

## Цели применения

Для применения тщательного мониторинга нескольких медицинских состояний (например, диабета, ожирения и т. д.), которые затрагивают большой процент населения мира.

В большинстве случаев интервенционные психологи просят людей вести подробный ручной учет ежедневно потребляемых блюд. Однако, обычно люди склонны недооценивать количество потребляемой пищи до 33%. (по “Стивен В. Лихтман, Кристина Писарска, Эллен Рейнс Берман, Мишель Пестоне, Хиллари Доулинг, Эстер Оффенбахер, Хоуп Вайзель, Стэнли Хешка, Дуайт Э. Мэтьюз и Стивен Б. Хеймсфилд. Расхождение между самооценкой и фактическим потреблением калорий и физическими упражнениями у лиц с ожирением. *Журнал медицины Новой Англии*, 327(27):1893–1898, 1992.”)

## Технология реализации

1) Модель: CNN (Convolutional Neural Networks). Т.к. она способна изучать сложные пространственные закономерности на изображениях.

2) Особенность: обнаружение и распознавание пищи как на обычных, так и на эгоцентрических зрительных изображениях (снятые на любительскую камеру/телефон с точки зрения пользователя).

3) Что умеет:

- алгоритм локализации объектов, связанных с пищей, который специально обучен различать изображения, содержащие обычную пищу, и имеет возможность предлагать несколько ограничивающих рамок, содержащих обычную пищу (без определенных классов) на одном изображении:
  - ✓ обучение бинарного классификатора пищи/не пищи CNN для обнаружения пищи.
  - ✓ генерация карты вероятности пищевых продуктов.
  - ✓ обнаруживает не только продукты питания, но и объекты, связанные с продуктами питания (например, бутылки, чашки и т. д.)

## Преимущество перед сегментацией с классификацией

Основная проблема сегментации с классификацией заключается в том, что она предполагает, что блюдо было ранее локализовано и поэтому оно находится в центре изображения. Вместо этого в контексте локализации продуктов питания, авторы статьи заинтересованы в поиске точных общих областей (или ограничивающих рамок) на изображении, где присутствует какой-либо вид пищи.

Открытым вопросом является то, как получить простые методы локализации объектов, которые одновременно обеспечивают высокую точность и отзыв.

Альтернативой методам локализации общих объектов являются методы, обученные локализовать набор предопределенных объектов, такие как Faster R-CNN. Авторы предлагают мощную сквозную CNN, оптимизированную для локализации набора из 20 определенных классов объектов. (ШАОЦИН РЕН, КАЙМИН ХЕ, РОСС ГИРШИК И ЦЗЯНЬ САН. Более быстрый r-cnn: на пути к обнаружению объектов в реальном времени с сетями предложений регионов. В *Достижения в области нейронных систем обработки информации*, страницы 91–99, 2015)

### **Методология**

- 1) Классификатор пищевых и непищевых товаров: первый шаг к получению универсального локализатора еды — обучить CNN для бинарной классификации еды.
- 2) Тонкая настройка для генерации FAM (рамок): как только у авторов появилась модель, способная различать изображения Food и Non Food.
- 3) Генерация ограничивающей рамки.

### **Выводы**

Сравнивая методы с точки зрения точности, можно заметить, что метод авторов статьи превосходит другие методы во всех случаях. Эту закономерность легко объяснить, учитывая, что любой общий метод локализации объектов (в данном случае выборочный поиск) обычно выдает несколько тысяч предложений на изображение, что приводит к получению большого количества FP (это предсказанные ограничивающие рамки, которые не существуют в истинных данных). Предложение авторов лучше, т.к. способно получить более сбалансированные результаты точности-полноты.

В итоге авторы статьи предложили первую методологию для одновременной локализации и распознавания пищи. Этот метод применим к обычным, так и эгоцентрическим изображениям.