

Применение методов машинного обучения для оценки количества биомассы по данным спутниковых снимков Sentinel-1 и Sentinel-2

Автор: Денис Петренко

Научный руководитель: Алексей Озерин

Краткий обзор данных, модели и метрик качества

Данные

Данные представляют из себя набор изображений для примерно 13.000 участков полученные с двух спутников. Sentinel-1 предоставляет радарные данные в четырех каналах, Sentinel-2 - мультиспектральные изображения (10 каналов и данные облачности).

Данные для обучения модели были семплированы: из каждого участка случайно выбирались 1 000 пикселей.

Предлагаемая модель

В качестве baseline решения используется модель на базе encoder-decoder архитектуры. Выбор такой архитектуры обусловлен тем фактом, что задачу можно рассматривать как задачу регрессии на последовательности значений взятых для каждого отдельного пикселя.

Метрики качества

Для обучения модели используется RMSE функция потерь. Лосс RMSE считается для всех пикселей в батче. Метрика качества модели представляет собой RMSE усредненную по изображениям.

$$Average\ RMSE = \frac{\sum_{i=0}^M \sqrt{\left(\frac{1}{N} \sum_{i=0}^N (y_i - \hat{y}_i)^2 \right)}}{M}, \text{ где}$$

\hat{y}_i – предсказание, y_i – таргет, N – количество пикселей в каждом изображении, M – количество изображений

Описание проведенных экспериментов

S1-baseline

Для решения задачи предлагался baseline подход на основе данных спутниковых снимков Sentinel-1. Причины выбора данного подхода заключались в том, что изображения Sentinel-1 выглядят аналогично изображениям Sentinel-2: виден контраст различных областей, аналогичных изображениям Sentinel-2 и при этом нет пропусков в данных.

К сожалению, нейронная сеть не училась и не смогла выдать результат значительно отличающийся от тривиального решения.

Детальный анализ устройства SAR снимков, проведенный в работе Mort Cauty [Detecting Changes in Sentinel-1 Imagery](#), продемонстрировал, что свойства SAR изображений отличаются от обычных с которыми мы привыкли работать.

S2-baseline

При исследовании причин неудовлетворительных результатов baseline модели было обнаружено, что ввиду природы данных снимков спутника Sentinel-1, их нельзя использовать напрямую без предобработки. Для проверки работоспособности текущей модели были использованы данные снимков Sentinel-2. Для отсутствующих снимков была симитирована экстремальная облачность: канал CLP был заполнен максимально допустимым значением 255, остальные каналы заполнены нулями. Данное восстановление пропусков позволило использовать модель S1-baseline без существенных изменений.

Анализ результатов

Переход на снимки Sentinel-2 принес значительное увеличение качества модели: loss уменьшился до 27.2, метрика до 42.8.

Детальный анализ снимков, сгенерированных моделью и сравнение с таргетом выявил системное занижение предсказаний. На рисунке ниже приведен пример целевого изображения и изображения, полученного от модели. Можно заметить, что модель систематически занижает прогнозы для зон с высоким уровнем таргета. Также регионы с нулевым таргетом на ответе модели имеют большую площадь чем

целевом изображении. Например, дороги существенно шире, чем на самом деле.

