

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЧОРНОМОРСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ПЕТРА
МОГИЛИ

Шеремет Костянтин Олександрович

УДК 004.8

**СИСТЕМА РОЗПІЗНАВАННЯ ТА ПЕРЕТВОРЕННЯ СИМВОЛІВ В
ТЕКСТ НА ОСНОВІ НЕЙРОННОЇ МЕРЕЖІ**

Галузь знань 12 «Інформаційні технології» за спеціальністю

122 «Комп'ютерні науки та інформаційні технології»

122 - ДР.А – 403.21610425

Автореферат

дипломної роботи на здобуття освітньої кваліфікації

«бакалавр комп'ютерних наук та інформаційних технологій»

Миколаїв – 2020

Дипломна робота є рукопис.

Робота виконана в Чорноморському національному університеті імені Петра Могили Міністерства освіти і науки України на кафедрі Інтелектуальних інформаційних систем

Науковий керівник: к.т.н., доцент, доцент каф. ПС О.В. Козлов

Рецензент: ст. викладач кафедри ІПЗ М.Л. Дворецький

Захист відбудеться «22» червня 2020 р. о 10³⁰ год. на засіданні екзаменаційної комісії (ауд. 2-406) у Чорноморському національному університеті імені Петра Могили за адресою: 54003, м. Миколаїв, вул. 68-ми Десантників, 10.

З дипломною роботою можна ознайомитись в бібліотеці Чорноморського національного університету імені Петра Могили за адресою: 54003, м. Миколаїв, вул. 68-ми Десантників, 10.

Автореферат представлений «___» червня 2020 р.

Секретар
екзаменаційної комісії,
Викладач

Скакодуб О.С.

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Актуальність теми визначається швидким розвитком систем штучного інтелекту та комп'ютерного зору. Машинне навчання, розпізнавання об'єктів і комп'ютерний зір є одними з найбільш перспективних в подальшому розвитку сучасної кібернетики. Розпізнавання образів використовується для ідентифікації і розпізнавання образів в даних, наприклад переводу печатного і рукописного тексту в формат електронного документу, виділення і розпізнавання нечітких або спотворених символів на фотографії, таких як автомобільні номери, розпізнавання тексту в робототехніці для прийняття рішень. Розпізнавання образів, в тому числі і символів, є на сьогоднішній день однією з актуальних задач, що виникають в різних областях людської діяльності.

Метою дипломної роботи є розробка та дослідження системи для розпізнавання символів та перетворення їх в текст.

Практичне значення отриманих результатів. За результатом отриманих практичних значень було створено систему для розпізнавання символів та перетворення їх в текст.

Структура дипломної роботи. Пояснювальна записка до дипломної роботи складається із вступу, 4 розділів, висновків, додатків. Загальний обсяг роботи складає 96 сторінок, 32 рисунків, 15 таблиць та 44 посилань на літературні джерела.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** було розглянуто актуальність даної теми, а також описані об'єкт, предмет і мета дослідження. Недоліки та проблеми візуальної класифікації. Важкості під час розпізнавання образу, які ускладнюють виконання поставленої задачі. Приклади проблем та недоліків при обробці зображення та можливі шляхи вирішення за допомогою різних методів та підходів створення нейронних мереж. Обрано кроки для досягнення поставленої мети. Було визначено практичну цінність роботи.

У **першому розділі** був проведений аналіз підходів до проектування штучних нейронних мереж. Розглянуто основні положення розпізнавання об'єкта методами машинного навчання, структуру штучного нейрона та нейронної мережі. Досліджено функції активації які визначають функціональні можливості нейронної мережі і метод навчання цієї мережі. Також було розглянуто зародження штучних нейронних мереж, основні архітектури штучних нейронних мереж призначених для розпізнавання і класифікації об'єктів та аналоги, адже на сьогоднішній день існує велика кількість методів та запропоновано безліч алгоритмів рішення задачі розпізнавання символів. Однак більшість ідей поступаються у швидкодії, простоті та точності результату штучним нейронним мережам. Розглянуто методології моделювання штучних нейронних мереж і використання їх у різних сферах життя, описано зародження штучних нейронних мереж. Було визначено постановку задачі.

У **другому розділі** були розглянуті моделі, методи та інструменти, що використовуються при розпізнаванні символів. Проводиться порівняльний аналіз різних архітектур в залежності від типу поставленого завдання, вибір більш оптимальної для вибраної цілі. Розглянуто рекурентні нейронні мережі з довгою короткочасною пам'яттю та згорткові нейронні мережі для розпізнавання і класифікації об'єктів з зображення. Досліджено метод зворотного поширення помилки, що широко застосовується в сучасних

штучних нейронних мережах, в якому використовується похідна від функції активації і який використовується при автоматичному оновленні ваг багатошарового перцептрона. Також відбувається описування різних шляхів покращення структури нейронної мережі. При аналізі шляхів покращення було виявлено, що не всі методи підходять для вирішення поставленої задачі. Було визначено, найкращим з методів шляху покращення архітектури мережі є batch-нормалізація. Також необхідно використовувати метод max pooling та метод зворотного поширення помилки для навчання нейромережі.

