

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Центр післядипломної освіти
(повна назва)

Кафедра програмної інженерії
(повна назва)

АТЕСТАЦІЙНА РОБОТА
Пояснювальна записка

другий (магістерський)
(рівень вищої освіти)

Дослідження методів ідентифікації
при розпізнаванні візуальних образів
(тема)

Виконав: студент 2 курсу, групи ІІЗмзд-18-1

Кірсанов А.С.
(прізвище, ініціали)

спеціальності 121 – Інженерія програмного забезпечення
(код і повна назва спеціальності)

Освітньо-наукової програми
(тип програми)

Інженерія програмного забезпечення
(повна назва освітньої програми)

Керівник доц. Назаров О.С.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____
(підпис)

Дудар З.В.
(прізвище, ініціали)

2020 р.

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ РАДІОЕЛЕКТРОНІКИ

Факультет _____ Центр післядипломної освіти _____

Кафедра _____ програмної інженерії _____

Рівень вищої освіти _____ другий (магістерський) _____

Спеціальність _____ 121 – Інженерія програмного забезпечення _____
(код і повна назва)

Тип програми _____ освітньо-наукова програма _____

Освітня програма _____ Інженерія програмного забезпечення _____

ЗАТВЕРДЖУЮ:

Зав. кафедри _____

(підпис)

« _____ » _____ 20 ____ р.

**ЗАВДАННЯ
НА АТЕСТАЦІЙНУ РОБОТУ**студентові _____ Кірсанов Андрій Сергійович _____
(прізвище, ім'я, по батькові)1. Тема роботи Дослідження методів ідентифікації при розпізнаванні візуальних образів

затверджена наказом університету від _____ 20__ р. № _____

2. Термін подання студентом роботи до екзаменаційної комісії
_____ 20__ р.3. Вихідні дані до роботи електронні ресурси за обраною тематикою, план дослідження, методи аналізу, алгоритми ідентифікації візуальних образів, платформа розробки Java/Android, Matlab4. Перелік питань, що потрібно опрацювати в роботі мета роботи, аналіз проблемної галузі і постановка задачі, огляд методів ідентифікації при розпізнаванні візуальних образів, методи ідентифікації при розпізнаванні візуальних образів, аналіз результатів дослідження

5 Консультанти розділів роботи

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата
Дослідження існуючих методів	доц. Назаров О.С.		

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка *
1	Аналіз проблемної галузі	10.02.2020	
2	Постановка задачі	13.02.2020	
3	Дослідження існуючих методів	28.03.2020	
4	Моделювання задачі	10.04.2020	
5	Підготовка бази	17.04.2020	
6	Проведення досліджень	28.04.2020	
7	Коригування досліджень	4.05.2020	
8	Програмна реалізація	8.05.2020	
9	Аналіз отриманих результатів	11.05.2020	
10	Підготовка пояснювальної записки	13.05.2020	
11	Підготовка презентації та доповіді	14.05.2020	
12	Нормоконтроль, рецензування	26.05.2020	
13	Занесення роботи до електронного архіву	26.05.2020	
14	Попередній захист	26.05.2020	
15	Допуск до захисту у зав.кафедри	26.05.2020	
* заповнюється вручну після виконання чергового пункту			

Дата видачі завдання _____ 2020 р.

Студент _____ Кірсанов Андрій Сергійович
(підпис)

Керівник роботи _____ доц. Назаров О.С.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ / ABSTRACT

Атестаційна робота магістра містить: 68 с., 21 рис., 2 додатки, 31 джерел.

ГЛУБОКЕ НАВЧАННЯ, МАШИННЕ НАВЧАННЯ, НЕЙРОННІ МЕРЕЖІ, РОЗПІЗНАВАННЯ ВІЗУАЛЬНИХ ОБРАЗІВ, ФІЛЬТРАЦІЯ, ШТУЧНИЙ ІНТЕЛЕКТ, ANDROID, JAVA

Метою роботи є дослідження впливу різних методів цифрової фільтрації візуальних зображень на якість роботи нейронної мережі для отримання більш якісних результатів її роботи[1].

В ході роботи були досліджені цифрові фільтри Вінера, Гауса, підвищення контрасту і їх вплив при обробці вхідних даних на роботу нейронної мережі.

В результаті роботи розглянуті методи фільтрації візуальних зображень, досліджені на практиці вплив деяких фільтрів на роботу нейронної мережі. Встановлено, що фільтр Гауса має позитивний вплив на її роботу. Результати роботи можуть бути використані в розробці програмного комплексу з автоматичного підбору пакету фільтрів для поліпшення роботи окремо взятої нейронної мережі.

ANDROID, ARTIFICIAL INTELLIGENCE, DEEP LEARNING, FILTERING, MACHINE LEARNING, NEURAL NETWORKS, RECOGNIZED VIZUAL IMAGES

The aim of the work is to study the influence of different methods of digital filtering of visual images on the quality of the neural network to obtain better results of its work.

In the course of the work, Wiener and Gauss digital filters were investigated, increasing the contrast and their influence when processing input data on the operation of a neural network.

As a result of the work, methods for filtering visual images are examined, and the effect of some filters on the operation of a neural network is investigated in practice. It is established that the Gaussian filter has a positive effect on its operation. The results of the work can be used in the development of a software package for automatic selection of a filter package to improve the operation of a single neural network.

ЗМІСТ

Вступ	6
1 Аналіз проблемної області	8
1.1 Аналіз предметної області	8
1.2 Алгоритми попередньої обробки зображень	11
1.3 Постановка задачі	19
1.4 Методи поліпшення зображень.....	20
2 Опис проведених досліджень	35
2.1 Формулювання задачі	35
2.2 Результати обробки зображень без попередньої обробки	37
2.3 Результати обробки зображень з попередньою обробкою	43
2.4 Аналіз результатів	59
Висновки.....	50
Перелік посилань.....	52
Додаток А Лістинг коду.....	55
Додаток Б Слайди презентації	60

ВСТУП

Розпізнавання візуальних образів за допомогою нейронних мереж в даний час застосовується в багатьох сферах нашого життя.[2] Навчання нейронної мережі з нуля займає величезну кількість тимчасових і фінансових ресурсів. Це пов'язано з необхідністю наявності великого обсягу бази навчальних зображень і її ручної класифікації. Також саме навчання має проводитися на потужній апаратурі або на хмарних ресурсах. І перше, і друге вимагає фінансових витрат. Поліпшити ситуацію може розробка методики використання вже існуючих навчених нейронних мереж для виконання поставлених завдань.

Часто буває, що в наявності є навчена нейронна мережа для виконання певного завдання. При цьому можливий варіант, що з її допомогою необхідно виконати завдання, які трохи відрізняються від первісної. Також можливий варіант, коли вже в процесі використання виявляється, що нейронна мережа недостатньо навчена для виконання своєї поставленого завдання (наприклад, через малий обсяг навчальної бази). При цьому перевчити або до навчити нейро-мережу вже немає можливості через відсутність часу, фінансів або просто відсутність початкових кодів. У цьому випадку може допомогти попередня обробка вихідного зображення методами цифрової фільтрації.

Метою роботи є дослідження впливу різних методів цифрової фільтрації візуальних зображень на якість роботи нейронної мережі, для отримання більш якісних результатів її роботи. Для попередньої обробки вихідних зображень необхідно вибрати кілька цифрових фільтрів, які за попередньою оцінкою будуть впливати на якість обробки. Змінюючи параметри обраних фільтрів необхідно знайти такі значення, які відіб'ються на результатах позитивним чином.

Розглянута методика попередньої обробки вихідних зображень дозволить використовувати наявні в наявності навчені нейронні мережі для виконання поряд поставлених завдань. Результати роботи можуть бути використані в розробці

програмного комплексу з автоматичного підбору пакету фільтрів для поліпшення роботи окремо взятої нейронної мережі.

1 АНАЛІЗ ПРОБЛЕМНОЇ ОБЛАСТІ

1.1 Аналіз предметної області

Розпізнавання образів - область науки, присвячена розробці принципів, побудови програмних і (або) апаратних систем, призначенням яких є визначення того, якому з заздалегідь заданих класів належить даний об'єкт. Об'єктом може бути будь-який сигнал, явище, реально існуючий предмет або абстракція і інше.

Об'єкт визначається кінцевою послідовністю ознак або властивостей. Ознаками називаються дані, отримані в результаті вимірювання об'єкта, для того, щоб використовувати їх в процесі класифікації. Безліч ознак одного об'єкта утворює так званий вектор ознак. Цей вектор ознак може бути визначений як точка в гіперпросторі ознак. Кожному способу підтверджено тільки один вектор ознак, також як і вектору ознак зіставлений тільки один образ. Співвіднесення об'єкта одному з класів називається класифікацією. Класифікація виконується із застосуванням еталонних об'єктів - об'єктів, клас яких вже відомий.

Образ – опис об'єкта, що використовується для вирішення задачі розпізнавання образів, співвіднесення об'єкта одному з класів. Для того, щоб було можливо вирішити задачу розпізнавання образів, необхідно що всі можливі образи об'єктів розбивалися на кінцеве безліч класів. Також має бути задано безліч еталонних об'єктів. Кожен еталонний об'єкт належить одному з класів. Існує також об'єкт, приналежність якого якого-небудь класу невідома. Завдання розпізнавання образів полягає в тому, щоб встановити, якого класу належить цей об'єкт. Або розробити алгоритм, який виконує це для будь-якого даного об'єкту. Чим цей алгоритм точніше визначає приналежність об'єкта класу, тим точніше система розпізнавання образів.[3]

Принципи побудови систем розпізнавання образів:

- принцип перерахування членів класу. Виконується послідовне порівняння образу даного об'єкту з усіма образами еталонних об'єктів. Об'єкт відносять класу

еталонного об'єкта, образ якого найбільше збігається з образом пред'являється. Наприклад, завдання розпізнавання букв одного алфавіту, надрукованих за допомогою різних шрифтів. Принцип легко реалізуємо, однак обчислювально складний для великого числа об'єктів, ознак одного об'єкта, класів.

- принципи схожості властивостей. Для кожного класу визначаються ознаки, характерні тільки для всіх еталонних об'єктів класу. Новий об'єкт буде віднесений класу з сукупністю ознак, найбільш подібних ознакам аналізованого об'єкта. У деяких випадках буває складно знайти весь набір загальних ознак, що характеризують об'єкти класу.

- принцип кластеризації. Кластер-група образів, які формують в просторі ознак компакту область. Образи розглядаються як точки в гіперпросторі ознак. Кластер включає в себе всі еталонні об'єкти одного класу. Образи одного кластера розташовані поблизу один від одного і в видаленні від інших образів. Новий об'єкт буде віднесений до одного класу за принципом найближчого сусіда. Виконується пошук еталонного образу, ближче всіх розташованого до даного об'єкту. Об'єкт буде віднесений до класу кластера цього найближчого образу. Принцип не завжди реалізуємо: образи одного класу можуть не утворювати кластер, всередині кластера можуть перебувати образи іншого класу.[4]

Основні під задачі проектування системи розпізнавання образів:

- отримання векторів ознак - всієї вимірюваної інформації про разі. У деяких випадках розглядають вектора як точки в багатовимірному просторі ознак.

