

ФА79-НС1-054

ФАКУЛТЕТ АВТОМАТИКА

Вх. № 13-6/05.06.2026г.

## Рецензия

От доц. д-р Денис Сафидинов Чикуртев

Върху дисертационен труд за придобиване на образователна и научна степен „доктор“

Професионално направление: **5.2. Електротехника, електроника и автоматика**

Докторска програма: “**Системи с изкуствен интелект**“

Автор на дисертационния труд: **маг. инж. Анастасия Владимировна Славова**

Тема на дисертационния труд: **СИСТЕМИ С ДЪЛБОКО ОБУЧЕНИЕ ПРИ  
АВТОНОМНИТЕ МОБИЛНИ РОБОТИ**

### Актуалност на проблема

Проблемът за навигация на автономни мобилни роботи в динамична, непозната среда без предварително изградени карти и GPS е изключително актуален както научно, така и индустриално. В дисертацията ясно е аргументирана значимостта на темата, включително чрез фокус върху използването на бюджетен LiDAR и дълбоко обучение с подсилване за намаляване на цената и сложността на системите.

Изследването е строго насочено към навигация на автономни мобилни роботи (AMP) в затворени, структурирани пространства в условия на неопределеност. Подчертава се ролята на AMP за оптимизация на вътрешната логистика, складовите операции и работата в опасни среди (пожароопасност, радиация).

### Степен на познаване състоянието на проблема и обща характеристика на труда

Кандидатът демонстрира много добро познаване на съвременното състояние на изследванията в областта на дълбокото обучение с подсилване за навигация на мобилни роботи, включително DQN, A2C, TRPO, PPO, както и ролята на сензорните системи (LiDAR, одометрия, камери). Литературният обзор е широк, систематичен и добре структуриран, като са обсъдени предимства, ограничения и проблеми на разглежданите подходи, включително скорост на обучение, стабилност и точност на моделите.

Авторът представя задълбочено разбиране за параметрите и характеристиките на автономната навигация при колесни мобилни роботи. В детайли разглежда проблемите при прилагане на навигационните алгоритми и анализира актуални решения.

Дисертационният труд е структуриран в 4 глави, в обем от 150 страници, съдържащ 80 фигури, 23 таблици и математически зависимости, 105 литературни източника и едно приложение. Общото впечатление е за добре организиран, логично изграден труд, в който

теоретичната част е последователно свързана с симулационните и експерименталните изследвания.

### **Целта на дисертацията**

Целта на дисертационния труд може да се формулира като разработване, обучение и експериментална валидация на система за управление на автономен мобилен робот, базирана на дълбоко обучение с подсилване и използваща основно LiDAR и одометрия, без предварителни карти и GPS. Тази цел е ясно мотивирана във въведението и е логично разпределена в отделните глави чрез конкретни подзадачи: избор и анализ на алгоритми, симулационно обучение в 2D и 3D среда, и пренос към реален робот.

Представена е една хипотеза, която се основана на визията, че подобряването на заложените алгоритми и архитектурни решения в метода PPO (модифицирана невронна архитектура, адаптивна скорост на обучение и оптимизирано използване на буфер за опит), би довело до подобряване на навигационните параметри и поведение на автономен мобилен робот.

Дефинирани са пет задачи, които са последователни и логически свързани. Изпълнението на задачите, води до успешно постигане на зададената цел и до потвърждение на поставената хипотеза.

### **Кратко съдържание по глави**

**В Глава 1** е представен подробен обзор на дълбокото машинно обучение при автономни мобилни роботи, с акцент върху RL/DRL, функциите на възнаграждане и основните алгоритми (DQN, A2C, TRPO, PPO), както и проблемите при тяхното приложение за навигация. Направен е задълбочен сравнителен анализ на четири от основните алгоритми в дълбокото обучение с подсилване (DRL) – DQN, A2C, TRPO и PPO. Изследваните алгоритми са съпоставени по ключови за автономната навигация показатели, включително стабилност на процеса на обучение, ефективност при използването на натрупания опит (*sample efficiency*), изисквания към изчислителните ресурси и чувствителност на моделите към настройката на хиперпараметрите. Разгледано е детайлно математическото описание на Марковските процеси, които стоят в основата на вземането на решения при управлението на автономни системи. Главата завършва с формулиране на целта и задачите на дисертацията, което създава добра логическа рамка за останалата част от труда.

