

# РОЗРОБКА ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ЖЕСТОВОГО КЕРУВАННЯ ВЕБ-ДОДАТКОМ НА ОСНОВІ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ

Вінницький національний технічний університет

## *Анотація*

*Досліджується інтелектуальна технологія жестового керування веб-додатком з використанням веб-камери та вбудованим в браузер Web API. Даний додаток дасть змогу керувати веб-сторінкою за допомогою жестів рук без використання інших видів пристроїв вводу.*

**Ключові слова:** веб-додаток, жести, веб-камера.

## *Abstract*

*The intelligent technology of gesture control of a web application using a webcam and a built-in Web API browser is investigating. This application will allow controlling the web page with hand gestures without using other types of input devices.*

**Keywords:** web application, gestures, webcam.

## Вступ

На сьогоднішній день користувачі ПК звикли до того, що взаємодія з комп'ютером відбувається за допомогою клавіатури та миші. Але бувають випадки, коли не зручно використовувати те чи інше. Наприклад, під час доповіді чи презентації. Тому почали задумуватись, чи є ще якісь альтернативні методи взаємодії з комп'ютером. Одним із них є жестове керування [1]. Ні для кого не секрет, що під час спілкування, люди звикли жестикулювати, тим сам вони проявляють свої емоції та акцентують увагу на деталях. Також до прикладу соціум прийшов до того, що було створено мову жестів, для взаємодії з людьми з обмеженими можливостями. Так чому таким же чином людина не може взаємодіяти з комп'ютером?

Розпізнавання жестів може полегшити керування технікою, у тих випадках, коли складно використовувати звичайні пристрої вводу і в цьому нам може допомогти комп'ютерний зір.

Комп'ютерний зір це доволі новий, але стрімко розвиваючий напрям в системах штучного інтелекту. Він дає змогу аналізувати та знаходити на зображеннях та відео певні образи та елементи [2,3]. Все частіше для задач комп'ютерного зору застосовують штучні нейронні мережі [4,5].

## Постановка задачі

Останнім часом клавіатура та мишка відіграють значну роль у взаємодії людини і комп'ютера. Однак, завдяки швидкому розвитку апаратного та програмного забезпечення, потрібні нові типи методів людино-комп'ютерної взаємодії (ЛКВ). Зокрема, такі технології, як розпізнавання мови та розпізнавання жестів, притягують велику увагу у сфері ЛКВ. Комп'ютерне розпізнавання жестів руки може забезпечити більш природний інтерфейс ЛКВ. До прикладу інтерактивні комп'ютерні ігри можуть бути покращені, якщо комп'ютер міг би розуміти жести рук гравців. Розпізнавання жестів може бути навіть корисним для управління побутовою технікою.

Програмні продукти, у яких використовується розпізнавання жестів вузько спеціалізовані, де інтелектуальний модуль має тільки декілька шляхів взаємодії, можна привести такі відомі приклади:

LeapMotion – це невеликий USB-пристрій, який слід розташувати зверху, тим самим створюється віртуальна область взаємодії об'ємом близько 227 дециметрів кубічних, тобто абстрактний куб зі стороною 61 см. В середині цієї області The Leap відстежує рух пальців і рук, олівців, ручок, паличок для їжі з великою точністю. Судячи з відгуків аудиторії користувачів, є велика розбіжність в точності розпізнавання, деякі надавали результат 70%, а інші користувачі надсилали звіти з 95% точністю

розпізнавання жестів.

Іншим прикладом є TableFX - це технологія обробки поверхні столу, яка перетворює звичайну горизонтальну поверхню на інтерактивні стіл. Використовуючи передову технологію камери та запатентоване програмне забезпечення для відстеження тіла, TableFX створює динамічну інтерактивну таблицю з управління жестами. Як заявляє представник департаменту розробки, точність розпізнавання жестів пристрою сягає 80-85%.

Усі вищенаведені комерційні рішення у тій чи іншій формі стикаються із проблемою точності розпізнавання. Тому стоїть поставлена задача створення інформаційної технології що в результаті підвищить достовірність розпізнавання.

*Метою дослідження є розробка інформаційної системи для жестового керування веб-додатком з підвищеною достовірністю розпізнавання жестів за рахунок використання згорткових нейронних мереж..*

*Об'єктом дослідження є процес розпізнавання жестів.*

*Предметом дослідження є програмне забезпечення, що реалізовує процес розпізнавання жестів..*

## Результати дослідження

Згорткові нейронні мережі схожі до нейронних мереж прямого поширення, в яких нейрони мають тренувальну вагу та упередження. Цей тип мережі, як правило використовують для обробки зображень та сигналів.

В порівнянні з мережею прямого поширення, де кожен вхідний нейрон з'єднується з вихідним нейроном в наступному шарі, в згорткових мережах для одержання вихідних значень використовуються згортки над кожним вхідним шаром. Для операції згортки застосовується матриця ваг малого розміру, яка «пробігається» по всьому поточному шарі, створюючи після кожного зсуву сигнал активації для нейрона наступного шару з аналогічною позицією. Така матриця отримала назву - ядро згортки.

