

Министерство образования и науки Российской Федерации

Федеральное государственное автономное  
образовательное учреждение высшего образования  
«Московский физико-технический институт  
(научно-исследовательский университет)»

Физтех-школа прикладной математики и информатики  
Кафедра когнитивных технологий

«Бинаризация изображений документов»

Выполнил:  
студент группы М05-215г  
Куланин Евгений

Москва 2023

## Введение.

Под задачей бинаризации изображения понимается задача присвоения каждому из пикселей исходного изображения одного из 2 возможных классов: объекта или фона. Данную задачу можно называть одним из частных случаев задачи сегментации изображений, в которой каждому из пикселей сопоставляется метка одного из нескольких классов.

В работе представлен код 4 различных методов бинаризации изображений:

1. Метод глобальной бинаризации Оцу;
2. Модифицированный метод Оцу для наравновесного случая;
3. Локальный метод бинаризации Ниблэка;
4. Многомасштабная модификация метода Ниблэка.

Так же в работе представлены несколько изображений, для которых была проведена ручная разметка классов, а также найден метод и параметры для каждого из изображений, с наивысшим показателем функции оценки качества F1. Все методы разрабатывались для использования их при бинаризации полутоновых изображений, т.е. изображений, имеющих не более 256 значений показателя яркости пикселя.

## Метод Оцу.

Метод бинаризации изображений Оцу – это представитель класса методов глобальной бинаризации, т.е. методов, в которых для всего изображения в целом находится один конкретный порог. Пиксели, которые по величине яркости являются меньше заданного порога, помечаются классом 0 (объект). Остальные же пиксели помечаются классом 255 (фон).

Суть метода заключается в минимизации значения внутриклассовой дисперсии (1). Однако было доказано, что вместо минимизации внутриклассовой дисперсии можно максимизировать межклассовую дисперсию (2), что позволяет немного сократить количество операции в одной итерации метода.

$$\sigma_w^2(k) = \varpi_0(k)\sigma_0^2(k) + \varpi_1(k)\sigma_1^2(k) \quad (1)$$

$$\sigma_B^2(k) = \varpi_0(k)\varpi_1(k)(\mu_0(k) - \mu_1(k))^2 \quad (2)$$

Алгоритм метода Оцу:

1. Вычислить гистограмму яркости изображения (столбчатую диаграмму, которая имеет 256 различных значений по оси x, в каждом столбце которой подсчитано количество пикселей, соответствующих конкретному значению);
2. В цикле по каждому возможному значению порога от 0 до 255 вычисляется доля пикселей, которые соответствуют каждому из двух классов, а также математическое

ожидание по каждому из классов. После чего подсчитывается межклассовая дисперсия по формуле (2).

3. Если значение межклассовой дисперсии по новому порогу будет превосходить максимальное найденное значение дисперсии, то порог меняется на новый, при котором было получено данное значение межклассовой дисперсии.
4. Все пиксели, имеющие значение ниже порога, помечаются как 0, а все остальные как 255.

Основные преимущества метода состоят в быстрой работе и отсутствии каких-либо настроечных параметров. Однако недостатками метода является неустойчивость к изменению яркости фона, а также сложности при классификации изображений, в которых количество пикселей фона сильно больше количества пикселей объекта.

### **Модифицированный метод Оцу для неравновесного случая.**

В первую очередь модификация метода Оцу для неравновесного случая решает проблему, когда количество пикселей объектов сильно меньше количества пикселей фона.

Для решения задачи с помощью данной модификации метода Оцу необходимо подсчитывать внутриклассовую дисперсию. После чего сам критерий поиска оптимального порога будет состоять в максимизации величины, вычисляемой по формуле (3).

$$Q(k) = \varpi_0(k) \log(\varpi_0(k)) + \varpi_1(k) \log(\varpi_1(k)) - \log(\sigma_w) \quad (3)$$

Основным преимуществом данного метода стала возможность работать с изображениями, описанными выше, при увеличении количества операций, необходимых для вычисления критерия максимизации. Так же, данный метод все еще не имеет настроечных параметров, что облегчает работу с ним. Повышается устойчивость к изменению яркости фона. Однако, главным недостатком метода все еще является глобальный порог.