**В Глава 2** е описан изборът на симулационна среда Flatland и ROS2, дефинирана е функцията на възнаграждане и са обучени и сравнени няколко DRL модела за навигация в 2D среда. Авторът успешно изгражда специализирана експериментална среда в симулатора Flatland, работещ на базата на Python библиотеките *PyMunk* и *Pygame*. За управлението и комуникацията между отделните модули на системата е внедрен интеграционният

софтуерен пакет *ROS2* Направен е сравнителен анализ на алгоритмите по критерии като скорост на обучение, стабилност и успеваемост, което позволява аргументиран избор на най-подходящ модел за по-нататъшно развитие.

**В Глава 3** се разработва 3D симулационна среда в *Gazebo* с 3D модел на робота, оборудван с *LiDAR*, и се изгражда система за обучение по метода *PPO*. Предложени и реализирани са специфични подобрения в архитектурата на използваната дълбока невронна мрежа (ДНМ). Извършено е прецизиране на скоростта на обучение (*learning rate*) и е оптимизиран буферът за съхранение на опита с цел постигане на по-бърза сходимост и стабилност на модела. Анализирани са различни архитектури на дълбоката невронна мрежа, хиперпараметри, буфер за опит и скорост на обучение, като е проведен аблационен анализ на подобрения модел. Резултатите от аблационния анализ експериментално доказват и оценяват индивидуалният принос на всяка една от направените модификации върху крайната ефективност на навигационната система.

**В Глава 4** се описва реалната експериментална постановка с *AMP Yahboom RDK X3* с механум колела, конфигуриран с *ROS2*, и се изследва поведението на обучените модели в реална затворена среда. Извършен е прецизен процес на калибриране на линейната и ъгловата скорост на платформата. Детайлно е описана конфигурацията на софтуерната архитектура в среда на *ROS2*, включително управлението на пространствените трансформации и координатните системи чрез библиотеката *tf2*. Основен акцент е поставен върху изследването на устойчивостта и поведението на робота при пренос на навигационния модел от симулация към реална физическа среда. Проведени са експерименти при различен брой препятствия и сценарии, анализирани са траектории, крайни състояния и успеваемост, както и ефектът от преминаване от симулация към реална среда. Направен е детайлен анализ на робастността на управлението при преминаване през трасета с различен брой препятствия, с което експериментално се доказва способността на робота да навигира безопасно при ограничена сензорна информация.

#### **Съответствие на предложената методика на изследване и поставените цел и задачи**

Предложената методика е последователна и добре съгласувана с целта и задачите на дисертацията: от 2D симулация и избор на алгоритъм, през 3D симулация и оптимизация на модела, до реални експерименти с физически робот. Използването на *Flatland* и *Gazebo*, заедно с *ROS2*, е адекватен и съвременен избор, който позволява контролирано обучение и сравнително плавен пренос към реална среда.

Като критична бележка може да се отбележи, че в някои части методиката би могла да бъде описана по-детайлно от гледна точка на възпроизводимост—например по-точни параметри на хардуера за симулации, времеви разход за обучение, както и по-формална статистическа

обработка на резултатите (повече повторения, доверителни интервали, тестове за значимост).

## Приноси

В дисертационния труд авторът е формулирал детайлен списък с приноси. Смятам, че представеният брой приноси е излишно разширен, като в него са включени рутинни инженерни задачи, етапи от софтуерната конфигурация като калибриране на платформи и настройка на координатни системи в ROS2 и последователни стъпки от симулационното моделиране. Тези дейности по своята същност представляват изпълнение на конкретните задачи по дисертацията, а не самостоятелни научни или научно-приложни постижения.