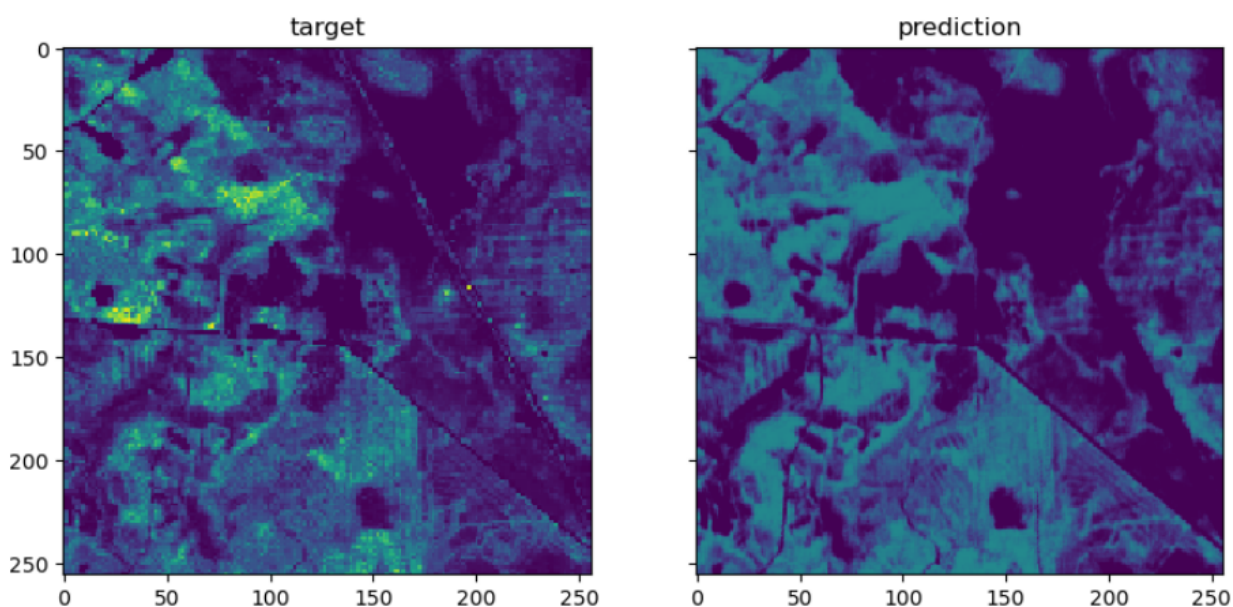


Рисунок 1. Пример целевого изображения и предсказания для участка fed6bb57

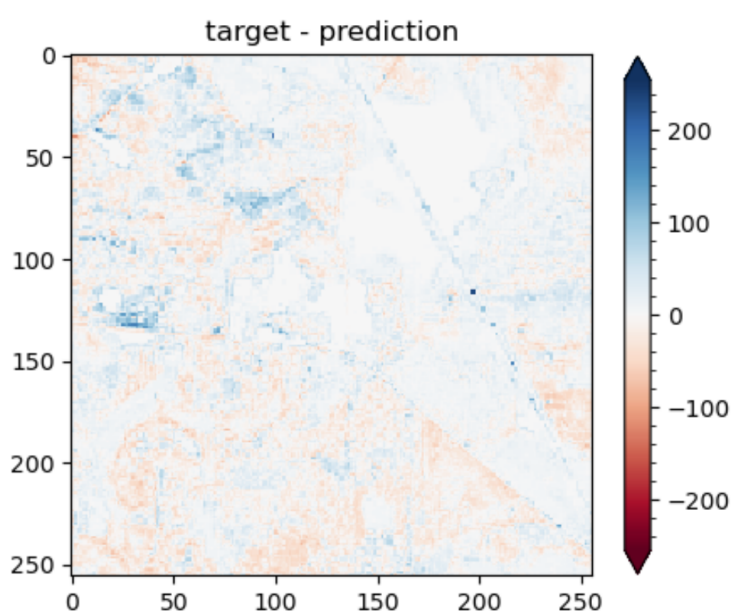


Рисунок 2. Пример разницы изображения и предсказания для участка fed6bb57

Озвученные выводы подтверждаются распределением таргета и прогноза. Мы наблюдаем, что количество пикселей со значениями близкими к нулю существенно выше у модели, чем у реального изображения.