- зниження розмірності векторів ознак. Повинні залишитися лише ознаки, що характеризують відмінності між класами.

- знаходження оптимальної вирішальною процедури (функції, правила). Вирішальна процедура $d(x)$ відносить будь-який вектор ознак x одному з класів. Вирішальну процедуру називають також класифікатором. Оптимальна вирішальна процедура $d^0(x)$ дає найменшу ймовірність помилкового отримання класу для всіх векторів ознак. Якщо перед побудовою системи розпізнавання відомі кордону класів в багатовимірному просторі ознак, оптимальна вирішальна процедура може бути точно визначена. Якщо ж заздалегідь кордону класів невідомі, то виконується

поетапна коригування оптимальної вирішальною процедури. Для цієї коригування використовується додаткова інформація про вимірних образах.

- чисельна оцінка якості роботи системи розпізнавання. Якість роботи може бути оцінений з точки зору точності призначення об'єкта класу.

Розрізняють розпізнавання з навчанням і без навчання. Якщо розпізнавання виконується із застосуванням наявного безлічі еталонних об'єктів, то завдання розпізнавання називають класифікацією з навчанням.

Якщо існує деякий набір векторів ознак, витягнутих з набору експериментально досліджуваних об'єктів, але невідомо, якого класу належить кожен із векторів ознак, то процедура розпізнавання називається класифікацією без навчання. В цьому випадку залишається невирішеною проблема розбиття всього набору об'єктів за принципом близькості їх розташування в гіперпросторі ознак. Така проблема називається кластеризацією.

Основні методи побудови систем розпізнавання образів:

- евристичний. Для віднесення об'єкта одному з класів використовується набір специфічних процедур, коректне працюють тільки для конкретної задачі розпізнавання. Наприклад, розпізнавання друкованих букв з набору штрихів, які характеризують букву.

- математичний. Вирішальна процедура задається з математичного формалізму. Існують детерміновані і статистичні методи. Детерміністські не використовують статистичні властивості аналізованих об'єктів. Статистичні вирішують задачу розпізнавання в термінах статистики.

- лінгвістичний. В рамках лінгвістичного підходу, образ - ієрархічна структура подобразів. Задано безліч найпростіших елементів, що застосовуються при описі об'єкта, а також правила композиції цих елементів. Ці правила називаються граматиною мови опису об'єктів. Таким про разом, розпізнавання полягає в розпізнаванні найпростіших елементів і наступний синтаксичний розбір опису. Наприклад, опис друкованого знака через послідовність елементарних векторів одиничної довжини.[5]

1.2 Алгоритми попередньої обробки зображень.

Передобробка включає: вирівнювання загального яркостного фону зображення, ліквідацію на оригінальному документі високочастотних перешкод, контрастування, бінарного та інших функціональних перетворень. Вихідна опис вихідного інформаційного образу має бути максимально пристосована для передачі і аналізу (прийняття рішення на основі цієї інформації). Однак, одним з найбільш важливих показників при всіх цих перетвореннях повинно бути збереження максимального числа елементів вихідної інформації.[6]

Термін просторова область належить до площини зображення як такої, и это понятие объединяет методы, базирующиеся на операциями с пикселями изображения. Способы обработки в частотной области базируются на преобразовании сигнала, формованого шляхом застосування до зображення перетворення Фур'є. Часто бувають корисними технології, які є синтезом цих двох категорій. На жаль не існує загальної методики поліпшення зображень, візуальне сприйняття результатів є індикатором того, наскільки добре діє конкретний метод. Візуальна оцінка якості зображення це суб'єктивний процес. Коли метою є обробка зображення для подальшого опрацювання комп'ютером, критерій оцінювання трішки простіше. Зокрема, при розпізнавання знаків найліпшим буде той спосіб обробки зображень, який дає більш точні результати розпізнавання. Тим не проте, навіть в обстановці, коли питання допускає поставити тверді мірила добротності, зазвичай необхідна якась кількість дослідів, поки буде обраний конкретний підхід до поліпшення зображень

Існують два основних види уявлень зображень - векторне і растрове. У векторному поданні зображення описується набором ліній (векторів), які містять координати початкових і кінцевих точок, кривизну ліній і іншими геометричними характеристиками, описуються також правила побудови різних областей і колірні характеристики.

Іншими словами для растрового представлення необхідне формування деякої математичної моделі. Тому векторне подання використовуються, в основному, при вирішенні задач синтезу зображень. Хоча деякі алгоритми розпізнавання зображень для своєї роботи вимагають саме векторне подання, яке необхідно отримати з вихідного зображення.

Растрове зображення являє собою одну або кілька матриць, що описують просторове розподіл характеристик зображення на деякій декартовій координатній сітці. У цьому випадку зображення будується з безлічі точок і має структуру растра.

Основним елементом растрового представлення зображення є піксель (скорочення від словосполучення «picture elements» - елементи зображення), що має координати в растровій системі координат і деякі атрибути (колір, яскравість, прозорість і т.п.). Число пікселів по координатам X і Y (по горизонталі і вертикалі) задає дозвіл (розмірність) представлення зображення. Колір пікселя задається глибиною - кількістю бітів, необхідне для завдання будь-якого кольору.

Класифікація алгоритмів попередньої обробки

Алгоритми попередньої обробки зображень підрозділяються на різні групи в залежності від класифікує ознаки. Всі алгоритми попередньої обробки або повинні покращувати в якому - то сенсі якість зображень, або перетворювати його до виду, найбільш зручному для подальшої обробки.

Алгоритми спрямовані на поліпшення колірної передачі зображення називаються алгоритмами корекції. У цю групу входять також алгоритми працюють з напівтоновими зображеннями змінюють їх яскравості і контрастні характеристики. Алгоритми спрямовані на обробку просторових характеристик зображень називаються алгоритмами просторової фільтрації. До цієї групи належать алгоритми придушення перешкод, алгоритми просторового згладжування і алгоритми просторового посилення, алгоритми придушення і посилення просторових частот. Алгоритми виконують геометричні операції з зображення називаються алгоритмами геометричній обробки. До них відносяться:

- кадріроване зображення - виділення з вихідного зображення деякої частини прямокутної форми;

- зміна розмірів зображення. Ці алгоритми застосовують різні методи інтерполяції, що дозволяють або коректно заповнити відсутні пікселі в збільшеному зображенні, або перерахувати значення пікселів при зменшенні зображення

- поворот зображення. Ці алгоритми здійснюють поворот вихідного зображення на заданий кут, коректно перераховуючи значення пікселів використовуючи різні методи інтерполяції.

Алгоритми виконують перетворення з однієї колірної системи в іншу називаються алгоритмами кольорового перетворення. До них відносяться також алгоритми перетворення кольорових зображень в напівтонові і алгоритми бінаризації, що переводять вихідне зображення в бінарне.

Алгоритми виділяють на оригінальному документі деякі області за різними, часто неформальним умов називаються алгоритмами сегментації. Прикладом такого алгоритму може наприклад служити алгоритм, який повинен виділяти на зображенні документа області тексту та графічної інформації або алгоритм виділяє в зображенні тексту області, що відносяться до окремих слів.

$$I_f(i, j) = \sum_{m=N_{11}}^{N_{21}} \sum_{n=N_{12}}^{N_{22}} I_m(im, jn)h(m, n)$$

де I_f – матриця відфільтрованого зображення,

N_{11} - нижня межа стовпців імпульсної характеристики,

N_{21} - верхня межа стовпців імпульсної характеристики,

N_{12} - ліва межа кордон рядів імпульсної характеристики,

N_{22} - права межа кордон рядів імпульсної характеристики,

I_m - матриця вихідного зображення,

h - матриця імпульсної характеристики фільтра.

Фільтри просторового диференціювання

Просторове диференціювання зображень призначений для виділення областей, що містять різкі переходи по яскравості чи кольоровості і придушення областей з приблизно однорідної інтенсивністю і кольоровістю. Іншими словами просторове диференціювання можна розглядати як обробку зображення за допомогою фільтра високих частот. Фільтри високих частот можна проектувати по різному, наприклад за допомогою класичних методів. Однак якщо їх використовувати для вирішення задачі знаходження контурів на зображеннях, то зазвичай сходять з досить простих міркувань, що ґрунтуються на методах обчислення градієнта.

Здійснення операцій диференціювання зображень за допомогою фільтрації є універсальним методом, що дозволяє одноманітно реалізовувати різні алгоритми і оператори диференціювання. Для цього достатньо реалізувати алгоритм обчислення згортки, а самі зображення і різні оператори передавати в цей алгоритм в якості параметрів .

Просторове згладжування

Просторове згладжування застосовується для зменшення впливу шуму на зображення. Основна ідея просторового згладжування - заміна значення пікселя деяким «середнім» значенням, який обчислюється за сусіднім пікселям знаходяться в околиці оброблюваного. Слід зазначити, що при будь-якому усередненні відбувається розмиття деталей і погіршення чіткості.

Операцію просторового згладжування можна розглядати як здійснення цифрової фільтрації за допомогою фільтра низьких частот.

Найпростіший згладжуваний фільтр називається усереднюючим. Кожен елемент імпульсної характеристики цього фільтра дорівнює $1 / (m * n)$, де m і n - розміри вікна фільтра. Найбільш поширені фільтри з вікном 3×3 . Для цього фільтра імпульсна характеристика буде

$$|1/9, 1/9, 1/9 \quad 1/9, 1/9, 1/9 \quad 1/9, 1/9, 1/9|.$$

Фільтр Гауса також є фільтром нижніх частот, але в порівнянні з усереднюючим фільтром він менше розмиває зображення. Центральний елемент маски цього фільтра має максимальне значення, відповідне піку розподілу Гауса (нормального розподілу).

Загальні принципи шумозаглушення в задачах обробки цифрових зображень
Алгоритми шумозаглушення (фільтрації шумів) зазвичай спеціалізуються на придушенні якогось конкретного виду шуму [15]. Багато шумів можна досить точно описати моделлю білого Гауссова шуму, тому більшість алгоритмів фільтрації орієнтоване на його придушення. До найбільш поширених методів фільтрації шумів відносяться згладжують фільтри, фільтри Вінера, медіанний фільтри і ранжирующие [16] фільтри.

При вирішенні задач придушення шуму Гауса можна застосовувати нелінійні і лінійні фільтри. Лінійний фільтр визначається вещественнозначной функцією (ядром фільтра), заданої на растрі. У разі лінійної згладжує фільтрації значення інтенсивності усередняється за певною масці в кожній точці В ході фільтрації яскравість кожної точки зображення в цифровому вигляді замінюються значенням яскравості, яке вважається не побите перешкодою.

Розділяють частотну і просторову фільтрацію. Частотні методи базуються на використанні перетворення Фур'є. Такий алгоритм допускає обробляти функцію в частотній області, а потім без втрати інформації повернутися до первинного стану. Просторові методи поліпшення зображень, як правило, застосовуються до зображень, представлених як двовимірні матриці.