### **Метод Ниблэка.**

Метод Ниблэка является представителем методов локальной бинаризации, когда порог для каждого пикселя вычисляется в зависимости от локальных особенностей пикселей изображения вокруг него. Для вычисления порогов необходимо находить среднее значение в окрестности вокруг пикселя. Данное действие можно выполнить с помощью интегрального изображения – матрицы размера изображения, в которой для каждого из пикселей вычисляется сумма всех пикселей, находящихся выше и левее исходного пикселя. Тогда среднее можно вычислить при помощи нее как сумму 4 элементов. Причем, вычислительная сложность не будет зависеть от размера ядра.

Алгоритм метода Ниблэка:

1. Найти интегральное изображение для исходного и изображения, в котором все пиксели соответствуют квадрату пикселей исходного изображения;
2. Найти среднее по окрестности для каждого из пикселей изображения, аналогично вычислить среднее для матрицы квадратов пикселей;
3. Найти дисперсию по формуле (4);

$$\sigma_R(x, y) = \sqrt{\mu_R(I^2(x, y)) - \mu_R^2(I(x, y))} \quad (4)$$

4. Найти пороги бинаризации для каждого из пикселей по формуле (5);

$$T(x, y) = \mu_R(I(x, y)) + k\sigma_R(x, y) + a \quad (5)$$

5. Бинаризовать каждый из пикселей в соответствии с его порогом.

Основные преимущества метода заключаются в высокой устойчивости к изменению яркости фона изображения, а также возможность подстроить параметры метода под каждое изображение индивидуально. Недостатками метода является количество параметров метода (размер ядра, коэффициент  $k$ , коэффициент  $a$ ). Так же метод может плохо работать в условиях со сложными локальными особенностями вокруг пикселей или большим количеством шума.

Для высокого уровня точности бинаризации изображения стоит серьезно относиться к подбору размера ядра, по которому находится среднее вокруг пикселя. Если ядро будет слишком маленьким, то не получится найти локальные особенности изображения. Если в окно попало слишком мало пикселей фона или объекта, то метод Ниблэка может давать неточные результаты.

### **Многомасштабная модификация метода Ниблэка.**

Данная модификация призвана избавиться от одного из недостатков исходного метода, а именно получения неточного значения порога бинаризации, когда в окно попало слишком мало пикселей одного из классов. Решается это с помощью увеличения размера ядра в случае, если в окрестности пикселя дисперсия имеет значение ниже заданного порога.

Все плюсы метода Ниблэка сохраняются, а также появляется возможность учитывать масштаб объекта, однако вычислительная сложность данного алгоритма сильно возрастает.

### **Метрики оценки бинаризации изображения.**

Важной задачей при изучении качества метода бинаризации изображения является оценка качества данного метода. С этой целью в программе был реализован метод оценки

качества бинаризации F1 score, который вычисляется по формуле (6). Для вычисления данной функции нужно посчитать количество пикселей, классифицированных правильно и неправильно. При этом, необходимо иметь маску разметки, которая будет иметь идеальную разметку.

$$F1 = \frac{2TP}{2TP + FP + FN} \quad (6)$$

### **Работа с программой.**

Для того, чтобы бинаризовать изображение с помощью данной программы следует сделать следующее:

В терминале следует запустить `python main.py <название_метода> <путь_до_исходного_файла> <расположение_файла_после_бинаризации> <параметры_метода>`.

В качестве названия метода можно выбрать:

1. `otsu` – для запуска метода Оцу (дополнительные параметры не нужны);
2. `otsunoneq` – для запуска модификации метода Оцу (дополнительные параметры не нужны);
3. `niblack` – для запуска метода Ниблэка (нужно задать размер крыла окна (т.е. для ядра  $n \times n$  задается значение, равное  $(n-1)/2$ ), коэффициент  $k$ , коэффициент  $a$ );
4. `niblackms` – для запуска метода Ниблэка (нужно задать размер крыла окна (т.е. для ядра  $n \times n$  задается значение, равное  $(n-1)/2$ ), коэффициент  $k$ , коэффициент  $a$ ).

### **Эксперимент по бинаризации.**

Для проведения эксперимента были выбраны 5 полутоновых изображений, созданных из цветных. Для каждого из них была создана полуавтоматически (бинаризации с помощью компьютера, а затем ручная доработка полученных изображений) маска, состоящая из пикселей 3 различных классов (0 - объект, 128 - игнорировать в статистике, 255 - фон). Для каждого из них был проведен эксперимент и найден лучший способ бинаризации и параметры по мере F1.