На базата на критичен преглед на проведените изследвания и реално постигнатите резултати, предлагам приносите на докторанта да бъдат прецизирани до следните три основни научно-приложни приноса, които съответстват на обективните резултати в труда:

1. Научно-приложен принос (Модификация и оптимизация на алгоритъма за управление): Разработена е модифицирана архитектура на DRL алгоритъма PPO (Proximal Policy Optimization) чрез архитектурни промени в дълбоката невронна мрежа, внедряване на адаптивна скорост на обучение и оптимизация на буфера за опит. Чрез детайлен аблационен анализ е доказано математически и експериментално, че модификациите значително ускоряват сходимостта и повишават стабилността на обучението.
2. Приложен принос (Етапно симулационно моделиране в ROS2): Създадена и софтуерно интегрирана е цялостна методология за етапно тестване на автономната навигация – от двумерна среда (*Flatland*) до реалистична триизмерна среда (*Gazebo*), базирана на симулирани данни от бюджетен LiDAR сензор (*MS200*) и одометрия.
3. Приложен принос (Успешна Sim2Real валидация): Реализиран и експериментално оценен е успешен пренос на обучените модели от симулация към реална физическа среда (Sim2Real) върху мобилен робот *Yahboom RDK X3* с меканум колела. Доказана е робастността на системата за управление при избягване на препятствия в затворени пространства в условия на неопределеност и при липса на GPS и предварително изградени карти.

## Публикации по дисертацията

Кандидатът е представил 7 научни публикации, свързани с темата на дисертационния труд, включително статии в списания и доклади на конференции. Прави добро впечатление, че всички публикации са реферирани в Scopus, което подчертава и засилва научния и научно-

приложния потенциал на дисертацията. Публикациите отразяват основните етапи от изследването—от симулационните модели до реалните експерименти—и показват последователно научно развитие по темата.

**Оценка на съответствието на автореферата с изискванията за изготвянето му, както и на адекватността на отразяване на основните положения и приносите на дисертационния труд.**

Представеният автореферат на маг. инж. Анастасия Владимировна Славова напълно съответства на нормативните изисквания за неговото изготвяне, като по структура, обем и съдържание следва стриктно логиката и номерацията на разработения дисертационен труд. Текстът е систематизиран в обем от 32 страници и включва всички задължителни компоненти: обща характеристика, подробно резюме по глави, научно-приложни и приложни приноси, списък на публикациите и абстракт на английски език.

**Мнения, препоръки и бележки**

Като цяло дисертационният труд на маг. инж. Анастасия Славова е разработен на много високо научно и научно-приложно ниво. Изследванията са проведени систематично, с прецизен преход от теоретични постановки към симулации и финална физическа реализация. Постигнатите резултати демонстрират, че докторантът притежава задълбочени теоретични познания и завидни практически умения в областта на роботиката, изкуствения интелект и софтуерното инженерство.

Въпреки отличните впечатления от дисертацията, с цел подобряване на бъдещата научноизследователска дейност на автора, мога да отправя следните препоръки и бележки:

- **Относно представянето на графичния материал:** В някои от главите графиките, изобразяващи резултатите от обучението на моделите (напр. кривите на функцията на възнаграждение), ще бъдат по-добре отчетени и разбрани ако бъдат по-детайлно разписни на легендите и скалите, за да бъдат по-лесно четими в самостоятелен контекст.
- **Бъдещо развитие:** Препоръчвам на докторанта да потърси възможности за тестване на разработената система в динамични среди с по-висока гъстота на непредвидими препятствия (например движещи се хора), за да се провери пределът на робастност на модифицирания PPO алгоритъм.

Бележки:

- Модифицираният от Вас PPO алгоритъм включва промени в архитектурата на дълбоката невронна мрежа. Как тези промени рефлектират върху изчислителната

натовареност (времето за обработка на един такт) при изпълнение на алгоритъма в реално време директно върху бордовия компютър на робота?

