

УДК 004.85:[004.896+004.94

DOI: 10.31673/2412-9070.2023.052934

В. Р. МИКОЛАЙЧУК<sup>1</sup>, ст. викладач;Р. А. МИКОЛАЙЧУК<sup>2</sup>, доктор техн. наук, доцент;К. П. СТОРЧАК<sup>1</sup>, доктор техн. наук, професор,<sup>1</sup> Державний університет інформаційно-комунікаційних технологій, Київ<sup>2</sup> Національний університет оборони України імені Івана Черняхівського, Київ

## ВИКОРИСТАННЯ ЗАСОБІВ МАШИННОГО НАВЧАННЯ ДЛЯ МОДЕЛЮВАННЯ СЕРЕДОВИЩА ПРОТИДІЇ МІЖ РОБОТИЗОВАНИМИ ЗАСОБАМИ ТА ОБ'ЄКТАМИ МОНІТОРИНГУ

*Репрезентовано розроблену методику моделювання середовища для задачі пошуку та переслідування рухомих об'єктів роботизованими засобами систем моніторингу місцевості в контексті навчання з підкріпленням. Методика інтегрує в себе низку етапів: від визначення цілей та задач до використання згорткових нейронних мереж для аналізу середовища та визначення його впливу на параметри роботизованих засобів. Особливу увагу приділено адаптації моделі до динамічних змін у середовищі та точності відтворення реальних умов. Результати дослідження показують, що запропонована методика може значно підвищити ефективність навчання роботизованих засобів, забезпечуючи їх здатність до адаптації та оптимальної поведінки в різноманітних сценаріях реального світу, з використанням технологій штучного інтелекту.*

**Ключові слова:** штучний інтелект; математичне моделювання; навчання з підкріпленням; штучні нейронні мережі; роботизовані засоби; сенсорні мережі; системи моніторингу; модель середовища; автоматизація.

### Вступ

**Аналіз останніх досліджень і публікацій.** Протягом останніх десятиліть спостерігаються значні досягнення в галузі штучного інтелекту (машинного навчання), зокрема у застосуванні методів навчання з підкріпленням (*Reinforcement Learning, RL*). Алгоритми *RL* використовуються в різних галузях, таких як маркетинг, робототехніка, ігрова індустрія, автомобільна промисловість, оброблення природної мови, безпека Інтернету речей, системи рекомендацій, фінанси та енергоменеджмент [1; 2]. Це свідчить про широкий спектр можливостей застосування *RL* у розв'язанні різноманітних завдань.

Досвід, набутий у сфері ігор [3-5], де *RL* використовується для створення інтелектуальних та адаптивних агентів, може бути вельми корисним для розвитку робототехнічних систем та систем моніторингу. Уміння ефективно функціонувати за реальних та невизначених умов надає *RL* потенціал стати ключовим елементом у керуванні робототехнічними системами, забезпечуючи їхню адаптивність та інтелектуальність у вирішенні різних завдань. З огляду на аналіз попередніх досліджень у галузі машинного навчання стає очевидним, що *RL* є важливим інструментом для розроблення високоефективних та адаптивних систем керування. Цей підхід дає змогу агентам здобувати досвід через взаємодію з навколишнім середовищем, а потім застосовувати набуті знання для вирішення викликів у реальному часі. Отже, *RL* відкриває перспективи для розроблення інтелектуальних систем, які можуть ефективно адаптуватися до змін у довкіллі та забезпечувати надійний та оптимальний рівень керування за умов невизначеності та мінливості.

**Постановка проблеми.** У сучасному світі велика кількість сенсорних та роботизованих систем активно взаємодіє з навколишніми об'єктами, виконуючи завдання відслідковування, контролю та моніторингу. Досягти оптимальної продуктивності та надійності в таких середовищах може бути складно через рухливість об'єктів та непередбачуваність їхньої поведінки.

