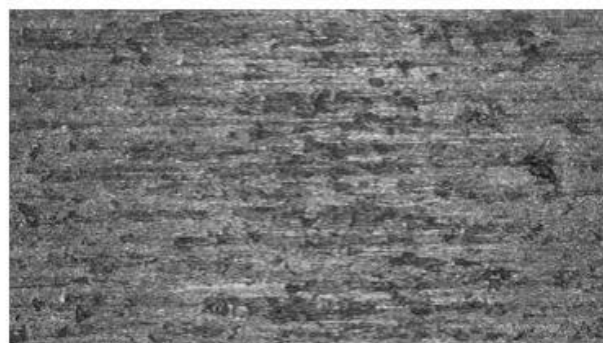
Тип технического процесса	Классы чистоты поверхности
Расточка	5, 6, 7
Строгание	4, 5, 6, 7
Точение	4, 5, 6, 7
Фрезерование торцевое	5, 6, 7
Фрезерование цилиндрическое	5, 6, 7
Шлифование внутреннее	7, 8, 9
Шлифование круглое	7, 8, 9
Шлифование плоское	6, 7, 8, 9



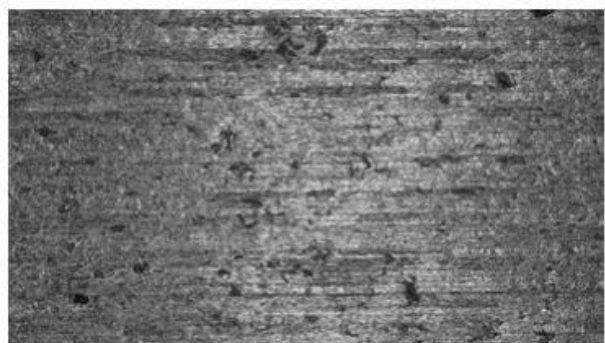
а



б



в



г

Рис. 1. Примеры изображений из первого набора обучающих данных:
а — строгание 4 кл.; б — строгание 5 кл.; в — строгание 6 кл.; г — строгание 7 кл.