На рисунку 1 зображений, один із етап під час якого мережа аналізує зображення частинами. Тобто вхідний сигнал надається частково, як фільтр. Це дає змогу мережі запам'ятовувати зображення порціями та проводити необхідні розрахунки. Ці обрахунки включають в себе перетворення зображення з колірної моделі RGB або HSI в сірий масштаб. Згодом, відмінності в значеннях пікселів допоможуть визначити границі зображень, які можна буде класифікувати на різні категорії.

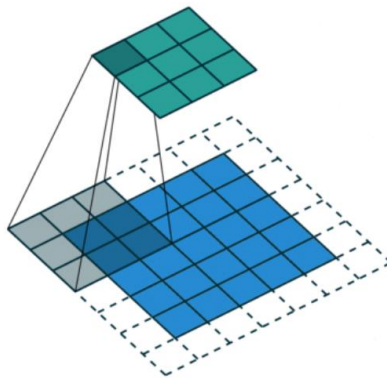


Рисунок 1 - Накладання фільтру згортки на зображення.

Архітектура згорткових мереж складається з різних шарів, що виконують перетворення вхідного об'єму даних. Кожен шар згорткової мережі складається з нейронів, з'єднаних з вузлами попередніх шарів, таким чином, вихід у даному вузлі для шару  $L$  є вихідною функцією вузлів у шарі  $L-1$ . На прикладі п'яти основних шарів даної мережі наведемо архітектуру, що пропонується для розгляду[6]:

1. Згортковий шар. Його параметри містять набори фільтрів (або ядер) для навчання мережі, які мають невелике рецептивне поле, що простягаються на всю глибину вхідного об'єму. Кожен фільтр здійснює згортку по ширині та висоті вхідного об'єму під час прямого проходу.

2. Шари підвибірки. Вони реалізовані нелінійними функціями для реалізації підвибірки, серед яких найпоширенішою є максимізаційна підвибірка. Вона розділяє вхідне зображення на набір прямокутників без перекриття, і для кожної такої підобласті виводить її максимум. Ідея створення цього шару полягає в тому, що якщо ознаку було знайдено, то її точне положення не так важливе, як її положення відносно інших ознак.

3. Шари активації - функції активації імітують поведінку аксона нейрона, що запускає сигнал при контакті з подразником. Ключовою особливістю згорткових мереж є функція активації ReLU.

4. Шари виключення, які генерують випадковий набір активацій, обнуляючи їх значення. Це означає, що мережа повинна надавати правильну класифікацію, навіть якщо деякий набір активацій виключено.

5. Повноз'язні шари слугують як провідники між шарами, де всі виходи попереднього шару пов'язані з усіма входами повноз'язних шарів.

На рисунку 2 наведено варіант архітектури, що спроектована на основі SSD архітектури.

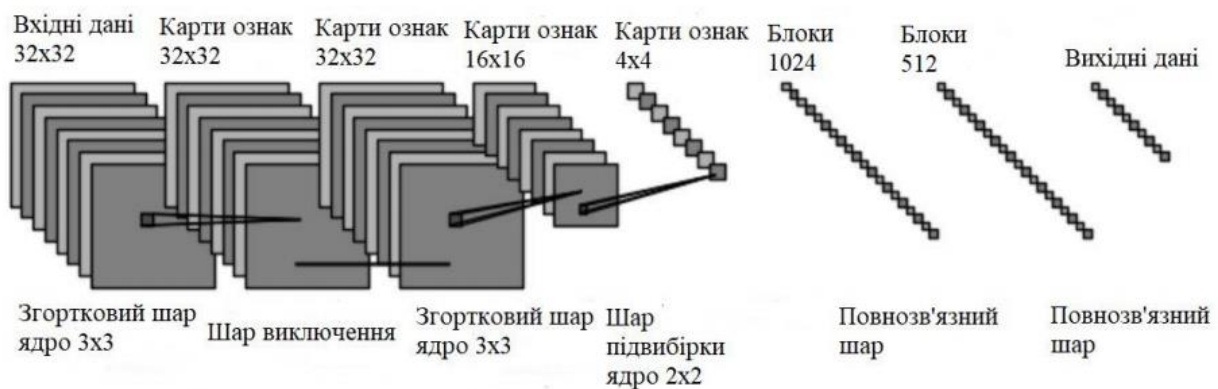


Рисунок 2 – Архітектура запропонованої згорткової нейронної мережі, що спроектована на основі SSD архітектури.

Досліджувану інтелектуальну технологію можна представити у вигляді UML-діаграми класів(рис. 3).. Схема складається з декількох секцій. Безпосередньо модуль розпізнавання жестів, який вхідними параметрами отримує відео потік з веб-камери та список заздалегідь заданих та описаних цільових жестів. Цей модуль має двосторонній зв'язок з інтерфейсом користувача, який представлений у вигляді таблиці з відео та іншими функціями, якими можна керувати жєстами, які були проаналізовані та надані модулем розпізнавання.