Метод просторових алгоритмів засновується в застосуванні віконних операторів (масок) до кожної точки вихідного зображення. Методи поліпшення, коли в якості маски використовується невеликий двовимірний масив, називають фільтрацією по масці.

При здійсненні лінійної фільтрації відгук маски задається сумою творів пікселів в області покриття фільтра. Як лінійного фільтра, що згладжує використовується усереднюючий фільтр, результатом якого є середнє значення по

околиці маски фільтра. Операції, що виконуються нелінійним просторовим фільтром, залежать від значінь частин матриці зображення, котрі знаходяться в аналізованій околиці.

При медіанній фільтрації значення пікселів представляють собою усереднене значення точок відповідної околиці. Часто при усуненні шуму більш ефективним є медіанний фільтр, ніж усереднюючий, тому що призводить до менших спотворень кордонів виділених об'єктів. Як маски при медіанній фільтрації використовується двовимірне вікно з центральною симетрією, при цьому центр знаходиться в поточній точці фільтрації. В основу адаптивної фільтрації покладено винеровський фільтр. Він є одним з типів лінійного фільтра для локальної адаптивної фільтрації зображення. Перевага адаптивного фільтра полягає ще й в тому, що він зберігає краю і інші фрагменти об'єктів зображення.

У той же час винеровський фільтр вимагає більшого часу для обчислень, ніж лінійний фільтр. Середнє значення яскравості обчислюється для центрального елемента маски, який містить значення яскравості оригінального зображення в покритої маскою локальної області зображення. Найпростіша ідея видалення шуму - усереднювати значення пікселів в просторової околиці. Для кожного елемента зображення аналізуються сусідні для нього пікселі, які розташовуються в прямокутному вікні навколо цього пікселя.

Медіанна фільтрація - стандартний спосіб придушення імпульсного шуму. Для кожного пікселя в деякій його околиці (вікні) визначається медіанне значення. Однак в чистому вигляді дрібні деталі розмиваються медіанним фільтром, величина яких менше розміру вікна пошуку медіани.

Метод головних компонент (МГК) дозволяє виокремити структуру в багатовимірному масиві даних. МГК застосовується в основному для стиснення або розпізнавання зображень з білим гаусовским шумом і заснований на знаходженні таких базисних векторів досліджуваного простору, які б найкращим чином відображали розташування вихідних даних в цьому просторі.

При цьому всі зображення розбивається спочатку на блоки, які обробляються незалежно і повинні розташовуватися з невеличким перекриттям, щоб ухилитись

«артефактів блочності» при їх стикуванні. МГК добре працює на текстурованих областях (після придушення шуму текстура залишиться як і раніше чіткої).

Остаточний результат аналізу багато в чому визначається якістю сегментації зображення, а ступінь деталізації виділяються характеристик залежить від конкретного завдання. В даний час не існує універсального методу або алгоритму для вирішення всіх типів завдань сегментації. У більшості випадків вибирається один чи кілька алгоритмів, які переробляються під специфічні умови задачі. Сегментація вирішує в загальному дві основні задачі: поділ зображення на частини з метою здійснення подальшого локального аналізу та зміна форми опису часток зображення, що дозволяє уявити точки як високо рівневі структури, що забезпечують ефективність подальшого аналізу зображення.

Принцип роботи методів сегментації цифрових зображень заснований на різниці яскравості елементів і фону зображення. Математичне обґрунтування цих методів полягає в обчисленні похідних, представлених для цифрових зображень у вигляді дискретних наближень градієнта.

Серед градієнтних методів виділяють оператори, що представляють собою матричні маски: перехресний оператор Робертса, оператор Превітта, оператор Собеля і оператор Канни. Розглянемо деякі особливості цих операторів.

Оператор Канни дозволяє, як правило, забезпечити: гарне виявлення (підвищення відносини сигнал / шум); хорошу локалізацію (правильне визначення положення кордону); єдиний відгук на одну кордон. Для зменшення чутливості алгоритмів сегментації до шуму застосовується фільтрація по першій похідній, кілька розмиває вихідне зображення (після обчисленням градієнта згладженого зображення в контурі кордону залишаються тільки точки максимуму градієнта зображення). Далі використовується дані про напрямлення кордону, щоб видаляти точки саме поруч з кордоном і не розривати кордон поблизу локальних максимумів градієнта.

Для визначення напрямку градієнта можна використовувати оператор Собеля, округлюються отримані значення напрямків до одного з чотирьох кутів (0, 45, 90 і 135 градусів) і видаляє за допомогою порогів слабкі кордону.

Диференціальні оператори більш високого порядку (наприклад оператор Лапласа) можна застосувати для вирішенню завдання відлення перепадів яскравості

$$\Delta^2 f(x_1, x_2) = \frac{\delta^2 f(x_1, x_2)}{\delta x_1^2} + \frac{\delta^2 f(x_1, x_2)}{\delta x_2^2}$$

Даний оператор виявляє кордону в місцях зміни знака похідної функції яскравості. В задачах сегментації роль Лапласіан зводиться до використання властивості перетину нульового рівня для виділення контуру і з'ясування, на темній або на світлій стороні контуру знаходиться розглянутий піксель $(x; y)$. Основним недоліком Лапласіан є висока чутливість до шумів і можливість появи розривів в контурі. До його переваг можна віднести те, що він нечутливий до орієнтації кордонів областей і має малу ресурсомісткість. Для зменшення впливу шумів часто використовують лапласіан в поєднанні зі згладжуванням за методом Гауса. Таке поєднання називають оператором лапласіан гауссіана (Laplacian of Gaussian - LoG). До його переваг можна віднести слабку чутливість до шумів і орієнтації кордонів областей, що дозволяє чітко виділяти зовнішній і внутрішні контури об'єкта.

Для обчислення дискретного градієнта можна застосовувати оператор Робертса. Оскільки для обчислення модуля градієнта можна використовувати різниці в будь-яких двох взаємно перпендикулярних напрямках, в операторі Робертса беруться діагональні різниці. До недоліків цього оператора відносяться висока чутливість до шумів і орієнтації кордонів областей, а також можливість появи розривів в контурі і відсутність явно вираженого центрального елемента. Оператор Робертса простий в реалізації і володіє високою швидкістю, однак сильно чутливий до перешкод.

На практиці для обчислення дискретних градієнтів найчастіше використовуються оператори Превітта і Собеля, що дають нульовий відгук на областях постійної яскравості. При наявності центрального елемента і малої ресурсоемності цим операторам властива висока чутливість до шумів і орієнтації кордонів областей, а також можливість появи розривів в контурі. Аналіз методів

фільтрації показує, що для імпульсних перешкод найбільш підходить медіанний фільтр, добре зберігає кордону елементів і володіє високою швидкістю. Градієнтні алгоритми виділення контурів елементів зображень прості в програмній реалізації, однак результати, одержані за їх допомогою, залежать від якості досліджуваних зображень.

1.3 Постановка задачі

Необхідно провести дослідження методів цифрової фільтрації візуальних зображень на предмет їх впливу на результати роботи нашої нейронної мережі.[17] У нас в наявності є нейронна мережа, навчена виділяти передній план на селф-фотографіях. Селфі-фотографія - це зображення, що містить одного або кількох людей і отримане в результаті зйомки на відстані близько одного метра від об'єктива камери. Передній план, в нашому випадку, це люди і, що знаходяться в безпосередньому контакті з людьми, предмети.

Наше завдання вибрати такі цифрові фільтри, які дозволять впевнено виділяти передній план на фотографіях, зроблених на відстані 2-3 метри від об'єктива камери. Акцент у виборі варто зробити на фільтрах, які мають властивості корекції якості і шуму, так як логічно припустити, що погіршення якості виділення переднього плану пов'язано з видаленням об'єктів від об'єктива камери.

Дослідження має включати в себе теоретичне вивчення методів, експериментальне застосування методів й аналіз результатів.

1.4. Методи поліпшення зображень

Під методами поліпшення зображення мається на увазі здійснення таких перетворень над вихідним зображенням, які призводять до отримання результату, більш підходящого з точки зору конкретного застосування. Візуальне оцінювання якості зображення являє собою вкрай суб'єктивний процес. В тому випадку, коли метою обробки зображення є їх подальше використання в системах машинного сприйняття, критерієм ефективності обробки вихідного зображення є отримання більш точних результатів машинного розпізнавання. Безліч підходів до поліпшення зображення розпадається на 2 категорії: методи обробки в просторовій області і методи обробки в частотній області. термін просторова область належить до площини зображення як такої, і дана категорія об'єднує підходи, засновані на прямому маніпулюванні пікселями зображення. Просторові методи являють собою процедури, оперують безпосередньо значеннями пікселів, і описуються рівнянням:

$$g(x, y) = T[f(x, y)],$$

де $g(x, y)$ - оброблене зображення,

T - оператор над f , певний в деякому околі точки (x, y) (під околицею точки розуміється квадратна або прямокутна область, яка є підмножиною зображення і центрована щодо даної точки),

$f(x, y)$ - вхідне зображення.

Лінійна просторова фільтрація зображень

Локальні перетворення оперують і з значеннями пікселів в окрузи, і з відповідними їм значеннями деякої матриці, що має ті ж розміри, що і округа. Таку матрицю прийнято називати фільтром, міським, ядром, шаблоном або вікном. Обчислення здійснюється шляхом переміщенні маски фільтра від точки до точки

зображення. В кожній точці (x, y) відгук фільтра обчислюється з використанням попередньо заданих зв'язків.[18]

Фільтрація зображення f , що має розміри $M \times N$, за допомогою фільтра розмірами $m \times n$ задається виразом загального вигляду:

$$g(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x + s, y + t),$$

де $g(x, y)$ - оброблене зображення,

a - рівно $(m - 1)/2$,

b - рівно $(n - 1)/2$,

$w(s, t)$ - маска,

$f(x, y)$ - вхідне зображення,

x - рівно $0, 1, 2 \dots M - 1$,

y - $0, 1, 2 \dots N - 1$.

Для маски розміром елемента 3×3 результат лінійної фільтрації зображення в точці (x, y) представлений виразом:

$$R = w(-1, -1)f(x - 1, y - 1) + w(-1, 0)f(x - 1, y) + w(0, 0)f(x, y) + \dots \\ \dots + w(1, 0)f(x + 1, y) + w(1, 1)f(x + 1, y + 1)$$

Таким чином, процедура лінійної фільтрації є процедуру згортки коефіцієнтів фільтра (маски) з отсчетами вихідного зображення. Якщо необхідне тільки значення відгуку R по масці $m \times n$, в якій точці (x, y) використовується такий вираз:

$$R = w_1 z_1 + w_2 z_2 + \dots + w_{mn} z_{mn} = \sum_{i=1}^{mn} w_i z_i$$

де R – відгук,

w_i – коефіцієнт маски,

z_i – значення пікселя, яке відповідає даному коефіцієнту,

mn - загальне число коефіцієнтів в масці.

Згладжуючі просторові фільтри

Згладжуючі фільтри застосовуються для расфокусування зображення і придушення шуму. Расфокусування застосовується для підготовки зображення, наприклад, для видалення дрібних деталей перед пошуком великих об'єктів, для усунення розривів в лініях або деталях.

