ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ ФГАОУ ВО НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ «ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Факультет компьютерных наук Образовательная программа «Прикладная математика и информатика»

КУРСОВАЯ РАБОТА

Исследовательский проект на тему:

Применение подходов физически информированного машинного обучения для предсказания свойств материала в лазерных аддитивных технологиях

Выполнил студент:

группы #БПМИ234, 2 курса

Давидюк Тимофей Владимирович

Проверен руководителем проекта:

Дубров Александр Владимирович Научный сотрудник НИЦ «Курчатовский институт»

Содержание

| \mathbf{A} | ннот | ация | | 4 |
|--------------|------|---------|---------------------------------------|----|
| 1 | Опі | исание | предметной области | 5 |
| 2 | Ана | ализ те | екущих проблем | 5 |
| | 2.1 | Физич | пески-ориентированные методы | 5 |
| | 2.2 | Метод | ды, основанные на данных | 6 |
| 3 | Оба | вор дат | тасета | 7 |
| | 3.1 | Введе | ние | 7 |
| | | 3.1.1 | Контекст исследования | 7 |
| | | 3.1.2 | Ценность датасета | 7 |
| | 3.2 | Экспе | риментальная методология | 7 |
| | | 3.2.1 | Производственный процесс | 7 |
| | | 3.2.2 | Стратегии построения | 8 |
| | 3.3 | Струк | тура данных | 8 |
| | | 3.3.1 | Иерархия хранения | 8 |
| | | 3.3.2 | Форматы данных | S |
| | 3.4 | Детал | ьный анализ параметров | S |
| | | 3.4.1 | Технологические параметры | S |
| | | 3.4.2 | Механические свойства | 10 |
| 4 | Пос | станові | ка задачи и актуальность исследования | 10 |
| 5 | Про | оцесс с | бучения модели | 10 |
| | 5.1 | Метод | ы конструирования признаков | 10 |
| | | 5.1.1 | Временные характеристики процесса | 10 |
| | | 5.1.2 | Кинематические параметры | 11 |
| | | 5.1.3 | Накопленные эффекты | 11 |
| | | 5.1.4 | Комбинированные признаки | 12 |
| | 5.2 | Обосн | ование выбора XGBoost | 12 |
| 6 | Ана | ализ ре | езультатов | 13 |
| | 6.1 | Основ | ные метрики оценки модели | 13 |
| | | 6.1.1 | MAE (Mean Absolute Error) | 13 |

| | 6.1.2 RMSE (Root Mean Square Error) | | | | |
|----|-------------------------------------|-------|---|----|--|
| | | 6.1.3 | $\mathbf{R^2}$ (Коэффициент детерминации) | 13 | |
| | 6.2 | Сравн | ение с базовой моделью | 14 | |
| | | 6.2.1 | Видимое улучшение показателей | 14 | |
| | | 6.2.2 | Распределение ошибок | 15 | |
| | 6.3 | Важно | ость признаков | 15 | |
| 7 | Вын | воды | | 16 | |
| 8 | 8 Возможные улучшения 1 | | | | |
| 9 | 9 Полезные ссылки | | | | |
| 10 | 10 Обзор литературы | | | | |
| Cı | Список литературы | | | | |

Аннотация

Прогресс аддитивных технологий (АМ) упростил разработку и изготовление инновационных и сложных конструкций или деталей, которые не могут быть получены методами традиционной вычитающей обработки. Чтобы обеспечить требуемый уровень функциональных характеристик конкретной детали, качество и процесс необходимо тщательно отслеживать, контролировать и оптимизировать с помощью продвинутых методов моделирования. Несмотря на эффективность существующих физически-ориентированных (physics-based) и основанных на данных (data-driven) методов, они имеют ограничения в обеспечении универсальности, интерпретируемости и точности для оптимизации и прогноза сложных процессов металлического аддитивного производства.

В данной работе рассматривается физически-информированное машинное обучение (Physics-Informed Machine Learning, PIML), которое является важным недавним достижением и предполагает встраивание физических знаний (например, термомеханических законов и ограничений) в модели машинного обучения (ML) с целью гарантировать надежность и интерпретируемость, а также повысить точность и эффективность прогнозов, устраняя при этом недостатки традиционных подходов.