Результаты вместе с изображениями представлены ниже.

ОТЧЕТ о научно-исследовательской работе за I семестр 2022/2023 учебного года	
ФИО студента	Будякина Елена Николаевна
Фамилия, имя, отчество	Будякина Елена Николаевна
Наименование организации, кафедры	Кафедра Информационных технологий
Тема НИР	Анализ эффективности методов бинаризации текстовых изображений
Текст отчета НИР за семестр (результаты работы и научные результаты)	В отчёте описаны классические методы бинаризации изображений, а также рассмотрены современные методы бинаризации текстовых изображений.
Итого НИР за семестр	Принятие участия в научных конференциях, симпозиумах (названия, темы, даты проведения и конференции, место проведения)
	Научные публикации (названия, авторы, даты и журналы)
	Участие в конкурсах на уровне НИР и за пределами (названия, даты проведения, результаты)
	Модели, программы, проекты, чертежи, схемы, рисунки, фотографии НИР и за пределами (названия, даты проведения, результаты)
	Прочие материалы на конкурс (названия, даты проведения, результаты)
	Прочие материалы (названия, даты проведения, результаты)
	Прочие материалы (названия, даты проведения, результаты)
	Прочие материалы (названия, даты проведения, результаты)
	Прочие материалы (названия, даты проведения, результаты)
	Прочие материалы (названия, даты проведения, результаты)
Материальная поддержка НИР студента за семестр (с указанием источников финансирования)	
План работы на следующий семестр	
Оценка научного руководителя	Анна Николаевна Будякина
Оценка НИР студента за семестр	ОЦ (10)
Студент	Будякина Елена Николаевна
Научный руководитель	Будякина Елена Николаевна
Зам. кафедры	

Рис. 1 Изображение 1 - рукописный текст и таблица.

Лучшим набором метода и параметров для бинаризации изображения 1 стал метод Ниблэка с параметрами: размер крыла ядра = 50,  $k = -1.8$ ,  $a = -10$ . При данных параметрах было получено значение меры  $F1 = 0.923$ .

ОТЧЕТ о научно-исследовательской работе за I семестр 2022/2023 учебного года	
ФИО студента	Будякина Елена Николаевна
Фамилия, имя, отчество	Будякина Елена Николаевна
Наименование организации, кафедры	Кафедра Информационных технологий
Тема НИР	Анализ эффективности методов бинаризации текстовых изображений
Текст отчета НИР за семестр (результаты работы и научные результаты)	В отчёте описаны классические методы бинаризации изображений, а также рассмотрены современные методы бинаризации текстовых изображений.
Итого НИР за семестр	Принятие участия в научных конференциях, симпозиумах (названия, темы, даты проведения и конференции, место проведения)
	Научные публикации (названия, авторы, даты и журналы)
	Участие в конкурсах на уровне НИР и за пределами (названия, даты проведения, результаты)
	Модели, программы, проекты, чертежи, схемы, рисунки, фотографии НИР и за пределами (названия, даты проведения, результаты)
	Прочие материалы на конкурс (названия, даты проведения, результаты)
	Прочие материалы (названия, даты проведения, результаты)
	Прочие материалы (названия, даты проведения, результаты)
	Прочие материалы (названия, даты проведения, результаты)
	Прочие материалы (названия, даты проведения, результаты)
	Прочие материалы (названия, даты проведения, результаты)
Материальная поддержка НИР студента за семестр (с указанием источников финансирования)	
План работы на следующий семестр	
Оценка научного руководителя	Анна Николаевна Будякина
Оценка НИР студента за семестр	ОЦ (10)
Студент	Будякина Елена Николаевна
Научный руководитель	Будякина Елена Николаевна
Зам. кафедры	

Рис.2. Маска бинаризации, созданная полуавтоматически.