Відгук або вихідний сигнал лінійного згладжує фільтра являє собою середнє значення елементів по околиці, покритої маскою фільтра, що забезпечує згладжує ефект. Такі фільтри іноді також називають низькочастотними фільтрами.

Заміною вихідних значень елементів зображення на середні значення по масці фільтра досягається скорочення різких переходів ступенів яскравості. Оскільки випадковий шум якраз характеризується різкими стрибками яскравості, найбільш очевидним застосуванням згладжування є придушення шуму. Однак негативною стороною застосування згладжуючих фільтрів є расфокусування контурів зображення, які також характеризуються різкими перепадами яскравості. Головне використання згладжують фільтрів полягає в придушенні несуттєвих деталей на зображенні, під якими розуміється сукупність пікселів зображення, які малі в порівнянні з розмірами маски фільтра.

$$R = \frac{1}{9} \sum_{i=1}^9 z_i$$

$$\frac{1}{9} * 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1 \qquad \frac{1}{16} * 1\ 2\ 1\ 2\ 4\ 2\ 1\ 2\ 1$$

Постійний множник перед маскою дорівнює одиниці, поділений на суму значень коефіцієнтів, множник необхідний для нормування. Перший фільтр має однакові коефіцієнти і називається однорідним усереднюються фільтром, вихідне значення даного фільтра є звичайне середнє значення по масці. Друга маска забезпечує на виході фільтра зважене середнє, коефіцієнт в центрі маски має найбільше значення, значення інших коефіцієнтів зменшуються в міру віддалення від центру маски. Основна стратегія застосування подібного роду фільтрів - зменшення расфокусировки при згладжуванні. Загальна формула фільтрації зображення розмірами $M \times N$ з допомогою фільтра зваженого середнього по околиці $m \times n$ визначається виразом:

$$g(x, y) = \frac{\sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x + s, y + t)}{\sum_{s=-a}^a \sum_{t=-b}^b w(s, t)}$$

Підбір оптимального розміру маски вимагає проведення декількох експериментів і визначається як розмірами вихідного зображення, так і основним призначенням виконання процедури згладжування зображення.

Просторові фільтри підвищення різкості зображень

Центральна мета підвищення різкості зображення полягає в тому, щоб підкреслити дрібні деталі або поліпшити ті деталі, які опинилися не в фокусі внаслідок помилок або недосконалості самого методу реєстрації зображень.

Раніше було показано, що расфокусировка зображення здійснювалася шляхом усереднення значень пікселів по околиці.

Підвищення різкості зображення може бути досягнуто за допомогою зворотного процедури - просторового диференціювання.

Диференціювання зображення дозволяє посилити перепади та інші розриви на зображенні, але також неминуче призводить і до посилення шумів на зображенні. Величина відгуку оператора похідною в точці зображення

пропорційна ступеню розривності зображення в даній точці. Похідні дискретної функції визначаються в термінах різниці.

Перша похідна повинна бути:

- рівний нулю на ділянках з незмінним рівнем яскравості;
- не рівний «0» на початку і в кінці сходинки або схилу яскравості;
- не рівний «0» на схилах яскравості.

Друга похідна повинна бути:

- рівний нулю на ділянках з незмінним рівнем яскравості;
- не рівний «0» на початку і в кінці сходинки або схилу яскравості;
- дорівнює нулю на схилах постійної крутизни.

Перша похідна одновимірної функції визначається як різниця значень сусідніх елементів:

$$\frac{\partial f}{\partial x} = f(x + 1) - f(x)$$

Також друга похідна визначається як різниця сусідніх значень першої похідної:

$$\frac{\partial^2 f}{\partial x^2} = f(x + 1) + f(x - 1) - 2f(x)$$

Відзначимо, що друга похідна є набагато більш ефективною при вирішенні завдань посилення різких переходів на зображенні, ніж перша похідна, проте друга похідна забезпечує набагато більше посилення дрібних деталей, включаючи шум.

Поліпшення зображень з використанням других похідних

В даному випадку розглядається використання оператора двовимірної другої похідної в задачах поліпшення зображення. Підхід зводиться до вибору дискретної

формулювання другої похідної і ось до чого побудови маски фільтра, заснованої на запропонованій формулюванні. У даних задачах розглядаються виключно ізотропні фільтри, відгук яких не залежить від напрямки неоднорідностей на зображення, яке обробляється.

Ізотропні фільтри мають властивість інваріантності до повороту зображення, тобто поворот зображення перед використанням фільтра дає той же результат, що і первісне використання фільтра з подальшим поворотом результату.

Найпростішим оператором, заснованим на похідних, є оператор Лапласа - лапласіан, який для функції двох змінних визначається як:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Для застосування оператора Лапласа в цифровій обробці зображень його необхідно висловити в дискретному вигляді. Найбільш часто використовується таке уявлення:

$$\frac{\partial^2 f}{\partial x^2} = f(x + 1, y) + f(x - 1, y) - 2f(x, y)$$

$$\frac{\partial^2 f}{\partial y^2} = f(x, y + 1) + f(x, y - 1) - 2f(x, y)$$

Дискретна формулювання двовимірного Лапласіан виходить об'єднанням цих двох складових:

$$\nabla^2 f = [f(x + 1, y) + f(x - 1, y) + f(x, y + 1) + f(x, y - 1) - 4f(x, y)]$$

Дане рівняння може бути реалізовано за допомогою наступної маски:

0 1 0 1 – 4 1 0 1 0

В силу того, що оператор Лапласа по факту є другою похідною, його використання підкреслює розриви рівнів яскравості на зображенні і пригнічує області зі слабкими змінами яркостей. З цієї причини виходить зображення, що містить сіруваті лінії на місці контурів і інших розривів, накладені на темний фон без особливостей. Для відновлення колишнього фону при збереженні ефекту підвищення різкості, що досягається використанням Лапласіан, досить скласти вихідне зображення і зображення, що отримується на виході оператора Лапласа.

Нерізде маскуваннн і фільтрація з підйомом високих частот

Процедура, яка використовується для підвищення різкості зображення, полягає в відніманні з вихідного зображення його розфокусовану копію, називається нерізким маскуванням і виражається наступною формулою:

$$f_s(x, y) = f(x, y) - \underline{f}(x, y),$$

де $f_s(x, y)$ - зображення з підвищеною різкістю, отримане методом нерізкого маскування,

$f(x, y)$ - розфокусована копія вихідного зображення $\underline{f}(x, y)$,

$\underline{f}(x, y)$ - вихідне зображення.

Узагальненої операцією нерізкого маскування є фільтрація з підйомом високих частот, яка задається в наступному вигляді:

$$f_{hb}(x, y) = Af(x, y) - \underline{f}(x, y),$$

де $f_{hb}(x, y)$ - зображення з підвищеною різкістю, отримане методом нерізкого маскування,

A - при збільшенні **A** більше 1 ефект збільшення різкості стає менше. $A \geq 1$.

$f(x, y)$ - розфокусована копія вихідного зображення $\underline{f}(x, y)$,

$\underline{f}(x, y)$ - вихідне зображення.

Типовим застосуванням фільтрації зображень з підйомом високих частот є випадок, коли вихідне зображення темніше, ніж це потрібно. Варіюванням постійного коефіцієнта A вдається домогтися загального збільшення яскравості зображення і отримати бажаний результат.

Медіанна фільтрація

Медіанний фільтр найбільш відомий як фільтр заснований на порядкових статистиках. Сутність цього фільтра полягає в заміні значення в оброблюваній точці зображення на медіану яскравостей в околиці цієї точки:

$$\hat{f}(x, y) = \text{med}_{(s,t) \in S_{xy}} \{g(s, t)\}$$

Медіанні фільтри відмінно справляються із завданням приборкання деяких видів непередбачених шумів і при цьому забезпечують менший розмиття зображень в порівнянні з лінійними згладжуючими фільтрами того ж розміру. Медіанні фільтри найбільше ефективні при наявності як біполярного, так і уніполярного імпульсного шуму.

Ранжуючий фільтр

Суть таких фільтрів полягає в застосуванні місцевих оцінок, адекватних визначенням контуру на зображенні і згладжування таких ділянок в мінімальній мірці. При імпульсному шумі частина пікселів на зображенні замінюється випадковими або фіксованими значеннями. Така модель шуму часто виникає при передачі зображень.

Імпульсний шум може бути виділений використовуючи спеціальний клас нелінійних фільтрів, побудованих на основі рангової статистики. Загальними

властивостями таких фільтрів є детектування позиції імпульсу і заміна його оцінним значенням, при збереженні інших пікселів зображення незмінними.

Ранжуючий фільтр, як і згладжуючий, використовує для перетворення зображення маску. Маска може включати чи не включати центральний піксель. Значення елементів, що потрапляють в маску, можна розташувати в упорядкований ряд і впорядкувати по зростанню (або зменшенню), і обчислити, наприклад, середнє значення інтенсивності та дисперсії і інші моменти цього ряду. Вихідним значенням фільтра, яким замінюється центральний відлік, є зважена сума інтенсивності центрального пікселя і медіани отриманого ряду. Коефіцієнти зазвичай пов'язані певної залежністю зі статистикою пікселів у вікні фільтра.

Фільтрація Вінера

Одним з найбільш ефективних методів відновлення зображення, заснованих одночасно на обліку властивостей спотворює функції і статистичних властивостей шуму в процесі відновлення, є винеровського фільтрація або фільтрація методом мінімізації середньоквадратичного відхилення. Даний метод заснований на розгляді зображень і шуму як випадкових процесів, і завдання формулюється в такий спосіб: необхідно знайти таку оцінку \hat{f} для неспотвореного зображення f , щоб середньоквадратичне відхилення цих величин один від одного було мінімальним. Середньоквадратичне відхилення e задається формулою:

$$e^2 = E\{(f - \hat{f})^2\}$$

де $E\{\cdot\}$ позначає математичне очікування свого аргументу.

Передбачається, що виконані наступні умови:

- шум і неспотворене зображення не корельовано між собою;
- або шум, або неспотворене зображення мають нульове середнє значення;
- оцінка лінійно залежить від спотвореного зображення.

При виконанні цих умов мінімум середньоквадратичного відхилення досягається для оцінює вихідне зображення функції, заданої в частотній області, в такий спосіб:

$$\hat{F}(u,v) = \left(\frac{H^*(u,v)S_f(u,v)}{S_f(u,v)|H(u,v)|^2 + S_\eta(u,v)} \right) G(u,v) = \left(\frac{1}{H(u,v)} \cdot \frac{|H(u,v)|^2}{|H(u,v)|^2 + S_\eta(u,v)/S_f(u,v)} \right) G(u,v)$$

де $\hat{F}(u,v)$ – Фур'є-перетворення викривленого зображення,

$H^*(u, v)$ – комплексне сполучення $H(u, v)$,

$S_f(u, v) = |F(u, v)|^2$ – енергетичний спектр неспотвореного зображення,

$H(u, v)$ – частотне уявлення спотворює функції,

$|H(u, v)|^2 = H^*(u, v) * H(u, v)$ – енергетичний спектр спотворює функції,

$S_\eta(u, v) = |N(u, v)|^2$ – энергетический спектр шума,

$G(u,v)$ – Фур'є-перетворення викривленого зображення.

