

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
ЧОРНОМОРСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ПЕТРА МОГИЛИ

**Шеремет Анастасія Олександрівна**

УДК 004.8

**СИСТЕМА РОЗПІЗНАВАННЯ ТА КЛАСИФІКАЦІЇ ОБ'ЄКТІВ НА ОСНОВІ  
MICROSOFT COGNITIVE TOOLKIT**

Напрямок підготовки 6.050101 – «Комп'ютерні науки»

ДР.ПЗ-401.10790908

Автореферат

дипломної роботи на здобуття освітньої кваліфікації

«Бакалавр комп'ютерних наук»

Миколаїв – 2019

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: кандидат технічних наук, доцент кафедри інтелектуальних інформаційних систем Кондратенко Галина Володимирівна.

Рецензент: старший викладач кафедри комп'ютерної інженерії, Бурлаченко Іван Сергійович.

Захист відбудеться «21» червня 2019 р. о 9<sup>30</sup> год. на засіданні екзаменаційної комісії (ауд. 2-403) у Чорноморському національному університеті імені Петра Могили за адресою: 54003, м. Миколаїв, вул. 68-ми Десантників, 10.

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

Автореферат представлений «20» червня 2019 р.

Секретар  
екзаменаційної комісії,  
ст.викл.



С.В.Дворецька

## ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

**Актуальність теми.** Останнім часом все більш популярними стають системи комп'ютерного зору. Розпізнавання образів застосовується для виявлення і розпізнавання закономірностей в даних. Моделі розпізнавання образів приймають зображення або шаблони в якості вхідних даних і видають вихідні значення, звані прогнозами. Комп'ютерний зір охоплює отримання цифрового зображення, його обробку, аналіз і розуміння зображень, використовуючи при цьому статистичні методи та моделі побудовані за допомогою фізики, статистики, геометрії та теорії статистичного навчання. Необхідність розпізнавання і класифікації візуальних об'єктів впливає з самої головної мети комп'ютерного зору – створювати системи сприйняття, подібні системі людського зору. Завдання розпізнавання об'єктів можуть бути застосовані до широкого кола проблем. Протягом останніх десятиліть комп'ютерний зір розвивався дуже швидко, і в наш час він використовується в різних областях, таких як робототехніка, автономні транспортні засоби, камери спостереження за рухом, медицина та багато інших. Здатність розпізнавати об'єкти сучасних нейронних мереж в деяких конкретних завданнях досягла майже людських можливостей.

Використання згорткових нейронних мереж необмежене у різних сферах життя від розпізнавання дорожніх знаків до анотацій захворювань по рентгенограмам. І з кожним днем виникають все нові потреби у згорткових нейронних мережах. Це дуже цікаве і популярне направлення з великими перспективами у майбутньому, які навіть складно уявити у нас час.

**Метою дипломної роботи** розробка, проектування і тестування вдосконаленої згорткової нейронної мережі для полегшення процедури розпізнавання і класифікації образів.

Для досягнення поставленої мети потрібно виконати такі завдання:

- дослідити предметну область;
- провести аналіз наявних методів та інструментів;

- розробити архітектуру мережі та навчити її розпізнавати та класифікувати об’єкти;
- розробити програмний застосунок.

При цьому визначені деякі вимоги, які повинна задовольняти розроблена система. Тож система повинна:

- мати високу точність розпізнавання;
- бути зрозумілою користувачу;
- бути більш ефективною ніж вже існуючі рішення;
- бути оптимізованою для швидкодії обчислювальних процесів.

Проаналізувавши предметну область і підходи до проектування нейронної мережі було зроблено постановку задачі. Задача полягає в розробці архітектури нейронної мережі, яка буде розпізнавати та класифікувати об’єкти на зображенні. Проаналізувати методи і шляхи покращення архітектури задля підвищення класифікаційної точності. Навчити готову модель на обраному наборі даних.

Створити програмний застосунок, який буде використовувати навчену модель штучного інтелекту, що має класифікаційну точність більше 90%.

Програма має надавати змогу користувачу завантажувати зображення для автоматичного розпізнавання. У графічному вигляді виводити інформацію у відсотках щодо належності завантажувального зображення до певного класу.

Зображення, що користувач хоче класифікувати повинно відображатися на головному вікні програмного застосунку.

Головне вікно програми повинно містити кнопку для завантаження зображення, яка буде називатися «Огляд». Після обрання файлу зображення буде відбуватися автоматичний процес розпізнавання об’єкту і його подальша класифікація.

