



А. Г. Казарян, О. П. Максимів, К. І. Ткачук, Р. В. Парцей, С. В. Теслюк

Національний університет "Львівська політехніка", м. Львів, Україна

ВИКОРИСТАННЯ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ У СИСТЕМІ УПРАВЛІННЯ МОБІЛЬНОЮ РОБОТОТЕХНІЧНОЮ ПЛАТФОРМОЮ

В епоху стрімкого технологічного розвитку, коли робототехніка й інтелектуальні системи стають невід'ємною частиною повсякденного життя, важливість розроблення систем управління мобільними робототехнічними платформами з використанням штучних нейронних мереж стає надзвичайно високою та актуальною. Для цієї галузі характерна не тільки істотна практична потреба, але й значний потенціал в інноваційному розвитку. Розвиток сучасної робототехніки та обчислювального інтелекту спонукав до створення ефективніших та адаптивніших мобільних робототехнічних систем. Розроблено систему та засоби для управління мобільними робототехнічними платформами із застосуванням штучних нейронних мереж (ШНМ). Імітуючи роботу нейронної системи, ШНМ дають змогу роботам не тільки реагувати на вхідні дані, але й вчитися вирішувати складні завдання та адаптуватися до змін у навколишньому середовищі. Однією з ключових проблем управління мобільними платформами є розроблення ефективних та інтуїтивно зрозумілих інтерфейсів, які б забезпечували зручну та надійну взаємодію користувача із робототехнічною системою. У цьому контексті використання жестів руки людиною є прогресивним та перспективним напрямом, оскільки це дає змогу створити максимально природний та ефективний спосіб управління. Основне завдання – створення ефективної та інтуїтивно зрозумілої системи, яка дає змогу операторові взаємодіяти із робототехнічною платформою за допомогою натуральних рухів та жестів. Розроблено програмне забезпечення із графічним інтерфейсом для розпізнавання жестів у реальному часі з використанням машинного навчання. Наукова новизна підходу полягає в інтеграції передових методів ШНМ для покращення якості управління та функціональності мобільних робототехнічних платформ. Основні аспекти наукової новизни охоплюють: інтеграцію зі штучним інтелектом; інтерактивність управління; розвиток мобільності робототехніки; адаптивність до різних завдань. Проблема, розглянута у цій роботі, полягає в необхідності розроблення ефективних та інтуїтивно зрозумілих систем управління мобільними робототехнічними платформами із використанням технологій розпізнавання жестів. На основі згорткових нейронних мереж розроблено програму, в результаті виконання якої визначаються положення руки та ідентифікація певних жестів, таких як рух вперед, назад, повороти направо та наліво, а також зупинка руху. Технологію можна використовувати у різних сферах людської діяльності (керування розумним будинком, створення технологічних рішень для людей з фізичними вадами, використання для підвищення інтерактивності в розважальних пристроях, удосконалення інтерфейсів взаємодії із технічним обладнанням).

Ключові слова: автоматизація, інтерфейси взаємодії, технологічні рішення, машинне навчання, ШНМ.

Вступ / Introduction

Разом зі стрімким розвитком робототехніки та обчислювального інтелекту управління мобільними робототехнічними платформами стає предметом інтенсивного вивчення та досліджень. Використання штучних нейронних мереж у цьому контексті відкриває нові можливості для синтезу автономних систем, які здатні ефективно функціонувати в різноманітних умовах і виконувати складні завдання у режимі реального часу.

Важливим аспектом управління мобільними роботами є їхня взаємодія з навколишнім середовищем. За-

вдяки штучним нейронним мережам (ШНМ) роботи можуть адаптуватися до змін у середовищі та ефективно взаємодіяти із непередбаченими факторами. Це дає змогу створювати гнучкіші та адаптивні системи, які можуть використовувати свої навички в різних ситуаціях.

Мобільні робототехнічні платформи (МРП), здатні взаємодіяти в різних середовищах, є ключовим інструментом для вирішення широкого спектра завдань у промисловості, транспорті, медицині та інших галузях. Проблема ефективного управління такими системами стає важливим фактором в їхньому подальшому розви-

тку. Використання технологій розпізнавання жестів як засобу комунікації і управління видається необхідним напрямом для вирішення цієї проблеми. Інтуїтивна природа жестів, які розпізнає мобільний робот, відкриває нові можливості для ефективної та комфортної взаємодії з ним. Відповідно актуальність дослідження полягає у прагненні до розширення можливостей керування мобільними роботами, забезпечуючи операторам простоту та натуральність взаємодії. Застосування розпізнавання жестів може не тільки підвищити ефективність управління, але й зробити цей процес доступним для ширшого кола користувачів, урахувавши людей із обмеженими можливостями щодо користування традиційними інтерфейсами. Отже, дослідження у галузі розпізнавання жестів для управління робототехнічними платформами відіграє ключову роль у забезпеченні подальшого розвитку і вдосконалення технологічних рішень, що істотно впливають на сучасне та майбутнє використання мобільних робототехнічних систем. Тому розроблення підсистеми управління мобільною робототехнічною системою з використанням жестової мови є актуальним завданням.

Об'єкт дослідження – процес управління мобільною робототехнічною платформою із використанням жестової мови.

Предмет дослідження – засоби управління мобільною робототехнічною платформою із використанням штучних нейронних мереж та мови жестів. Завдання – створення ефективної та інтуїтивно зрозумілої системи, яка дає змогу операторові взаємодіяти із робототехнічною платформою за допомогою натуральних рухів та жестів. Предмет проектування охоплює розроблення алгоритмів розпізнавання жестів, програмне забезпечення для оброблення отриманих сигналів та керування платформою, а також інтеграцію цих елементів для розроблення функціональної системи управління мобільною робототехнічною платформою.

Мета роботи полягає у розробленні та оптимізації системи управління мобільною робототехнічною платформою на основі згорткових нейронних мереж для розпізнавання жестів. Основний акцент зроблено на інтеграції передових алгоритмів машинного навчання для забезпечення високої точності розпізнавання жестів та адаптації системи до змін умов експлуатації. Ця мета передбачає створення інтуїтивно зрозумілого інтерфейсу для управління платформою, що полегшує її використання навіть людям із обмеженими можливостями.

Для досягнення зазначеної мети визначено такі основні завдання дослідження:

- аналіз та вибір технологій розпізнавання жестів: проаналізувати сучасні методи розпізнавання жестів; визначити оптимальні технології для використання в контексті мобільної робототехнічної платформи;
- розроблення алгоритмів розпізнавання жестів: розробити ефективні алгоритми для розпізнавання жестів оператора; враховувати можливі різні варіації жестів та їх комбінації;
- реалізація програмного забезпечення: створити програмне забезпечення для оброблення сигналів від камер, а також для керування мобільною робототехнічною платформою;

- інтеграція апаратного забезпечення: здійснити інтеграцію розробленого програмного забезпечення з апаратними компонентами робототехнічної платформи;
- тестування та валідація: виконати серію тестів для перевірки ефективності та точності розпізнавання жестів; валідувати роботу системи у різних умовах реального використання;
- оптимізація та вдосконалення: виявити можливості оптимізації алгоритмів та програмного забезпечення; розглянути можливості для подальшого вдосконалення системи управління;
- вивчення можливостей практичного впровадження.