Наведений результат був вперше отриманий Н. Вінером, і метод відомий як оптимальна фільтрація за Вінером. Фільтр, представлений функцією передачі всередині дужок, називається винеровским фільтром. Аналізуючи наведене вираз, необхідно зауважити, що для винеровского фільтра відсутня проблема поводження в нуль спектра спотворює функції, за винятком того випадку, коли функції $H(u, v)$ і $S_\eta(u, v)$ звертаються в деяких точках в нуль одночасно.

Відновлення зображення в просторової області досягається застосуванням зворотного перетворення Фур'є до оцінки $\hat{F}(u, v)$. Відзначимо, що якщо шум дорівнює нулю, то і його енергетичний спектр звертається в нуль, і в цьому випадку винеровского фільтрація зводиться до інверсної фільтрації. У тому випадку, якщо на зображенні присутній білий шум, спектр якого $|N(u, v)|^2$ є постійною функцією, то вираз для винеровского фільтра спрощується. Однак в більшості випадків спектр неспотвореного зображення рідко буває відомий. У тих випадках, коли спектри шуму і неспотвореного зображення невідомі і не можуть бути оцінені, часто

використовується підхід, що складається в апроксимації вихідного вираження наступним виразом:

$$\hat{F}(u, v) = \left(\frac{1}{H(u, v) \cdot \frac{|H(u, v)^2|}{|H(u, v)^2| + K}} \right) G(u, v),$$

де K – певна константа.

Вінерівська фільтрація дозволяє отримати результати відновлення зображення дуже близькі з вигляду до вихідного зображення.

Збільшення контрастності

У більшості методів збільшення контрастності растрового зображення яскравість пікселя перетвореного зображення $z'(x, y)$ залежить від яскравості тільки відповідного пікселя вихідного зображення $z(x, y)$, що дозволяє перед обробкою сформуванню масив відповідності яскравості відповідно до заданого алгоритму $z'(x, y) = f(z)$.

Такий масив значно скорочує час обробки, тому що після його формування власне обробка поточного пікселя вихідного зображення зводиться до присвоєння відповідного пікселя перетвореного зображення яскравості, що дорівнює значенню елемента масиву з номером, рівним яскравості поточного пікселя вихідного зображення.

При функціональному перетворенні яскравості пікселів перетвореного зображення не залежать від гістограми яскравості вихідного зображення і формування значень z' масиву відповідності яскравості проводиться послідовною підстановкою цілочислових значень яскравості z в функцію перетворення шкали яскравості $f(z)$.

При гістограмном перетворенні вид функції $f(z)$ залежить від значень гістограми яскравості вихідного зображення і перед формуванням масиву

відповідності яскравостей необхідно побудувати нормалізовану гістограму яскравості вихідного зображення.

Зображення, отримані в несприятливих умовах освітленості, часто не використовують весь діапазон можливих градацій яскравості, що зумовлює їх низьку інформативність. Контрастність зображення, яскравості елементів якого розташовані у вузькому проміжку можливих значень, низька.

В результаті сусідні на зображенні фрагменти мають надто близьку яскравість, що й обумовлює їх погану розрізнення. Один з методів підвищення якості таких зображень полягає в збільшенні на максимально можливу величину різниці значень яскравості сусідніх фрагментів, наприклад, лінійним розтягуванням шкали яскравості. Перетворення шкали має вигляд:

$$z' = \frac{z - z_{min}}{C_{om}}$$

де z' – яскравість пікселя після перетворення,

z – яскравість пікселя до перетворення,

z_{min} - мінімальна яскравість вихідного зображення,

C_{om} - відносна контрастність зображення.

Видно, що істотне збільшення розрізнення фрагментів досягається тільки при низькій контрастності вихідного зображення. У порівнянні з щільною гістограммою вихідного зображення, гістограма перетвореного зображення виявляється розтягнутою і в ній з'являються пропуски, але самі значення гістограми не змінюються, а змінюються тільки яскравості, до яких вони належать.

Саме наявність в будь-якої частини гістограми пропусків (розрідженість), тобто відсутність на зображенні пікселів деяких йдуть підряд яркостей, є непрямою ознакою контрастності фрагментів, що відносяться по яскравості до цієї частини гістограми. При C_{om} близькою до 1 збільшення контрастності лінійним перетворенням всієї шкали яскравості неефективно. Але якщо відомо, що погано помітні фрагменти розташовані у вузькому діапазоні яскравості, то можна провести

усічення шкали яскравості до необхідного діапазону з подальшим шматочколінійним перетворенням.

Нелінійне перетворення шкали яскравості або її отриманої шкали. В результаті пікселі, яскравість яких знаходиться вище або нижче обраного діапазону, отримують відповідно максимально і мінімально можливу яскравість, а виділений діапазон буде лінійно розтягнутий.

Гамма-корекція при Com близькому до 1 дозволяє збільшити контрастність в локальному діапазоні яскравості за рахунок зниження контрастності в інших діапазонах. Наприклад, якщо відомо, що погано помітні фрагменти мають високу яскравість, то можна застосувати статичне перетворення шкали яскравості

$$z' = z_m z_{om}^n$$

де $z_{om} \frac{z}{z_m}$ – відносна яскравість.

Якщо яскравості пікселів вихідного зображення займають весь можливий діапазон, то z_{om} змінюється в межах $0 \dots 1$.

Значення n вибирається експериментально, як правило, з діапазону $n = 2 \dots 10$. В результаті перетворення шкала яскравості в зоні, близької до нуля, виявляється стислій, а в зоні близької до z_m - розтягнутої, що призводить до підвищення контрастності світлих фрагментів за рахунок зниження контрастності темних фрагментів. Зворотний результат дає перетворення при діапазоні $n=0,1 \dots 0,9$.

Аналогічні результати можна отримати, змінюючи шкалу яскравості логарифмічним або експоненціальним перетворенням з відповідними нормуються коефіцієнтами.

Одночасно збільшити контрастність темних і світлих фрагментів зображення за рахунок часткового злиття фрагментів середньої яскравості можна, перетворити до виду

$$z' = 0.5 z_m (1 + (2z_{om} - 1)^n) \text{ для } n > 1 \text{ непарного ступеня.}$$

Іншим методом, що дозволяє одночасно збільшити розрізнення, як темних, так і світлих фрагментів є соляризація, при якій перетворення шкали яскравості має вигляд

$$z' = 4 z_{om}(z_m - z).$$

Функція, що описує дане перетворення, є квадратичною параболою, параметри якої підбрані так, щоб $z' = z_m$ для $z = 0,5 z_m$.

В результаті фрагменти, що мають близьку яскравість, як в темній, так і в світлій області шкали, розносяться по шкалі на достатню для збільшення їх розрізнення відстань, а фрагменти середньої яскравості набувають яскравість близьку до максимальної і частково зливаються. Метод називається «соляризація», тому що створюється враження, що світлі фрагменти «засмагли» на сонці.

Нерівномірне висвітлення об'єктів сцени може привести до того, що на зображенні з'являться як зайве, так і недостатньо висвітлені фрагменти.

Для компенсації нерівномірності освітлення рекомендується наступна послідовність дій. На початку зображення обробляється усереднюючим фільтром, наприклад, гауссіаном, з великим розміром апертури. При правильному підборі розміру апертури результат обробки $z_G(x, y)$ є наближеним розподілу освітленості по зображенню. Далі для кожного пікселя (x, y) вихідного зображення проводиться розподіл його яскравості на відповідне значення освітленості, тобто

$$z'(x, y) = C_{norm} \frac{z(x, y)}{z_G(x, y)}$$

Нормуючий коефіцієнт C_{norm} підбирається таким чином, щоб отримані значення $z'(x, y)$ займали весь діапазон можливих значень яскравості (0 - 255).

Слід зазначити, що для підвищення розрізнення деталей часто доводиться застосовувати послідовну обробку декількома різними методами і експериментально підібрати параметри обробки.

2 ОПИС ПРОВЕДЕНИХ ДОСЛІДЖЕНЬ

2.1 Формулювання задачі

Розпізнавання візуальних образів є дуже широким поняттям і для проведення експериментів необхідно виділити вузьку спрямованість. Для цього я вибрав напрямок - видалення фону в зображеннях або по іншому - знаходження переднього плану. Видалення фону - завдання, яке легко виконати вручну, або майже вручну (наприклад, Photoshop), якщо ви використовуєте якийсь «маркер» і технологію виявлення кордонів. Однак повністю автоматизоване видалення фону - досить складне завдання. Який фон ми будемо видаляти? Це питання виявилось важливим, оскільки чим більше конкретна модель з точки зору об'єктів, кутів і іншого, тим вищою буде якість поділу фону і переднього плану. Тому зосередимося на Селфі і портретах людей.

Селфі - це зображення з характерним і орієнтованим переднім планом (один або кілька «людинок»), яке гарантує нам гарний поділ між об'єктом (особа + верхня частина тіла) і фоном, а також з постійним кутом і завжди одним і тим же об'єктом (людиною). Для експериментів будемо використовувати модель нейронної мережі, навчену в TensorFlow.

TensorFlow - це опенсорсна бібліотека, створена Google, яка використовується при розробці систем, що використовують технології машинного навчання. Ця бібліотека включає в себе реалізацію безлічі потужних алгоритмів, розрахованих на рішення поширених завдань машинного навчання, серед яких можна відзначити розпізнавання образів і прийняття рішень. Тензор (tensor) - це стандартний спосіб представлення даних в системах глибокого навчання. Тензори - це багатовимірні масиви, розширення двовимірних таблиць (матриць) для представлення даних, що мають більш високі розмірності. Простіше кажучи, тензор - це n-вимірний масив.

На першому етапі дослідження ми розглянемо, як нейронна мережа справляється зі своїм завданням виділення переднього плану. На другому етапі

дослідження ми проведемо обробку вихідних даних різними фільтрами і з'ясуємо, як від цього змінюються результати виділення переднього плану. Попередню обробку вихідних даних будемо проводити в середовищі Matlab.