Один із підходів до вирішення цих проблем — використання *RL* для моделювання та оптимізації середовища протидії роботизованих засобів та рухомих об'єктів. *RL* дає змогу агентам, зокрема роботам чи сенсорним системам, вчитися взаємодіяти зі змінною динамікою, пристосовуватися до різних сценаріїв та робити більш інтелектуальні рішення.

Розвиток технологій машинного навчання, зокрема глибокого навчання та навчання з підкріпленням (*RL*), відкриває нові можливості для створення інтелектуальних систем роботизованих засобів (*P3*) та об'єктів моніторингу. Ефективне моделювання динамічних середовищ, де *P3* взаємодіють із рухомими об'єктами (*PO*) моніторингу, має значний потенціал у цивільних та військових застосуваннях, від безпілотного транспорту до автоматизованої розвідки та нагляду.

Динамічні середовища, характерні для реального світу, охоплюють безліч викликів: змінні умови, непередбачувану поведінку інших агентів, протидію досягненню цілей системи, що ставить вимогу до

© В. Р. Миколайчук, Р. А. Миколайчук, К. П. Сторчак, 2023

швидкого реагування та адаптації до змін. Моделювання таких середовищ потребує складних математичних та алгоритмічних розв'язків. Використання RL для моделювання поведінки РЗ та РО може значно підвищити їх здатність до самостійного ухвалення рішень у реальному часі та адаптації до змін середовища, що робить цю сферу вкрай актуальною для досліджень.

**Формулювання мети статті.** Метою цієї наукової статті є розроблення моделі середовища для моделювання протидії роботизованих засобів та об'єктів моніторингу з використанням методів машинного навчання, зокрема навчання з підкріпленням. Розглядувані в статті питання спрямовано на побудову математичної моделі середовища протидії, яка враховує фізичні особливості простору моніторингу, включно з параметрами руху, різноманітними завадами та взаємодією між роботизованими агентами.

### Основна частина

Створення математичної моделі середовища для використання в галузі навчання з підкріпленням потребує кількох кроків.

#### 1. Визначення цілей та задач.

**Визначення цілей:** що саме мають робити роботизовані засоби? Наприклад, відшукувати та слідкувати за рухомими об'єктами. **Задачи:** які задачі мають виконувати роботи? Наприклад, навігація до об'єкта, уникнення перешкод, ідентифікація об'єктів.

#### 2. Моделювання середовища.

**Просторове моделювання:** створення карти місцевості, яка може містити перешкоди, зони видимості, різні типи поверхонь тощо. **Динамічні елементи:** визначення параметрів рухомих об'єктів, таких як швидкість, напрямок, поведінка. **Сенсорні моделі:** моделювання давачів роботів, які використовуються для виявлення об'єктів та навігації (наприклад, лідари, камери, ультразвукові давачі).

#### 3. Визначення станів, дій та нагород.

**Стани:** визначення можливих станів середовища, які можуть спостерігати роботи. **Дії:** визначення набору дій, які можуть виконувати роботи (наприклад, рух вперед, поворот, захоплення об'єкта). **Нагороди та покарання:** створення системи нагород, яка мотивує роботів виконувати бажані дії, та покарань за помилки або неефективні дії.

Далі наведемо вербальне визначення цілей та задач системи.

**Цілі роботизованих засобів.** РЗ мають за ціль виявлення та слідкування за рухомими об'єктами. Основна мета полягає в тому, щоб мінімізувати час та відстань, потрібні для досягнення дистанції, достатньої для ідентифікації РО. Це охоплює здатність РЗ рухатися місцевістю, що впливає на швидкість і маневреність, та використовувати свої сенсори для точного визначення місцезнаходження РО.

**Задачи РЗ:**

1. Навігація до РО, зважаючи на місцевість та перешкоди.
2. Уникнення перешкод та інших РЗ під час руху.
3. Ідентифікація РО, коли РЗ наближаються на достатню відстань.

**Цілі рухомих об'єктів:** РО намагаються уникнути зближення з РЗ, тобто зміна швидкості, напрямку руху та використання місцевості для приховування свого місцезнаходження.

