

Ограничения для разрешений моделей:

Архитектура	Допустимый максимум	Примечания
U-Net	4096x4096	Требует деления на степени двойки (512, 1024). Для вытянутых изображений - тайлинг по ширине
ResNet/EfficientNet	1024x1024	Полносвязные слои ограничивают входной размер
Vision Transformer	2048x2048	Потребление памяти растет квадратично из-за self-attention
Swin-Transformer	4096x4096	Иерархическая структура снижает нагрузку на GPU
YOLOv8	1600x1600	При детекции объектов высокие разрешения приводят к потере мелких деталей

Для очень вытянутых изображений ($N > 10 \times M$) стоит нарезать на квадратные патчи $M \times M$ или $M \times k$ (где k кратно 32 для YOLO).

Вытянутые изображения - например вертикальные срезы породы дерева с разрешением 2000x200 пикселей.

Итого:

- Большинство архитектур работают с изображениями до 1024-4096 пикселей по одной оси. Для больших размеров > 10000 пикс требуется тайлинг или даунсэмплинг
- Изображения > 100 Мб будет сложно обработать без разбиений на патчи
- Для вытянутых изображений ($N \gg M$) оптимально резать на квадраты $M \times M$ или применять адаптивное масштабирование

Локальные бинарные паттерны

Возможное применение:

- Классификация пород (однородные текстуры)
- Детекция трещин и разломов (контрастные границы трещин могут создавать специфические паттерны)
- Сравнение образцов деревьев (быстрая оценка схожести текстур разных срезов)

Плюсы:

- Быстрая работа прямо на CPU
- Инвариантность к освещению (нет строгой привязки к свету, контрасту)

Минусы:

- Неустойчив к шумам. Необходимо использовать фильтры
- Уступает по качеству результатам обработки нейросетям

Выводы

Если есть ресурсы (GPU) и время (на подготовку данных и обработку), и требуется максимальная точность, оптимальнее может быть использование нейросетей. Если изображения не содержат сложных структур, то имеет смысл использовать локальные бинарные паттерны, как минимум, чтобы проверить качество обработки. Так, если трещины крупные, то метод LBP может быть оптимальным и наоборот, если трещины мелкие, то требуется большая точность, а значит необходимо

применение CNN. В любом случае, нужно пробовать оба метода, а начинать, возможно, стоит с локальных бинарных паттернов.