



СУББОТНИКИ  
Мини-конференции. Омск

# Сегментация и обработка чеков

Дмитрий Вальчук  
backend developer  
7bits

# Задача

Сегментировать (найти) чек на фотографии и сделать его “красивеньким”,  
затем обработка текста на фотографии чека

# Зачем?

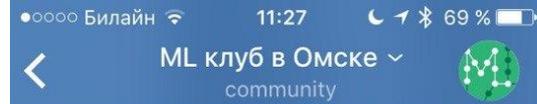
- сделать удобный учет расходов
- изучение рынка
- саморазвитие
- ...

# Подзадачи

- Поиск данных
- Создание модели машинного обучения для сегментации чека на фотографии
- На основе предсказаний модели необходимо выделить чек и преобразовать его в прямоугольник
- Бинаризация изображения и всякие другие улучшения изображения
- Определение местоположения символов (скорее всего для каждого блока отдельно)
- Распознавание символов
- Улучшение полученных результатов с помощью языковых моделей

# Поиск данных

## В интернете данных мало — собирай свои



Сын уронил мой айфон  
и слегка повредил  
экран. Если кого-то  
интересует - могу  
подарить: ему 3 года,  
русый, глаза карие.



Message



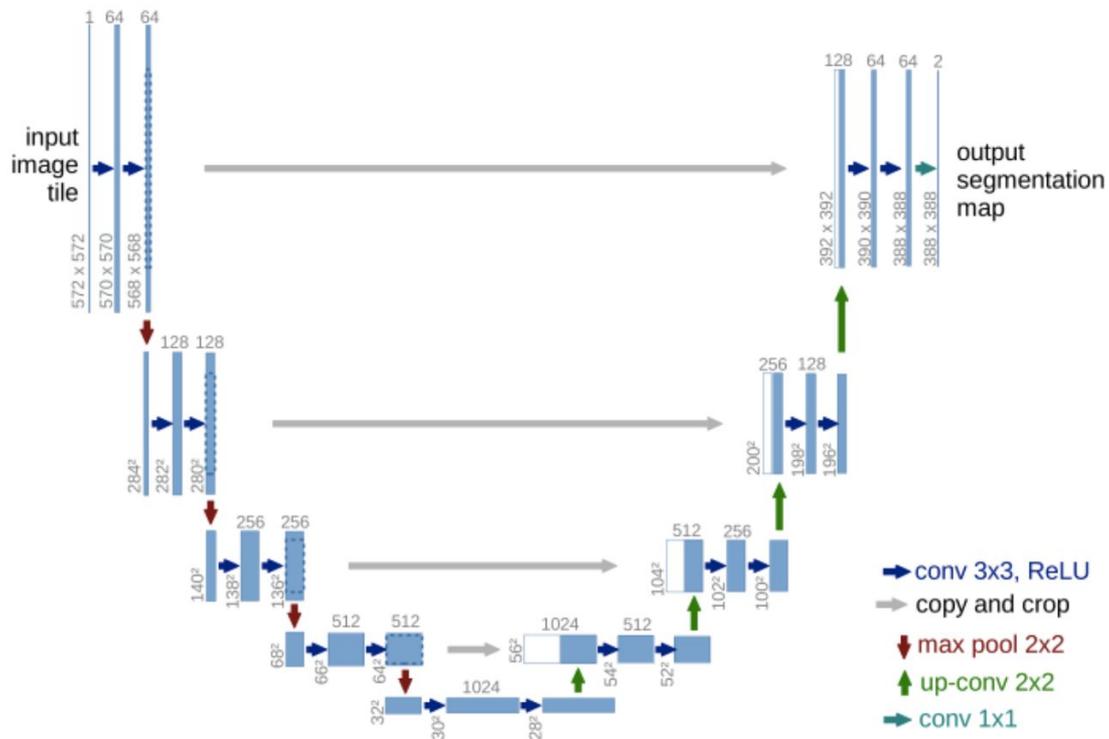
# Разметка данных

Для разметки данных использовался бесплатный инструмент labelbox.io

С его помощью на данный момент размечено около 300 фотографий.