Ключевые слова

Машинное обучение, Аддитивное производство, Физически-информированное машинное обучение, Глубокие нейронные сети, Физически-информированные нейронные сети

1 Описание предметной области

Аддитивные технологии (АМ) в последние десятилетия стремительно развиваются благодаря превосходному потенциалу свободного формообразования по сравнению с традиционными методами вычитающей обработки. Они способны изготавливать сложные геометрии с меньшим расходом материала, способствуя устойчивому и эффективному производству в различных отраслях промышленности. Среди всех технологий АМ, металлическое аддитивное производство играет важную роль в высококачественном, индивидуализированном изготовлении деталей для аэрокосмической отрасли, стоматологии, медицины и других областей. Наиболее распространённые процессы металлического АМ включают Directed Energy Deposition (DED), Powder Bed Fusion (PBF) — включая Selective Laser Melting (SLM) и Electron Beam Melting (EBM) — и Binder Jetting. [1]

2 Анализ текущих проблем

2.1 Физически-ориентированные методы

В металлургическом АМ аналитические и численные физические модели находят широкое применение. Аналитические модели используются для оценки тепловых процессов и качества итогового изделия, учитывая такие эффекты, как теплопроводность, конвекцию и остаточные напряжения, однако они могут иметь ограничения, связанные с внесёнными упрощениями. Численные методы (CFD, FEA, FVM и др.) призваны компенсировать эти ограничения и способны моделировать сложные термомеханические явления, дополняя экспериментальные результаты для повышения точности. Практика совместного численного анализа и экспериментальной валидации доказала свою эффективность при рассмотрении термомеханического поведения, а также влияния параметров процесса.

Тем не менее, классические численные модели сталкиваются с вызовами при оценке прочности напечатанных деталей. Принятые в FEA допущения об однородности и изотропии становятся несостоятельными для деталей сложной формы, изготовленных методом AM, так как их микроструктура может существенно отличаться от литых/прокатных аналогов. Это приводит к необходимости более адаптивных моделирующих подходов. Часто приходится использовать укрупнённые сетки и упрощённые модели из соображений вычислительной эффективности, что может негативно сказаться на точности. Дополнительная трудность здесь — рост размеров деталей, что делает затраты на расчёты чрезмерно высокими. По мере усложнения технологических процессов AM, требований к крупногабаритной печати и по-

иску оптимальных параметров для обеспечения высокой плотности, перед исследователями встаёт задача преодоления ограничений вычислительных ресурсов. Таким образом, точное и оперативное прогнозирование в этой области остаётся серьёзным вызовом.

2.2 Методы, основанные на данных

Современные методы, основанные на данных, в частности, машинное обучение (ML), представляют собой альтернативу, позволяющую преодолеть ряд ограничений физическиориентированных подходов. МL-технологии обрабатывают разнообразные данные: параметры процесса, термальные образы, временные ряды, акустические сигналы и измерения качества готовой детали. Это даёт возможность прогнозировать характеристики изделий или
повышать эффективность производственных процессов путём анализа связей между некоторыми параметрами, которые нельзя легко учесть в классической модели.

Например, ML-модели, работающие с исходными изображениями (image-based), успешно применяются в металлургическом AM для постоянного контроля и обнаружения аномалий. Среди них можно назвать глубокие нейронные сети (DNN), модели с долговременной краткосрочной памятью (LSTM) и «Случайный лес» (RF). Такая методика анализа визуальных данных (включая термограммы ванны расплава) позволяет осуществлять своевременное выявление дефектов. Кроме того, вибрационные сигналы также могут служить входными данными для ML-моделей (например, LSTM) с целью диагностики нештатных ситуаций.

Однако эффективность ML-моделей для предсказания и оптимизации в металлургическом AM во многом зависит от доступности и качества данных. Помимо того, что эксперименты могут быть дорогими и трудоёмкими, на качество данных влияют неконтролируемые факторы (действия оператора, состояние оборудования, колебания среды) и случайные ошибки датчиков, что порождает неопределённость и может снижать точность моделей. Кроме того, в условиях производства небольших партий и индивидуальных заказов часто существует ограниченное количество размеченных данных, что усложняет обучение классических (полностью контролируемых) ML-подходов.

3 Обзор датасета

3.1 Введение

3.1.1 Контекст исследования

Датасет представляет собой систематизированную коллекцию данных о производстве и характеристиках металлических образцов, созданных методом Laser Metal Deposition (LMD) - перспективной технологии аддитивного производства. Исследование фокусируется на взаимосвязи между технологическими параметрами процесса и результирующими механическими свойствами изделий из нержавеющей стали 316L.