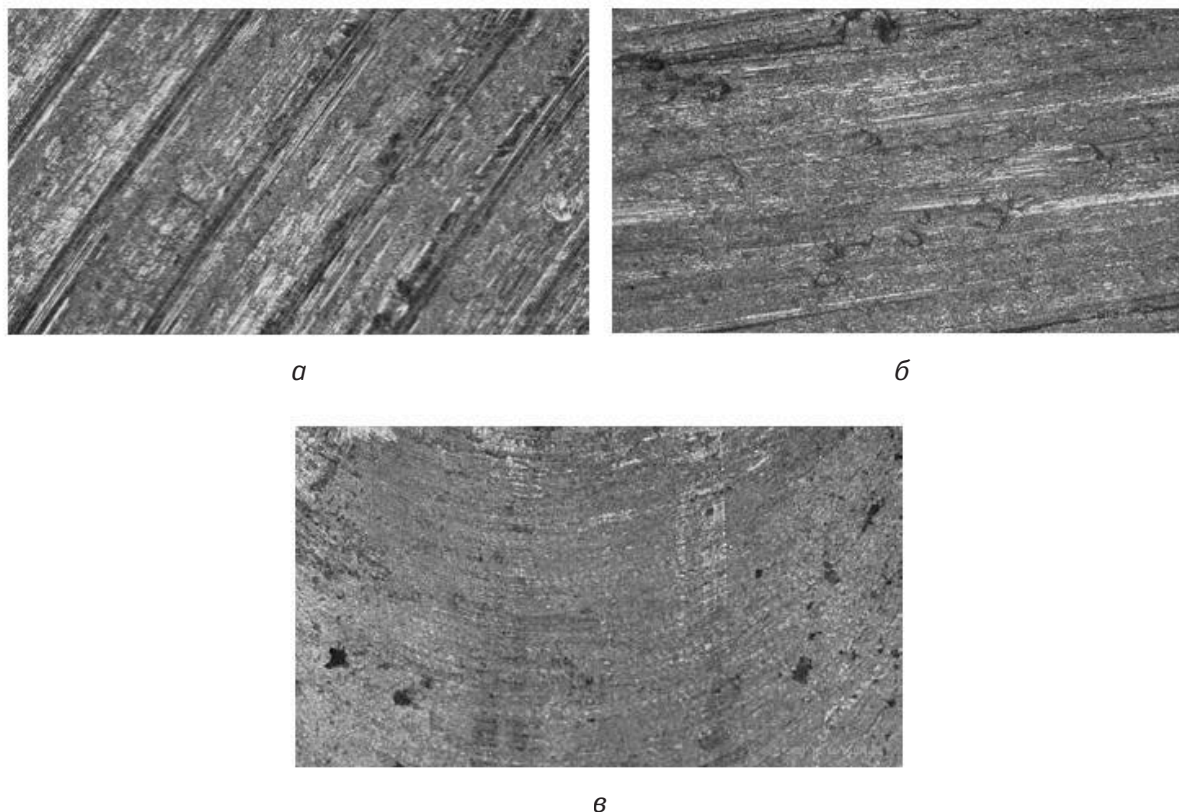


Рис. 2. Примеры изображений из второго набора обучающих данных:
а — торцевое фрезерование; *б* — ленточное пиление; *в* — плоское шлифование

В силу того, что данные были сбалансированы по классу, качество работы моделей компьютерного зрения оценивается с помощью метрики точности — доли правильно классифицированных изображений.

Величины метрик лучших версий моделей, полученные на отложенной выборке, представлены в табл. 2.

Таблица 2

Точность предсказаний на тестовой выборке

Модель	Набор данных № 1	Набор данных № 2
ResNet18	0,85	0,88
MobileNetV3Small	0,88	0,99
ResNeXt-50 (32x4d)	0,90	0,98
EfficientNetV2-S	0,92	0,96
ViT-B/16	0,93	0,99
Swin-B	0,95	0,99

Из результатов экспериментов видно, что точность предсказаний на тестовой выборке данных тем выше, чем более сложные модели классификации используются. Также стабильный рост качества говорит о достаточном количестве собранных данных для обучения.

Закключение. Использование изображений обработанных поверхностей для их классификации по степени шероховатости ускоряет процесс классификации и позволяет экономить ресурсы. Несмотря на незначительное снижение точности, использование изображений может быть оправдано в условиях ограниченных ресурсов и необходимости быстрого получения результатов, особенно при необходимости массового контроля качества.

Литература

- [1] ГОСТ 2789–73. *Шероховатость поверхности. Параметры и характеристики*. Москва, Изд-во стандартов, 1973, 7 с.
- [2] ГОСТ Р ИСО 3855–2013. *Фрезы. Термины и определения*. Москва, Стандартинформ, 2014, 19 с.
- [3] Karabagli B., Simon T., Orteu J.J. A new chain-processing-based computer vision system for automatic checking of machining set-up application for machine tools safety. *International Journal of Advanced Manufacturing Technology*, 2016, vol. 82, pp. 1547–1568. <https://doi.org/10.1007/s00170-015-7438-y>
- [4] Кротова Н.А., Пушков Р.Л. Определение геометрических параметров и ориентации заготовки на станке с применением алгоритмов технического зрения. *Вестник МГТУ Станкин*, 2021, № 2 (57), с. 8–12.
- [5] Siswanto T.W.A., Hong C.B., Ibrahim R., Rahim Z., Khoo N. Design and Performance Analysis of Developed Computer Vision Measurement Application for End Mill Cutting. *International Journal of Engineering Research and Technology*, 2020, vol. 13, no. 12, pp. 5373–5378.
- [6] Demircioglu P., Seçkin A.Ç., Torgersen J., Bogrekeci I., Durakbasa N. Metal Surface Texture Classification with Gabor Filter Banks and XAI. *DAAAM International Scientific Book*, 2023. <https://doi.org/10.2507/daaam.scibook.2023.03>
- [7] Alaa T., Kotb M., Zakaria A., Diab M., Gomaa W. Automated Detection of Defects on Metal Surfaces using Vision Transformers. *Arxiv preprint arXiv:2410.04440*, 2024.

Automation of surface analysis using artificial intelligence methods

Kasiutin Ilia Alexeevich

Spasenov Aleksey Yurievich (*scientific advisor*)

BMSTU, Moscow, Russia

This study addresses the task of automating the analysis of metal surfaces using machine learning methods. The objective is to classify surface types and assess roughness. The work includes a mathematical formulation of the classification problem, a review of existing approaches to metal roughness estimation, a description of the data collection process for the training dataset, as well as the training and comparison of computer vision models for the classification task.

Keywords: classification, optimality criteria, neural networks, computer vision, technological processing of materials