- Постигната ли е пълна синхронизация с честотата на опресняване на данните от ROS2?

### **Заклучение**

Представеният дисертационен труд отговаря на изискванията за придобиване на образователната и научна степен „доктор“ в професионално направление 5.2 „Електротехника, електроника и автоматика“. Трудът съдържа ясно формулирана цел, последователна методика, оригинални научно-приложни и приложни приноси, подкрепени със симулационни и реални експериментални резултати. В заключение давам **положителна оценка** на дисертационни труд.

Убедено препоръчван на научното жури да присъди на **маг. инж. Анастасия Владимировна Славова** образователната и научна степен „доктор“ в професионално направление **5.2. Електротехника, електроника и автоматика**.

Дата: 05.06.2026 г.

Член на журито:

*121*

(доц. д-р Денис Чикуртев)

*Върно с оригинала*



ФД 79 - НС1 - 054

ФАКУЛТЕТ АВТОМАТИКА

Вх. № 136/05.06.2026г.

## REVIEW

By: Assoc. Prof. Denis Safidinov Chikurtev, PhD

**On the Dissertation Thesis for acquiring the educational and scientific degree "Doctor" (PhD)**

**Professional Field:** 5.2. Electrical Engineering, Electronics and Automation

**Doctoral Program:** Artificial Intelligence Systems

**Author of the Dissertation:** Mag. Eng. Anastasia Vladimirovna Slavova

**Dissertation Topic:** DEEP LEARNING SYSTEMS IN AUTONOMOUS MOBILE ROBOTS

### Relevance of the Problem

The problem of autonomous mobile robot navigation in dynamic, unknown environments without pre-built maps and GPS is highly relevant from both a scientific and industrial perspective. The dissertation clearly argues the significance of the topic, specifically focusing on the use of budget-friendly LiDAR and deep reinforcement learning to reduce the cost and complexity of such systems.

The research is strictly focused on the navigation of Autonomous Mobile Robots (AMRs) within enclosed, structured spaces under conditions of uncertainty. The role of AMRs in optimizing internal logistics, warehouse operations, and performance in hazardous environments (e.g., fire hazards, radiation) is strongly emphasized.

### Degree of Knowledge of the State of the Problem and General Characteristics of the Thesis

The candidate demonstrates an excellent grasp of the current state of research in the field of deep reinforcement learning for mobile robot navigation, including DQN, A2C, TRPO, PPO, as well as the role of sensor systems (LiDAR, odometry, cameras). The literature review is comprehensive, systematic, and well-structured, discussing the advantages, limitations, and challenges of the examined approaches, including training speed, stability, and model accuracy.

The author presents a profound understanding of the parameters and characteristics of autonomous navigation in wheeled mobile robots. The challenges of applying navigation algorithms are analyzed in detail alongside up-to-date solutions.

The dissertation is structured into 4 chapters, spanning 150 pages, containing 80 figures, 23 tables and mathematical dependencies, 105 literary sources, and one appendix. The overall impression is that of a well-organized, logically constructed work in which the theoretical part is consistently linked with simulation and experimental studies.

### Aim of the Dissertation

The aim of the dissertation can be formulated as the development, training, and experimental validation of a control system for an autonomous mobile robot based on deep reinforcement learning, utilizing primarily LiDAR and odometry without pre-built maps and GPS. This goal is clearly motivated in the introduction and logically distributed across individual chapters through

specific subtasks: the choice and analysis of algorithms, simulation training in 2D and 3D environments, and transfer to a real robot.

A central hypothesis is presented, based on the vision that improving the core algorithms and architectural solutions within the PPO method (modified neural architecture, adaptive learning rate, and optimized experience replay buffer utilization) would lead to an improvement in the navigation parameters and behavior of an autonomous mobile robot.

Five sequential and logically connected tasks are defined. The execution of these tasks leads to the successful achievement of the specified goal and confirms the stated hypothesis.