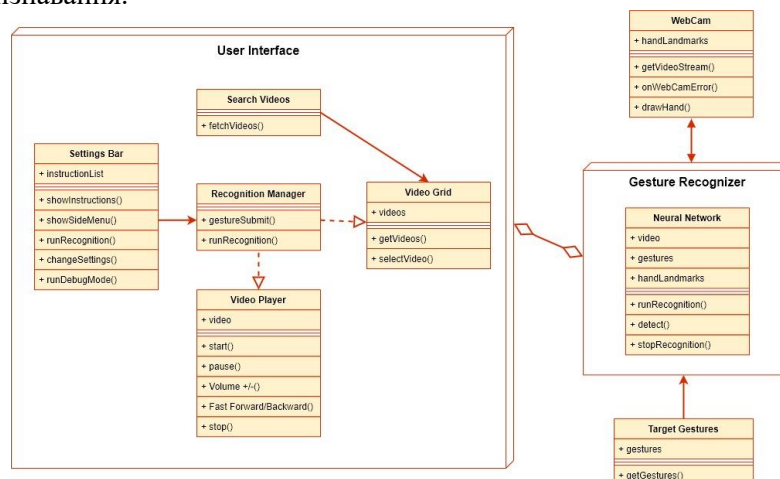


Рисунок 3 –UML-діаграма інтелектуальної технології керування веб-додатком

У ході дослідження було досліджено область розпізнавання рухів та жестів. Запропоновано альтернативний підхід взаємодії з комп'ютером, а саме управління пристроєм із допомогою жестів.

Досліджено шляхи реалізації та оптимізації для використання цієї інтелектуальної системи у веб-додатку.

### Висновки

1. Проаналізовано відомі аналоги і методи жестового керування різноманітними пристроями.
2. Проаналізовано нейромережевий метод та обґрунтовано його вибір для розв'язання задачі.
3. Досліджено структуру, математичну модель та порядок функціонування згорткової нейронної мережі.
4. Запропоновану архітектуру згорткової нейронної мережі.
5. Запропоновану схему функціонування інтелектуальної технології у вигляді UML-діаграми класів

### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Sonia Schechter, Beck Besecker. «What is gesture recognition? Gesture recognition defined». URL: <https://www.marxentlabs.com/what-is-gesture-recognition-defined/>.
2. Zhi-hua Chen, Jung-Tae Kim, Jianning Liang, Jing Zhang, Yu-Bo Yuan. «Real-Time Hand Gesture Recognition Using Finger Segmentation». URL: <https://www.hindawi.com/journals/tswj/2014/267872/>.
3. Maydaniuk V. P. Increasing the Speed of Fractal Image Compression Using Two-Dimensional Approximating Transformations / V. P. Maydaniuk, I. R. Arseniuk, O. O. Lishchuk // Journal of Engineering Sciences. – Sumy : Sumy State University, 2019. – Volume 6, Issue 1. – P. E16 – E20.
4. V. P. Kozemiako ; O. K. Kolesnytskyj ; T. S. Lischenko ; W. Wojcik and A. Sulemenov " Optoelectronic spiking neural network ", Proc. SPIE 8698, Optical Fibers and Their Applications 2012, 86980M (January 11, 2013); doi:10.1117/12.2019340; http://dx.doi.org/10.1117/12.2019340
5. O. K. Kolesnytskyj, I. V. Bokotsey, S. S. Yaremchuk Optoelectronic Implementation of Pulsed Neurons and Neural Networks Using Bispin-Devices // Optical Memory & Neural Networks (Information Optics), 2010, Vol.19, №2, pp.154-165.
6. Буц Вікторія Віталіївна, Магістерська дисертація, «Способи організації засобів нейромережевого розпізнавання об'єкта на зображенні» [Електронний ресурс] – Режим доступу: [https://ela.kpi.ua/bitstream/123456789/23184/1/Buts\\_magistr.pdf](https://ela.kpi.ua/bitstream/123456789/23184/1/Buts_magistr.pdf).

**Каліцінський Владислав Вікторович** – студент кафедри комп'ютерних наук, Вінницький національний технічний університет, м. Вінниця. e-mail: [kalitsinskij.46@gmail.com](mailto:kalitsinskij.46@gmail.com).

**Колесницький Олег Костянтинович** – канд. техн. наук, доцент кафедри КН, Вінницький національний технічний університет, м. Вінниця. e-mail: [kolesnytskyi@vntu.edu.ua](mailto:kolesnytskyi@vntu.edu.ua).

**Kalitsynskiy Vladyslav Viktorovich** – student of Computer Science Department, Vinnytsia National Technical University, Vinnytsia, e-mail: [kalitsinskij.46@gmail.com](mailto:kalitsinskij.46@gmail.com).

**Kolesnytskyi Oleh Konstantinovich** – Cand. Sc. (Eng), Associated Professor of Computer Science, Vinnytsia National Technical University, Vinnytsia, e-mail: [kolesnytskyi@vntu.edu.ua](mailto:kolesnytskyi@vntu.edu.ua).