З погляду математичної формалізації введемо наведені далі позначення та функції.

• **Позначення:**  $x_t^{(r)}$  — положення РЗ у момент часу  $t$ ;  $x_t^{(o)}$  — положення РО у момент часу  $t$ ;  $v_t^{(r)}$  — швидкість РЗ у момент часу  $t$ ;  $v_t^{(o)}$  — швидкість РО у момент часу  $t$ ;  $a_t^{(r)}$  — дія, вибрана РЗ у момент часу  $t$ ;  $d_{\text{ид}}$  — відстань, достатня для ідентифікації РО РЗ;  $S(x)$  — функція стану місцевості, що впливає на параметри руху.

• **Цільова функція для РЗ:** мінімізувати  $T$ , де  $T$  — час, потрібний для досягнення умови  $\|x_t^{(r)} - x_t^{(o)}\| \leq d_{\text{ид}}$ .

• **Динаміка РЗ:**

$$\begin{aligned} x_{t+1}^{(r)} &= x_t^{(r)} + v_t^{(r)}\Delta t + 1/2 S(x_t^{(r)})a_t^{(r)}\Delta t^2; \\ v_{t+1}^{(r)} &= v_t^{(r)} + S(x_t^{(r)})a_t^{(r)}\Delta t. \end{aligned} \quad (1)$$

• **Динаміка РО:**

$$x_{t+1}^{(o)} = x_t^{(o)} + v_t^{(o)}\Delta t + \text{стратегія втечі}; \quad (2)$$

$$V_{t+1}^{(o)} = \text{функція від стратегії втечі}. \quad (3)$$

• **Функція втрат для РЗ:**

$$J(a_t^{(r)}) = \sum_{t=0}^T \gamma^t C(x_t^{(r)}, x_t^{(o)}, a_t^{(r)}), \quad (4)$$

де  $C$  — вартість дії, що містить у собі відстань до РО та витрати на маневрування;  $\gamma$  — фактор дисконтування.

• **Проблема оптимізації для РЗ:**

$$\min_{a_t^{(r)}} J(a_t^{(r)}) \quad (5)$$

з обмеженнями на динаміку РЗ та умовами безпеки.

Запропонована формалізація охоплює моделювання руху РЗ та РО, зважаючи на вплив місцевості на параметри руху та ідентифікації. Вона також враховує стратегічну взаємодію між РЗ та РО, де РО намагаються уникнути зближення, а РЗ прагнуть мінімізувати відстань до РО для їх ідентифікації.

Далі запропоновано результати детального моделювання середовища.

**Просторове моделювання.** Просторове моделювання охоплює створення віртуальної карти місцевості, яка відображає реальні умови якомога точніше. Це може бути зроблено за допомогою таких кроків:

- **збір даних** — використання наявних картографічних сервісів (наприклад, Google Maps, OpenStreetMap) для отримання базової карти місцевості;
- **доповнення даними** — інтеграція даних із нерухомих сенсорів, зокрема камер та давачів руху, для додавання інформації про перешкоди, зони видимості та інші важливі особливості;
- **анотація карти** — ручне або автоматизоване додавання метаданих до карти, таких як типи поверхонь, висота об'єктів та інші фізичні характеристики.

**Динамічні елементи.** До динамічних елементів середовища належать рухомі об'єкти, які можуть бути змодельовані за допомогою таких параметрів:

- **швидкість  $v$**  — скалярна або векторна величина, що визначає, наскільки швидко об'єкт переміщується в просторі;
- **напрямок  $\theta$**  — вектор, що показує напрямок руху об'єкта;
- **поведінка** — набір правил або алгоритмів, які визначають, як об'єкт реагує на зміни в середовищі.

Відповідно до наведених кроків проведемо математичну формалізацію середовища протидії роботизованих засобів та об'єктів моніторингу.

**1. Карта місцевості  $M$ .** Таку карту можна подати як матрицю або граф, де кожен елемент або вузол відповідає певній ділянці в просторі. Кожен елемент містить інформацію про тип поверхні  $S$ , наявність перешкод  $O$  та інші характеристики.

