

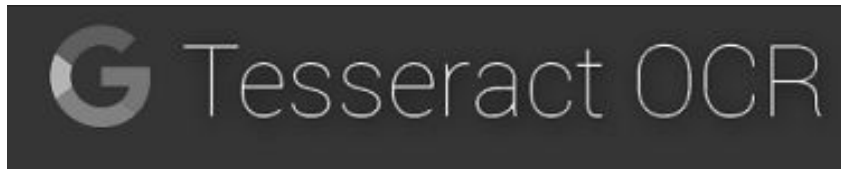
Модуль розпізнавання зображень чеків

Автор: студент 4-го курсу
групи КА-43
Люнда Володимир
Дипломний керівник:
Дідковська М.В.

Актуальність роботи

- ведення особистої фінансової звітності
- збір інформації про покупців великих торгових мереж
- аналіз купівельних трендів за межами окремих торгових мереж

Існуючі підходи до розв'язання даної задачі



- + працює в режимі **offline**
- + відкрите програмне забезпечення
- + працює на мобільних пристроях
- висока чутливість до будь-якого шуму



- + **state of the art** точність
- + немає необхідності портувати систему на мобільні пристрої
- необхідність підключення до мережі Інтернет
- висока вартість

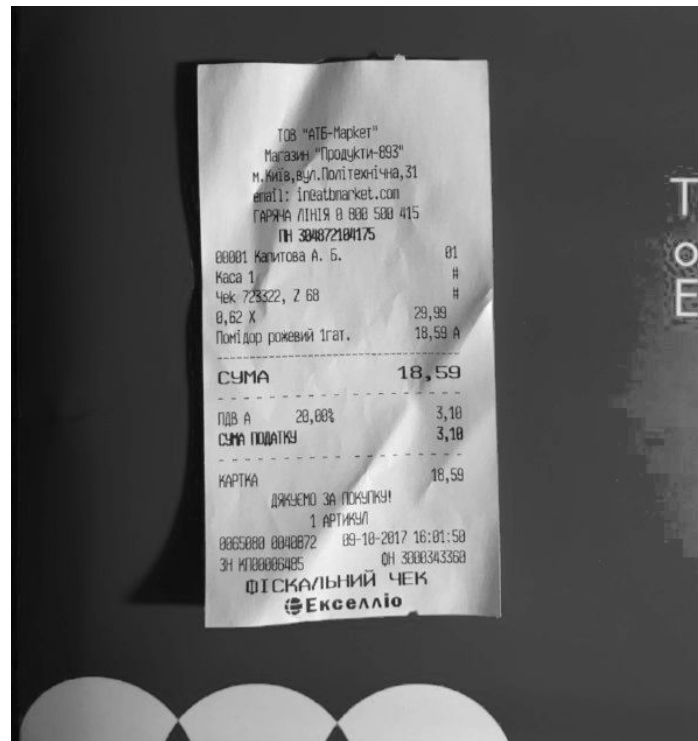
Постановка задачі дипломної роботи

1. Розробити підхід до мінімізації шумів на фотографіях чеків
2. Розробити підхід до розпізнавання тексту чеків на попередньо оброблених фотографіях
3. Забезпечити можливість портувати систему на мобільні пристрої
4. Розробити програмний продукт що ефективно працює з фотографіями зробленими на мобільних пристроях

Задача **пониження шуму** на фотографіях чеків

Проблеми:

- крім чеку на фотографії присутні інші об'єкти
- Чек повернутий відносно горизонталі
- Освітлення нерівномірно розповсюджено по поверхні чеку