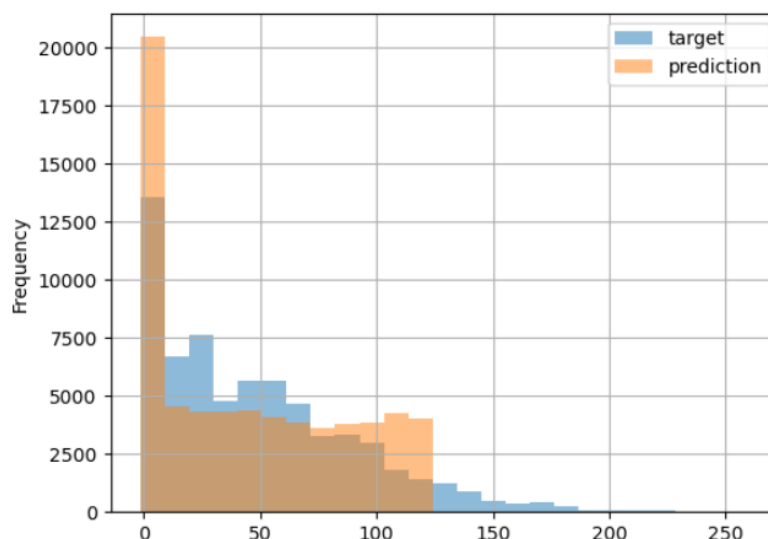


Рисунок 3. Распределение таргета и предсказания для участка fed6bb57

S2-baseline + стратификация выборки

При анализе результатов работы s2-baseline модели мы обратили внимание на то, что прогнозы модели ограничены сверху. Одной из причин такого поведения может быть тот факт, что пикселей с небольшими значениями существенно меньше, поэтому при формировании батчей для обучения и валидации у нас отдается предпочтение пикселям с небольшими значениями. Для того чтобы поднять качество предсказания пикселей со значительными значениями мы пересобрали батчи для обучения и валидации со стратификацией пикселей по значению таргета. Для каждого участка проводится анализ распределения пикселей по бинам размера 100 в диапазоне от 0 до 1000. Пиксели со значениями больше 1000 помещаются в отдельный бин. Для каждого бина определяется количество пикселей, которое будет выбрано в батч для обучения или валидации.

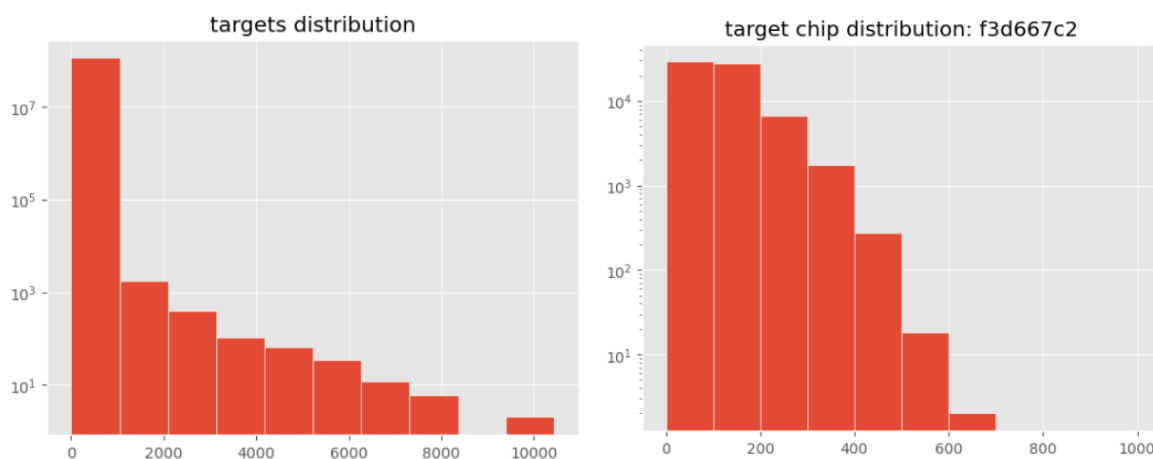


Рисунок 4. Распределение таргета на трейне (слева) и распределение таргета для участка f3d667c2 (справа)

Анализ результатов

Добавление стратификации по таргету привело к увеличению качества модели: loss уменьшился до 23.1, а итоговая метрика уменьшилась до 36.6.

Детальный анализ снимков, сгенерированных моделью и сравнение их с таргетом показал, что модель стала лучше предсказывать пиксели с большим значением. На рисунке ниже приведен пример целевого изображения и изображения, полученного от модели. Можно заметить, что значения модели все еще ниже, чем целевое изображение. Остальные артефакты остались без изменений.