### **Brief Content by Chapters**

**In Chapter 1**, a detailed overview of deep machine learning in autonomous mobile robots is presented, focusing on RL/DRL, reward functions, core algorithms (DQN, A2C, TRPO, PPO), and the problems of their application in navigation. A thorough comparative analysis of four main algorithms in deep reinforcement learning (DRL) – DQN, A2C, TRPO, and PPO – is conducted. The examined algorithms are compared based on indicators critical to autonomous navigation, including training process stability, experience replay efficiency (sample efficiency), computational resource requirements, and model sensitivity to hyperparameter tuning. A detailed mathematical description of Markov processes, which form the basis of decision-making in the control of autonomous systems, is covered. The chapter concludes with the formulation of the dissertation's aim and tasks, establishing a solid logical framework for the rest of the thesis.

**In Chapter 2**, the choice of the Flatland simulation environment and ROS2 is described, the reward function is defined, and several DRL models are trained and compared for navigation in a 2D environment. The author successfully builds a specialized experimental environment in the Flatland simulator, which operates using the Python libraries Pymunk and Pygame. The integration software package ROS2 is deployed for system module management and communication. A comparative analysis of the algorithms is performed based on criteria such as training speed, stability, and success rate, allowing for a well-reasoned selection of the most suitable model for further development.

**In Chapter 3**, a 3D simulation environment in Gazebo is developed featuring a 3D robot model equipped with LiDAR, and a training system using the PPO method is built. Specific improvements in the architecture of the utilized Deep Neural Network (DNN) are proposed and implemented. The learning rate is refined, and the experience replay buffer is optimized to achieve faster convergence and model stability. Various DNN architectures, hyperparameters, experience buffers, and learning rates are analyzed, culminating in an ablation analysis of the enhanced model. The results of the ablation analysis experimentally prove and evaluate the individual contribution of each modification to the final efficiency of the navigation system.

**In Chapter 4**, the real experimental setup using the AMR Yahboom RDK X3 with Mecanum wheels, configured with ROS2, is described, and the behavior of the trained models in a real indoor environment is investigated. A precise calibration process of the platform's linear and angular velocity is performed. The software architecture configuration within the ROS2 environment is described in detail, including the management of spatial transformations and coordinate systems via the tf2 library. Main focus is placed on investigating the robot's stability and behavior during the transfer of the navigation model from simulation to a real physical environment (Sim2Real). Experiments are conducted across a variety of obstacle counts and scenarios, analyzing trajectories, final states, and success rates, as well as the effects of moving from simulation to

reality. A detailed analysis of control robustness when traversing paths with varying obstacle densities is performed, experimentally proving the robot's ability to navigate safely with limited sensory information.

### **Compliance of the Proposed Research Methodology with the Stated Aim and Tasks**

The proposed methodology is consistent and well-aligned with the aim and tasks of the dissertation: from 2D simulation and algorithm choice, through 3D simulation and model optimization, to real-world experiments with a physical robot. The use of Flatland and Gazebo, alongside ROS2, represents an adequate and modern choice that enables controlled training and a relatively smooth transfer to a real environment.

As a critical note, it can be mentioned that certain parts of the methodology could have been described in greater detail regarding reproducibility—for instance, more precise simulation hardware parameters, training time consumption, and a more formal statistical processing of the results (more iterations, confidence intervals, significance tests).

### **Contributions**

In the dissertation thesis, the author has formulated a detailed list of contributions. I consider the presented number of contributions to be excessively expanded, as it includes routine engineering tasks, software configuration stages—such as platform calibration and coordinate system setup in ROS2—and sequential steps of simulation modeling. By their nature, these activities represent the execution of specific tasks within the dissertation rather than independent scientific or scientific-applied achievements.