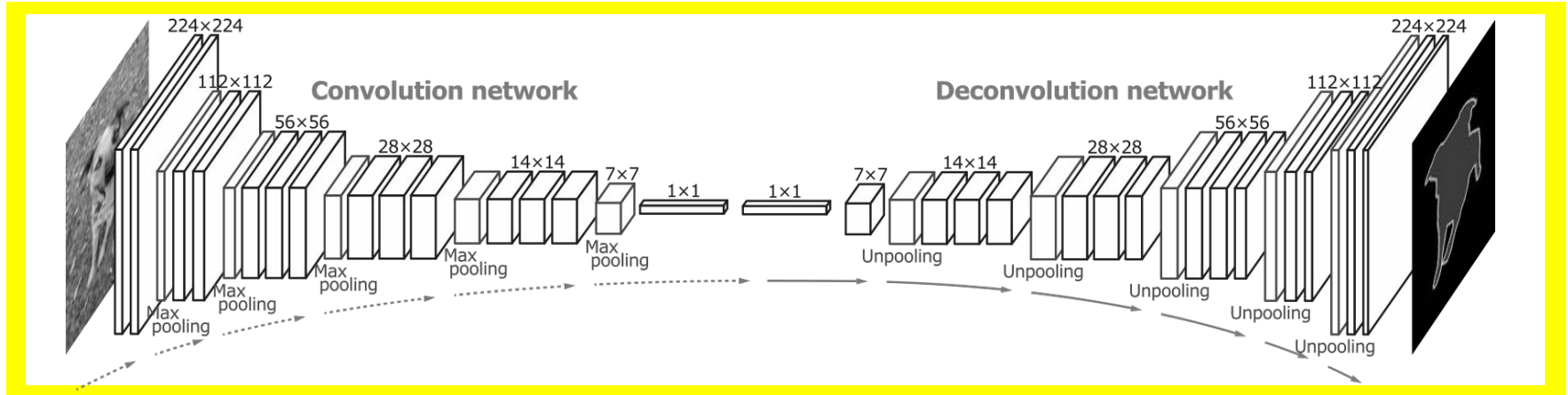


Алгоритми сегментації для відділення фону



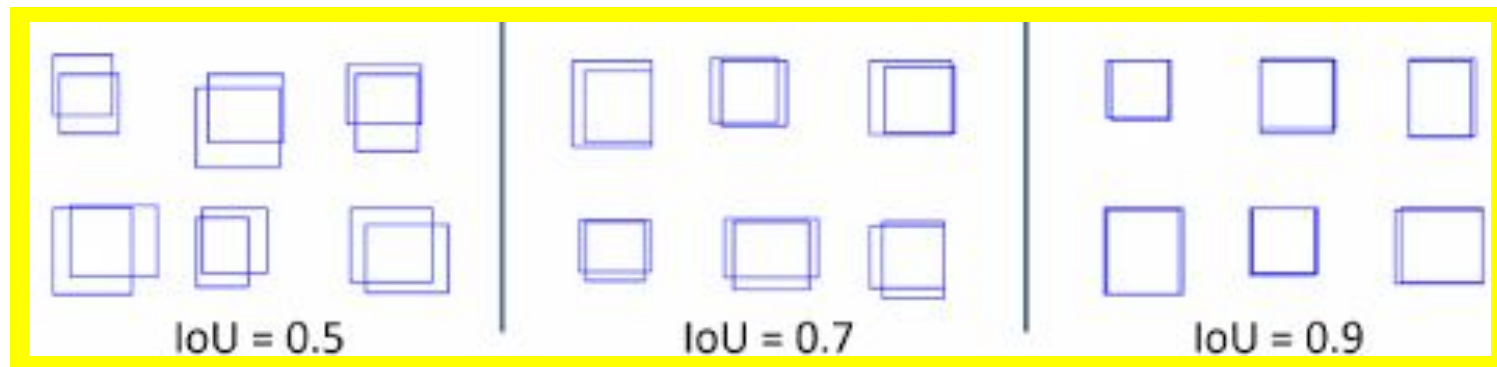
Алгоритм **GrabCut** працює з зображенням як з графом областей і порівнюючи їх між собою видаляє ті, які не відносяться до переднього плану

Алгоритми сегментації для відділення фону



Алгоритм **U-net** розроблений для задач сегментації мікроскопічних фотографій клітин використовує згортковий автоенкодер з додатковими зв'язками для збереження деталізації