Матеріали та методи дослідження. Для розроблення системи керування мобільною робототехнічною платформою на основі жестової мови та штучних нейронних мереж (ШНМ) використано комплекс матеріалів та методів. Мова програмування Python та фреймворки TensorFlow і OpenCV забезпечили опрацювання даних і стали основою для розроблення алгоритмів керування платформою, що дало змогу обробляти зображення, ідентифікувати ключові точки руки та будувати нейромереві моделі. Для того, щоб відділити руку від фону, колір шкіри був відкалібрований у колірному просторі RGB, що дало змогу успішно відокремити руку від навколишнього середовища. Розроблена згорткова нейромерева модель складається з кількох шарів, кожен з яких виконує специфічні функції для розпізнавання жестів руки. Початковий шар відповідає за оброблення вхідного зображення та виділення основних контурів руки. Наступні шари поступово аналізують отримані ознаки, інтегруючи інформацію про положення пальців та орієнтацію руки в просторі. На виході модель генерує класифікацію жестів, таких як “вперед”, “назад”, “поворот” і “зупинка”. Ця структура дає змогу ефективно адаптуватися до змін умов освітлення та фону, що підвищує точність і надійність системи управління робототехнічною платформою. Після навчання та тестування моделі програмне забезпечення було інтегровано з апаратним забезпеченням мобільної платформи, завдяки чому система працювала в режимі реального часу та забезпечила високу точність розпізнавання жестів.

Аналіз останніх досліджень та публікацій. Використання нейронних мереж у робототехніці розглянуто у низці науково-технічних статей. Наприклад, у праці “A Novel Supertwisting Zeroing Neural Network With Application to Mobile Robot Manipulators”, авторами якої є Д. Чен, С. Лі та Ц. Ву, запропоновано новий метод керування мобільними роботами-маніпуляторами [1]. Цей метод поєднує переваги алгоритмів суперскручування та нейронних мереж, які обнулюють, для досягнення збіжності за кінцевий час та стійкості до зовнішніх збурень.

Метод складається із двох основних компонентів: суперскручувальної нейронної мережі, яка обнулює, та керування ковзним режимом. STZNN використовують для оцінювання невизначених динамік робота та компенсації їх впливу, а керування ковзним режимом – для забезпечення стійкості системи керування та збіжності до заданої траєкторії.

Як свідчать результати моделювання, за допомогою запропонованого методу можна ефективно керувати мобільними роботами-маніпуляторами з невизначеними динаміками. Він забезпечує точне відстеження траєкторії та стійкість до невизначеностей. До недоліків методу можна зарахувати складність реалізації та потребу у значних обчислювальних ресурсах [1].

У статті “Artificial Intelligence and Internet of Things for Robotic Disaster Response” [2] розглянуто використання штучного інтелекту (ШІ) та Інтернету речей (IoT) для покращення роботизованого реагування на катастрофи. Автори стверджують, що поєднання цих технологій може допомогти зробити роботизовані системи ефективнішими та автономнішими під час рятувальних операцій.

Стаття висвітлює різні типи роботів, які можуть бути корисними під час катастроф, зокрема мобільних роботів, безпілотні підводні апарати та робочі станції. Розглянуто, як можна використовувати ШІ для надання роботам можливостей сприйняття, прийняття рішень та управління рухом в умовах невизначеності. IoT розглядають як спосіб спілкування роботів між собою та з іншими пристроями, із обміном даними та координацією їхніх дій.

Наведено приклади того, як можна використати поєднання ШІ та IoT для конкретних завдань реагування на катастрофи, таких як пошук та порятунок жертв, оцінка збитків та доставка допомоги.

У статті “Multi-Sensor Fusion Method Based on Artificial Neural Network for Mobile Robot Self-Localization” С. Е. Magrin і Е. Todt пропонують метод самолокалізації для мобільних роботів, оснований на злитті даних із декількох сенсорів за допомогою штучної нейронної мережі. Метод використовує дані різних сенсорів, таких як сонар, для забезпечення точнішої та надійнішої локалізації робота в динамічному середовищі. Він ґрунтується на тому, що дані різних сенсорів перед обробленням ШІМ нормалізуються та масштабуються. Нейронна мережа навчається на наборі даних, що містить дані сенсорів та відповідні позиції робота. Ознайомившись зі статтею, можна отримати інформацію про виконані симуляції, які показали, що запропонований метод забезпечує точнішу локалізацію порівняно з використанням окремих сенсорів. Крім того, метод стійкий до шуму та відсутності даних з деяких сенсорів.

Стаття пропонує перспективний метод самолокалізації для мобільних роботів, оснований на злитті даних з декількох сенсорів за допомогою ШІМ. Метод може бути корисним для різноманітних застосувань, де точна та надійна локалізація є критично важливою [3].

У статті “Selected Topics of Artificial Intelligence in Robotics” Л. Кельблінг розглядає застосування штучного інтелекту в робототехніці, зосереджуючись на кількох конкретних темах:

- Глибоке навчання: проаналізовано потенціал глибокого навчання для різних робототехнічних задач, таких як аналіз зображень та планування маршруту.
- Навчання: наголошено на важливості ефективних методів навчання для робототехнічних систем, зокрема оптимізації.

- Планування маршруту: розглянуто використання ШІ для планування оптимальних маршрутів для мобільних роботів.
- Виявлення аномалій: стаття описує, як можна використати ШІ для виявлення аномалій у даних датчиків роботів, що дає змогу виконувати попереднє технічне обслуговування.

У цій роботі подано огляд застосування ШІ в різних робототехнічних завданнях. Вона висвітлює потенціал цієї технології для підвищення продуктивності та можливостей роботів, але також згадує про певні проблеми, такі як: висока вартість розроблення та впровадження; складність алгоритмів; потенційні проблеми безпеки, які необхідно вирішити [4].

Також можна зазначити, що використання нейромережних структур у системах управління дронами стає дедалі популярнішим, адже ШІМ можуть істотно підвищити автономність, гнучкість та ефективність дронів. Детальніше розглянемо ключові сфери, де вплив нейромереж значний:

1. Навігація та планування маршрутів: нейромережі можуть навчатися на великих наборах даних карт, зображень та датчиків, щоб генерувати оптимальні маршрути для дронів, уникаючи перешкод та досягаючи цілей із максимальною ефективністю.

Алгоритми глибокого навчання, такі як згорткові нейронні мережі (CNN) та рекурентні нейронні мережі (RNN), можуть використовуватися для аналізу даних з камер та датчиків дронів у режимі реального часу, дозволяючи їм адаптуватися до мінливих умов середовища та уникати непередбачених перешкод.