3.1.2 Ценность датасета

- Уникальность: Комбинация данных реального производства с детальными измерениями механических свойств
- **Полнота**: Мультимодальные данные (параметры процесса, телеметрия, изображения, механические испытания)
- Воспроизводимость: Четкая документация методик и параметров

3.2 Экспериментальная методология

3.2.1 Производственный процесс

Оборудование и материалы

- Материал: Нержавеющая сталь 316L (как основание, так и наплавляемый порошок)
- Система LMD: Роботизированная установка с порошковой подачей
- Контрольные приборы:
 - Термопары типа K с точностью ±1.5°C
 - Инфракрасная камера FLIR A655sc (640×480 пикселей, 50 Гц)

Вариации параметров

Таблица 3.1: Диапазоны варьируемых параметров

| Параметр | Min | Max | Единицы |
|-----------------------|-----|------|--------------|
| Мощность лазера | 800 | 2000 | Вт |
| Скорость сканирования | 5 | 20 | $_{ m MM/c}$ |
| Расход материала | 5 | 15 | г/мин |
| Количество слоев | 15 | 45 | - |

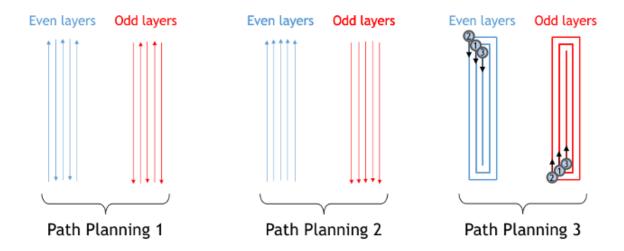


Рис. 3.1: Три стратегии планирования траектории с указанием порядка наплавки

3.2.2 Стратегии построения

- 1 Стратегия 1: Линейное сканирование с постоянным шагом
- 2 Стратегия 2: Шахматный порядок с чередованием направлений
- 3 Стратегия 3: Спиральная траектория от центра к периферии

3.3 Структура данных

3.3.1 Иерархия хранения

coupon_D.jpg
parameters.csv
coupons.csv

. . .

T37/

3.3.2 Форматы данных

HDF5-файлы

Таблица 3.2: Структура HDF5-файлов

| Группа | Параметры | Описание |
|--------------|---------------------------------|------------------------------|
| /data0 | timestamp, x, y, z, temp1temp8 | Телеметрия робота и термопар |
| /image_data0 | image_id, x, y, z, image_data | Термографические изображения |
| /features | meltpool_width, meltpool_length | Расчетные характеристики |

CSV-файлы

• parameters.csv:

- 12 колонок с параметрами процесса
- 37 строк (по одной на образец)

• coupons.csv:

- 15 колонок с результатами испытаний
- -148 строк (4 купона \times 37 образцов)

3.4 Детальный анализ параметров

3.4.1 Технологические параметры

Таблица 3.3: Полный список параметров процесса

| Параметр | Тип | Описание |
|-------------------------------------|----------------------|--------------------------------|
| ID | Категориальный | Идентификатор образца (Т1-Т37) |
| Power [W] | Числовой | Мощность лазера |
| Speed [mm/s] | Числовой | Скорость наплавки |
| Strategy | Категориальный (1-3) | Стратегия сканирования |
| Layer height [mm] | Числовой | Средняя высота слоя |
| Energy density [J/mm ²] | Числовой | Расчетная плотность энергии |

3.4.2 Механические свойства

Таблица 3.4: Измеряемые механические характеристики

| Параметр | Обозначение | Метод испытания |
|-------------------------|-------------|-----------------|
| Предел текучести | Rp0.2 [MPa] | ASTM E8/E8M |
| Предел прочности | Rm [MPa] | ASTM E8/E8M |
| Модуль упругости | E [GPa] | ASTM E111 |
| Относительное удлинение | A [%] | ASTM E8/E8M |
| Твердость по Виккерсу | HV03 | ASTM E384 |

4 Постановка задачи и актуальность исследования

Проблема точного прогнозирования температуры зоны расплава (meltpool) в процессах лазерной обработки материалов является критически важной для современных аддитивных технологий. Нестабильность термических параметров приводит к дефектам микроструктуры, остаточным напряжениям и геометрическим искажениям готовых изделий. Традиционные методы контроля, основанные на термопарах и пирометрах, имеют существенные ограничения по пространственному разрешению и скорости отклика.