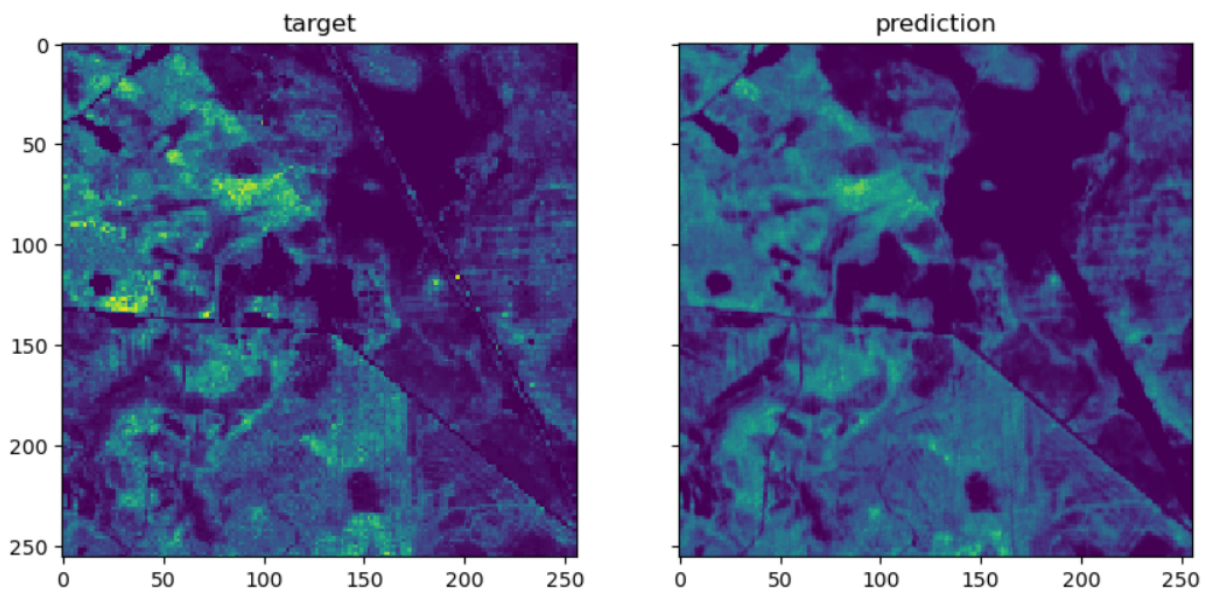


Рисунок 5. Пример целевого изображения и предсказания для участка fed6bb57

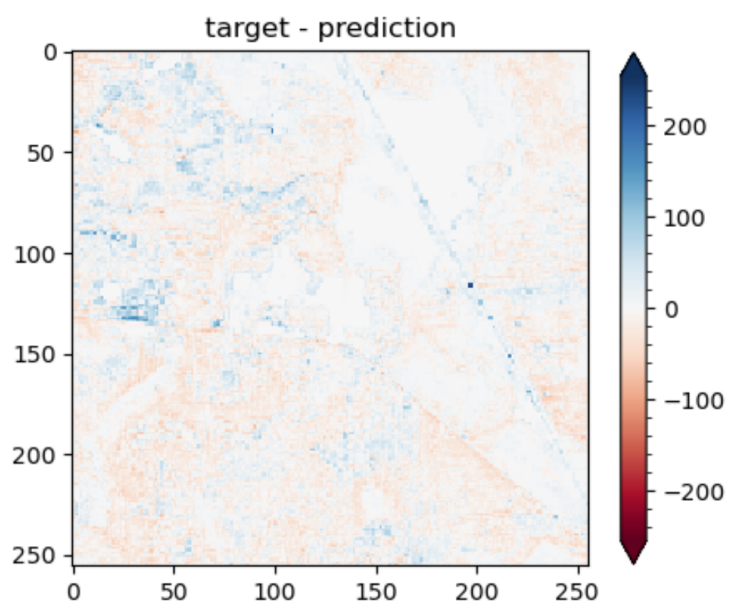


Рисунок 6. Пример разницы изображения и предсказания для участка fed6bb57

Озвученные выводы подтверждаются распределением таргета и прогноза. Мы наблюдаем, что количество пикселей со значениями близкими к нулю все еще выше таргета, но значительно ниже, чем у s2-baseline модели.