2. Розпізнавання об'єктів та відстеження: нейромережі можуть використовуватися для ідентифікації та відстеження об'єктів у полі зору дрона, що робить їх корисними для таких завдань, як розвідка, пошук, рятування та моніторинг.

Алгоритми комп'ютерного зору, основані на нейромережах, можуть розпізнавати людей, транспортні засоби, будівлі та інші об'єкти з високою точністю, навіть у складних умовах освітлення та за наявності перешкод.

3. Управління роєм: нейромережі можна використовувати для координації дій кількох дронів, що робить їх корисними для таких завдань, як картографування, моніторинг та пошуково-рятувальні операції.

Алгоритми ройового інтелекту, основані на нейромережах, дають змогу дронам самоорганізовуватися, співпрацювати та обмінюватися інформацією, щоб виконувати складні завдання, нездійсненні для одного дрона.

4. Автономне прийняття рішень: алгоритми машинного навчання, основані на нейромережах, можуть дати змогу дронам аналізувати ситуації в режимі реального часу, оцінювати ризики та вибирати найкращий курс дій без втручання людини.

5. Кібербезпека: нейромережі можуть використовуватися для захисту дронів від кібератак, що стає важливим питанням з огляду на зростання залежності від дронів у критично важливій інфраструктурі.

Алгоритми машинного навчання, основані на нейромережах, можуть виявляти та блокувати кібератаки, шифрувати дані та захищати системи управління дронів від несанкціонованого доступу.

Підсумовуючи, зазначимо, що використання нейромережових структур у системах управління дронами все ще перебуває на ранній стадії розвитку, проте потенціал для значного покращення можливостей та гнучкості дронів величезний. З розвитком технологій нейромереж можна очікувати, що дрони стануть ще автономнішими, ефективнішими та здатними виконувати складні завдання, які раніше вважалися нездійсненними.

У деяких роботах запропоновано використовувати згорткові нейронні мережі в робототехніці. Розглянемо і проаналізуємо їх основні архітектури [5, 6]:

- модель R-CNN – перша і найповільніша модель для виконання операцій класифікації зображень на основі окремих фрагментів;
- хоча конструкції Fast R-CNN і Faster R-CNN є вдосконаленими версіями попередньої моделі, вони непридатні для робочих навантажень у реальному часі;
- дизайн YOLO – один із найшвидших типів, який, на відміну від інших, зосереджений на виявленні положення об'єкта за один прохід;
- архітектурний підхід SSD схожий на попередній, але використовує ідею VGG16 [7]. Оснований на найкращих аспектах YOLO, а також на новій інформації.

Детальніше розглянувши архітектуру мережі SSD, можна зробити висновок, що:

- конструкція SSD мережі працює гірше на крихтих об'єктах, ніж Fast R-CNN. Зрештою, їх можна виявити лише на перших рівнях, де роздільна здатність вхідних даних ще висока. Однак складність полягає у тому, що на цьому етапі можна класифікувати лише ознаки низького рівня, такі як кути, краї та кольорові плями, які дають мало інформації для ідентифікації об'єкта;
- чим більше межових фрагментів використовується, тим вищою може бути точність детектора. Проте збільшення кількості меж може призвести до зростання обчислювальної складності та повільнішого оброблення зображень;
- використання ефективних фільтрів для оброблення зображень може підвищити точність детектора, допомагаючи зменшити шум та виявляти об'єкти точніше;
- архітектура SSD має менше помилок локалізації, ніж R-CNN, але за однакового введення може виникати більше помилок класифікації.

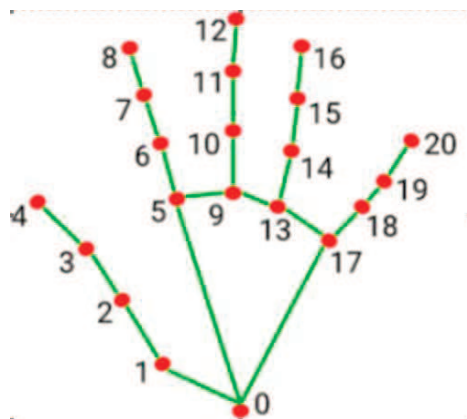


Рис. 1. Екстракція ключових точок в області долоні / Extraction of key points in the palm area [9]

Підсумуємо: SSD може мати деякі обмеження у виявленні дрібних об'єктів та швидкості оброблення, він може бути ефективним завдяки своїй здатності до точної локалізації об'єктів та порівняно високій точності. SSD можна навчати крок за кроком для підвищення точності. SSD надає більше прогнозів і вищий ступінь покриття з погляду розташування, розміру та співвідношення сторін. SSD може знизити роздільну здатність вхідного зображення до 256×256 пікселів без зниження точності завдяки попереднім удосконаленням. Запропонована модель може працювати в системах реального часу і виконувати завдання класифікації і розпізнавання значно швидше, ніж найсучасніший Fast R-CNN після видалення призначеної області та використання зображень із нижчою роздільною здатністю.

Виконавши аналіз, ми вирішили використовувати у дослідженні твердотілі накопичувачі, найкращою альтернативою є застосування нової платформи MediaPipe від Google, випущеної в 2020 р.

Ця структура містить набір інструментів для розв'язання різноманітних проблем категоризації та ідентифікації об'єктів [8]: обличчя і його контури; зіниць очей, рук, позиції тіла, волосся, інших об'єктів.

Модуль MediaPipe Hands використовує конвеєр ML, який складається із кількох моделей, що працюють разом. Це модель виявлення долоні, яка працює з повним зображенням і повертає орієнтовану обмежувальну рамку руки, а також модель орієнтира руки, яка працює з обрізаною областю зображення, визначеною детектором долоні, й повертає високоякісні тривимірні ключові точки рук. Ця стратегія схожа на те, що застосовано в іншому рішенні MediaPipe Face Mesh, яке використовує детектор обличчя разом із моделлю орієнтира обличчя.

Надання точно обрізаного зображення руки для моделі орієнтира руки різко зменшує потребу в доповненні даних (наприклад, обертання, трансляції та масштабування), натомість дає змогу мережі більшу частину своїх можливостей використати для точності прогнозування координат. Крім того, у конвеєрі культури (в контексті MediaPipe Hands конвеєр культури належить до процесу динамічного визначення області зображення, яка повинна бути оброблена для подальшого аналізу) також можуть бути створені на основі орієнтирів рук, визначених у попередньому кадрі, і лише тоді, коли модель орієнтира більше не може ідентифікувати наявність руки, запускається виявлення долоні для переміщення руки.

Конвеєр реалізовано у вигляді графіка MediaPipe, який використовує підграф відстеження орієнтирів вручну з модуля ручного орієнтира і відтворює за допомогою спеціального підграфа ручної візуалізації. Підграф відстеження орієнтирів руки використовує підграф орієнтира руки з того самого модуля та підграф визначення долоні з модуля виявлення долоні.