Метрики якості сегментації



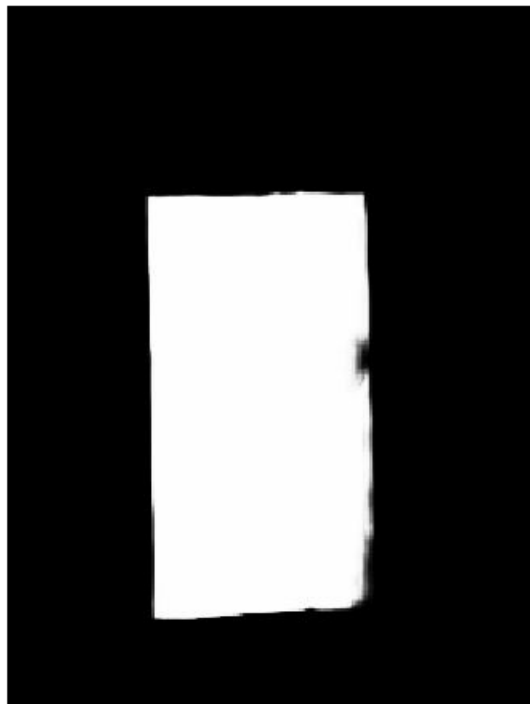
як метрику якості сегментації був обраний **коефіцієнт Жаккара** (англ. Intersection over Union)

Порівняння роботи алгоритмів сегментації

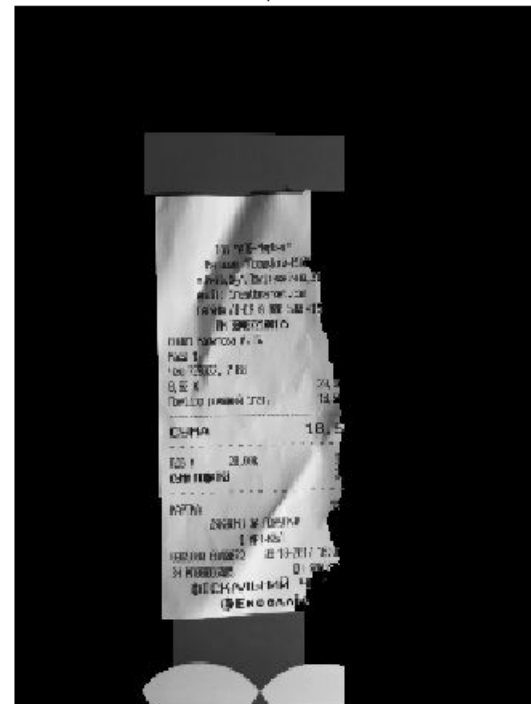
original



U-net



GraphCut



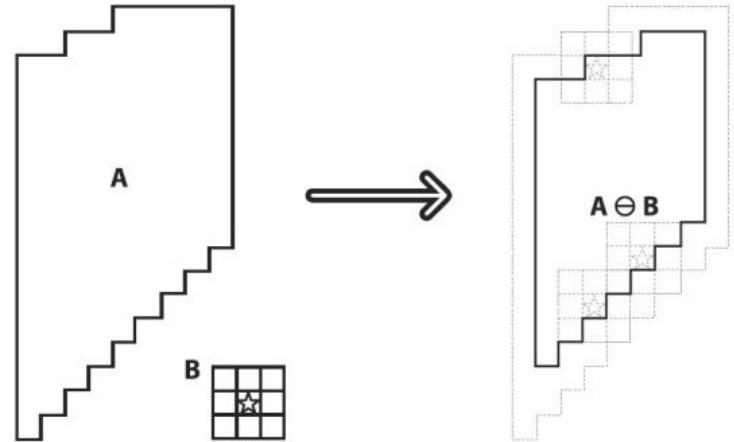
Порівняння роботи алгоритмів сегментації

алгоритм	коефіцієнт Жаккара	швидкодія (зобр./сек.)
U-net	0.9835	8.47
GrabCut	0.8797	1.32

Частотна фільтрація та морфологічні перетворення для боротьби з шумом

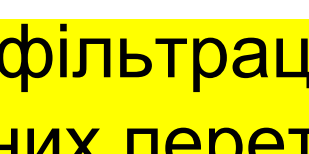


для фільтрації низькочастотних шумів було використано **медіанний фільтр** з розміром вікна (3,3)