У третьому розділі були розглянуто та визначено необхідні інструменти такі як OpenCV і Tesseract та їх принцип роботи для виконання поставленої задачі. Було обрано середовище Visual Studio для розробки програмного застосунку та мову програмування C#. Використано необхідні traineddata для поставленої задачі та вбудовано модель в програмний застосунок. Оформлення інтерфейсу зрозумілого і зручного для користувача, а також розробка й налаштування програмного забезпечення. Проведено тестування створеного застосунку на різних зображеннях.

У розділі з охорони праці були розглянуті безпечні та здорові умови праці на робочому місці програміста та у робочому приміщенні в цілому, шляхом визначення оптимальних умов праці програміста, розрахунку освітленості приміщення, організації робочого місця програміста. Першим ділом описано приміщення де людина буде проводити роботу над створенням дипломної роботи відповідної установи (компанії, підприємстві тощо). Для даної роботи достатньо однієї людини, для якої надано робоче місце з стаціонарним комп'ютером. Були викладені вимоги до робочого місця розробника програмного забезпечення. Відбувся аналіз умов праці в розглянутому робочому приміщенні, напруженості праці користувача ПЕОМ, час регламентованих перерв при роботі на комп'ютері, шкідливі чинники, що діють на працівника при роботі за комп'ютером, який показав, що умови праці з ПЕОМ відповідають вимогам. Була наведена схема, розміри приміщення та наведено значення температури, вологості й рухливості

повітря, необхідна кількість і потужність ламп та інші параметри, значення яких впливає на умови праці робітника, та знайдено рішення для зменшення впливу часу перебування перед комп'ютером на зір працівників. Визначено норми мікроклімату робочої зони об'єкту та норми подачі свіжого повітря в приміщення, де розташовані комп'ютери. Розглянуто ергономіку робочого місця та характеристики робочого місця. В результаті проведеної роботи було зроблено аналіз умов праці, шкідливих та небезпечних чинників, з якими стикається робітник. Було визначено параметри і певні характеристики приміщення для роботи над запропонованим проектом написаному в дипломній роботі, описано які заходи потрібно зробити для того, щоб дане приміщення відповідало необхідним нормам і було комфортним і безпечним для робітника. Приведені рекомендації щодо організації робочого місця.

ЗАГАЛЬНІ ВИСНОВКИ

В даній дипломній роботі розроблено систему розпізнавання та перетворення символів в текст. Під час виконання дипломної роботи було розглянуто та проаналізовано різні підходи до проектування штучних нейронних мереж, модель рекурентної нейронної мережі з довгою короткочасною пам'яттю, методи та моделі створення архітектури згорткової нейронної мережі, досліджено інструменти для розпізнавання об'єкту на зображенні. Обґрунтовано актуальність використання машинного навчання та систем штучного інтелекту.

Практична цінність роботи полягає в тому, що архітектура, яка призначається для вирішення поставленої задачі є універсальною для багатьох наборів даних зображень символів і має кращі показники продуктивності, такі як затрати пам'яті, час навчання мережі, а головне вище значення точності розпізнавання об'єктів. Велику увагу приділено підходу до розпізнавання символів за допомогою штучних нейронних мереж та методів їх навчання.

Проаналізувавши наявні штучні нейронні мережі було визначено, що для аналізу зображень і їх класифікації підходять п'ять типів мереж: перцептрон, згорткові нейронні мережі, рекурентні нейронні мережі, мережі адаптивного резонансу, мережа радіально-базисних функцій, з яких тільки дві краще підходять для створення власної моделі, це згорткові нейронні мережі та рекурентні нейронні мережі, а саме нейронна мережа з довгою короткочасною пам'яттю. Основним критерієм відбору при обранні нейромережі для створення власної навченої моделі є точність та швидкість навчання.

В роботі розглянуті загальні теоретичні принципи роботи рекурентної нейронної мережі з довгою короткочасною пам'яттю та згорткової нейронної мережі, методи та шляхи покращення структури нейромережі. При аналізів шляхів покращення було виявлено, що не всі методи підходять для

вирішення поставленої задачі. Було визначено, найкращим з методів шляху покращення архітектури мережі є batch-нормалізація. Також необхідно використовувати метод max pooling та метод зворотного поширення помилки для навчання нейромережі.

Методи розпізнавання об'єктів можуть бути застосовані до широкого кола проблем. Користувачами системи розпізнавання об'єктів може бути як і звичайна людина для задоволення своїх потреб, наприклад, розпізнавання тварин, квітів, рукописного тексту, так і великі підприємства, які можуть використовувати систему у комерційних цілях, таких як розпізнавання якості ресурсів, дорожніх знаків, стану обладнання, аналізу знімків медичних карт. Навіть в масштабах цілої країни система може використовуватися для розпізнавання обличчя в громадських місцях або дорожніх правопорушень.