На лістингу нижче представлений код відновлення змазаних зображень не у фокусі за допомогою фільтра Вінера. Деконволюція по Вінеру може бути ефективно використана, коли частотні характеристики зображення і адитивний шум відомі, щонайменше, до певної міри. При расфокусовці спотворююча система добре апроксимується циліндричної функцією розсіювання точки (ФРТ) радіуса r .

```

PATHNAME='D:\MATLAB\jpg\orig\';
FILENAME='orig.jpg';
s=[PATHNAME, FILENAME] ;
f = imread(s);
figure(1); imshow(f);
h = fspecial('disk',4);
figure(2); imshow(h);
J = deconvwin(f,h);
figure(4); imshow(J);
imwrite(J,'D:\MATLAB\jpg\orig\orig_vinner_4.jpg');

```

Для підвищення контрастності зображень високочастотна і низькочастотна фільтрація повинні використовуватися одночасно. Додаємо вихідне зображення до обробленої за допомогою "top-hat" - фільтра і віднімаємо зображення, оброблене "bottom-hat" - фільтром.

```

PATHNAME='D:\MATLAB\jpg\orig\';
FILENAME='orig.jpg';
s=[PATHNAME, FILENAME] ;
I=imread(s);
imshow(I), title('Original');
se=strel('disk', 3);
J=imsubtract(imadd(I, imtophat(I, se)), imbothat(I, se));
figure, imshow(J), title('Contrast filtered');

```

```
imwrite(J, 'D:\MATLAB\jpg\orig\orig_contrast.jpg');
```

2. 2 Результати обробки зображень без попередньої обробки

Результати обробки при різному фоні, освітленні і кількості людей представлені нижче. На рисунках 2.1-2.5 ми бачимо гарні і цілком задовільні результати обробки. Вхідними даними були зображення з різним фоном і освітленістю. Характерним для цих зображень було близька відстань до об'єктива камери і добре помітні кордони об'єктів. Гарний результат обробки швидше за все отримано завдяки тому, що навчання моделі нейронної мережі відбувалося саме на подібних зображеннях.

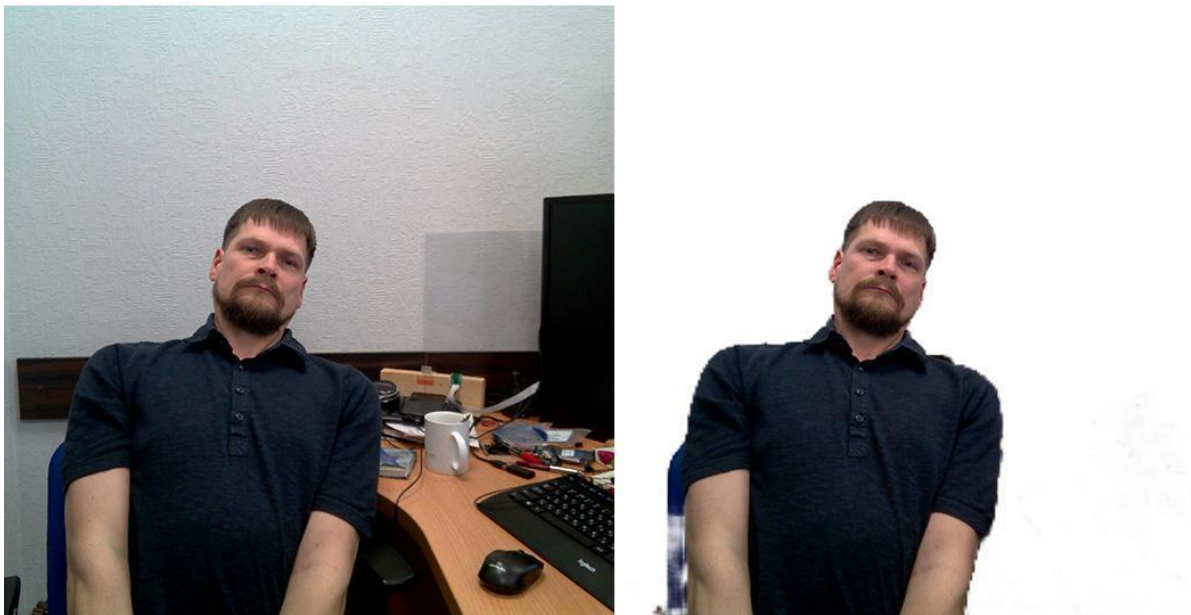


Рисунок 2.1 – Видалення фону одиничного об'єкта з гарним результатом

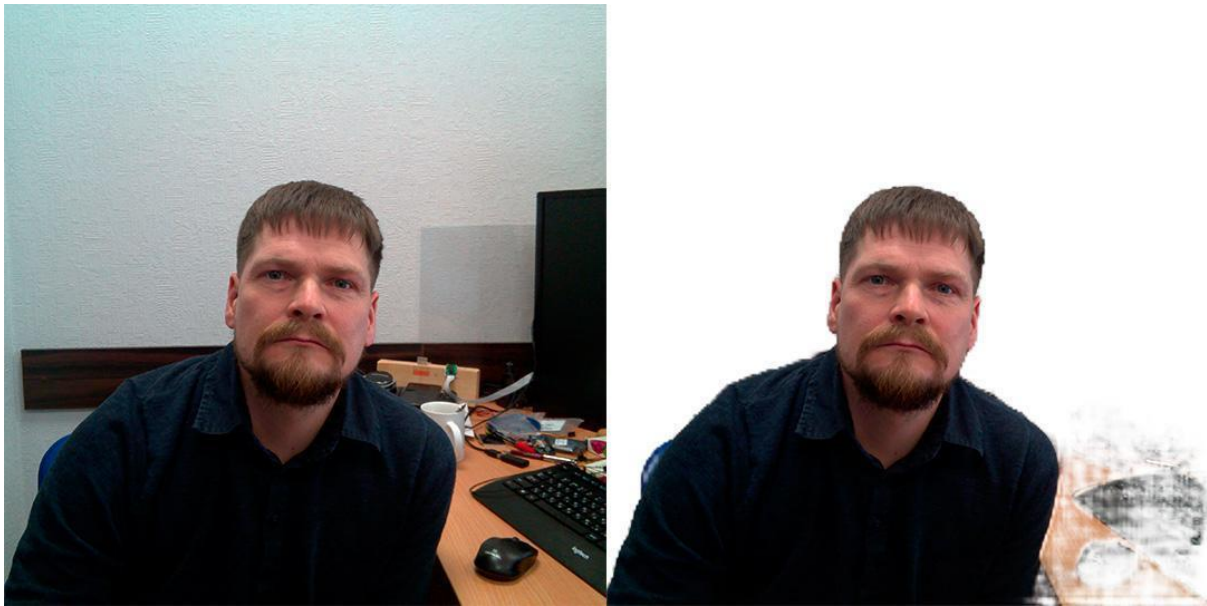


Рисунок 2.2 – Видалення фону одиничного об'єкта з частковим гарним результатом

На рисунку 2.2 ми бачимо невпевнене видалення фону з лівого боку. По розташуванню і за кольором нейросеть можливо визначила цей об'єкт як руку.

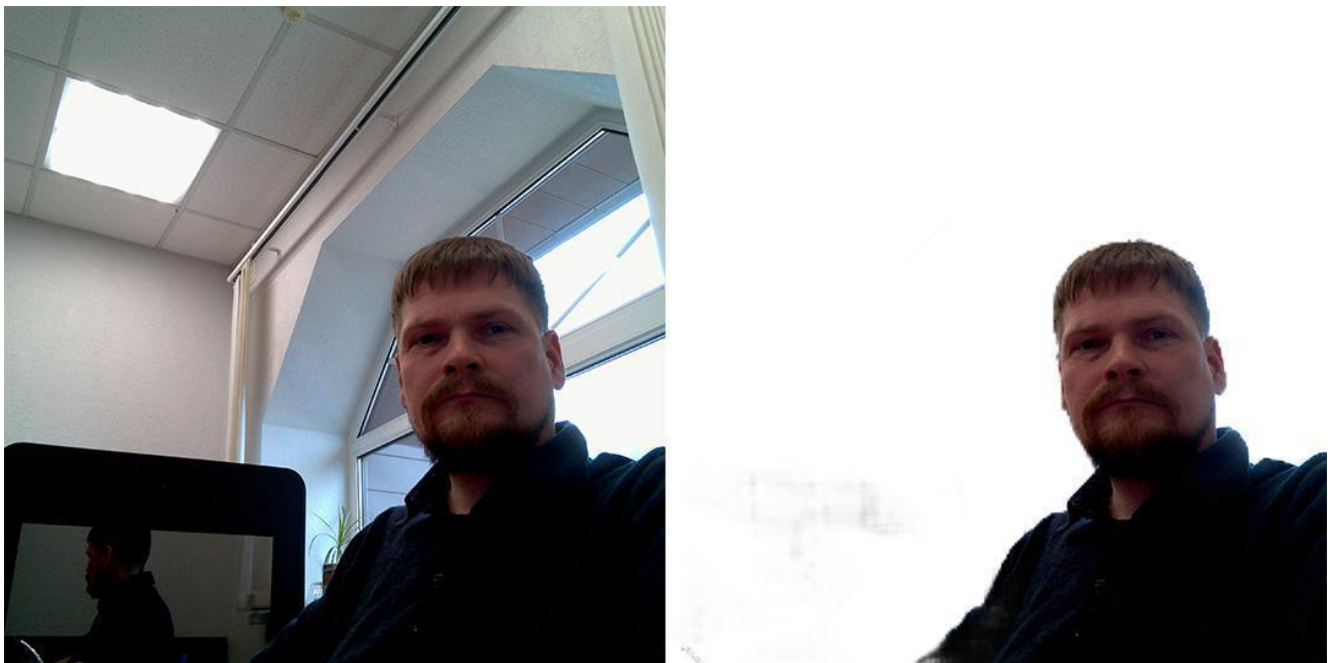


Рисунок 2.3 – Видалення фону одиничного об'єкта з гарним результатом

А на рисунку 2.4 нейросеть навпаки не визначила руку як об'єкт через часткову видимість і розташування руки.



Рисунок 2.4 – Видалення фону одиничного об'єкта з задовільним результатом

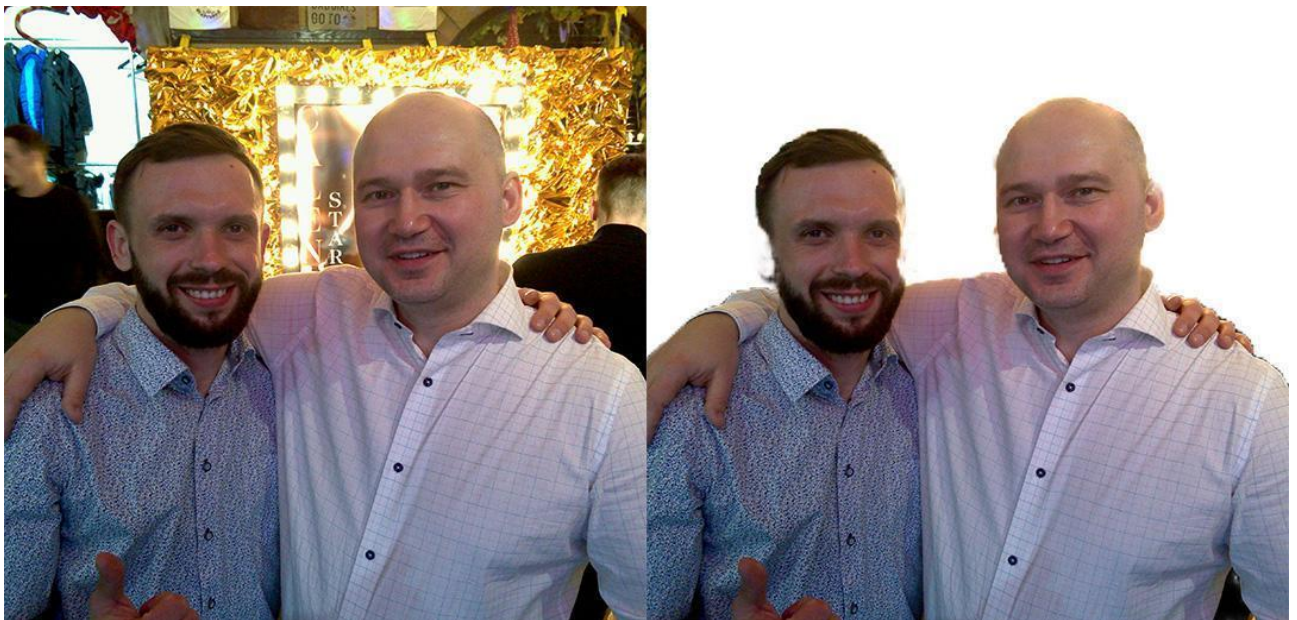


Рисунок 2.5 – Видалення фону кількох об'єктів з гарним результатом

На рисунках 2.6-2.7 представлені незадовільні результати обробки. Причини, які призвели до цього можуть бути зливання фону і об'єкта а також більш гіршу якість деталізації об'єктів через віддалення від об'єктива камери.