# U-Net

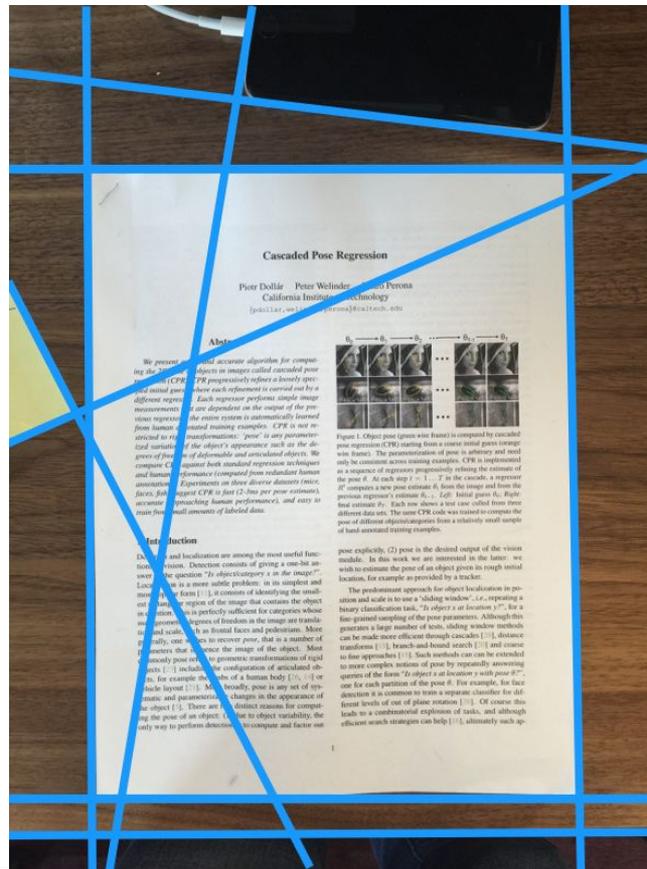




# Преобразование Хафа

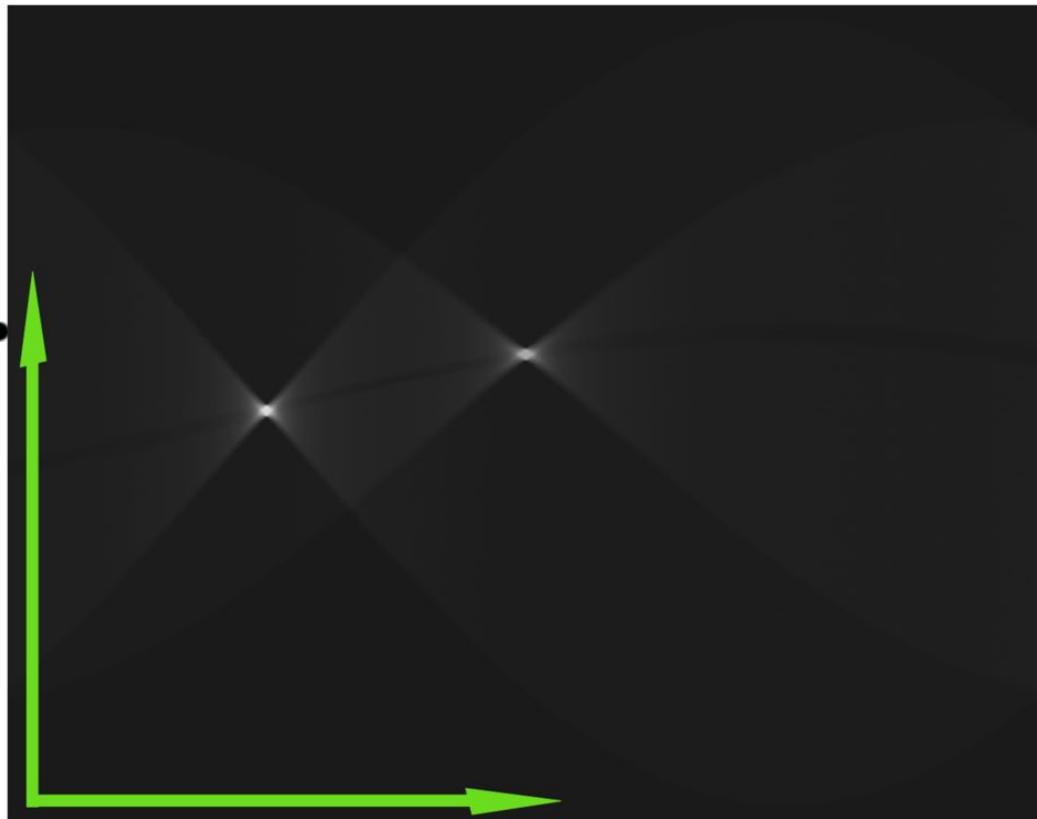
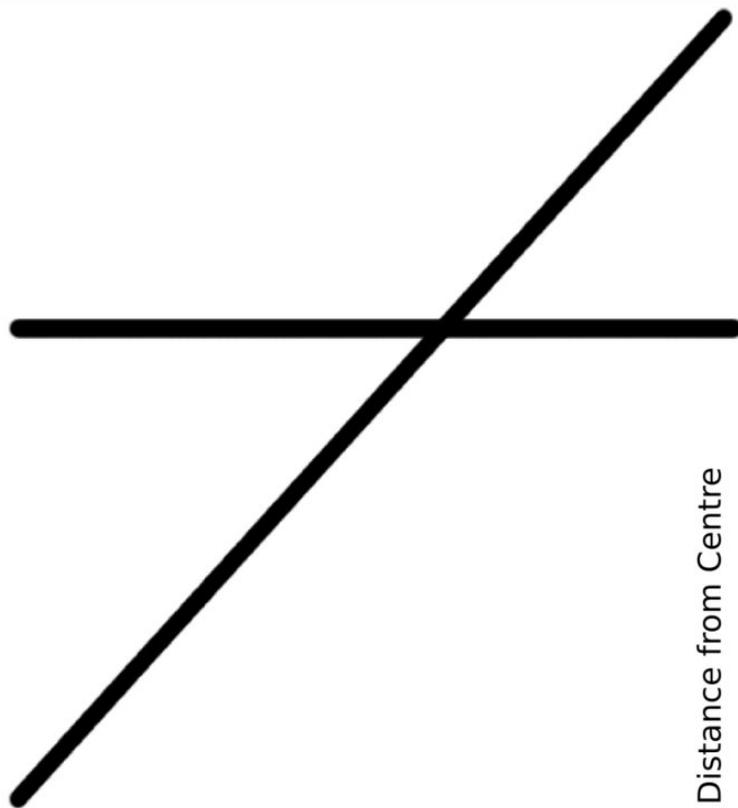
$$y = mx + b$$

$$y = \begin{pmatrix} -\cos \theta \\ \sin \theta \end{pmatrix} x + \begin{pmatrix} r \\ \sin \theta \end{pmatrix}$$