Після виявлення долоні на всьому зображенні модель орієнтира руки виконує точну локалізацію 21 ключової точки тривимірної координати кисті всередині виявлених областей кисті за допомогою регресії, тобто прямого прогнозування координат. Модель вивчає послідовне внутрішнє подання пози рук і стійка навіть до частково видимих рук і перекритих іншою ділянкою тіла.

Щоб краще охопити можливі пози рук і забезпечити додатковий контроль за природою геометрії руки, також відтворюється високоякісна синтетична модель руки на різних фонах. Після завершення другого етапу нейронна мережа проіндексовує долоню та 21 важливу координату в тривимірному просторі, як показано на рис. 1, для кращої ітерації зі зчитаними даними.

На рис. 1 бачимо, що у моделі для визначення відображення кожного пальця використано по чотири ключові точки та одна точка використовується для позначення зап'ястка.

Отже, виконавши аналіз, можна зробити висновок про необхідність вдосконалення моделей, методів та засобів щодо управління мобільною робототехнічною системою із використанням жестової мови. В запропонованій роботі для отримання результату використовуємо згорткові нейронні мережі та програмне рішення MediaPipe.

Результати дослідження та їх обговорення / Research results and their discussion

Розпізнавання жестів ґрунтується на використанні математичних алгоритмів оброблення зображень та машинного навчання. Під час роботи над системою використовують такі математичні моделі:

1. Модель класифікації жестів за допомогою нейронних мереж. Використання штучних нейронних мереж для розпізнавання основних жестів рук, таких як махання, стищення та розкриття кулака, показ вказівника тощо.

2. Модель аналізу просторових координат жестів. Визначення руху руки в тривимірному просторі для точного розпізнавання жестів та їх подальшого інтерпретування.

Математичні моделі використовують для підвищення ефективності алгоритмів керування рухом мобільної робототехнічної платформи, а саме:

1. Оптимізація шляху руху. Використання алгоритмів пошуку шляху, таких як алгоритм A*, для знаходження оптимального маршруту в зазначеному просторі.

2. Регулювання швидкості та поворотів. Математичні моделі для регулювання швидкості та кута повороту платформи забезпечують точне виконання команд, отриманих від модуля розпізнавання жестів.

Особливості алгоритмів і підходу машинного навчання та обчислювального інтелекту до розпізнавання жестів

Для розв'язання задачі розпізнавання жестів із застосуванням машинного навчання використовують згорткові нейронні мережі [10]. Завдяки внутрішній побудові вони спеціалізуються на розпізнаванні та класифікації об'єктів для завантаженого зображення або за допомогою комп'ютерного зору.

Нейронні мережі такого типу відрізняються від інших тим, що можуть опрацювати складніші зображення з меншою залежністю від якості, центрування та інших факторів. Традиційні нейронні мережі добре розпізнають числа, але не дуже добре класифікують речі. Фундаментальною причиною такої невідповідності є використання згорткових шарів для фільтрації [5]. Спочатку необхідно розглянути, як згортковий шар обчислює рівень, не покладаючись на порівняння мозок / нейрон. Параметри цього шару складаються із кількох мільйонів фільтрів, для налаштування яких потрібен тривалий час. Кожен з цих фільтрів має дуже

маліу ширину і висоту, але покриває всю глибину вхідного зображення на виділеному наборі даних.

Наприклад, звичайний фільтр у першому шарі CNN можна подати у форматі $5 \times 5 \times 3$ (тобто фільтр буде налаштований на 5 пікселів у висоту і 5 пікселів у ширину, а останнім числом передають кількість кольорових каналів – 3 у цьому випадку). Під час прямого проходу фільтри застосовують до ширини та висоти вхідного зображення, вони переміщуються по зображенню, обчислюючи локальні результати. Це означає, що фільтри виконують операції на невеликих ділянках зображення, порівнюючи значення пікселів у кожному місці з відповідними значеннями в фільтрі. Результати цих обчислень з'єднуються для формування виходу згорткового шару. В результаті оброблення отримуємо двовимірну карту активації, яка відображає результати застосування фільтрів до кожної просторової позиції зображення. Наступний етап – переміщення фільтра по ширині та висоті вхідного зображення. На першому шарі нейронна мережа аналізує фільтри, які реагують на конкретні візуальні характеристики, такі як контури або кольорові плями. Це дає змогу виявляти основні елементи, такі як краї об'єктів або певні кольорові деталі. На наступних шарах мережа комбінує ці елементи для розпізнавання складніших структур, таких як повні шаблони або об'єкти загалом.

Кожен рівень згорткового шару матиме власний набір фільтрів, кожен з яких створюватиме унікальну двовимірну карту активації. Щоб отримати початковий обсяг, ці карти активації будуть створені вздовж вимірювання глибини. Основна роль згорткового шару – представляти невелику групу пікселів, використовуючи лише один піксель. Цей підхід зменшує кількість даних (пікселів, наявних до опрацювання), що надходять у нейронну мережу [5].

У результаті виконання операції отримуємо карту активації після проходження шару згортки, яку необхідно нормалізувати за допомогою функції активації:

$$\text{ReLU}(x) = \max(x, 0), \quad (1)$$

де x – це дійсне число. Ця формула визначає випрямлену лінійну одиницю (ReLU), яка є типом функції активації, що використовується в нейронних мережах.

Функція ReLU працює так: якщо x більше або дорівнює 0, то $\text{ReLU}(x) = x$; якщо x менше від 0, то $\text{ReLU}(x) = 0$.

Інакше кажучи, ReLU просто встановлює всі негативні входи в 0. Це може допомогти нейронним мережам навчатися швидше, оскільки не потрібно опрацювати негативні входи. ReLU також може допомогти нейронним мережам краще узагальнювати, оскільки ReLU є нелінійною функцією.

Нелінійні функції важливі для функціонування нейронних мереж, оскільки дають змогу навчатися складним моделям даних.

Розпізнавання жестів за допомогою ключових координатних точок: для переведення в подальші команди зображення руки, обличчя і тіла необхідно перевести їх у набір тривимірних координат у просторі. Ці координати повинні зберігати розташування руки та окремих пальців на ній.

Для розпізнавання жестів треба застосовувати підхід із використанням координат точок для пошуку точного просторового розташування векторів. Необхідно виконати достатньо обчислень за формулою (2), щоб ідеально визначити довжину вектора на показаному жесті:

$$AB = \sqrt{(x_b - x_a)^2 + (y_b - y_a)^2 + (z_b - z_a)^2}, \quad (2)$$

де $x_a, y_a, z_a, x_b, y_b, z_b$ – це відповідні координати початку й кінця вектора АВ.