**2. Вплив місцевості на параметри руху.** Коефіцієнт швидкості  $k_v(S)$  змінює швидкість об'єкта залежно від типу поверхні. Коефіцієнт маневреності  $k_m(S)$ , який впливає на здатність об'єкта змінювати напрямки на різних типах поверхонь.

**3. Динаміка рухомих об'єктів.** Положення об'єкта  $x_t$  в часі  $t$  може бути визначено як  $x_{t+1} = x_t + k_v(S) \times v_t \cdot \Delta t$ , де  $\Delta t$  — крок часу.

Напрямок руху об'єкта може бути оновлено з огляду на коефіцієнт маневреності та вхідних дій  $\theta_{t+1} = \theta_t + k_m(S) \cdot \Delta \theta$ , де  $\Delta \theta$  — зміна напрямку.

Ця формалізація дає змогу створити математичну модель середовища, яку можна використовувати для симуляції та аналізу поведінки роботизованих засобів та рухомих об'єктів за різних умов.

**Поняття станів середовища.** Нехай  $S_t$  позначає стан середовища в час  $t$ , який містить у собі положення та швидкість усіх роботизованих засобів  $R_t$  і рухомих об'єктів  $O_t$ , а також стан місцевості та інших елементів середовища  $M_t$ . Тоді  $S_t = \{R_t, O_t, M_t\}$ .

**Рівняння переходу** описує, як стан середовища  $S_t$  змінюється з часом до  $S_{t+1}$ , зважаючи на дії роботизованих засобів та поведінку рухомих об'єктів, а також стан місцевості, тобто:

- для роботизованих засобів — положення:  $x_{t+1}^{(r)} = x_t^{(r)} + k_v(M_t) \cdot v_t^{(r)} \cdot \Delta t$ ; швидкість:  $v_{t+1}^{(r)} = v_t^{(r)} + a_t^{(r)} \cdot \Delta t$ , де  $a_t^{(r)}$  — прискорення, яке може залежати від керування та властивостей місцевості  $M_t$ ;
- для рухомих об'єктів — положення:  $x_{t+1}^{(o)} = x_t^{(o)} + k_v(M_t) \cdot v_t^{(o)} \cdot \Delta t$ ; швидкість:  $v_{t+1}^{(o)} = v_t^{(o)} + a_t^{(o)} \cdot \Delta t$ , де  $a_t^{(o)}$  — прискорення, яке може бути результатом автономної поведінки або взаємодії з роботизованими засобами;

- для стану місцевості та інших елементів середовища  $M$  — оскільки стан місцевості  $M_t$  містить інформацію про перешкоди, типи поверхонь, зони видимості тощо, то зміни в стані місцевості можуть відбуватися через зовнішні впливи або дії роботизованих засобів:  $M_{(t+1)} = f(M_t, R_t, O_t)$ , де  $f$  — функція, що описує динаміку місцевості.

Коефіцієнти  $k_v(M_t)$  та множини дій  $a_t^{(r)}, a_t^{(o)}$  можуть бути функціями стану місцевості  $M_t$ , які визначають, як середовище впливає на рухомість та поведінку об'єктів. Наприклад, коефіцієнт швидкості може зменшуватися в зонах із високою щільністю перешкод або на складних поверхнях.

Запропонована модель дає змогу симулювати динамічну взаємодію між роботизованими засобами, рухомими об'єктами та змінним середовищем, що є ключовим для розроблення алгоритму.

Для створення математичної моделі, яка використовується в навчанні з підкріпленням RL, потрібно визначити стани, дії та нагороди, які використовуватимуться в нашому разі агентами (роботами) для навчання.

Стани середовища  $S$  — це всі можливі спостереження, які робот може зробити. До них належать:

- положення робота  $p_r$  та рухомих об'єктів  $p_o$ ;
- швидкість робота  $v_r$  та рухомих об'єктів  $v_o$ ;
- стан перешкод  $M$  та інших елементів середовища;
- інформація про останню відому позицію рухомих об'єктів.