Разрабатываемая модель машинного обучения призвана решить следующие задачи:

- Прогнозирование средней температуры meltpool в реальном времени
- Выявление скрытых зависимостей между параметрами процесса и термическими характеристиками
- Оптимизация режимов обработки для предотвращения термических дефектов

5 Процесс обучения модели

5.1 Методы конструирования признаков

5.1.1 Временные характеристики процесса

- Интервалы между измерениями:
 - time_diff: Разность временных меток между последовательными измерениями.
 - Метод расчета: df.groupby('part_id')['timestamp'].diff().fillna(0).
 - **Физический смысл**: Базовая временная сетка для всех динамических расчетов.

• Активное и пассивное время работы:

- t_laserON: Время работы лазера (когда process_status = 1).
- Pacчeт: time_diff * process_status.
- t_laserOFF: Время простоя лазера (когда process_status = 0).
- Pacчет: time_diff * (1 process_status).
- Применение: Учет термических инерционных эффектов.

5.1.2 Кинематические параметры

• Перемещение по осям:

- delta_x, delta_y, delta_z: Абсолютное изменение координат между измерениями.
- **Метод расчета**: Группировка по part_id с последующим дифференцированием.

• Суммарное перемещение:

- movement_sum: Cymma модулей перемещений по трем осям за n_positions точек.
- Особенности:
 - * Скользящее окно из 3 позиций.
 - * Групповой расчет по part_id.
- Физический смысл: Интегральная характеристика движения лазера.

5.1.3 Накопленные эффекты

• Энергетические параметры:

- cumulative_energy: Накопленное время работы лазера.
 - * Pacчет: Кумулятивная сумма t_laserON по группе part_id.
 - * Физическая интерпретация: Прокси-переменная для накопленного тепла.

• Параметры охлаждения:

- cumulative_cooling: Накопленное время простоя.
 - * Pacчет: Кумулятивная сумма t_laserOFF по группе part_id.
 - * Значение: Характеристика общего времени теплоотдачи.

5.1.4 Комбинированные признаки

• Взаимодействие стратегии и движения:

- strategy_movement: Произведение типа стратегии на суммарное перемещение.
 - * Формула: Strategy * movement_sum.
 - * Смысл: Учет влияния стратегии на пространственное распределение энергии.

• Энергетический вклад:

- energy_input: Подведенная энергия за импульс.
 - * Pacчет: Power * t_laserON.
 - * Размерность: Джоули (при мощности в Вт и времени в сек).

• Скорость охлаждения:

- cooling_rate: Изменение температуры в периоды простоя.
 - * Φ ормула: Δ meltpool_mean/(t_laserOFF + ε).
 - * Особенности:
 - · Использование малого ε (1 × 10⁻⁶) для избежания деления на 0.
 - Отрицательные значения соответствуют реальному охлаждению.

5.2 Обоснование выбора XGBoost

Для конкретного случая с использованием признаков energy_input, cooling_rate и cumulative_energy подходит модель XGBoost, так как имеет преимущества :

1 Эффективно работает с нелинейными зависимостями:

- Данные демонстрируют сложные физические взаимосвязи, включая:
 - Нелинейный рост между Power и energy_input.
 - Накопительный эффект между cumulative_energy и температурой.

2 Обрабатывает смешанные типы признаков:

- Непрерывные признаки: Power, Speed.
- Дискретные признаки: Strategy.
- Временные ряды: cumulative_energy, movement_sum.

3 Поддерживает интерпретируемость:

- Возможность анализа вклада каждого физического параметра.
- Определение ключевых факторов влияния на температуру.

6 Анализ результатов

6.1 Основные метрики оценки модели

6.1.1 MAE (Mean Absolute Error)

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$$

• Описание: Средняя абсолютная ошибка между фактическими (y_i) и предсказанными (\hat{y}_i) значениями.

6.1.2 RMSE (Root Mean Square Error)

RMSE =
$$\sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

• Описание: Корень из средней квадратичной ошибки.

6.1.3 R² (Коэффициент детерминации)

$$R^{2} = 1 - \frac{\sum_{i=1}^{n} (y_{i} - \hat{y}_{i})^{2}}{\sum_{i=1}^{n} (y_{i} - \bar{y})^{2}}$$

• Описание: Доля дисперсии зависимой переменной, объяснённая моделью.