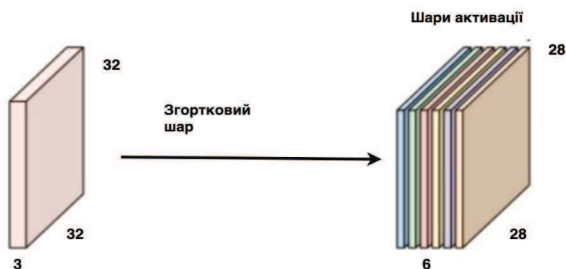


Рис. 2. Схема, яка зображає вихід після проходження через згортковий шар нейронної мережі / Schema illustrating the output after passing through a convolutional layer of the neural network [9]

У дослідженні застосовано метод навчання нейронної мережі – “навчання з вчителем” [11]. Тобто для того, щоб навчити програму розпізнавати жести, потрібно надати набір зображень жестів, які вона повинна розпізнавати. Надалі система почне навчання на підставі отриманих даних.

Процес навчання повинен бути неперервним. У системі немає процесу донавчання – тобто для того, щоб розпізнавати новий жест, потрібно заново запускати навчання. Проте цей недолік можна усунути в майбутньому. Якщо налаштування передбачає фіксування 30 кадрів на один жест, у середньому на навчання одного жесту потрібно виділити 10 хвилин. Тому цей показник варто брати до уваги, щоб довести процес навчання до кінця, оскільки немає можливості поставити його на паузу.



Рис. 3. Результат вдалого калібрування шкіри / Result of successful skin calibration

Процес взаємодії передбачає такі етапи: отримання відеоряду з вебкамери; за допомогою відокремлення навколишнього середовища від об'єкта отримується масив координатних точок регіону руки; на зображенні необхідно побудувати вектори і визначити відносне положення точок; відбувається виведення на екран 3D-хмари точок на об'єкті; за допомогою датасету, який міститься в словнику жестів, система визначає жест. У результаті дані про відповідний жест передаються через послідовний (UART) порт на модуль Arduino.

Модуль Arduino отримує відповідний сигнал, який трансформує та передає на керування двигунами платформи (рис. 5).

За достатнього освітлення, правильної пози руки на момент калібрування та правильних значень відхилення допустимого кольору шкіри можливе успішне визначення шкіри. Якісне калібрування забезпечує чіткий контур руки, як показано на рисунку.

Жести кулака і відкритої долоні, які належать до шести жестів, розпізнаваних системою, потребують вертикального положення руки, тобто положення, паралельного до вертикальної осі об'єктива камери. Це потрібно, щоб обмежувальний прямокутник, утворений навколо форми руки, мав сторони, які б точно відповідали співвідношенням ширини та висоти жесту.

На рис. 4 показано рекомендоване положення руки в об'єктиві камери для визначення усіх шести рухів.

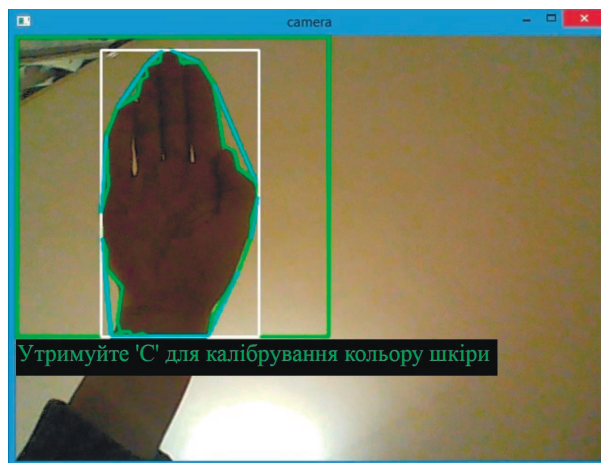


Рис. 4. Рекомендоване положення руки для розпізнавання жестів / Recommended hand position for gesture recognition

Із шести жестів, які може розпізнавати система, жести кулака і відкритої долоні потребують вертикального положення руки, тобто положення паралельно до вертикальної осі об'єктива камери. Це необхідно для того, щоб сторони обмежувального прямокутника, побудованого навколо контуру руки, правильно відображали відношення ширини і висоти показаного жесту. Передавання відповідного коду жесту під час розпізнавання виконується без затримок у роботі системи. Після виконання певної підготовчої роботи користувач може керувати платформою жестами.

Розроблений програмний продукт дає можливість використовувати пристрій запису відео для визначення ділянки руки і розпізнавання таких жестів; рухатись вперед, назад, поворот направо, наліво та зупинка руху.

За допомогою програмної обробки отримані з веб-камери зображення перетворюють з RGB формату подання кольору в формат BGR для досягнення високої роздільності між кольором шкіри руки та навколишнім фоном. В комп'ютерному аспекті значення кольору шкіри, отримані в форматі BGR, істотно відрізняються від значень навколишнього середовища.

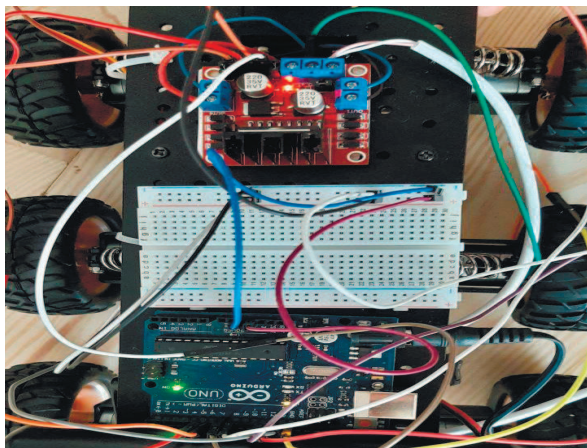


Рис. 5. Робот підключений до джерела живлення / The robot is connected to the power source



Рис. 6. Бінарне зображення з виділеним фрагментом руки / Binary image with a highlighted fragment of a hand

Внаслідок переведення вхідного, необробленого зображення в просторі кольорів RGB у формат BGR отримують зображення, на якому дуже швидко можна визначити фрагмент руки.

Варто зазначити, що для успішного переведення кольорів з одного формату в інший необхідно в режимі реального часу виконати ручне калібрування, після завершення якого система отримує зображення у форматі HSV, яке потім перетворюється на бінарне зображення силуету руки. На рис. 6 подано результат оброблення зображення руки після перетворень.

Як результатівне отримано чорно-біле зображення руки. Білим кольором виділено саме фрагмент долоні, а навколишній фон зафарбовано чорним кольором.

Бінарне зображення є основою для знаходження контуру руки й отриманої області середовища навколо неї.

Хоч камера, на яку здійснено запис, може мати набагато вищу роздільну здатність, у фрагмента, отриманого внаслідок перетворень, завжди фіксована розмірність – 340 на 340 пікселів.

Розроблення основних алгоритмів програмного забезпечення.

У розробленні системи розпізнавання жестів можна виділити три головні етапи, які передбачають виконання кількох алгоритмів:

- отримання граничних значень кольору шкіри з кожного фрагмента тіла користувача;
- перетворення зображення в бінарне подання;
- розпізнавання жесту на основі обробленого зображення.