Результат виконання нейронної мережі буде виведений на головне вікно у вигляді графіку належності об’єкту на зображенні до класів. Кожний рядок діаграми відобразатиме відсоток належності до певного класу, при наведенні на який буде відображатися назва класу. Вгорі має знаходитися клас до якого

об'єкт належить з більшим відсотком, нижче – класи по зменшенню відсотка. Всього на вікно буде виведено не більше 5 класів.

**Практична цінність** роботи полягає в тому, що архітектура, яка призначається для вирішення поставленої задачі є універсальною для багатьох наборів даних зображень і має вищі показники продуктивності, такі як затрати пам'яті, час навчання мережі, кількість параметрів при розпізнаванні, а головне краще значення точності розпізнавання об'єктів у порівнянні з уже наявними моделями згорткових нейронних мереж з тією ж кількістю згорткових шарів. Велику увагу приділено підходу до розпізнавання образів за допомогою штучних нейронних мереж та методів їх навчання.

**Структура дипломної роботи.** Пояснювальна записка до дипломної роботи складається із вступу, 4 розділів, висновків, додатків. Загальний обсяг роботи складає \_\_\_ сторінок, \_\_\_ рисунків, \_\_\_ таблиць та \_\_\_ посилання на літературні джерела.

## **ОСНОВНИЙ ЗМІСТ РОБОТИ**

**У вступі** подано загальну характеристику досліджуваної теми, обґрунтовано актуальність дипломної роботи, сформульовано мету, завдання розробки, визначено основні етапи роботи над дипломним проектом.

**У першому розділі** було розглянуто теорію розпізнавання об'єкту, різні архітектури ШНМ та наявні на міжнародному ринку аналоги. Адже на сьогоднішній день існує велика кількість методів та запропоновано безліч алгоритмів рішення задачі розпізнавання об'єктів. Однак більшість ідей поступаються у швидкодії, простоті та точності результату штучним нейронним мережам. В даний час найбільш успішними моделями візуального розпізнавання є глибокі нейронні мережі.

Глибокі нейронні мережі – це нейронні мережі, що складаються з декількох шарів. Їх глибина дозволяє їм вивчати більш глибоке уявлення даних, що призводить до приголомшливої продуктивності в порівнянні з іншими методами машинного навчання. Більшість сучасних нейронних мереж

побудовані на основі архітектури згорткового типу, як правило, таких як когнітрон і неокогнітрон.

Внаслідок дедалі більшого розвитку глибоких згорткових мереж здобуваються суттєві досягнення у розпізнаванні об'єктів. Ефективність і стрімкий розвиток зумовлено гібридним підходом до архітектурних рішень, розвитком методів навчання, додаткових методів захисту від перенавчання.

Також у розділі було розроблено постановку задачі.

**У другому розділі** було виконано більш детальний аналіз поставленої задачі з точки зору її рішення та розглянуто основні математичні методи її вирішення. Огляд існуючих методів та бібліотек для покращення роботи нейромережі, показав, які з них дають гарний результат при навчанні моделі.

Аналіз показав, що не всі методи підходять для вирішення поставленої задачі, деякі з них замість покращення результату роботи архітектури погіршили її працездатність. Найкраще для навчання моделі підходить CNTK. Бібліотека надає безліч добре оптимізованих компонентів, ефективна при використанні ресурсів і показала хорошу продуктивність під час розробки програмного застосунку. При використанні для поставленого завдання Keras і PyTorch не так функціональні як TensorFlow і CNTK, і дають менше опцій для управління мережевими з'єднаннями. CNTK має кращу інтеграцію з додатком C# ніж Tensorflow.

У розділі проходив аналіз і порівняння хмарних віртуальних машин, за результатами якого було обрано віртуальну машину Standard D4 v2 компанії Azure. Вибір зупинився на ній адже вона найкраще підходить через наявність простої інтеграції з CNTK. І по цінній характеристиці вона є достатньо оптимальною

Проаналізувавши альтернативи було визначено, що найбільш оптимальною типом мережі є згорткова штучна нейрона мережа, адже вона краще відповідає вимогам. Основним критерієм відбору при обранні нейромережі для поставленої задачі була точність. Інші нейронні мережі поступаються в ефективності для поставленого завдання. Перцептрон має

нижчу точність ніж ЗНМ через відсутність згорткових шарів. Для навчання мережі радіально-базисних функцій необхідно великий обсяг даних, що значно збільшує ресурсозатратність, але при цьому не дає більшу точність ніж ЗНМ. Основним недоліком мережі адаптивного резонансу є недостатня надійність збереження інформації. Так, в разі «втрати» одного образу руйнується вся пам'ять.

Також була виявлена необхідність використовувати метод max pooling та зворотного поширення помилки для навчання нейромережі

Найкращим з методів шляху покращення архітектури мережі було обрано batch-нормалізація. Порівняння шляхів покращення відбувалося методом тестування їх у розробленій архітектурі. Пакетна нормалізація допомогла підвищити точність навченої моделі при цьому не збільшивши час навчання нейронної мережі.