Під час виконання дипломної роботи було розглянуто та проаналізовано різні підходи до проектування штучних нейронних мереж, модель рекурентної нейронної мережі з довгою короткочасною пам'яттю, методи та моделі створення архітектури згорткової нейронної мережі, досліджено інструменти для розпізнавання об'єкту на зображенні.

Для створення системи було обрано середовище Visual Studio та мову програмування C#. Досліджено принцип роботи розпізнавання та класифікації об'єктів такими інструментами як OpenCV і Tesseract і використано їх для зручного розпізнавання символів.

Одержані результати мають практичне значення і можуть використовуватися для комерційних підприємств з надання послуг розпізнавання тексту або для звичайної людини, яка хоче не витратити багато часу для перенесення тексту із зображення в текстовий вигляд.

Цінність роботи полягає в тому, що створена система є універсальною для багатьох наборів даних зображень та гарно справляється з вирішенням поставленої задачі, має не погані показники продуктивності.

АНОТАЦІЯ

Шеремет К. О. Система розпізнавання та перетворення символів в текст на основі нейронної мережі. – На правах рукопису.

Дипломна робота на здобуття освітньої кваліфікації «бакалавр комп'ютерних наук та інформаційних технологій» в галузі знань 12 «Інформаційні технології» за спеціальністю 122 «Комп'ютерні науки та інформаційні технології»

Чорноморський національний університет імені Петра Могили,
Миколаїв

Дипломна робота присвячена розробці програмного застосунку на основі нейронної мережі для розпізнавання та перетворення символів в текст.

Об'єкт дослідження – система обробки зображень з метою виділення, класифікації та подальшого використання текстової інформації.

Предмет дослідження – методи, засоби та алгоритми машинного навчання для розпізнавання тексту.

Метою дипломної роботи є розробка та дослідження системи для розпізнавання символів та перетворення їх в текст.

В роботі описано основні принципи проектування штучних нейронних мереж та інструменти для вирішення задач розпізнавання та класифікацій зображень. Спираючись на аналіз різних архітектур нейронної мережі, було обрано й оптимізовано власну модель, в основі якої знаходиться композиція згорткової нейронної мережі. Згорткова нейронна мережа є багаторівневою архітектурою, що самостійно визначає найважливіші характеристики на вхідних даних.

Дипломна робота складається з фахового розділу і спеціальної частини з охорони праці. Пояснювальна записка дипломної роботи складається зі вступу, трьох розділів, висновків та додатків.

В першому розділі розглядаються теорія розпізнавання об'єкта методами машинного навчання, основні архітектури та методології моделювання штучних нейронних мереж та використання їх у різних сферах життя.

В другому розділі проводиться порівняльний аналіз різних архітектур в залежності від типу поставленого завдання, вибір більш оптимальної для вибраної цілі. Також відбувається описування різних шляхів покращення структури нейронної мережі.

У третьому розділі розглядаються необхідні інструменти та їх принцип роботи для виконання поставленої задачі, а також розробка й налаштування програмного забезпечення.

Дипломна робота містить: сторінок – 92, рисунків – 32, таблиць – 15, джерел –44.

Ключові слова: комп'ютерний зір, рекурентна нейронна мережа, згорткова нейронна мережа, Tesseract, розпізнавання символів.

ABSTRACT
to bachelor's thesis

Topic: "Neural network based character recognition and conversion system"

Student: Sheremet Kostiantyn Oleksandrovykh

Head: Candidate of Technical Sciences, Associate Professor, Associate
Professor of the Department IIS Kozlov Oleksii Valeriiovych

Thesis is devoted to the development of a software application based on a neural network for recognizing and converting characters into text.

The object of research – system for image processing in order to select, classify and further use textual information.

The subject of research methods, tools and algorithms of machine learning for text recognition.

The purpose of the thesis is development and research of the system for character recognition and conversion into text.

The paper describes the basic principles of designing artificial neural networks and tools for solving problems of image recognition and classification. Based on the analysis of different neural network architectures, we chose and optimized our own model, which is based on the composition of the convolutional neural network. The convolutional neural network is a multilevel architecture that independently determines the most important characteristics on the input data.

Thesis consists of a professional section and a special part on labor protection. The explanatory note of the thesis consists of an introduction, three sections, conclusions and appendices.

The first section discusses the theory of object recognition by machine learning methods, the basic architectures and methodologies for modeling artificial neural networks and their use in various spheres of life.

The second section provides a comparative analysis of different architectures depending on the type of task, choosing the most optimal for the selected purpose. There are also descriptions of various ways to improve the structure of the neural network.

The third section discusses the necessary tools and their principle of operation to perform the task, as well as software development and configuration.

Thesis contains: pages - 92, figures - 32, tables - 15, sources - 44.

Keywords: computer vision, recurrent neural network, convolutional neural network, Tesseract, character recognition.