Щоб успішно виконати калібрування, необхідно тримати свою руку паралельно до камери. Якщо почекати деякий час, то на руці користувача з'являться шість синіх прямокутників, що означатиме готовність до переходу на наступний етап.

Після цього система виведе на екран повідомлення про виконання необхідної для калібрування дії.

У результаті оброблення одержуємо середнє значення кольору в колірному просторі HSV з кожної з шести квадратних ділянок зображення розміром 15 на 15 пікселів для вхідних даних. Відповідно до кожного з квадратів на основі цих значень, а також значень відхилень нижньої та верхньої меж для кожного із каналів HSV встановлюються шість кольорних меж. Колірні обмеження – важлива інформація для виділення силуету руки. Блок-схему алгоритму калібрування кольору шкіри зображено на рис. 7.

Щоб побудувати двійкове зображення силуету руки, вхідне зображення фільтрують із видаленням пікселів, колір яких не відповідає певним обмеженням, тобто з вхідного зображення вибирають лише ті пікселі, які потрапляють за межі.

Функція бібліотеки OpenCV `convvar2.inRange` використовується для відкидання пікселів, які не підходять під обмеження кольору шкіри. Цей метод створює двійкове зображення з усіма пікселями, виставляється на 255 (білий) у заданих межах, а всі інші пікселі про- ставляються зі значенням 0 (чорний). Результат фільтрації зображень додається до кінцевого подання, відображеного на рис. 6, за допомогою побітової операції XOR для кожного з шести граничних значень. На рис. 8 зображено блок-схему методу побудови бінарного зображення силуету руки.

Функція бібліотеки OpenCV `cv2.bitwise_exclusiveor` використовується для виконання порозрядної операції “або”, яка створює двійкове зображення внаслідок накладання двох зображень, які надаються як вхідні дані. Наприклад, на рис. 9 побудовано контур силуету області кисті, є дефекти опуклості для виявлення дії. Якщо кут біля дефекту опуклості менше ніж 90 градусів, вважається, що дефект є точкою між пальцями. Зважаючи на це, кількість видимих пальців розраховують на основі кількості дефектів опуклості з кутом менше ніж 90 градусів. Наприклад, чотири такі дефекти розглядають як п'ять пальців. У цьому випадку кожен дефект представляє простір між пальцями. Це пов'язано з тим, що один дефект може вказувати на кілька пальців, що частково перекриваються.



Рис. 7. Блок-схема алгоритму отримання граничних значень кольору шкіри з кожного фрагмента / Diagram illustrating the extraction of boundary values of skin color from each fragment

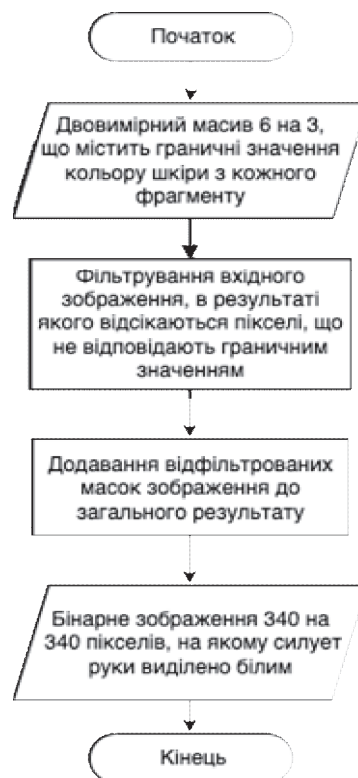


Рис. 8. Блок-схема алгоритму створення бінарного зображення силуету руки / Diagram of the algorithm for creating a binary silhouette image of the hand

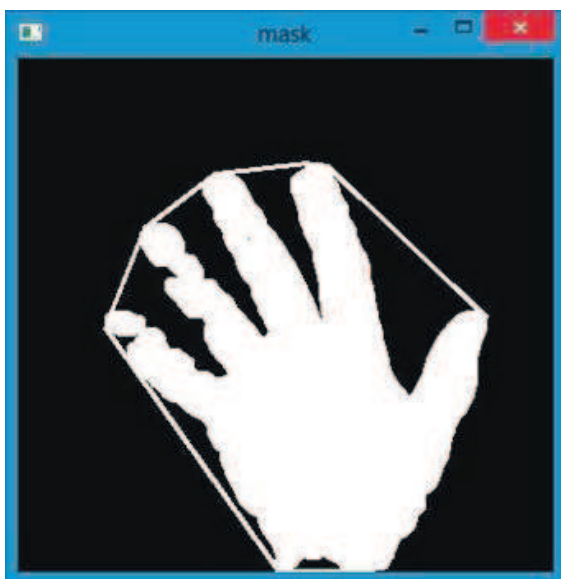


Рис. 9. Отримана область навколо силуету руки / The obtained area around the silhouette of the hand

Якщо значення усіх дефектів дорівнюють нулю, вважається, що на зображенні кулак, долоня або один палець. Функціонування програми полягає в читанні та аналізі кожного кадру із відеопотоку камери, подія розпізнавання запускається одночасно під час читання кожного нового зображення, що зумовлює постійне відображення міток для певного жесту.

Особливості та обмеження розробленої системи розпізнавання жестів. Використовуючи розроблену

систему розпізнавання жестів на основі машинного навчання, треба враховувати, що існують певні особливості та обмеження щодо її роботи.

Оскільки функціонування системи для розпізнавання жестів залежить від відносного розташування частин руки, які виконують жест, результат не відповідає очікуванню для певних комбінацій вхідних даних.

Формат граничного прямокутника контуру руки втрачає сенс, коли рука розташована на відстані більше ніж 60 см від камери. У цьому сценарії рух стиснутого кулака буде неправильно розпізнано, як бачимо на рис. 10.

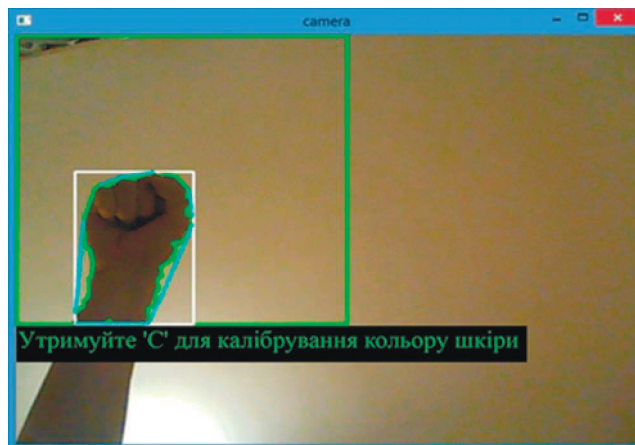


Рис. 10. Результат з помилкою після розпізнавання жесту кулака за віддалення руки від камери / Result with an error after recognizing the fist gesture when moving the hand away from the camera

Розроблений підхід використовує співвідношення площі форми руки й опуклої оболонки для ідентифікації жесту виставлення одного пальця з долоні або кулака.