**У третьому розділі** детально розглянуто створений відповідно до поставленої мети програмний застосунок, надано пояснення та продемонстровано скріншоти кожного кроку тестування роботи зі створеною системою. Також було описано основні можливості системи, основні обмеження та продемонстровано працездатність застосунку і результати його використання на прикладі тестових даних, підібраних близько до реальних.

**У четвертому розділі** були викладені вимоги до робочого місця розробника програмного забезпечення. Відбувся аналіз умов праці в розглянутому робочому приміщенні 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. Запропоновані світлодіодні світильники мають строк служби 50 тисяч годин, що значно краще ніж у люмінесцентних ламп, де строк рівний 10 - 20 тисяч годин, і крім того залежить від кількості переключень. Ці лампи створюють оптимальні умови для зорової роботи інженера-програміста, а порівняно невисока температура нагрівання підвищує рівень пожежної безпеки.

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 дипломній роботі розглянуто різноманітні підходи до вирішення задачі розпізнавання образів, наведені результати розробленої комп'ютерної системи, вказані переваги та недоліки модифікованої нейронної мережі. Запропоновано архітектуру згорткової нейронної мережі, що формує нейромережеву систему розпізнавання об'єктів на зображенні, використовуючи власний підхід до класифікації.

Усі поставлені завдання на початку дипломної роботи були успішно виконані. Результати роботи розробленого застосунку цілком задовільні. Архітектура згорткової штучної нейронної мережі була добре оптимізована, пакетна нормалізація покращила точність навчання завдяки чому навчена модель розпізнає і класифікує об'єкти на зображенні з точністю навіть більшою ніж передбачалося.



## АНОТАЦІЯ

**Шеремет Анастасія Олександрівна.** Система розпізнавання та класифікації об'єктів на основі Microsoft Cognitive Toolkit. – На правах рукопису.

Дипломна робота на здобуття освітньої кваліфікації «Бакалавр комп'ютерних наук». – Чорноморський національний університет імені Петра Могили, Миколаїв, 2019.

Дипломна робота присвячена розробці архітектури й проектуванню моделі нейронної мережі для розпізнавання і класифікації об'єктів та подальшого впровадження в програмний застосунок.

В роботі описано основні принципи проектування штучних нейронних мереж для вирішення задач класифікації зображень. Спираючись на аналіз різних архітектур нейронної мережі, було створено 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.

Фахова частина включає вступ, три розділи, висновки та додатки до дипломної роботи. Спеціальна частина включає розділ про охорону праці та безпеку на підприємстві.

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 відбір фреймворку, хмарної віртуальної машини та шляху покращення структури нейронної мережі. В результаті аналізу було обрано архітектуру згорткові нейронні мережі з пакетною нормалізацією, яка буде реалізована за допомогою фреймворку Microsoft Cognitive Toolkit.

У третьому розділі розглядається функціональна структура розробленої системи, а також розробка й налаштування програмного забезпечення.

Ключові слова: комп'ютерний зір, згорткова нейронна мережа, CNTK, розпізнавання об'єкта.

## ABSTRACT

**Sheremet Anastasia Aleksandrovna.** The system of recognition and classification of objects based on Microsoft Cognitive Toolkit. – On the rights of the manuscript.

Diploma work on obtaining an educational qualification "Bachelor of Computer Science". - Black Sea National University named after Petro Mohyla, Nikolaev, 2019.

Thesis is devoted to the development of architecture and design of the model of the neural network for the recognition and classification of objects and subsequent implementation in the software application.

The paper describes the basic principles of designing artificial neural networks for solving problems of image classification. Based on the analysis of various architectures of the neural network, an own model was created and optimized, based on the composition of the convolutional neural network. The rolling neural network is a multi-level architecture that independently defines the most important characteristics of the input data.

The professional part includes an introduction, three sections, conclusions and appendices to the thesis. The special section includes a section on occupational safety and security at the enterprise.

The first chapter deals with the theory of object recognition by methods of machine learning, basic architectures and methodologies for the simulation of artificial neural networks and their use in various spheres of life.

In the second section a comparative analysis of different architectures is carried out, depending on the type of task, the choice is more optimal for the selected target. In addition, there is a selection of the framework, a cloud virtual machine, and a way to improve the structure of the neural network. As a result of the analysis, the architecture of the convolutional neuron of the network with packet normalization, which will be implemented using the framework Microsoft Cognitive Toolkit.

The third section deals with the functional structure of the developed system, as well as the development and configuration of software.

Key words: computer vision, convolutional neural network, CNTK, object recognition.