Кондратюк Сергій Сергійович, асистент кафедри теоретичної кібернетики, Київський національний університет імені Тараса Шевченка. Назва дисертації: &laquo;Моделювання та розпізнавання жестів української дактильної абетки за допомогою кросплатформених технологій&raquo;. Шифр та назва спеціальності 01.05.03 математичне та програмне забезпечення обчислювальних машин і систем. Спецрада Д26.001.09 Київського національного університету імені Тараса Шевченка

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

ІМЕНІ ТАРАСА ШЕВЧЕНКА

Кваліфікаційна наукова

праця на правах рукопису

Кондратюк Сергій Сергійович

УДК 004.912

ДИСЕРТАЦІЯ

МОДЕЛЮВАННЯ ТА РОЗПІЗНАВАННЯ ЖЕСТІВ УКРАЇНСЬКОЇ

ДАКТИЛЬНОЇ АБЕТКИ ЗА ДОПОМОГОЮ

КРОСПЛАТФОРМЕНИХ ТЕХНОЛОГІЙ

Спеціальність 01.05.03 – математичне та програмне забезпечення

обчислювальних машин і систем

Подається на здобуття наукового ступеня кандидата технічних наук

Дисертація містить результати власних досліджень. Використання ідей,

результатів і текстів інших авторів мають посилання на відповідне джерело.

С. С. Кондратюк

Науковий керівник: Крак Юрій Васильович,

член-кореспондент НАНУ,

доктор фізико-математичних наук, професор

Київ – 2021

Зміст

АНОТАЦІЯ....................................................................................................................................... 2

Список опублікованих праць за темою дисертації ................................................................... 5

Зміст ................................................................................................................................................... 9

Перелік умовних скорочень ........................................................................................................ 11

Список рисунків ............................................................................................................................ 12

Список таблиць.............................................................................................................................. 14

Вступ ................................................................................................................................................ 15

Розділ 1. Огляд підходів до моделювання та розпізнавання жестової мови...................... 22

1.1. Огляд досліджень з побудови кросплатформеної інформаційної технології для

моделювання та розпізнавання жестової мови......................................................................... 22

1.1.1. Модель скелету руки для демонстрації жестової інформації ................................... 22

1.1.2. Підходи, які базуються на 27 ступенях свободи моделі руки людини .................... 28

1.2. Підходи до моделювання та аналізу рухів руки людини ................................................. 28

1.2.1. Моделювання руки........................................................................................................ 32

1.3. Підходи до розпізнавання дактилем................................................................................... 34

1.3.1. Концепції обробки зображень...................................................................................... 39

1.3.2. Підходи машинного навчання для класифікації зображень...................................... 41

1.4. Використання нейромережевих технологій для задач розпізнавання ........................... 42

1.5. Кросплатформені засоби розробки технології моделювання та розпізнавання жестів. 44

1.6. Висновки до Розділу 1.......................................................................................................... 45

Розділ 2. Моделі та методи опису тривимірного скелету та моделі руки, моделювання та

розпізнавання жестів за допомогою глибокого навчання..................................................... 47

2.1. Моделювання скелету руки людини................................................................................... 47

2.1.1. Модель скелету руки для моделювання жестової інформації .................................. 47

2.1.2. Обмеження суглобів скелету........................................................................................ 48

2.2. Моделювання руки............................................................................................................... 50

2.2.1. Моделювання жестів за допомогою просторової моделі руки ................................. 50

2.2.2. Параметрична модель руки .......................................................................................... 52

2.2.3. Модель з текстурою ...................................................................................................... 54

2.3. Адаптивність моделі руки ................................................................................................... 56

2.4. Математична формалізація процесу подання даних для розпізнавання дактилем........ 57

2.4.1. Просторові дескриптори ............................................................................................... 59

2.4.2. Попередня обробка даних............................................................................................. 60

2.4.3. Просторово-часове подання даних .............................................................................. 61

2.5. Модель розпізнавання на основі глибокої згорткової нейронної мережі....................... 64

2.5.1. Математична формалізація структури нейронної мережі......................................... 65

2.5.2. Використання даних у просторово-часовому вимірі ................................................. 67

2.5.3. Удосконалення моделі нейронної мережі тривимірними згортками для

використання даних у просторово-часовому вимірі............................................................ 67

2.5.4. Тренування нейронної мережі...................................................................................... 70

2.5.5. Проблема перенавчання та трансферу навчання........................................................ 72

2.6. Висновки до Розділу 2.......................................................................................................... 74

Розділ 3. Кросплатформена інформаційна технологія моделювання та розпізнавання

жестів за допомогою тривимірних згорток .............................................................................. 77

3.1. Інфологічна модель .............................................................................................................. 77

3.2. Моделювання ........................................................................................................................ 81

3.2.1. Аналіз та завантаження жестових конфігурацій ........................................................ 81

3.2.2. Побудова тривимірної сцени з моделлю руки заданої полігональності.................. 82

3.2.3. Анімація жестів.............................................................................................................. 83

10

3.3. Розпізнавання (тренування)................................................................................................. 83

3.3.1. Збір набору даних.......................................................................................................... 83

3.3.2. Обробка даних ............................................................................................................... 83

3.3.3. Розширення (аугментація) даних та розподіл............................................................. 85

3.3.4. Задання набору конфігурацій архітектур.................................................................... 87