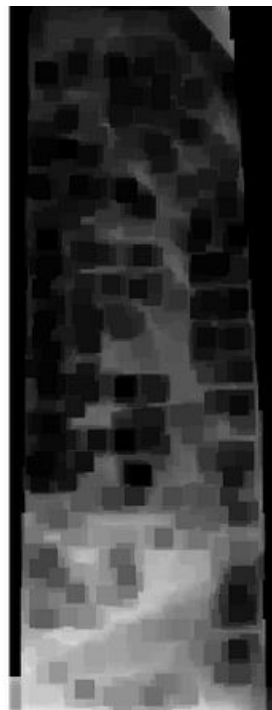
Для боротьби з високочастотними шумами викорисовується морфологічне перетворення **ерозія** з розміром вікна (3, 3)

Результати фільтрації та морфологічних перетворень



ТОВ "АВІА-Сервіс"	
"АВІА" "Авиасервіс"	
Львів, Дзержинська вулиця, 90	
тел.: (0322) 407100	
ГРПР № 4104 000 500 410	
ID: 404210415	
0005 Даремні Г. С.	81
Код Є	8
№К 116380, 2 бл	8
Тракт. вул. Миколаївська	1,40 А
5,854 X	6,29
Копіюва 1 шт	24,21 А
Висл. 1 л. Різдво казарне	47,00Г
Прод. дес. чашки 50шт	21,50 А
Саше 190 г Барс	11,70 А
Омлетська	59,50
Чашка 1шт	4,80 А
Пакети дес. ората 20 шт	9,90 А
Sunny 45	8
round	8
Огарки Стоунд. West	26,50Г
25 Red TP	0,01Г
Знижка	0,01Г
СУМА	157,40
ПДВ А 20,00%	12,42
ПДВ Б 20,00%	10,63
ЗСП ПУ-6 6,00%	3,62
СІМА ПІДАТРИ	25,42
ГОТІВКА	200,00
РЕМТА	62,00
Дякуємо за покупку!	
в Україні	

інверсія



медіанне розмиття



ерозія

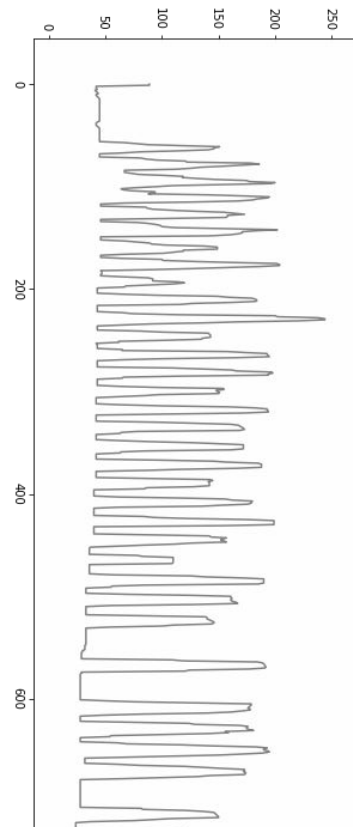


ТОВ "АВІА-Сервіс"	
"АВІА" "Авиасервіс"	
Львів, Дзержинська вулиця, 90	
тел.: (0322) 407100	
ГРПР № 4104 000 500 410	
ID: 404210415	
0005 Даремні Г. С.	81
Код Є	8
№К 116380, 2 бл	8
Тракт. вул. Миколаївська	1,40 А
5,854 X	6,29
Копіюва 1 шт	24,21 А
Висл. 1 л. Різдво казарне	47,00Г
Прод. дес. чашки 50шт	21,50 А
Саше 190 г Барс	11,70 А
Омлетська	59,50
Чашка 1шт	4,80 А
Пакети дес. ората 20 шт	9,90 А
Sunny 45	8
round	8
Огарки Стоунд. West	26,50Г
25 Red TP	0,01Г
Знижка	0,01Г
СУМА	157,40
ПДВ А 20,00%	12,42
ПДВ Б 20,00%	10,63
ЗСП ПУ-6 6,00%	3,62
СІМА ПІДАТРИ	25,42
ГОТІВКА	200,00
РЕМТА	62,00
Дякуємо за покупку!	
в Україні	

результат

Задача розбивки чеку на окремі рядки тексту

Для цієї задачі було розроблено модифікацію алгоритму **імітації випалювання**, яка шукає мінімуми на гістограмі фільтрованого зображення