ОТЧЕТ о научно-исследовательской работе за I семестр 2022/2023 учебного года	
ФИО студента	Халикина Евгения Александровна
Факультет, группа	ФПИ-ИИ, ИИ-24-2
Наименование кафедры	Кафедра Информационных технологий
Тема НИР	Искусственный интеллект в обработке изображений
Тематика системы НИР за семестр (проектная работа и промежуточные результаты)	Разработка прототипа системы классификации изображений на основе нейронных сетей; анализ эффективности различных архитектур нейронных сетей в задачах классификации изображений.
Иные НИР за семестр	Доклады на научных конференциях, семинарах (адрес, название документа и конференция, место проведения)
	Публикации в научных изданиях (адрес, название работы и издания)
	Участие в конкурсах на уровне НИР и выставках (адрес, название работы и конкурсов (выставки и выставки))
	Материалы, доклады, презентации, отчеты и т.д. на конкурсах на уровне НИР и на выставках (адрес, название работы и конкурсов (выставки и выставки), дата участия)
	Проекты, поданные на конкурсы грантов (адрес, название и код гранта)
	Публикации в печати (адрес, название и код гранта)
	Другие формы и способы деятельности по области интеллектуальной собственности, включая патенты на изобретения, полезные модели, программы для ЭВМ и т.д.
	Материальные средства НИР студента за семестр (с указанием источников финансирования)
	Иные работы за семестр
	Отзыв научного руководителя
Оценка НИР студента за семестр	ОЦ (10)

Студент Халикина Е.А. дата составления отчета 26.12.2022

Научный руководитель Гарайшина И.Р.

Зав. кафедрой \_\_\_\_\_

Рис.3. Результат бинаризации изображения 1.

Федеральное агентство по образованию  
Российский государственный университет информационных технологий, связи и массовых коммуникаций  
Федеральное государственное образовательное учреждение высшего образования  
«Московский государственный университет информационных технологий, связи и массовых коммуникаций»  
(МГУИТ, ФПИИТ)  
119899, Москва  
№ 263902

Почтовый адрес: 119701, Москва (г.м.),  
г. Дзержинский, Мухоморовский пер., д. 4  
Тел.: 8 (495) 476-1340; факс: 8 (495) 476-0140

**СПРАВКА**

Дана Халикина Евгения Александровна, дата рождения 04.02.2000 г., в чем она является обучающимся 1 курса магистратуры МФ ФПИИТ-ИИТ (Информационные технологии и Информатика) очной формы обучения по основной профессиональной образовательной программе (специальности 09.04.01 Информатика и вычислительная техника).

Принят и зачислен в институт № 4476-4 от 12.08.2022 г.

Дата начала обучения 01.09.2022 г.

Предполагаемая дата окончания обучения 31.08.2024 г.

Справка дана для предъявления по месту требования.

Начальник учебного управления \_\_\_\_\_

Гарайшина И.Р.

Рис.4. Изображение 2 - документа с печатью.

Министерство науки и высшего образования  
Российской Федерации  
Федеральное государственное автономное  
образовательное учреждение  
«Московский физико-технический институт»  
(национальный исследовательский университет)  
(МФТИ, Физтех)  
13.09.2022 г.  
№ 16392  
Почтовый адрес: 141701, Московская обл.,  
г. Долгопрудный, Институтский пер., д. 9  
Тел.: 8 (495) 408-57-00, факс: 8 (495) 408-68-69

#### СПРАВКА

Дана Кузнецову Евгению Игоревичу, дата рождения 04.02.2000 г., в 1  
что он является обучающимся 1 курса магистратуры МФ  
Физтех-школы Прикладной Математики и Информатики очной фор  
обучения по основной профессиональной образовательной программе.  
Обучается за счет бюджетных ассигнований федерального бюджета  
направленно 09.04.01 Информатика и вычислительная техника.  
Приказ о зачислении в институт № 4476-4 от 12.08.2022 г.  
Дата начала обучения 01.09.2022 г.  
Предполагаемая дата окончания обучения 31.08.2024 г.  
Справка дана для предъявления по месту требования.

Начальник учебного  
управления



Гарайшина И. Р.

Рис.5. Маска бинаризации изображения 2.

Министерство науки и высшего образования  
Российской Федерации  
Федеральное государственное автономное  
образовательное учреждение  
«Московский физико-технический институт»  
(национальный исследовательский университет)  
(МФТИ, Физтех)  
13.09.2022 г.  
№ 16392  
Почтовый адрес: 141701, Московская обл.,  
г. Долгопрудный, Институтский пер., д. 9  
Тел.: 8 (495) 408-57-00, факс: 8 (495) 408-68-69

#### СПРАВКА

Дана Кузнецову Евгению Игоревичу, дата рождения 04.02.2000 г., в 1  
что он является обучающимся 1 курса магистратуры МФ  
Физтех-школы Прикладной Математики и Информатики очной фор  
обучения по основной профессиональной образовательной программе.  
Обучается за счет бюджетных ассигнований федерального бюджета  
направленно 09.04.01 Информатика и вычислительная техника.  
Приказ о зачислении в институт № 4476-4 от 12.08.2022 г.  
Дата начала обучения 01.09.2022 г.  
Предполагаемая дата окончания обучения 31.08.2024 г.  
Справка дана для предъявления по месту требования.