Дії — це набір можливих кроків, які робот може виконати, наприклад, рух вперед, поворот, захоплення об'єкта.

Функція нагороди визначається на основі зменшення дистанції між РЗ та РО порівняно з «ідеальною» дистанцією, яка була б, якби обидва об'єкти рухалися назустріч один одному з максимальною швидкістю без перешкод:

$$\begin{aligned} R_t^{(r)} &= k(d_{t-1}^* - d_t); \\ R_t^{(o)} &= -R_t^{(r)}, \end{aligned} \quad (6)$$

де  $d_t$  — фактична дистанція між РЗ та РО;  $d_{t-1}^*$  — «ідеальна» дистанція;  $k$  — позитивний коефіцієнт масштабування.

Моделювання завершується, коли:

- РЗ досягає РО ( $d_t \leq \epsilon$ );
- РО виходить за межі визначеного середовища ( $p_o \notin \text{Environment Bounds}$ );
- РЗ потрапляє в ділянку, де він не може функціонувати ( $M(p_r) = \text{«untraversab»}$ ).

Отже, наведена модель дає змогу створити чітку систему нагород, яка мотивує РЗ ефективно переслідувати РО, а РО — уникати переслідування, та визначає чіткі умови для завершення симуляції.

Тепер з використанням CNN визначимо коефіцієнти  $k$  впливу середовища на параметри РЗ та РО. Сформуємо графічне відображення середовища. Розширимо точку знаходження РЗ (РО) до певної ділянки. Створимо набір даних типових ділянок середовища із заданими значеннями параметрів. Виконаємо доповнення набору даних через змішування у випадковій пропорції типових ділянок із відповідною апроксимацією значень параметрів. Далі створюємо CNN для визначення значень параметрів у будь-якій ділянці середовища. Надамо детальний опис, алгоритм та математичну формалізацію методики визначення впливу середовища на параметри РЗ (РО).

Для визначення коефіцієнтів  $k$  впливу середовища на параметри роботизованих засобів та рухомих об'єктів за допомогою конволюційних нейронних мереж (CNN), можна використати запропонований далі підхід.

1. *Створення графічного відображення середовища.* Спершу потрібно створити графічне відображення середовища, яке може бути подано як карти з різними текстурами, що відображають типи поверхонь (трава, пісок, вода тощо).

2. *Розширення точки перебування РЗ (РО) до певної ділянки.* Кожен РЗ або РО може впливати на ділянку середовища, яка його оточує. Це може бути подано як радіус дії або як зона впливу навколо кожного об'єкта.

3. *Створення набору даних типових ділянок середовища.* Потрібно зібрати набір даних, який містить зображення типових ділянок середовища з відомими значеннями параметрів, зокрема шорсткістю, кутом нахилу, вологістю тощо.

4. *Доповнення набору даних.* Для збільшення різноманітності набору даних можна використати техніку аугментації даних, такі як змішування зображень ділянок у випадковій пропорції та апроксимація значень параметрів для цих нових змішаних ділянок.

5. *Створення CNN для визначення параметрів ділянки середовища.* CNN може бути навчена на основі зібраного набору даних для визначення значень параметрів будь-якої ділянки середовища залежно від її зображення.

Грунтуючись на даному підході, було розроблено алгоритм та математичну формалізацію методики.

1. Попереднє оброблення даних:

• збір даних:  $D = \{(x_i, y_i)\}_{i=1}^N$ , де  $x_i$  — зображення ділянки, а  $y_i$  — вектор параметрів ділянки;

• аугментація даних: створення нових зображень  $x_{\text{new}}$  через комбінування зображень  $x_i$  та апроксимація відповідних  $y_{\text{new}}$ .