Рисунок 2.6 – Видалення фону кількох об'єктів з поганим результатом



Рисунок 2.7 – Видалення фону кількох об'єктів з поганим результатом

Також присутні не зовсім зрозуміли суперечливі результати обробки. Наприклад, на рисунку 2.8 близько від камери знаходяться 3 об'єкти на однаковій відстані з добре виділеними межами, але відпрацьований добре по фону тільки один об'єкт.

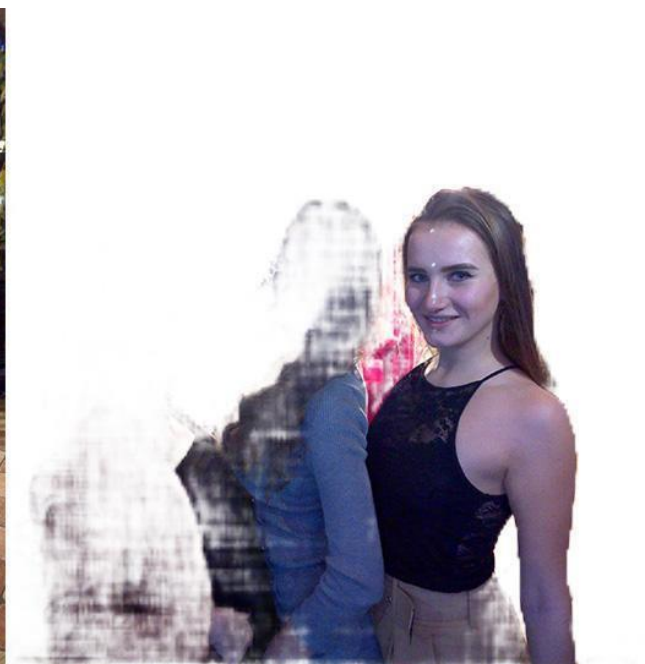


Рисунок 2.8 – Видалення фону кількох об'єктів з поганим результатом

При цьому на рисунках 2.9 і 2.10 об'єкти є темними з підсвічується фоном, але якість видалення фону досить задовільний.

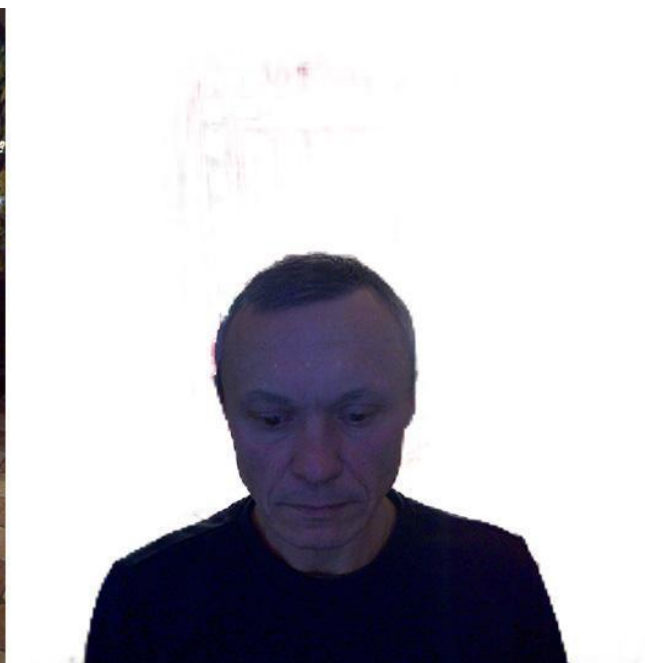
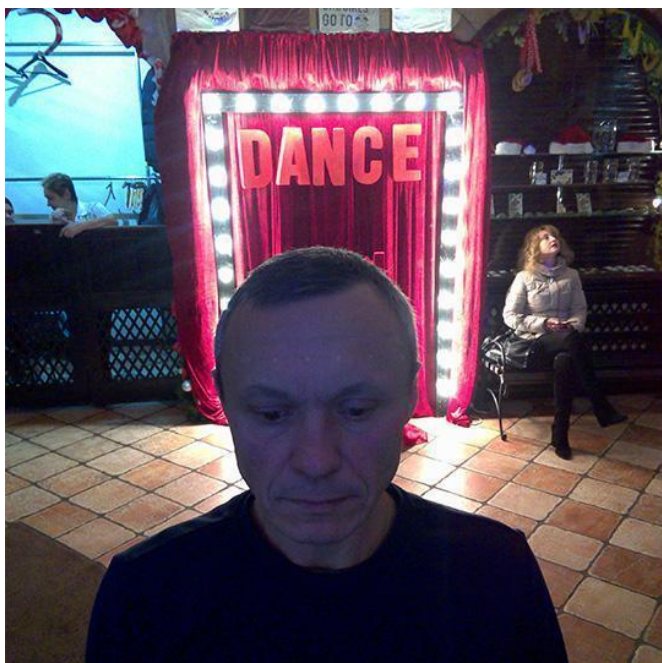


Рисунок 2.9 – Видалення фону темного об'єкту з гарним результатом



Рисунок 2.10 – Видалення фону темного об'єкту з гарним результатом

На рисунку 2.11 ми бачимо 2 об'єкти на відстані від об'єктива камери. Самі об'єкти виділені задовільно, але в результуюче зображення потрапила значна частина фону.



Рисунок 2.11 – Видалення фону з поганим результатом

Це прогнозований результат для даної моделі нейромережі, але на рисунку 2.12 ми бачимо 2 об'єкти, один з яких близько від камери, але результат залишає бажати кращого.



Рисунок 2.12 – Видалення фону з поганим результатом

2.3 Результати обробки зображень з попередньою обробкою

Для аналізу впливу попередньої обробки виберемо вихідне зображення з великою кількістю об'єктів - права частина рисунка 2.7. Велика кількість об'єктів дозволить в цілому прийняти рішення про вплив попередньої фільтрації на якість обробки. На рисунках 2.13-2.18 представлені результати виділення переднього плану з попередньою обробкою фільтром Гауса. В обробці зображень Гаусове розмиття (також відомі як Гаусові згладжування) є результатом розмиття зображення за допомогою Гаусових функції. Це широко використовується, як правило, щоб зменшити шум зображення і зменшити деталі. Візуальний ефект цього методу розмиття є гладким розмиванням. Ліва сторона малюнка відповідає

вихідним даним після обробки фільтром Гауса з відповідним коефіцієнтом. Коефіцієнти фільтра Гауса візьмемо, починаючи зі значення 0,7 і будемо поступово збільшувати і аналізувати результат. На рисунку 2.13 представлений еталонний результат без попередньої обробки вихідного зображення. Саме щодо цього початкового результату ми будемо судити про вплив попередньої фільтрації.



Рисунок 2.13 – Результат без попередньої обробки



Рисунок 2.14 – Результат з попередньою обробкою фільтром Гауса 0.7

Порівнюючи результати, видно, що при збільшенні коефіцієнта фільтра Гауса поступово проявляються все більше нових деталей, які раніше не розпізнавати нейронною мережею як частина об'єкта.



Рисунок 2.15 – Результат з попередньою обробкою фільтром Гауса 2



Рисунок 2.16 – Результат з попередньою обробкою фільтром Гауса 3



Рисунок 2.17 – Результат з попередньою обробкою фільтром Гауса 4

Але це поліпшення спостерігається до певного значення коефіцієнта фільтра Гауса, після якого якість обробки починає падати. Як видно з результатів, піком поліпшення якості обробки є значення коефіцієнта 4 - рисунок 2.17. Починаючи зі значення 5 якість обробки незначно починає погіршуватися. Проявляється це в зникненні з зображення деталей, що належать об'єкту. Для наочності на рисунку 2.18 ми бачимо вже значне погіршення якості обробки при значенні коефіцієнта 7.



Рисунок 2.18 – Результат з попередньою обробкою фільтром Гауса 7

Визначивши пік поліпшення обробки зображення після фільтрації, опрацюємо інше вихідне зображення, результат обробки якого спочатку було гарним. Це необхідно для того, щоб перевірити відсутність погіршення впливу на спочатку гарні результати обробки. Для цього візьмемо зображення, позитивний результат обробки якого представлений на рисунку 2.9 і опрацюємо його фільтром Гауса з коефіцієнтом 4. Результат виділення переднього плану після фільтра представлено на рисунку 2.19. Як видно з отриманого результату, ознак погіршення після фільтра ми не спостерігаємо.

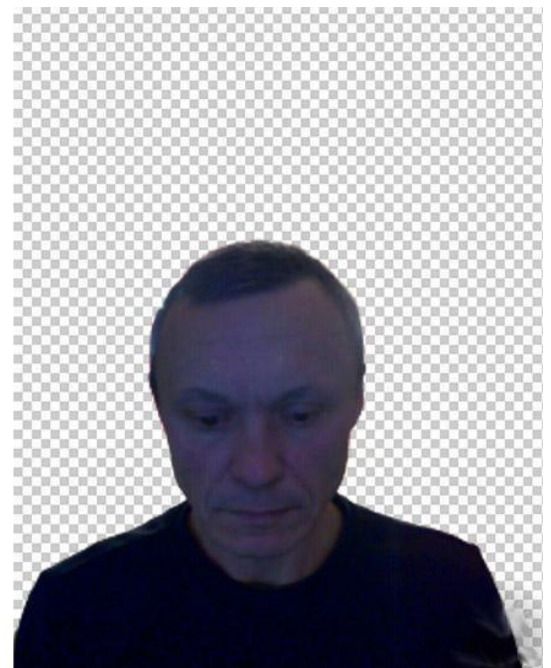
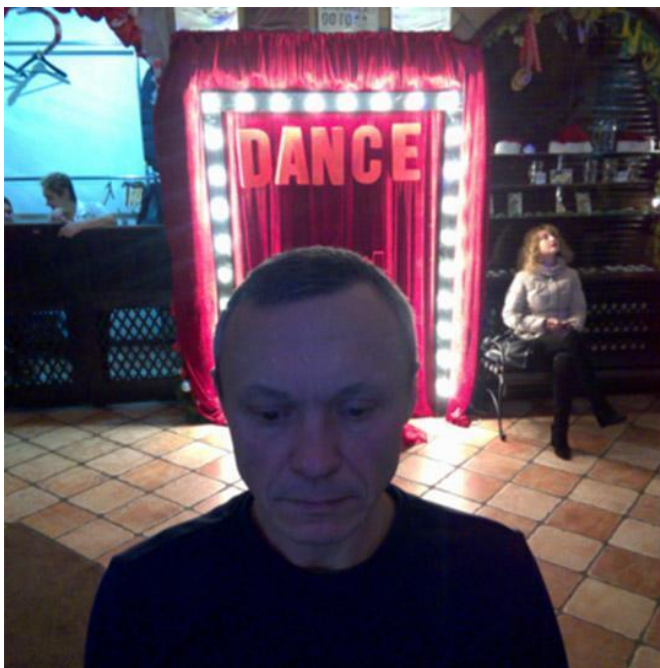


Рисунок 2.19 – Результат з попередньою обробкою фільтром Гауса 4

Далі перейдемо до аналізу інших способів фільтрації. Визначимо вплив фільтра Вінера на якість обробки. При відсутності фокусу спотворююча система добре апроксимується циліндричною функцією розсіювання точки (ФРТ) радіуса r . На рисунку 2.20 представлений результат з попередньою обробкою фільтром Вінера з ФРТ радіусом 4. Як видно з результату, фільтрація даним типом фільтру не привела до позитивного результату.