Начальник учебного  
управления



Гарайшина И. Р.

Рис.6. Результат бинаризации модификацией метода Оцу.

Для изображения 2 лучшим методом бинаризации стала модификация метода Оцу  
для несбалансированных классов. Мера  $F1 = 0,95$ .



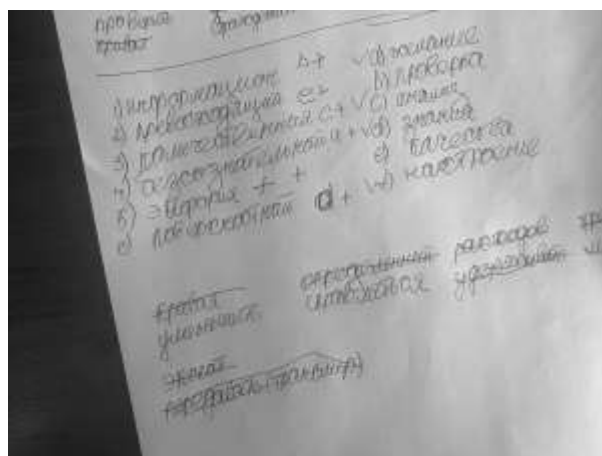


Рис.7. Изображение 3 – рукописный текст.

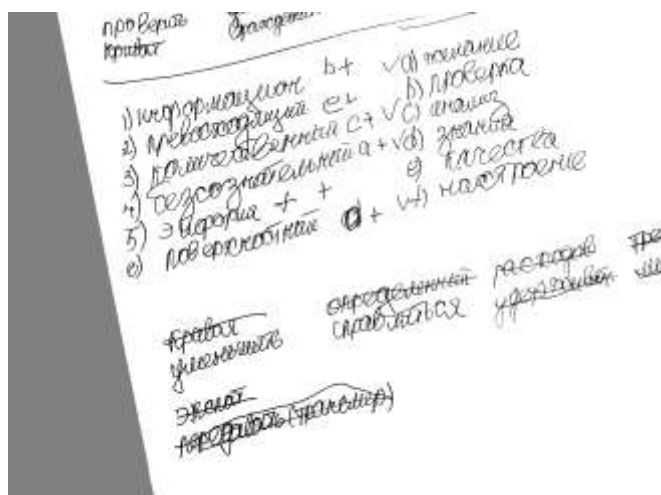


Рис.8. Маска изображения 3.

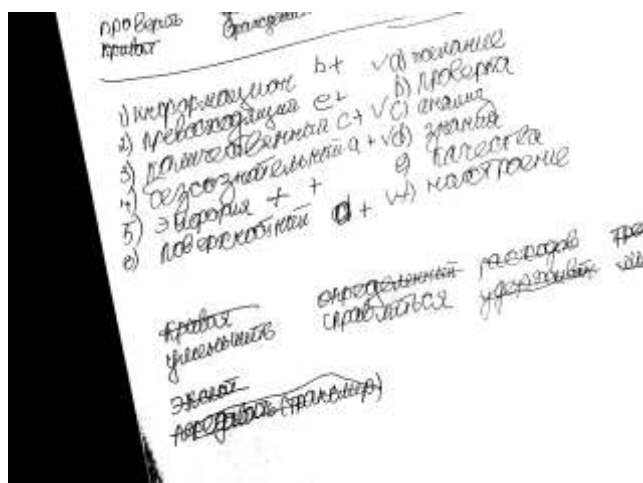


Рис.9. Результат бинаризации модификацией метода Оцу.

Для изображения 3 лучшим методом бинаризации стала модификация метода Оцу для несбалансированных классов. Мера F1 равна 0,95.