Input Image

Rendering of Transform Results





(a) Image



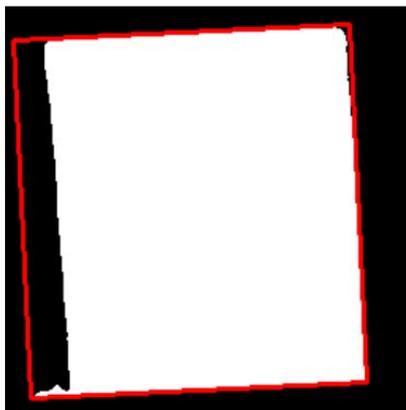
(b) FCN Prediction



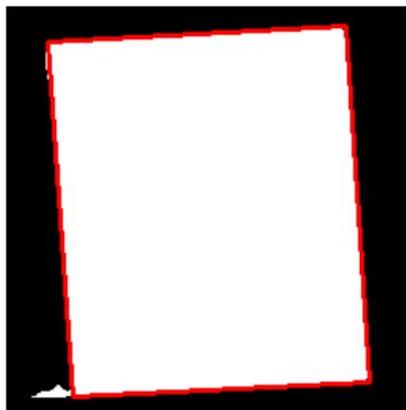
(c) Largest Connected Component



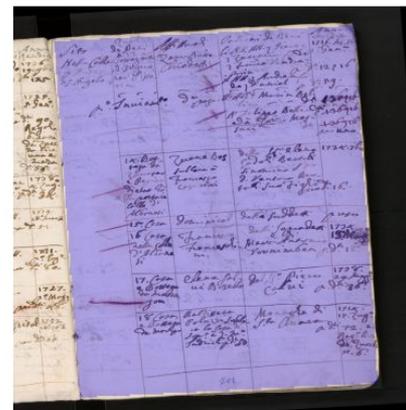
(d) Fill Holes



(e) Oriented Bounding Rectangle



(f) Quadrilateral Minimizing IoU



(g) Result

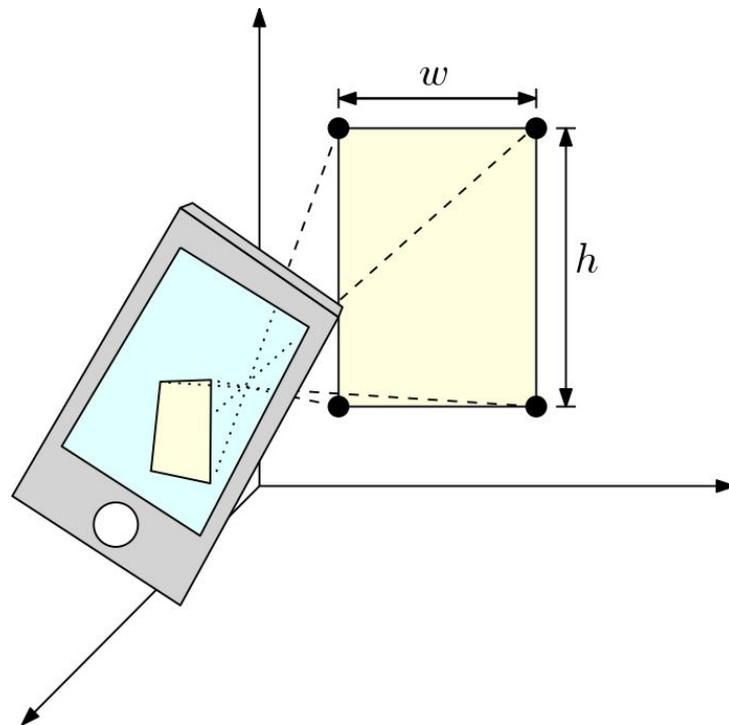
И что дальше?

# И что дальше?

Преобразование четырехугольника в  
прямоугольник!

# И что дальше?

Преобразование четырехугольника в  
прямоугольник!



# Калибровка камеры

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A \begin{bmatrix} R & T \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

$$A = \begin{bmatrix} \alpha_x & \gamma & u_0 \\ 0 & \alpha_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

# Бинаризация изображения

$$\sum_{x,y} \|J(x,y) - 255\|^2,$$

# Бинаризация изображения

$$\sum_{x,y} \|\nabla J(x,y) - \nabla I(x,y)\|^2.$$

# Бинаризация изображения

$$\operatorname{argmin}_J k_1 \sum_{x,y} \|J(x,y) - 255\|^2 + k_2 \sum_{x,y} \|\nabla J(x,y) - \nabla I(x,y)\|^2$$

# OCR (Optical Character Recognition)

Плюсы от того, что мы работаем с чеками:

- Моноширный шрифт

# OCR (Optical Character Recognition)

Плюсы от того, что мы работаем с чеками:

- Моноширный шрифт
- Более-менее устоявшийся словарь

# OCR (Optical Character Recognition)

Плюсы от того, что мы работаем с чеками:

- Моноширный шрифт
- Более-менее устоявшиеся сокращения слов
- Ограниченный набор слов в целом

# Давайте это использовать!

\*\*\*\*\*  
ООО "АТАК"  
г. Москва ул. Бибиревская г. 9  
тел. (495) 660-98-70  
Открыт: 9:00-22:00 без выходных  
\*\*\*\*\*

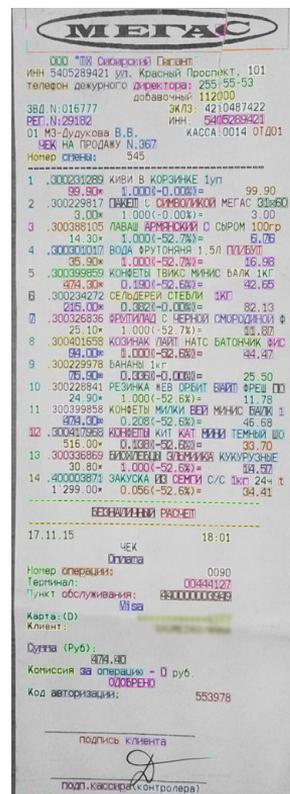
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1  
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2  
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3  
4 4 4 4 4 4 4 4 4 4 4 4 4 4 4  
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5  
6 6 6 6 6 6 6 6 6 6 6 6 6 6 6  
7 7 7 7 7 7 7 7 7 7 7 7 7 7 7  
8 8 8 8 8 8 8 8 8 8 8 8 8 8 8  
9 9 9 9 9 9 9 9 9 9 9 9 9 9 9

# Все бы хорошо...

1.	300331872	ПАКЕТ С СИМВОЛИКОЙ МЕГАС БОРД		
	7,50*	1.000(-0.0%)=		7,50
2.	300325660	КАРТОФЕЛЬ П/ПАК СВЕЖИЙ СТАНДАР		
	80,30*	1.000(-0.0%)=		80,30
3.	400006967	НАБОР Д/УХИ ИЗ ФОРЕЛИ ОХЛ П/Ф		
	199,00*	0,520(-0.0%)=		103,46
4.	400008707	ГОЛЕНЬ ЦЫПЛЕНКА КОПЧЕНАЯ 1кг 4		
	349,00*	0,322(-0.0%)=		112,38
5.	400000531	ВИНЕГРЕТ ОВОЩНОЙ 1кг 12ч t+2-6		
	189,00*	0,328(-0.0%)=		61,99

# OCR (Optical Character Recognition)

Использование findContours из OpenCV с последующим объединением близкорасположенных компонент (чтобы всякие ё, й, : объединились)



# Ошибки в распознавании символов

Ограниченный набор словаря пригодится нам



\*\*\*\*\*  
Ресторан "Grisha" ----- 000 <Рестор>  
ИНН 5504135503  
\*\*\*\*\*  
ГОСТЕВОЙ СЧЕТ

Зал: Зал Стол: 15  
Открыт: 22.02.2018 18:59 Заказ № 7080  
Официант: Максименко Иван

Наименование	Кол-во	Сумма
Барри сэт(4 шота)	1	700,00
Тонка Май Тай	2	700,00
Бургер сырная прелесть с мраморным мясом	1	380,00
1MediumRege(с кровью)		
Бланш де Брюссель	1	290,00
Штрудель яблочный	4	600,00
Сидр Любельски 0,75	3	900,00
Коллинз	1	250,00
Лондон Портер бут.	1	330,00
Зомби	5	1 750,00
Эклер с клубнично-базел .кремом	1	150,00
Черновар темное	2	520,00
Bloody Bull	1	250,00
Гри-Гри спиритц	1	300,00
Фруктовая тарелка	1	240,00
СЕРДЦЕ АЛТАЯ	1	130,00

# Спасибо за внимание и мотивацию