Based on a critical review of the conducted research and the actual results achieved, I propose that the doctoral candidate's contributions be refined into the following three main scientific-applied contributions, which correspond to the objective results in the work:

- **Scientific-Applied Contribution (Modification and optimization of the control algorithm):** A modified architecture of the DRL algorithm PPO (Proximal Policy Optimization) has been developed through structural changes in the deep neural network, the introduction of an adaptive learning rate, and the optimization of the experience buffer. Through a detailed ablation analysis, it is proven mathematically and experimentally that these modifications significantly accelerate convergence and increase training stability.
- **Applied Contribution (Stage-by-stage simulation modeling in ROS2):** A comprehensive methodology for stage-by-stage autonomous navigation testing has been created and software-integrated—spanning from a two-dimensional environment (Flatland) to a realistic three-dimensional environment (Gazebo), based on simulated data from a budget LiDAR sensor (MS200) and odometry.
- **Applied Contribution (Successful Sim2Real validation):** A successful transfer of the trained models from simulation to a real physical environment (Sim2Real) has been implemented and experimentally evaluated on the mobile robot Yahboom RDK X3 with Mecanum wheels. The robustness of the control system in avoiding obstacles within enclosed spaces under conditions of uncertainty, and in the absence of GPS and pre-built maps, has been proven.

## **Publications on the Dissertation**

The candidate has presented 7 scientific publications related to the topic of the dissertation thesis, including journal papers and conference proceedings. It makes a good impression that all publications are indexed in Scopus, which emphasizes and strengthens the scientific and scientific-applied potential of the dissertation. The publications reflect the main phases of the research—from simulation models to real-world experiments—and show continuous scientific development on the topic.

## **Assessment of the compliance of the abstract with the requirements for its preparation, as well as the adequacy of the reflection of the main points and contributions of the dissertation.**

The presented dissertation abstract (autoreferat) by Mag. Eng. Anastasia Vladimirovna Slavova fully complies with the regulatory requirements for its preparation, strictly following the logic and numbering of the developed dissertation thesis in terms of structure, volume, and content. The text is systematized within a volume of 32 pages and includes all mandatory components: a general characteristic, a detailed summary by chapters, scientific-applied and applied contributions, a list of publications, and an abstract in English.

## **Opinions, Recommendations, and Notes**

Overall, the dissertation thesis of Mag. Eng. Anastasia Slavova is developed at a very high scientific and scientific-applied level. The research was conducted systematically, with a precise transition from theoretical propositions to simulations and final physical realization. The achieved results demonstrate that the doctoral candidate possesses deep theoretical knowledge and enviable practical skills in the fields of robotics, artificial intelligence, and software engineering.

Despite the excellent impressions of the dissertation, with the aim of improving the author's future research activities, I can offer the following recommendations and notes:

**Regarding the presentation of graphical material:** In some of the chapters, the graphics depicting the model training results (e.g., the reward function curves) would be better evaluated and understood if the legends and scales were written out in greater detail, making them easier to read in an independent context.

**Future development:** I recommend that the doctoral student seek opportunities to test the developed system in dynamic environments with a higher density of unpredictable obstacles (e.g., moving people) in order to verify the limits of robustness of the modified PPO algorithm.

### **Notes:**

1. Your modified PPO algorithm includes changes to the architecture of the deep neural network. How do these changes reflect on the computational workload (the processing time per step) when executing the algorithm in real-time directly on the robot's onboard computer?
2. Has full synchronization with the data refresh rate of ROS2 been achieved?

## Conclusion

The presented dissertation thesis meets the requirements for acquiring the educational and scientific degree "Doctor" (PhD) in the professional field 5.2 "Electrical Engineering, Electronics and Automation". The work contains a clearly formulated aim, a consistent methodology, and original scientific-applied and applied contributions supported by simulation and real experimental results. In conclusion, I give a **positive evaluation** of the dissertation thesis.

I confidently recommend to the Scientific Jury to award Mag. Eng. Anastasia Vladimirovna Slavova the educational and scientific degree "Doctor" (PhD) in the professional field 5.2. Electrical Engineering, Electronics and Automation.

Date: 05.06.2026

Jury Member: *1021*

(Assoc. Prof. Denis Chikurtev, PhD)

*Взято с оригинала*