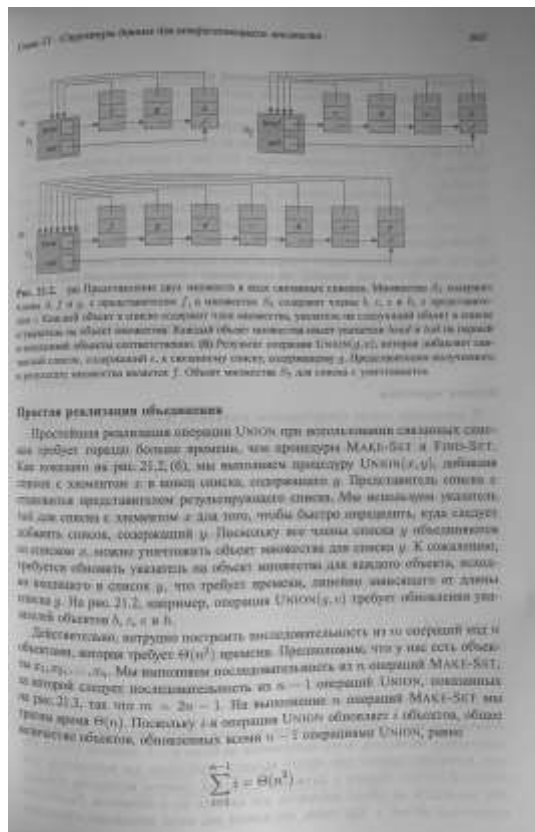


Рис.10. Изображение 4 - страница текста с рисунком.

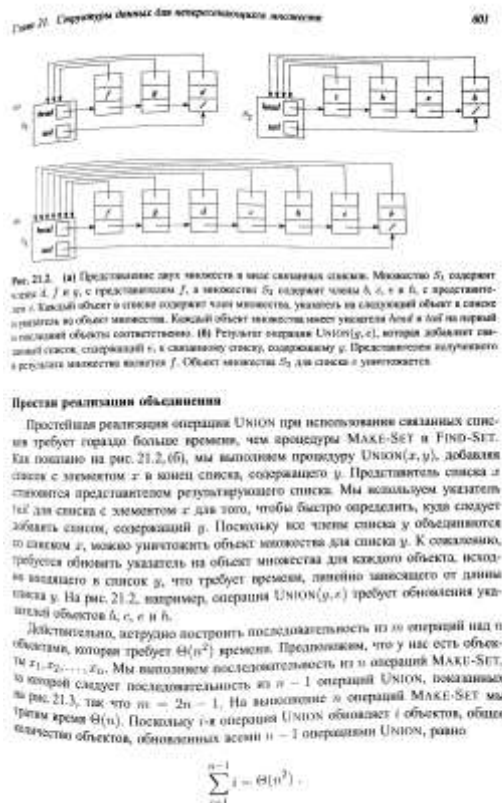


Рис.11. Маска изображения 4.

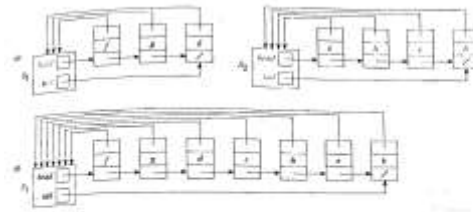


Рис. 21.2. (а) Представление двух множеств в виде связанных списков. Множество  $S_1$  содержит элементы  $a, b, c, d, e$ , с представителем  $f$ , а множество  $S_2$  содержит элементы  $b, c, e, f$ , с представителем  $a$ . Каждый элемент в списке содержит адрес множества, указатель на следующий объект в списке и указатель на объект множества. Каждый объект множества имеет указатель назад и тем самым на первый объект множества. (б) Результат операции  $Union(x, y)$ , которая добавляет элементы списка, содержащего  $e$ , в указанный список, содержащий  $y$ . Представителем полученного результата множества является  $f$ . Объект множества  $S_2$  для списка  $x$  уничтожается.

#### Простая реализация объединения

Простейшая реализация операции  $Union$  при использовании связанных списков требует гораздо больше времени, чем процедуры  $Make-Set$  и  $Find-Set$ . Как показано на рис. 21.2, (б), мы выполняем процедуру  $Union(x, y)$ , добавляя список с элементами  $x$  в конец списка, содержащего  $y$ . Представитель списка  $x$  становится представителем результирующего списка. Мы используем указатель  $tail$  для списка с элементом  $x$  для того, чтобы быстро определить, куда следует добавлять список, содержащий  $y$ . Поскольку все члены списка  $y$  объединяются со списком  $x$ , можно уничтожить объект множества для списка  $y$ . К сожалению, требуется обновить указатель на объект множества для каждого объекта, входящего в список  $y$ , что требует времени, линейно зависящего от длины списка  $y$ . На рис. 21.2, например, операция  $Union(y, x)$  требует обновления указателей объектов  $b, c, e$  и  $f$ .

