

Текущий план работ по ВКР

“Классификация архитектурных стилей”

1. Завершение формирования итогового датасета. Приведение фото к единому формату, очистка от явно нерелевантных изображений (например, слишком мелкие здания, что встречается). Нужно сбалансировать выборку по стилям в разумных пределах через undersampling крупных классов. Разбить данные на обучающую, валидационную и тестовую части с учетом географического разнообразия.

На текущий момент:

<https://www.kaggle.com/dumitru/architectural-styles-dataset>

<https://www.kaggle.com/tompaulat/modern-architecture-100k-small-images>

<https://zenodo.org/records/5166987>

и дополнения.

2. Построение и обучение базовой модели классификации по глобальному изображению. В базе использовать предобученные ResNet-50, EfficientNet-B0/B3, дообучая только несколько последних слоёв. На этом этапе планируется реализация простой постановки - одноклассовая классификация (ито есть один доминирующий стиль на изображение). Качество оценивается по метрикам accuracy, balanced accuracy, F1 по классам.
3. Внедрение модуля сегментации фасадов. То есть запустить предобученную модель сегментации (SegFormer), чтобы получить грубую разметку фасадных элементов (небо, здание, окна, двери и так далее). Задача этапа - получить устойчивую карту хотя бы базовых классов: здание, стеклянные области, глухие стены, крыша, растительность.
4. На базе сегментации из пред. пункта сформировать признаки фасада. Выделить из сегментации простые числовые характеристики: доля пикселей каждого класса (стекло, стена, крыша, растительность), отношение площади остекления к площади стены, наличие/отсутствие крыши определённого типа и др. Вероятно, добавить доли материалов (кирпич, камень, штукатурка, дерево, металл, стекло) и их цветовые характеристики (например, средний оттенок). Эти признаки сохранить в табличной форме как дополнительное описание каждого изображения.
5. Построение расширенного классификатора стилей с интеграцией разных уровней признаков. Модель, в которой глобальные признаки изображения (последний сверточный/трансформерный слой) соединены с табличными структурными и цветовыми признаками фасада. Скорее всего в простейшем варианте - заморозить большую часть CNN, извлечь для каждого изображения вектор признаков из предпоследнего слоя, затем обучить XGBoost на объединенном пространстве CNN-вектор + структурные признаки.

6. Добавить элементы объяснимости в модель, карты важности (Grad-CAM) - показывать, какие области фасада использует модель при принятии решения о стиле.
7. Сравнение всех вариантов моделей. Для каждого стиля оценить точность, полноту, F1. Вероятно, на основе карт Grad-CAM и статистики признаков - интерпретировать типичные ошибки (например, путаница готики и неоготики, барокко и колониального барокко).
8. Реализовать веб-сервис для демонстрации результатов. На основе полученной лучшей модели создать прототип веб-приложения, позволяющий загрузить фотографию фасада, выполнить предобработку (масштабирование, при необходимости выпрямление), сегментацию, классификацию стиля и отобразить пользователю:
 - предсказанный стиль с вероятностями по нескольким ближайшим классам,
 - карту сегментации фасада,
 - визуализацию важности (Grad-CAM) / ключевые числовые признаки.