3.4. Розпізнавання (прогнозування)........................................................................................... 89

3.4.1. Виділення вхідних даних.............................................................................................. 89

3.4.2. Обробка даних ............................................................................................................... 89

3.4.3. Розпізнавання та отримання результатів..................................................................... 90

3.5. Кросплатформена розробка ................................................................................................. 91

3.5.1. Моделювання руки........................................................................................................ 91

3.5.2. Розпізнавання жестів..................................................................................................... 92

3.6. Вдосконалена для розпізнавання дактилем архітектура MobileNetv2 ............................ 93

3.7. Висновки до Розділу 3.......................................................................................................... 94

Розділ 4. Експериментальна програмна реалізація та результати тренувань і тестувань

моделей розпізнавання дактилем............................................................................................... 96

4.1. Користувацький інтерфейс .................................................................................................. 96

4.2. Структура модулів і класів програмної реалізації ............................................................ 97

4.2.1. Структура модулів програмної реалізації................................................................... 97

4.2.2. Структура класів програмної реалізації ...................................................................... 98

4.2.3. Структура бази даних.................................................................................................. 100

4.3. Набір даних ......................................................................................................................... 101

4.4. Експериментальні тренування та досліди навченої моделі............................................ 104

4.5. Вибір архітектури моделі................................................................................................... 105

4.6. Висновки до Розділу 4........................................................................................................ 110

Висновки по роботі...................................................................................................................... 111

Література..................................................................................................................................... 113

Додатки.......................................................................................................................................... 124

Висновкипороботі

Уданомудисертаційномудослідженнібулирозглянутіпитання

присвяченівирішеннюпроблеммоделюванняірозпізнаваннямімічноїскладової

проявівемоційВрезультатідослідженнябулозапропонованошляхивирішення

поставленоїзадачідисертаційногодослідженнящополягалиурозробці

математичнихмоделейметодівалгоритмівтаїхреалізаційзокрема

Створенотехнологіющоскладаєтьсяздвохосновнихмодулівмоделювання

жестівтарозпізнаванняжестівяківикористовуютьбазуданихіз

специфікаціямижестівщозберігаютьсяуформатіубазіданих



Створенотехнологіюякавпершереалізуємоделюванняжестівта

розпізнаванняжестівдляукраїнськихжестівдактильноїабеткизадопомогою

інструментівкросплатформеноїрозробкиМоделюванняжестівбуло

реалізованозадопомогоюфреймворкуякийєкросплатформенимта

демонструєзадовільнупродуктивністьнарізнихплатформахмобільнихвебта

настільнихдемонструючиреалістичнутривимірнумодельрукиКількість

полігонівтаанімаційнийкрокпереходівжестуможнарегулюватизадля

виконанняроботи

Буловпершезібранонабірданихізпонадзображеньвикористовуючи

різніумовитарукилюдейрізноговікутастатіНабірданихдоповненоза

допомогоюметодикаугментаціїакінцевийнабірданихскладаєтьсяз

зображеньМодульрозпізнаванняжестівбувреалізованийзадопомогою

фреймворкущозабезпечуєможливістьфункціонуваннямоделіна

різнихплатформахбезбудьякихмодифікаційбазиданихтамоделіабомодулю

длятренуваннямоделі

БуловдосконаленопідхіддорозпізнаваннядактилемВякостімоделідля

розпізнаванняжестівбулаобранаархітектураякмодельз

найкращимкомпромісомміжрозмірамитаточністюособливона

низькопродуктивнихплатформахтакихякмобільнітавебМодельбула

підготовленаназібраномунаборіданихукраїнськоїдактильноїмовиЗавдяки



вдосконаленнюзадопомогоютехнологіїтривимірнихзгортокмодельпоказала

ультрасучаснийрівеньякості

Проведеноекспериментальнівипробуваннянаосновіякихбулаобрана

оптимальнаархітектурамоделізметоюзбереженнянайкращогорівняшвидкодії

знайменшимможливимрозміроммоделі

Булипоказаніпорівняльнірезультатирізнихархітектурізтривимірними

згорткамитабезнихотриманівходіекспериментуПродуктивністьмоделі

порівнюваласязіншимипідходамитапродемонструвалааналогічніабовищі

значенняякіпотенційнозбільшуютьсяіззбільшеннямрозмірунаборуданих

Запропонованутехнологіюкомунікаціїжестамиможнадодатково

доповнитиіншимижестамимовамийіншимикросплатформенимимодулями

Результатидисертаційногодослідженнямаютьяктеоретичнетакі

практичнеспрямуванняРезультатиокремихелементівданогодослідженнябули

використанівнауковихдослідженняхвнауковихтемахізвітах

Значеннярезультатівдослідженняполягаєузастосовностіїхврізних

галузяхНайбільшочевиднимєвикористанняззалученнямтехнологій

розпізнаванняімоделюваннядактилемпринавчанніжестовоїмовилюдейякі

маютьвоточенніособузвадамислухудлярозробкиінформаційнихтехнологій

дляінклюзивноїосвітидляшкільнихідошкільнихзакладівізвивченням

жестовоїмовидлясамонавчаннябатьківдітейізвадамислухуабосоціальних

працівниківмедицинанаціональнаполіціятаіншіслужбищоконтактуютьіз

людьмизвадамислуху