Розпізнавання тексту

- Використовувались нейронні мережі архітектури LSTM
- якісь розпізнавання можна істотно покращити використовуючи словники

```
In [266]: api = PyTessBaseAPI(psm=tesseract.PSM.SINGLE_LINE, lang='ukr')
_ = api.SetVariable("load_system_dawg", "0")
```

```
In [267]: text = []
for l in tqdm(im_lines):
    api.SetImage(preprocess(l))
    text.append(api.GetUTF8Text().lower())
```

100% 23/23 [00:04<00:00, 4.68it/s]

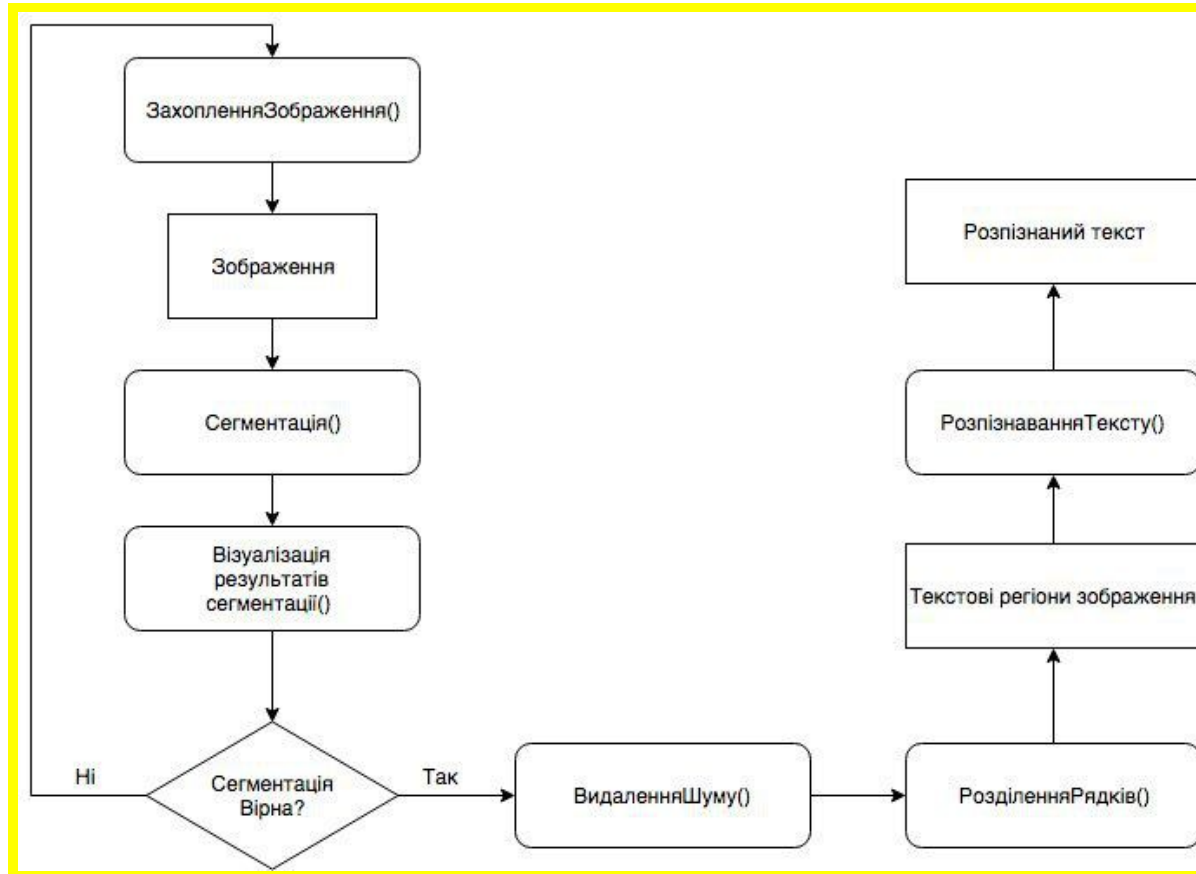
```
In [268]: print(''.join(text))
```

```
тов "атб-маркете
пагазин. "продукти-яз»
м. київ, вул. політехнічна зи
і аряча лінія в во ой де
елай!: іпоаколагкеьсоі
ті з0187216д175
вовоу оніценко г, 1. де
каса й я
чек м17ібоу, 7 об і
нектар 8,5 л біола 7
мультивітамін 12.ва
круасан 85 г розумний й
вивір з аб ца
счма 16,58
пва 20,005 2,5
си податку 2
готівка 16,50
дякуємо за покупку!
2 вртикулів
ф2589 02щ8127 18-04-2018 10:10:28
зі клоодоблви он 3000343361
о вііскальний чек
нин
```