Действительно, нетрудно построить последовательность из  $n$  операций над  $n$  объектами, которая требует  $\Theta(n^2)$  времени. Предположим, что у нас есть объекты  $x_1, x_2, \dots, x_n$ . Мы выполним последовательность из  $n-1$  операций  $Union$ , показанных на рис. 21.3, так что  $m = 2n - 1$ . На выполнение  $n$  операций  $Make-Set$  мы тратим время  $\Theta(n)$ . Поскольку  $i$ -я операция  $Union$  объединяет  $i$  объектов, общее количество объектов, объединенных всеми  $n-1$  операциями  $Union$ , равно

$$\sum_{i=1}^{n-1} i = \Theta(n^2).$$

Рис.12. Результат бинаризации многомасштабной модификацией метода Ниблэка.

Для изображения 4 лучшим методом бинаризации стала модификация метода Ниблэка с параметрами: размер крыла ядра – 28, порог увеличения ядра – 0,  $k=-2.8$ ,  $a=-8.1$ . Мера F1 равна 0,937.

N	Вариант	АПВ	Правильность	ПРП	Время	Ф1
1.	Средний	С-Т	5.0	7.0	2	+
2.	Короткий	С-Т	2.5	3.3	2	+
3.	Длинный	Орг	0.9	1.0	3	+
4.	Короткий	Орг	0.01	0.01	4	+
5.	Длинный	С-Т	0.02	0.03	2	+

В час вышло на проверку 0.01 10.04-1  
 ⇒ Всп. не удалось из-за ошибки ПРП  
 5.77 + 2.5/5.3 + 0.02/10.03 \* 2.139 > 1 ⇒  
 Всп. не удалось из-за ошибки ПРП  
 Выбор: Как старое качество из-за ошибки  
 проверит результат ПРП, но качество не  
 будет. Качество не будет 1 и 2 сильно  
 превышает 1, а следовательно Всп. не  
 удастся из-за ошибки проверки  
 качества. В итоге > 10.04.1  
 и будет проверка.

Рис.13. Изображение 5 – рукописный текст на тетрадном листе.

N	Вариант	ЛПВ	Факт. разн.	ПРК	Класс	Вывод
1.	Грозный	С-Т.	5.0	7.0	2	+
2.	Вирота	С-Т.	2.5	3.3	2	+
3.	Шко	Орг.	0.9	1.0	3	+
4.	Картинки	Орг.	0.01	0.01	4	+
5.	Шко	С-Т.	0.02	0.03	2	+

В нас видно по картинке 0.01 / 0.01 = 1

→ Всп. не проходит по правилам ПРК

$$5/7 + 2.5/3.3 + 0.02/0.03 \approx 2.139 > 1 \Rightarrow$$

Всп. не проходит по правилам ПРК

Вывод: Хотя структура каждого из вариантов проходит правило ПРК, но суммарное количество веществ класса 01 и 2 сильно превышает 1, а следовательно Всп. не проходит для использования картины в целом ⇒ не отнесена к первой категории безопасности

Рис.14. Маска изображения 5.

N	Вариант	ЛПВ	Факт. разн.	ПРК	Класс	Вывод
1.	Грозный	С-Т.	5.0	7.0	2	+
2.	Вирота	С-Т.	2.5	3.3	2	+
3.	Шко	Орг.	0.9	1.0	3	+
4.	Картинки	Орг.	0.01	0.01	4	+
5.	Шко	С-Т.	0.02	0.03	2	+