Рисунок 2.20 – Результат з попередньою обробкою фільтром Вінера

Наступний аналізований фільтр - фільтр підвищення контрастності. У нашому випадку він імовірно може дати позитивний результат, так як одним з його ефектів є більш чітке виділення меж об'єктів. На рисунку 2.21 представлений результат обробки при використанні фільтра Вінера. Як видно з результату - очікуваного поліпшення обробки нами не досягнуто.



Рисунок 2.21 – Результат з попередньою обробкою підвищення контрастності

2.4 Аналіз результатів

Аналізуючи результати після попередньої обробки фільтром Гауса візуально видно, що з підвищенням коефіцієнта, якість виділення переднього плану поступово поліпшується, але після 4 помітно погіршення. Найкращий показник видно на рисунку 2.17 при значенні коефіцієнта 4. На рисунку 2.19 перевірено відсутність погіршення при попередній обробці в порівнянні з рисунком 2.9, де вхідне зображення без попередньої обробки. Погіршення не помічено. На малюнку 2.20 і 2.21 вхідні зображення були оброблені фільтром поліпшення контрастності і фільтром поліпшення різкості Вінера. Як видно - попередня обробка не призвела до поліпшення відділення переднього плану нейронною мережею.

ВИСНОВКИ

В результаті виконання роботи були розглянуті різні методи розпізнаванні візуальних образів. Для експериментів був обраний метод, заснований на обробці зображення за допомогою нейронної навченої мережі. Навчання нейронної мережі проводилося на вибірці селфі-фотографій. Як видно з результатів, на якість обробки дуже сильно впливають параметри освітлення, видалення від об'єктива камери, характер заднього фону. Коли задній фон був за кольором і інтенсивності схожий з областями переднього плану, в цьому випадку області заднього фону могли залишатися на результуючому зображенні або навпаки з переднього плану віддалилися частини об'єкта. Також освітлення має сильну складову впливу на якість обробки зображення.

Для покращення якості обробки зображень в нашому випадку можна вибрати два напрямку. Перший спосіб - це додатково додати модель нейронної мережі великою кількістю еталонних зображень, що призведе до додаткових фінансових і тимчасових витратах. Крім того при зміні з часом зовнішніх чинників результати обробки також можуть погіршитися.

Другий спосіб - це зробити попередню обробку вхідних зображень. В якості такої обробки спробували наступні фільтри: низькочастотний фільтр Гауса, фільтр підвищення контрастності, високочастотний фільтр Вінера (різкість).

Результатом атестаційної магістерської роботи є аналіз результатів роботи нейронної мережі з використанням попередньої обробки вхідних даних, на підставі яких став зрозумілий подальший план робіт і досліджень. Встановлено, що найкращим чином на роботу використовуваної нейронної мережі за визначенням переднього плану впливає низькочастотний фільтр Гауса, проте ідеального результату добитися не вдалося. Подальшим шляхом досліджень може бути комбінація різних фільтрів, але для цього необхідно вже автоматичний перебір комбінацій і як наслідок більш потужні апаратні засоби.

В процесі роботи над магістерською роботою використовувалися навички і знання з наступних областей: чисельні методи та високопродуктивні обчислення, теорія прийняття рішень, інформаційні технології в системах машинного зору.

В кінці дослідження були зроблені висновки щодо ефективності запропонованих методів і методики, був зроблений комплексний аналіз всього дослідження та коригування методу обробки зображення.

ПЕРЕЛІК ПОСИЛАНЬ

1. Мокшін В. В. та ін. Розпізнавання образів транспортних засобів на основі евристичних даних і машинного навчання // Вісник Казанського технологічного університету. 2016. – Т. 19. – №. 5.
2. Заенцев, І. В. Нейронні мережі. Основні моделі / І. В. Заенцев. – Воронеж: ВГУ, 1999. – 76 с.
3. Шуригін А.М. Прикладна стохастика: робастної, оцінювання, прогноз. – М.: Фінанси і статистика, 2000..
4. Ту, Дж. Принципи розпізнавання образів / Дж. Ту, Р. Гонсалес. – М.: Світ, 1978. – 412 с.
5. Шапіро, Л. Комп'ютерне зір / Л. Шапіро, Дж. Стокман. – М.: Біном. Лабораторія знань, 2006. – 752 с.
6. Гонсалес, Р. Цифрова обробка зображень / Р.Гонсалес, Р. Вудс. – М.: Техносфера, 2005. – 1007 с.
7. С. Т. Zahn (1971): «Graph-theoretical methods for detecting and describing gestalt clusters», IEEE Transactions on Computers, pp. 68-86, Vol. 20, No.
8. Ross Girshick. Fast R-CNN / Ross Girshick // URL: <https://arxiv.org/abs/1504.08083> (дата звертання: 3.03.2020)
9. R-CNN, Fast R-CNN, Faster R-CNN, YOLO – Object Detection Algorithms. // URL: <https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms36d53571365e> (дата звертання: 22.03.19)
10. Selective Search for Object Recognition / Uijlings, J.R.R., van de Sande, K.E.A., Gevers, T. et al. Int J Comput Vis (2013) 104: 154 [Електронний ресурс]. - URL: <https://doi.org/10.1007/s11263-013-0620-5> (дата звертання: 5.04.2020)
11. SSD: Single Shot MultiBox Detector / Wei Liu, Dragomir Anguelov, Dumitru Erhan [и др.] // UNC Chapel Hill 2Zoox Inc. Google Inc. University of Michigan, Ann Arbor. – URL: <https://arxiv.org/abs/1512.02325> (дата звертання: 23.04.2020)

12. Traffic flow estimation with data from a video surveillance camera / Fedorov A., Nikolskaia K., Ivanov S., Shepelev V., Minbaleev A. Journal of Big Data. 2019. Vol. 6. Article 73. DOI: 10.1186/s40537-019-0234-z

13. What's new in YOLO v3? // URL: <https://towardsdatascience.com/yolo-v3-object-detection-53fb7d3bfe6b> (дата звертання: 23.03.20)

14. You Only Look Once: Unified, Real-Time Object Detection / Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi // University of Washington, Allen Institute for AI, Facebook AI Research. – URL: <https://arxiv.org/abs/1506.02640> (дата звертання: 23.03.2020)

15. Метод неймережевой ОБРОБКИ зашумлення просторових даних <file:///C:/Users/ankir/Downloads/metod-neyrosetevoy-obrabotki-zashumlenn-h-prostranstvenn-h-dann-h.pdf>

16. Апальков И.В., Хрящев В.В. Удаление шума из изображений на основе нелинейных алгоритмов с использованием ранговой статистики. — Ярославский государственный университет, 2007.

17.Круглов, В. В. Штучні нейронні мережі. Теорія та практика. / В.В. Круглов, В. В. Борисов. – М.: Горяча лінія-Телеком, 2002. – 382 с.

18. Canny, J., A Computational Approach To Edge Detection, IEEE Trans. Pattern Analysis and Machine Intelligence, 1986, С. 679–698.

19. A personalized view of the history of PACS in the USA. / J. Samuel, I. Dwyer. // Spie, 2000.- С. 2-9.

20. Pattern Classification and Scene Analysis./R. Duda, P. Hart. – John Wiley and Sons, 1973.- 482 с.

21. Прэтт Э. Цифровая обработка изображений: Пер. с англ, - М.: Мир, 1982.—Кн.1-312 с., ил.

22. Преобразование Хафа (Hough transform). – Режим доступа: <http://cgm.computergraphics.ru/content/view/36>. – Загл. с экрана.

23. Рябченко И.К. Программный комплекс для обработки потока изображений треков частиц в плазменной струе: МР 230100.503М.05 ПЗ. Алтайский государственный университет, Барнаул, 2012.- 85 с.: илл.

24. Определение контура объекта на примере изображения клеток. – Режим доступа: <http://masters.donntu.org/2005/kita/tribrat/diss/index.htm>. – Загл. с экрана.
25. Максфилд К. Проектирование на ПЛИС. Курс молодого бойца. – М.: Издательский дом «Додэка-XXI», 2007. – 408 с.: илл.
26. N.N.Aizenberg, I.N.Aizenberg "CNN-like networks based on multi-valued and universal binary neurons: learning and application to image processing", Proc. of the Third IEEE International Workshop on Cellular Neural Networks and their Applications, Rome, pp.153-158, 1994
27. C. Rekeczky, T. Roska and A. Ushida "CNN Based Self-Adjusting Nonlinear Filters" Proc. of the Fourth IEEE International Workshop on Cellular Neural Networks and their Applications, Seville, Spain, pp. 309-314, 1996.
28. Aizenberg I.N. "Multi-valued non-linear filters and their implementation on cellular neural networks", Frontiers in Artificial Intelligence and Applications, vol. 41 "Advances in Intelligent Systems" (F.C. Morabito - Ed.) IOS Press, Amsterdam-Berlin-OxfordTokyo-Washington DC, 1997, pp. 135-140.
29. N.N. Aizenberg, I.N. Aizenberg, G.A. Krivosheev "Multi-Valued Neurons: Mathematical model, Networks, Application to Pattern Recognition", Proc. of the 13 Int.Conf. on Pattern Recognition, Vienna, August 25-30, 1996, Track D, IEEE Computer Soc. Press, pp. 185-189, 1996.
30. Н.Н. Айзенберг, И.Н. Айзенберг, Г.А. Кривошеев. Нейросети на многозначных нейроэлементах: обучение и применение в обработке и распознавании изображений // Компьютерная оптика, 1995, Том 14-15, часть 1, стр. 179-186.
31. Основи розпізнавання образів. Розпізнавання мовних сигналів: Навч. посібник. Автори Работягов А. В. Д. 2016