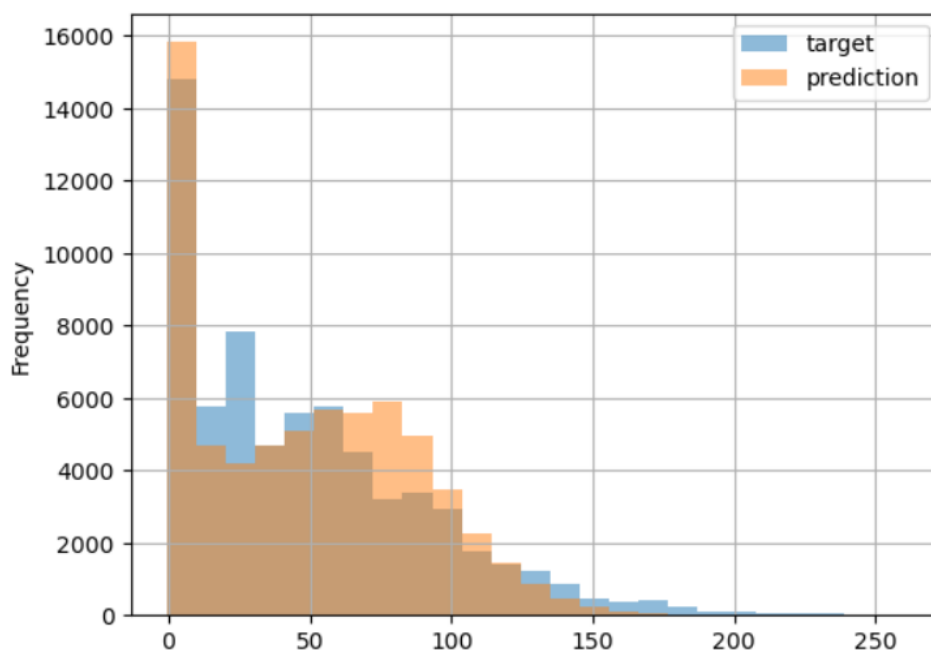


Рисунок 7. Распределение таргета и предсказания для участка fed6bb57

Magnitude-Signal модель

Несмотря на то, что добавление стратификации улучшило модель, мы наблюдаем, что распределение таргета и распределение предсказания все еще имеют отличия.