2. Архітектура CNN:

• визначення архітектури CNN:  $f_{\theta}(x)$ , де  $\theta$  — параметри мережі, які підлягають навчанню;

• втрати:  $L(\theta) = \frac{1}{2} \sum_{i=1}^N \text{loss}(f_{\theta}(x_i), y_i)$ , де  $\text{loss}$  — функція втрат, наприклад, середньоквадратична помилка.

3. Навчання мережі:

• використання алгоритму зворотного поширення помилки для оптимізації  $\theta$  з метою мінімізації  $L(\theta)$ .

4. Використання CNN для визначення  $k$ :

- для нового зображення  $x_{\text{test}}$  CNN визначає параметри ділянки  $\hat{y}_{\text{test}} = f_{\theta}(x_{\text{test}})$ ;
- коефіцієнти  $k$  визначаються на основі  $\hat{y}_{\text{test}}$ , наприклад через лінійну регресію або іншу модель залежності.

Такий підхід дає змогу автоматизувати процес визначення впливу середовища на параметри РЗ та РО, забезпечуючи адаптацію до різноманітних умов середовища.

Загальну методику моделювання середовища для задачі пошуку та переслідування роботизованими засобами рухомих об'єктів в контексті навчання з підкріпленням можна сформулювати в такий спосіб.

1. *Визначення цілей та задач.* Установлення конкретних цілей для РЗ та РО. Формулювання задач, які мають виконувати РЗ для досягнення цілей.

2. *Моделювання середовища.* Створення графічного відображення середовища, включно з перешкодами, зонами видимості, різними типами поверхонь. Визначення динамічних елементів, зокрема параметрів руху РО.

3. *Визначення станів, дій та нагород.* Опис можливих станів середовища  $S$ , які містять положення та швидкість РЗ і РО, а також стан перешкод  $M$ . Визначення набору дій  $A$ , які можуть виконувати РЗ. Розроблення системи нагород, яка мотивує РЗ виконувати бажані дії, та покарання за помилки.

4. *Формулювання функції переходу станів.* Визначення рівнянь переходу, які описують, як середовище переходить з одного стану в інший відповідно до дій РЗ.

5. *Визначення функції нагороди.* Формулювання функції нагороди, яка базується на зменшенні дистанції між РЗ та РО та інших критеріях.

6. *Критерії завершення моделювання.* Установлення умов, за яких моделювання завершується, зокрема досягнення РО, вихід РО за межі середовища або неможливість функціонування РЗ.

7. *Використання CNN для визначення впливу середовища.* Збір та аугментація даних про середовище. Навчання CNN на основі зібраних даних для визначення коефіцієнтів  $k$ , які впливають на параметри РЗ та РО. Використання CNN для аналізу нових ділянок середовища та прогнозування впливу на РЗ та РО.

8. *Інтеграція моделі в RL.* Упровадження моделі середовища в алгоритми RL. Використання отриманих даних для оптимізації стратегій поведінки РЗ та РО.

Ця методика дає можливість створити детальну та адаптивну модель середовища, яку можна використовувати для тренування РЗ за різноманітних умов та підвищення ефективності їхніх дій у реальному світі.

### Висновки

У процесі проведеного дослідження було розроблено комплексну методику моделювання середовища для задачі пошуку та переслідування рухомих об'єктів роботизованими засобами в межах навчання з підкріпленням. Методика охоплює визначення цілей та задач, моделювання середовища, визначення станів, дій та нагород, формулювання функції переходу станів, визначення функції нагороди, критерії завершення моделювання, використання згорткових нейронних мереж для аналізу середовища та інтеграцію моделі в алгоритми навчання з підкріпленням.

Особливістю розробленої методики є її здатність до адаптації під різноманітні умови середовища та динамічні зміни, що відбуваються в ньому. Використання графічного відображення середовища уможливує детальне відтворення реальних умов, в яких діють роботизовані засоби, включно з перешкодами, зонами видимості та різними типами поверхонь.