Порівняння систем розпізнавання

Система розпізнавання	Точність (Левенштейн)	Швидкодія (зобр./сек.)	Вартість (\$/1000 зобр.)	Необхідність підключення до мережі Інтернет
tesseract-legacy	537	3.3	~0	-
tesseract-LSTM	231	4.1	~0	-
Google CloudVision	147	1.3	1.50	+

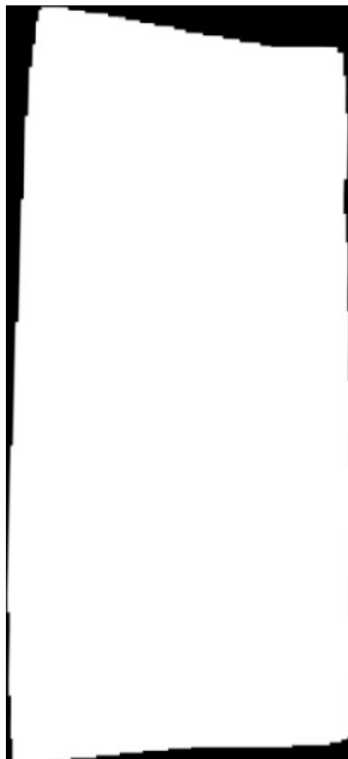
Повна блок-схема процесу



Результати роботи



оригінал



маска і обрізане зображення



Результати роботи

00004 Оніщенко Г. І. 01
Каса 4 #
Чек 1174594, 2 264 #
Нектар 0,5 л Біола
Мультивітамін 12,20 А
Круасан 65 г Розумний
вивір з аб 4,30 А
СУМА 16,50
ПДВ А 20,00% 2,75
СУМА ПОДАТКУ 2,75
ГОТІВКА 16,50
ДЯКУЄМО ЗА ПОКУПКУ!
2 АРТИКУЛІВ
0412589 0248127 18-04-2018 10:10:28
ЗІ КЛОДОБЛВИ ОН 3000343361
ПІДСКАДЬИИ ЧЕК

бінаризоване зображення

каса 4	01
Чек 1174594, 2 264	#
Нектар 0,5 л Біола	"
Мультивітамін	12,20 А
Круасан 65 г Розумний	"
вивір з аб	4,30 А
СУМА	16,50
ПДВ А 20,00%	2,75
СУМА ПОДАТКУ	2,75
ГОТІВКА	16,50
ДЯКУЄМО ЗА ПОКУПКУ!	
2 АРТИКУЛІВ	
0412589 0248127 18-04-2018 10:10:28	
ЗІ КЛОДОБЛВИ ОН 3000343361	

виділені рядки

каса й я
чек м17ібоу, 7 об і
нектар 8,5 л біола 7
мультивітамін 12.ва
круасан 85 г розумний й
вивір з аб ца
счма 16,50
пва 20,005 2,5
си податку 2
готівка 16,50
дякуємо за покупку!
2 вртикулів
2589028127 18-04-2018 10:10:28
зі клоодблви он 3000343361

розпізнаний текст

Висновки

- запропоновано використання модифікації **алгоритму сегментації U-net** яка на 10% точніша і в 6.4 разів швидша за класичний алгоритм GraphCut
- запропоновано архітектуру системи фільтрації шуму для зображень чеків
- запропоновано ефективний алгоритм розділення тексту чеку на рядки
- програмно реалізовано модуль розпізнавання та створено користувацький інтерфейс у вигляді телеграм-бота

Шляхи подальшого розвитку

- перетренувати tesseract під шрифт окремої торгової мережі
- використовувати нелінійні перетворення для кращого виділення області чеку
- натренувати мовну модель для тексту чеків
- розробити алгоритм вирівнювання слів в рядку тексту

Дякую за увагу