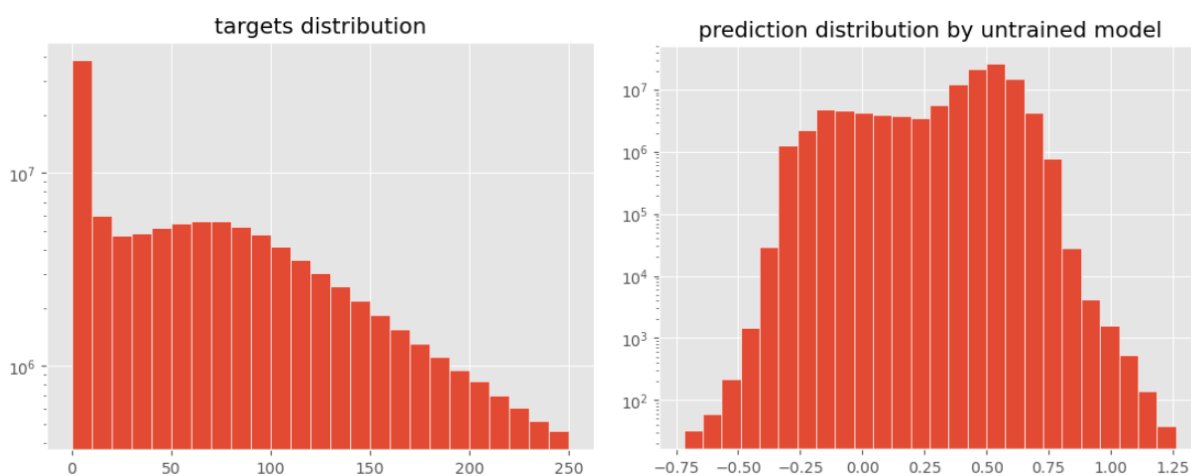


Рисунок 8. Распределение таргета(слева) и предсказания необученной модели

Распределение целевой переменной лежит в широком диапазоне с ярко выраженным пиком в нуле, а необученная модель регрессии

имеет значимо отличное распределение предсказаний (лежит в области нуля, отсутствует пик в нуле и нет хвоста в области больших значений). Это означает, что модели будет затруднительно прийти к состоянию способному выдавать распределение предсказаний близкое к распределению целевой переменной. Для того чтобы решить подобные задачи достаточно часто применяют специально подобранные константы внутри модели. Или, чтобы обеспечить большую гибкость в дальнейшей работе можно собирать предсказание из нескольких компонент. В данной работе предлагается использовать два линейных слоя на выходе: один для определения уровня ответа(амплитуда), а второй для его корректировки(сигнал). Тогда ответ модели будет суммой амплитуды и сигнала.

$$magnitude = Linear(embedding)$$

$$signal = Linear(embedding)$$

$$\hat{y} = \exp(magnitude) + signal$$

Первоначально веса слоя амплитуды инициализированы нулями, чтобы избежать экстремальных значений градиентов.

Анализ результатов

Применение описанной модели привело к ухудшению качества модели: loss увеличился до 24.5, а итоговая метрика увеличилась до 37.9.