В нас видно по картинке 0.01 / 0.01 = 1

→ Всп. не проходит по правилам ПРК

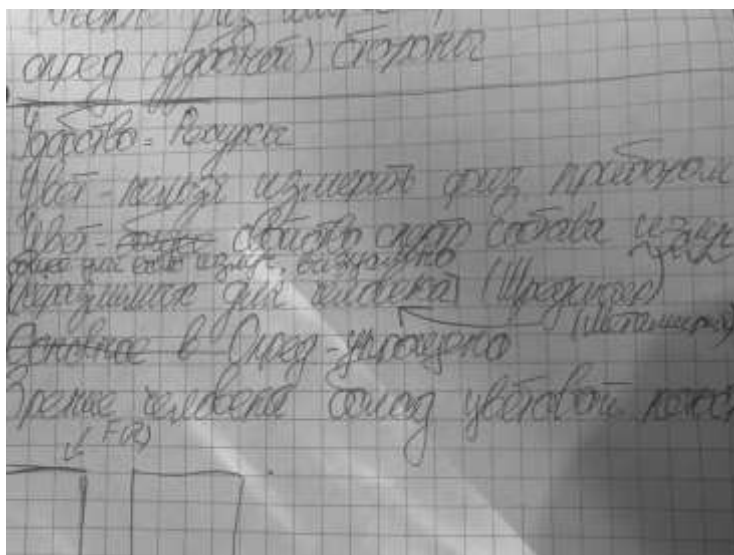
$$5/7 + 2.5/3.3 + 0.02/0.03 \approx 2.139 > 1 \Rightarrow$$

Всп. не проходит по правилам ПРК

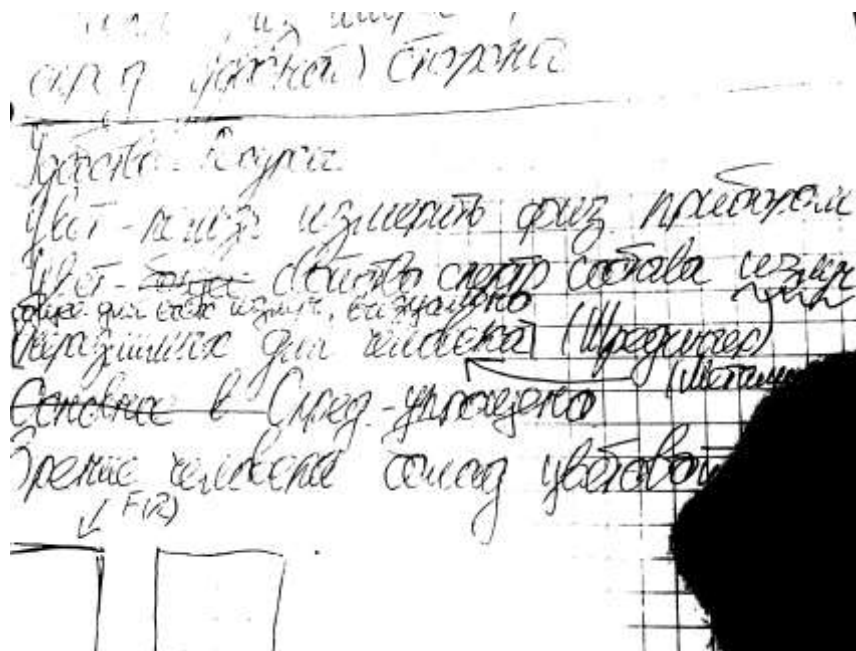
Вывод: Хотя структура каждого из вариантов проходит правило ПРК, но суммарное количество веществ класса 01 и 2 сильно превышает 1, а следовательно Всп. не проходит для использования картины в целом ⇒ не отнесена к первой категории безопасности

Рис.15. Результат бинаризации модификацией метода Ниблэка.

Для изображения 3 не было смысла использовать модификацию метода Оцу для неравновесного случая, так как на изображении видно, что количество пикселей текста(объектов) сопоставимо с количеством пикселей фона.



Для рисунка 6 нет смысла использовать Методы Оцу, так как они дают погрешность, если на изображении яркость фона меняется. Результат бинаризации изображения 6 методом Оцу представлен на рисунке 17.



Все представленные изображения находятся в директории вместе с кодом.

### **Заключение.**

В результате работы были изучены реализованы 4 метода бинаризации изображений, функция оценки качества бинаризации, а также был проведен эксперимент с бинаризацией 5 различных документов.

### **Список использованных источников.**

1. Otsu N. A threshold selection method from gray-level histograms // IEEE transactions on systems, man, and cybernetics. — 1979. — Vol. 9, no. 1. — P. 62–66.
2. Kurita T., Otsu N., Abdelmalek N. Maximum likelihood thresholding based on population mixture models // Pattern recognition. — 1992. — Vol. 25, no. 10. — P. 1231–1240.