

ПРИМЕНЕНИЕ СРЕДСТВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ФОРМИРОВАНИЯ СЕГМЕНТАЦИОННЫХ МАСОК

Лантев В.В.¹, Гергет О.М.²

¹ Сибирский государственный медицинский университет, техник-программист, email: vy139@tpu.ru

² Томский политехнический университет, ИШИТР, профессор, email: gerget@tpu.ru

Введение

Одной из основных задач обработки и анализа изображений является сегментация. Выделение на **изображении** области согласно определенному критерию позволяет не только обнаружить интересный объект на изображении, но и получить значимые характеристики, такие как: границы объекта, форма, площадь и т.п.

Существует большое множество алгоритмов сегментации и их можно подразделить на группы методов, основанные на: энтропии; кластеризации; оттенках серого и гистограммах; локальных порогах; пиксельных корреляциях, атрибутах; искусственных нейронных сетях [1].

В настоящее время наибольший интерес представляют системы, основанные на алгоритмах с использованием искусственных нейронных сетей [2]. Именно эта группа обладает таким понятием как обобщающая способность, что позволяет применять их на более широком спектре задач. Но стоит отметить, что для получения такой системы сегментации объектов высокой точности требуется большое количество обучающих данных. Под обучающими данными подразумевается набор изображений и аннотаций к ним (описание анализируемого изображения). Как правило составлением аннотации к данным занимаются эксперты в данной области, а под аннотацией сегментационной модели понимается маска изображения. Пример маски изображения приведен на рисунке 1.



Рис. 1. Пример изображения с маской: А) исходное изображение; В) маска изображения

За частую именно сбор данных и составление аннотаций занимает наибольшее количество времени при разработке систем компьютерного зрения. Причем составление сегментационной маски изображения является самым трудоемким из них.

Целью данной работы является разработка метода предварительного формирования сегментационных масок изображений, используя более простые модели анализа данных.

Описание алгоритма

Сегментация изображений является сложной задачей и требует наличия большого количества данных, подготовка которых занимает весомое количество времени. В статье предлагается метод, упрощающий данную задачу. Идея метода состоит в следующем: перейти от задачи сегментации изображений к классификации и использовать карту признаков, как сегментационную маску.

Задача классификации – это получение категориального ответа на основе набора признаков, например, определить есть ли на фотографии кот или присутствует на изображении лицо человека, болен ли пациент раком и т.д. Подготовка данных для получения модели классификации высокого качества, как правило, не занимает много времени, как и обучение самой модели. Используя данное преимущество, за короткий срок, нами была подготовлена модель классификации глазного дна на предмет патологии. В основе классификационной модели лежит энкодер Xception. Данный энкодер обучен на 350 миллионах изображений Imagenet и хранит веса для определения 17000 классов [3]. Для адаптации к описанной выше задаче, выходной слой был заменен на два полносвязных слоя размерами 2048 и 1. Активационная функция выходного слоя Sigmoid. Во время обучения модели, вычисление

градиента и изменение весов коснулось только последних 2-х слоев, что позволяет сохранить тенденцию модели к выделению характеристических особенностей изображения и повысить качество формирования конечной матрицы признаков. Данный подход именуется как трансферное обучение [4].

Выходная матрица признаков модели Xception имеет размер $7*7*2048$. Следующим этапом нашей работы является объединение сформированных карт признаков и активационных значений вектора признаков результирующего класса. Заключительным этапом является интерпретация полученного результата. Реализация данной задачи осуществлялась с использованием следующего набора изображений, приведенный на рисунке 2.

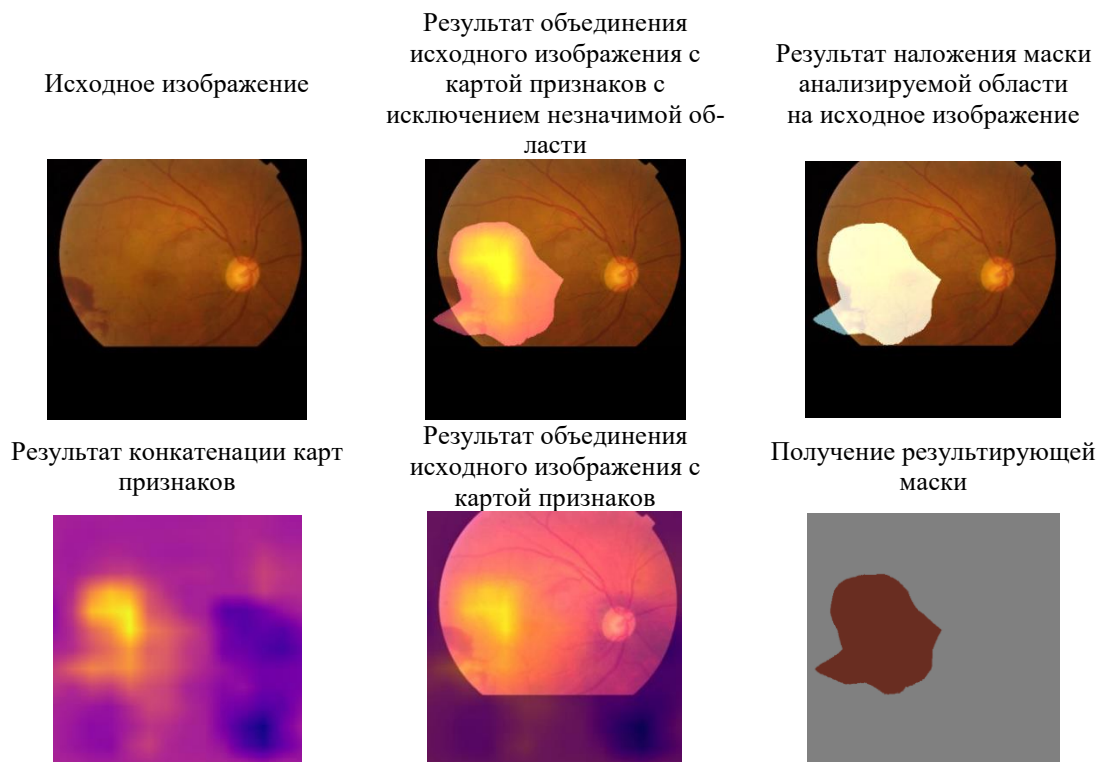


Рис. 2. Визуализация результатов эксперимента

Заключение

Представленный в данной работе метод формирования сегментационной маски изображения значительно сокращает время создания датасета для обучения модели сегментации изображений. Создаваемые таким образом маски могут быть использованы одним из следующих способов: 1) обучение модели сегментации на полученных масках с последующим краткосрочным обучением с низкой скоростью обучения на небольшом количестве данных, составленном экспертами; 2) улучшение качества полученных масок экспертами с последующим обучением модели сегментации.

Благодарности

Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации, проект Наука FSWW-2023-0007.

Список использованных источников

1. Панченко Д.С., Путятин Е.П. Сравнительный анализ методов сегментации изображений // Радиоэлектроника и информатика. – 1999. – №. 4 (9). – С. 109–114.
2. Немировский В. Б., Стоянов А. К. Сегментация изображений с помощью рекуррентной нейронной сети // Известия Томского политехнического университета. Инжиниринг георесурсов. – 2012. – Т. 321. – №. 5. – С. 205–210.
3. Chollet F. Xception: Deep learning with depthwise separable convolutions // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2017. – С. 1251–1258.
4. Torrey L., Shavlik J. Transfer learning // Handbook of research on machine learning applications and trends: algorithms, methods, and techniques. – 2010. – P. 242–264.