Використання згорткових нейронних мереж для аналізу середовища є інноваційним підходом, який дає змогу автоматизувати процес визначення впливу середовища на параметри роботизованих засобів та рухомих об'єктів. Навчання CNN на зібраних даних про середовище забезпечує точне прогнозування коефіцієнтів впливу на різних ділянках, оптимізуючи стратегію поведінки роботів залежно від змін умов.

Загалом, розроблена методика моделювання середовища здатна не тільки точно відтворити реальні умови, в яких діють роботизовані засоби, а й забезпечити гнучкість та адаптивність у разі зміни умов середовища. Це відкриває широкі перспективи для тренування та вдосконалення алгоритмів навчання з підкріпленням, підвищуючи ефективність виконання роботами своїх задач у різних сценаріях реального світу.

## Список використаної літератури

1. *A Systematic Study on Reinforcement Learning Based Applications* / K. Sivamayil [et al.] // *Energies*. 2023. Vol. 16, no. 3. P. 1512. URL: <https://doi.org/10.3390/en16031512>.
2. *Advances in Machine Learning for Sensing and Condition Monitoring* / S.-I. Ao [et al.] // *Applied Sciences*. 2022. Vol. 12, no. 23. P. 12392. URL: <https://doi.org/10.3390/app122312392>.
3. *AI Soldier using Reinforcement Learning* / S. Bhojwani [et al.] // *International Journal of Innovative Science and Research Technology*. 2019. Vol. 4, no. 4. P. 1122–1125. URL: <https://ijisrt.com/wp-content/uploads/2019/06/IJISRT19AP623.pdf>.
4. Almeida P., Carvalho V., Simões A. *Reinforcement Learning Applied to AI Bots in First-Person Shooters: A Systematic Review* // *Algorithms*. 2023. Vol. 16, no. 7. P. 323. URL: <https://doi.org/10.3390/a1607032>.
5. Souchleris K., Sidiropoulos G. K., Papakostas G. A. *Reinforcement Learning in Game Industry—Review, Prospects and Challenges* // *Applied Sciences*. 2023. Vol. 13, no. 4. P. 2443. URL: <https://doi.org/10.3390/app13042443>.

V. R. Mykolaichuk, R. A. Mykolaichuk, K. P. Storchak

**THE USE OF MACHINE LEARNING TOOLS FOR MODELING THE COUNTERMEASURE ENVIRONMENT OF ROBOTIC MEANS AND MOBILE OBJECTS OF MONITORING**

In this article, an analysis of research related to the problems and areas of application of machine learning was carried out and it becomes clear that reinforcement learning is an important tool for the development of highly effective and adaptive control systems. This approach opens perspectives for the development of intelligent systems that can effectively adapt to changes in the environment and provide a reliable and optimal level of control under conditions of uncertainty and variability. In the course of the research, a complex environment modelling technique was developed for the task of searching and chasing moving objects by robotic means as part of reinforcement learning. The methodology includes defining goals and objectives, modelling the environment, defining states, actions, and rewards, formulating the state transition function, defining the reward function, modelling completion criteria, using convolutional neural networks to analyse the environment, and integrating the model into reinforcement learning algorithms. A feature of the developed methodology is its ability to adapt to various environmental conditions and dynamic changes occurring in it. The use of graphical representation of the environment allows you to reproduce in detail the real conditions in which the robotic means operate, including obstacles, areas of visibility and different types of surfaces. The use of convolutional neural networks for environmental analysis is an innovative approach that allows you to automate the process of determining the impact of the environment on the parameters of robotic vehicles and moving objects. CNN training on the collected data about the environment provides accurate prediction of influence coefficients in different areas, which allows optimizing strategies of robot behaviour depending on changes in conditions. In general, the developed environment modelling technique allows not only to accurately reproduce the real conditions in which robotic means operate, but also provides flexibility and adaptability when environmental conditions change. This opens up vast opportunities for training and improving reinforcement learning algorithms, increasing the effectiveness of robots in performing their tasks in various real-world scenarios.

**Keywords:** artificial intelligence; mathematical modeling; reinforcement learning; artificial neural networks; robotics; sensor networks; monitoring systems; environment model; automation.