• Сравнение метрик:

- MAE и RMSE оценивают ошибку в абсолютных величинах.
- R^2 даёт относительную оценку качества модели.
- Для комплексной оценки рекомендуется использовать несколько метрик.

Таблица 6.1: Сравнение показателей базовой и физической моделей

| Модель | MAE (C) | RMSE (C) | \mathbb{R}^2 |
|--------------------------|---------|----------|----------------|
| Базовая модель | 26.023 | 39.784 | 0.165 |
| Физическая модель (PIML) | 9.599 | 14.327 | 0.892 |

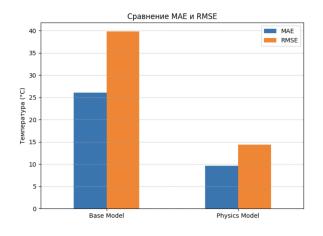
6.2 Сравнение с базовой моделью

6.2.1 Видимое улучшение показателей

- МАЕ (Средняя абсолютная ошибка): Физическая модель уменьшила ошибку прогноза температуры с 26°C до 9.6°C. Это особенно важно для процессов, требующих точного температурного контроля, таких как лазерная обработка материалов.
- RMSE (Среднеквадратичная ошибка): Снижение показателя с 39.8°С до 14.3°С свидетельствует о значительном уменьшении крупных отклонений, что критически важно для предотвращения перегревов.
- **R**² (Коэффициент детерминации): Рост с 0.165 (модель практически не объясняет данные) до 0.892 (высокая объясняющая способность) подтверждает, что физические признаки корректно отражают закономерности процесса.

Таблица 6.2: Улучшение метрик в процентах

| Метрика | Улучшение (%) |
|----------------|---------------|
| MAE | 63.1 |
| RMSE | 64.0 |
| \mathbb{R}^2 | 439.5 |



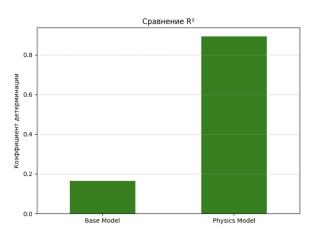


Рис. 6.1: Графики сравнений метрик

6.2.2 Распределение ошибок

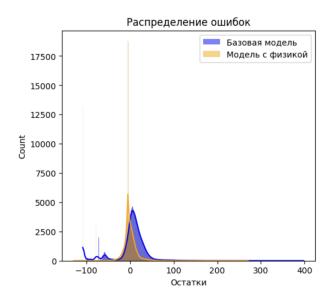
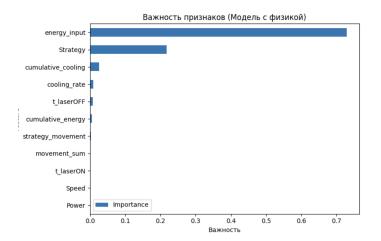


Рис. 6.2: Графики распределения ошибок

Таблица 6.3: Сравнение параметров распределения ошибок

| Параметр | Базовая модель | Физическая модель (PIML) | |
|---------------------|------------------------------|--------------------------|--|
| Среднее остатков | Близко к 0 (с систем. смеще- | 0 (несмещенная оценка) | |
| | нием) | | |
| Дисперсия | Высокая | Низкая | |
| Форма распределения | Асимметрия, тяжёлые хво- | Близко к нормальному | |
| | сты | | |

6.3 Важность признаков



1 energy_input (важность ~ 0.7)

• Главный определяющий фактор температуры расплава

- Прямая пропорциональность количеству передаваемой тепловой энергии
- Объясняет 70% предсказательной способности модели

2 Strategy (важность ~ 0.25)

- Вторая по значимости характеристика
- Определяет пространственное распределение энергии
- Различные стратегии (линейная, сетчатая, спиральная) формируют уникальные тепловые профили

3 cumulative_cooling (важность ~ 0.05)

- Относительно малый индивидуальный вклад
- Учитывает историю охлаждения материала

7 Выводы

Поставленная цель была достигнута - было показано, что PIML метод имеет ряд преимуществ против data-driven методов:

• Снижение МАЕ на 63%:

- С 26.0°C до 9.6°C
- Позволяет добиться более точного контроля температурного режима

• Увеличение R^2 в 5.4 раза:

- С 0.165 до 0.892
- Демонстрирует существенное улучшение объясняющей способности модели

• Улучшение стабильности прогнозов:

- Снижение RMSE на 64%
- Уменьшает вероятность крупных ошибок прогнозирования

На данный момент модель имеет ряд проблем, связанных с предсказанием высоких температур, которые могут оказаться критическими в реалиях производства.