У такій ситуації буде досягнуто неправильного результату (рис. 11), який створюється під кутом повороту кисті 45 градусів у разі виставляння одного пальця.

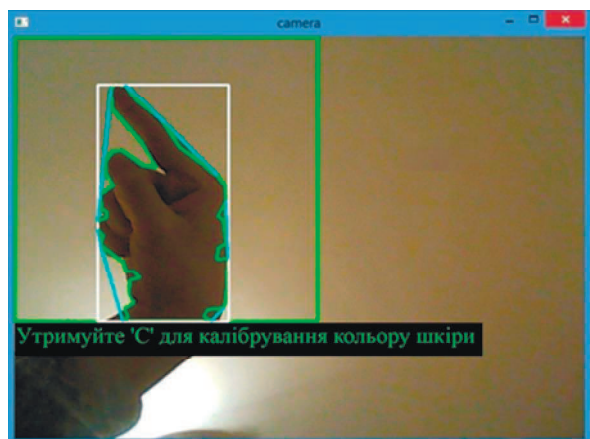


Рис. 11. Неправильне жестове розпізнавання одного пальця у разі обертання руки на 45 градусів / Incorrect gesture recognition of one finger when rotating the hand by 45 degrees

Для уникнення цього під час розпізнавання нового жесту ідентифікація відбувається лише у разі використання жесту, який відрізняється від попереднього. Для отримання найточніших результатів потрібно успішно пройти калібрування кольору шкіри, тримаючи руку на відстані 30–40 см від екрана під кутом 90 градусів, щоб система змогла відокремити від навколишнього середовища саме область руки користувача.

Результати показали, що програмне рішення може розпізнавати жести з високою точністю.

Обговорення результатів дослідження Розроблений програмний продукт дає можливість використовувати вебкамеру для виділення ділянки руки та розпізнавання фіксованого набору жестів, на основі яких виконується керування відповідним рухом мобільної платформи. Порівняння з іншими підходами показало, що запропонований метод перевершує аналогічні рішення за точністю розпізнавання та стійкістю до шуму вхідних даних. Результати навчання нейромережі були валідовані на різних наборах даних, що підтвердило її здатність ефективно працювати в реальних умовах експлуатації. Порівняно із традиційними підходами, модель продемонструвала кращу адаптивність до змін умов середовища та вищу швидкість опрацювання даних, що є ключовими факторами для застосування в мобільних робототехнічних системах.

Наукова новизна отриманих результатів дослідження полягає у впровадженні згорткових нейронних мереж для розпізнавання жестів, що забезпечує високу точність та надійність управління мобільною робототехнічною платформою. Запропонований підхід дає змогу досягти значного покращення в адаптивності системи до змін умов середовища та забезпечує інтуїтивно зрозумілий інтерфейс для користувачів. Такий підхід відкриває нові можливості для розвитку інтерактивних робототехнічних систем, що ґрунтуються на природній взаємодії з користувачем.

Практична значущість результатів дослідження полягає у розробленні програмного забезпечення для управління мобільними робототехнічними платформами, яке можна застосовувати в різних сферах, таких як промисловість, медицина, транспорт, розумний будинок та інші галузі, де необхідна автоматизація. Розроблене програмне забезпечення дає змогу здійснювати управління за допомогою жестів, що істотно спрощує взаємодію із робототехнічними системами, робить її інтуїтивно зрозумілою та доступною для користувачів з різним рівнем підготовки, зокрема для людей з обмеженими можливостями. Це сприяє підвищенню ефективності використання робототехніки, забезпечуючи точність, швидкість реакції та адаптивність до змін середовища. Впровадження таких систем у реальних умовах може істотно знизити витрати на трудові ресурси, підвищити продуктивність, а також безпеку виконання завдань, що потребують взаємодії з технічними пристроями.

Висновки / Conclusions

У статті висвітлено дослідження та розроблення програмного забезпечення для мобільних робототехнічних платформ із використанням штучних нейронних мереж. Система використовує сенсорну інформацію для розпізнавання жестів користувача та перетворює їх на команди для управління робототехнічною платформою.

Розроблено структуру, алгоритм та програмне забезпечення, які дають можливість визначати положення руки і класифікувати певні жести, а визначивши жест, передавати відповідний сигнал на мобільну платформу та виконувати рух.

Розроблена модель – це нейронна мережа, призначена для класифікації дій на основі вхідних даних із камери та ключових точок.

Дозволені межі кольору шкіри в колірному просторі RGB калібрують, щоб вибрати фрагмент зображення, який відповідає руці. Калібрування, яке ініціює користувач, дає можливість вибрати частину руки за різних умов освітлення, використовуючи програмне рішення MediaPipe.

Для пришвидшення розпізнавання жесту розроблено алгоритми переведення зображення руки в масив значень координат точок у просторі та переведення цих точок у бінарне зображення.

Основою алгоритму розпізнавання є побудова області навколо контуру руки і знаходження дефектів опуклості, розміщених між пальцями руки, за допомогою пошуку лише відмінностей у дефектах точок регіону руки, кут яких менший від або дорівнює 90 градусам.

Перспективи подальших досліджень: розширення набору розпізнаваних жестів; розроблення методів розпізнавання динамічних жестів; дослідження можливостей використання розробленого програмного комплексу для керування іншими типами робототехнічних систем.

References

1. Chen, D., Li, S., & Wu, Q. (2021). A novel supertwisting zeroing neural network with application to mobile robot manipulators. *IEEE Transactions on Neural Networks and Learning Systems*, 32(4), 1776–1787. <https://doi.org/10.1109/TNNLS.2020.2991088>