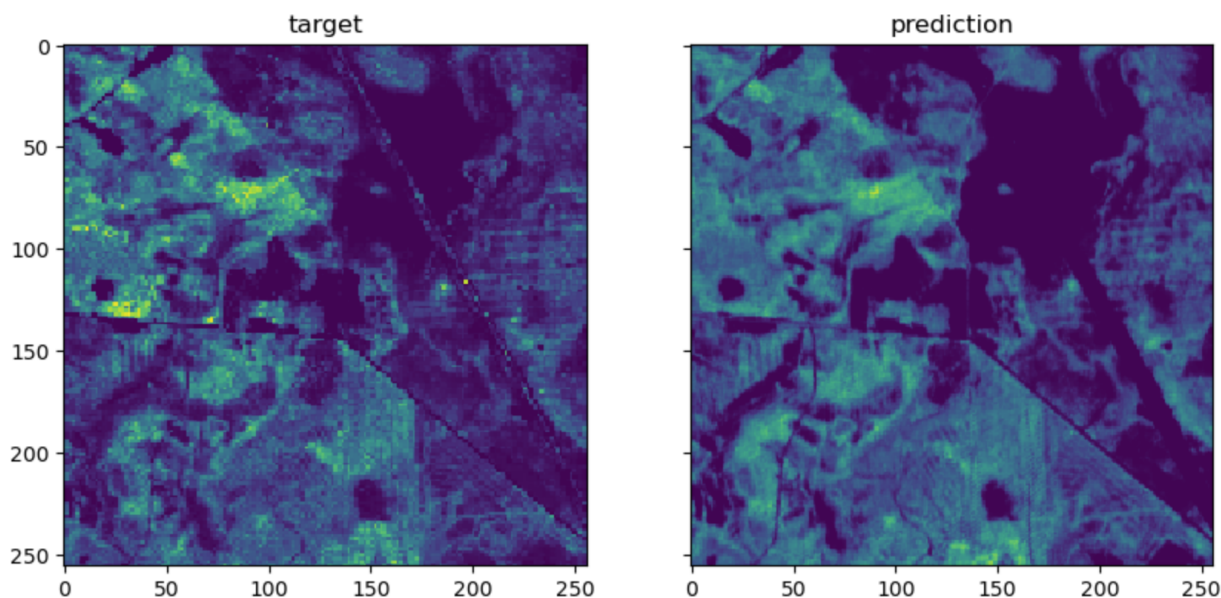


Рисунок 9. Пример целевого изображения и предсказания для участка fed6bb57

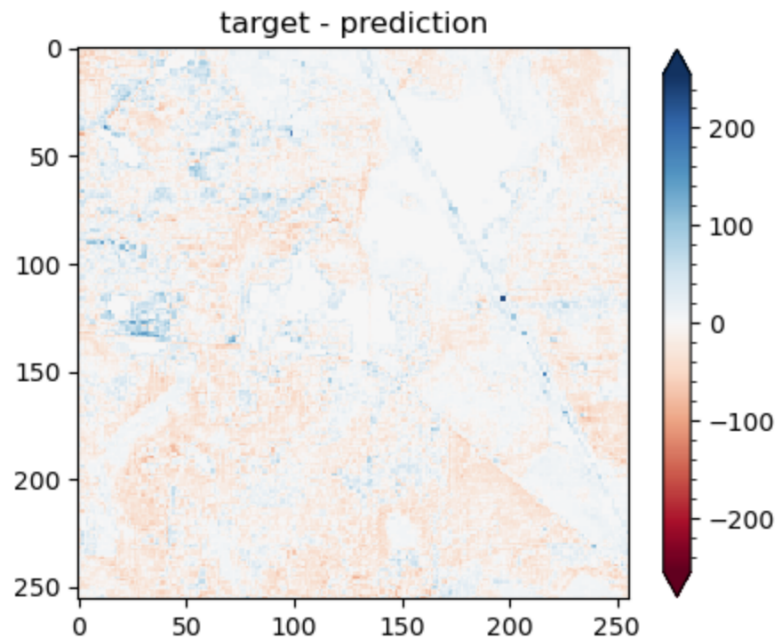


Рисунок 10. Пример разницы изображения и предсказания для участка fed6bb57

Анализ полученных изображений позволяет сделать вывод о том, что такая модель строит заниженные прогнозы для небольших значений таргета.

Распределение таргета и прогнозных значений иллюстрирует, что в области небольших значений целевой переменной модель отдает предпочтение прогнозам близким к нулю. Это приводит к тому, что в области значений таргета от 10 до 30 распределение ответов модели значительно ниже распределения целевой переменной. При этом следует отметить, что распределение предсказаний достаточно хорошо повторяет распределение таргета для диапазона значений от 100 до 150.

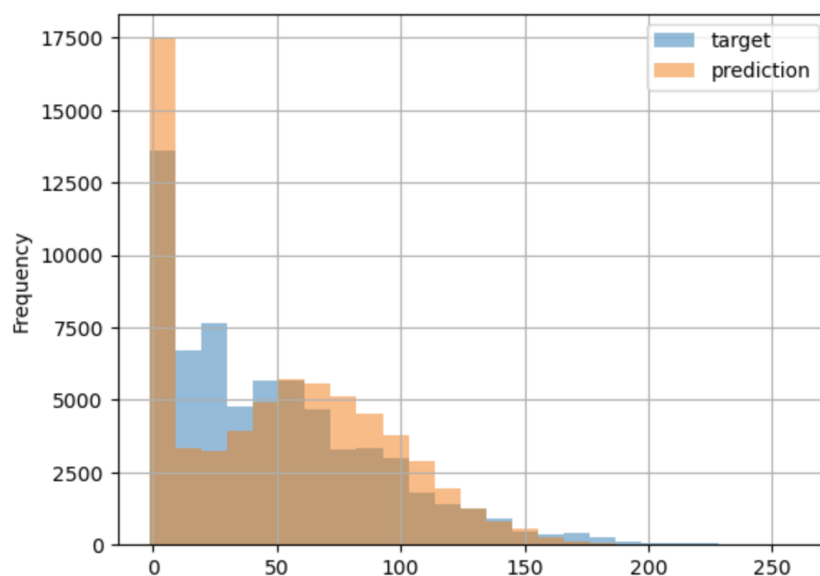


Рисунок 11. Распределение таргета и предсказания для участка fed6bb57