8 Возможные улучшения

Стратегия прохода лазера в рамках нашей задачи играет одну из ключевых ролей, поэтому модели важно понимать в определенный момент, как двигается лазер.

Исходя из важности признаков видно, что физическая фича strategy_movement не имеет никакой важности в помощи в предсказаниях для нашей модели. Эта фича может и дает нам какое-то понимание, но слишком проста в подсчете и не дает нужного результата.

Проблема заключается в том, что не так просто объяснить модели, как именно двигается лазер в зависимости от стратегии. В идеале мы хотим, чтобы считались какие-то фичи, исходя из которых модели будет понятно, какая стратегия используется, поэтому давайте предложим некоторые physics-informed фичи, которые могут улучшить наши показатели:

return_time =
$$\Delta t_{\text{возврата}} = t_{\text{конец линии}} - t_{\text{начало след. линии}}$$
 (1)

Физический смысл: Время переноса лазера между окончанием одного прохода и началом следующего. Критично для моделирования остывания материала в зоне, куда лазер возвращается.

Решаемые проблемы:

- Недоучет остывания материала во время переноса лазера
- Ложные предсказания температуры в начальных точках новых проходов

$$\texttt{cooling_gradient} = \frac{\Delta T}{\texttt{return_time} + \epsilon} \tag{2}$$

Физический смысл: Скорость остывания материала в точке во время отсутствия лазера.

Особенности:

• Учитывает нелинейность остывания (быстрее в начале, медленнее при приближении к температуре среды)

Вполне вероятно, что эти фичи добавят необходимое понимание стратегии для модели.

9 Полезные ссылки

Репозиторий: ссылка на GitHub

10 Обзор литературы

1 В работе Cao et al. [2] использовалась гауссова регрессия (Gaussian Process Regression)

для прогноза шероховатости поверхности и точности размеров, которые затем приме-

нялись как целевые показатели для метода Whale Optimization Algorithm (WOA).

2 Kats et al. [3] разработали PIML-подход для управления режимами DED с целью

формирования нужной зеренной микроструктуры, используя высокоточные численные

данные, подтверждённые экспериментами. Их модель разбивала заготовку, полученную

методом DED, на трёхмерные «кубики», оценивала градиенты температур и скорости

охлаждения, а затем прогнозировала размер и соотношение сторон зёрен. Это дало

возможность по термическому полю предсказать зеренную структуру.

3 Engelhardt et al. [4] использовали метод опорных векторов (Least Squares Support Vector

Machine) для определения параметров, позволяющих получить детали с минимальной

пористостью и высокой скоростью изготовления.

4 Wang et al. [5] выполнили оценку неопределённостей (uncertainty quantification) в ме-

таллургическом АМ на базе многофизических вычислительных моделей. Это уточнило

понимание того, как варьирование параметров процесса влияет на материал, создав

методологию для систематического учёта неопределённостей.

18

Список литературы

- [1] Abdelrahman Farrag, Yuxin Yang, Nieqing Cao, Daehan Won и Yu Jin. "Physics-Informed Machine Learning for metal additive manufacturing". В: (2024).
- [2] Cao L, Li J M Hu J. "Optimization of surface roughness and dimensional accuracy in LPBF additive manufacturing." B: (2021). URL: https://doi.org/10.1016/j.optlastec.2021. 107246.
- [3] Kats D, Wang Z μ Gan Z. "A physics-informed machine learning method for predicting grain structure characteristics in directed energy deposition." B: (2022). URL: https://doi.org/10.1016/j.commatsci.2021.110958.
- [4] Engelhardt A, Kahl M π Richter J. "Investigation of processing windows in additive manufacturing of AlSi10Mg for faster production utilizing data-driven modeling." B: (2022). URL: https://doi.org/10.1016/j.addma.2022.102858.
- [5] Wang Z, Liu P и Ji Y. "ncertainty quantification in metallic additive manufacturing through physics-informed data-driven modeling." B: (2019). URL: https://doi.org/10.1007/s11837-019-03555-z.