2. Lee, M.-F. R., & Chien, T.-W. (2020). Artificial intelligence and internet of things for robotic disaster response. *2020 International Conference on Advanced Robotics and Intelligent Systems (ARIS)*, Taipei, Taiwan, 1–6. <https://doi.org/10.1109/ARIS50834.2020.9205794>
3. Magrin, C. E., & Todt, E. (2019). Multi-sensor fusion method based on artificial neural network for mobile robot self-localization. 2019 Latin American Robotics Symposium (LARS), 2019 Brazilian Symposium on Robotics (SBR), and 2019 Workshop on Robotics in Education (WRE), Rio Grande, Brazil, 138–143. <https://doi.org/10.1109/LARS-SBR-WRE48964.2019.00032>
4. Kwiecień, J. (2022). Selected topics of artificial intelligence in robotics. *2022 7th International Conference on Mechanical Engineering and Robotics Research (ICMERR)*, Krakow, Poland, 40–44. <https://doi.org/10.1109/ICMERR56497.2022.10097788>
5. Ostrovka, D., Stasenko, D., & Teslyuk, V. (2022). Autonomous intelligent control system for mobile robotic system. In *Proceedings of the IEEE 17th International Conference on Computer Science and Information Technologies*, Lviv, Ukraine, 206–209. <https://doi.org/10.1109/CSIT56902.2022.10000459>
6. Tsmots, I., Tkachenko, R., Teslyuk, V., Optyak, Y., & Rabyk, V. (2022). Hardware components for nonlinear neuro-like data protection in mobile smart systems. In *Proceedings of the IEEE 17th International Conference on Computer Science and Information Technologies*, Lviv, Ukraine, 198–202. <https://doi.org/10.1109/CSIT56902.2022.10000636>
7. Posture and gesture recognition for human-computer interaction (n. d.). *IntechOpen*. <https://www.intechopen.com/chapters/8712> (Accessed: 22.10.2023).
8. Pavlovic, V. I., Sharma, R., & Huang, T. S. (1997). Visual interpretation of hand gestures for human-computer interaction: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7), 683–695. <https://doi.org/10.1109/34.598226>
9. Recognition and classification of sign language (n. d.). *Scielo*. Retrieved from: [http://www.scielo.org.mx/scielo.php?script=](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-55462018000100271)
[sci_arttext&pid=S1405-55462018000100271](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-55462018000100271) (Accessed: 11.11.2023).
10. Evergreens (n. d.). Convolutional neural networks: Explained in simple words. Retrieved from: <https://evergreens.com.ua/ua/articles/cnn.html>
11. Makantasis, K., Karantzas, K., Doulamis, A., & Doulamis, N. (2015). Deep supervised learning for hyperspectral data classification through convolutional neural networks. In *Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Milan, Italy, 4959–4962. <https://doi.org/10.1109/IGARSS.2015.7326945>
12. Zinko, R., Teslyuk, V., & Seneta, M. (2023). The evaluation of factors that influence the route formation of the mobile rescue robot. *Proceedings of the 7th International Conference on Computational Linguistics and Intelligent Systems*. Kharkiv, Ukraine. In *CEUR Workshop Proceedings*, 3403, 570–581.
13. Borkivskyi, B. P., & Teslyuk, V. M. (2023). Application of neural network tools for object recognition in mobile systems with obstacle avoidance. *Scientific Bulletin of UNFU*, 33(4), 84–89. <https://doi.org/10.36930/40330412>
14. Teslyuk, V., Borkivskyi, B., & Alshawabkeh, H. A. (2022). Models and means of object recognition using artificial neural networks. *Proceedings of the 4th International Workshop MoMLeT&DS 2022*. Leiden–Lviv, the Netherlands–Ukraine. In *CEUR Workshop Proceedings*, Vol. 3312, 241–251.
15. Zinko, R. V., Korendiy, V. M., Tesliuk, V. M., Demchuk, I. B., Kazymyra, I. Ya., & Ostrovka, D. V. (2022). Frequency analysis of the shaft of the electromechanical drive of the small mobile robot. *Industrial Process Automation in Engineering and Instrumentation*, 56, 27–38. <https://doi.org/10.23939/istcipa.2022.56.027>
16. Zinko, R. V., Teslyuk, V. M., Kazymyra, I. Ya., & Ostrovka, D. V. (2022). A model for improving the strength characteristics of the electromechanical drive of a mobile robot. *Ukrainian Journal of Information Technology*, 4(2), 80–85. <https://doi.org/10.23939/ujit2022.02.080>

A. G. Kazarian, O. P. Maksymiv, K. I. Tkachuk, R. V. Parcej, S. V. Tesliuk

Lviv Polytechnic National University, Lviv, Ukraine

ARTIFICIAL NEURAL NETWORKS IMPLEMENTATION IN MOBILE ROBOTIC PLATFORM CONTROL SYSTEM

In the era of rapid technological advancement, when robotics and intelligent systems are becoming an integral part of everyday life, the importance of developing control systems for mobile robotic platforms using artificial neural networks becomes extremely high and relevant. This field not only has significant practical needs but also holds considerable potential for innovative development. The evolution of modern robotics and computational intelligence has necessitated the creation of more efficient and adaptive mobile robotic systems. A system and tools for controlling mobile robotic platforms using artificial neural networks (ANNs) have been developed in this work. By simulating the workings of a neural system, ANNs enable robots not only to react to input data but also to learn to solve complex tasks and adapt to changes in their environment.

One of the key challenges in mobile platform control is the development of effective and intuitive interfaces that provide convenient and reliable interaction between the user and the robotic system. In this context, the use of hand gestures by humans represents a progressive and promising direction as it allows for the creation of the most natural and efficient means of control. The main task is to create an effective and intuitively understandable system that enables the operator to interact with the robotic platform using natural movements and gestures. As a result, software with a graphical interface for real-time gesture recognition using machine learning has been developed.

The scientific novelty of the approach is the integration of advanced ANNs methods to improve the quality of control and functionality of mobile robotic platforms. The main aspects of scientific novelty include integration with artificial intelligence, interactivity of control, development of robotics mobility, and adaptability to various tasks. The problem addressed in this work lies in the need to develop effective and intuitive control systems for mobile robotic platforms using gesture recognition technologies.

A program based on convolutional neural networks has been developed, which determines the position of the hand and identifies specific gestures such as forward, backward, right and left turns, as well as stopping movement. The implemented technology can be used in various fields of human activity (smart home control, technological solutions for people with physical disabilities, enhancing interactivity in entertainment devices, improving interfaces for interacting with technical equipment).

Keywords: automation, interaction interfaces, technological solutions, machine learning, ANNs.

Інформація про авторів:

Казарян Артем Геннадійович, канд. техн. наук, асистент, кафедра автоматизованих систем управління.

Email: artem.kazarian@gmail.com; <https://orcid.org/0000-0002-6883-0233>

Максимів Олег Петрович, магістр, кафедра автоматизованих систем управління. Email: oleh.maksymiv.mknus.2022@lpnu.ua

Ткачук Катерина Ігорівна, аспірант, кафедра автоматизованих систем управління.

Email: kateryna.i.tkachuk@lpnu.ua; <https://orcid.org/0009-0002-1788-8715>

Парцей Руслан Володимирович, аспірант, кафедра автоматизованих систем управління. Email: vasylteslyuk@gmail.com

Теслюк Софія Василівна, аспірант, кафедра систем штучного інтелекту. Email: sofiia.tesliuk.mknus.2021@lpnu.ua;

<https://orcid.org/0009-0005-6512-4447>

Цитування за ДСТУ: Казарян А. Г., Максимів О. П., Ткачук К. І., Парцей Р. В., Теслюк С. В. Використання штучних нейронних мереж у системі управління мобільною робототехнічною платформою. *Український журнал інформаційних технологій*. 2024, т. 6, № 2. С. 30–40.

Citation APA: Kazarian, A. G., Maksymiv, O. P., Tkachuk, K. I., Parcei, R. V., & Tesliuk, S. V. (2024). Artificial neural networks implementation in mobile robotic platform control system. *Ukrainian Journal of Information Technology*, 6(2), 30–40.

<https://doi.org/10.23939/ujit2024.02.030>