

45. JUPITER KONFERENCIJA
sa međunarodnim učešćem

45th JUPITER CONFERENCE
with foreign participants

ZBORNİK RADOVA
PROCEEDINGS



UVODNI RADOVI

Beograd, jun 2026.

UVODNI RADOVI
INTRODUCTORY PAPERS

Проф. др Зоран Миљковић ЕМПИРИЈСКО УПРАВЉАЊЕ АУТОНОМНИХ РОБОТСКИХ СИСТЕМА - СОПСТВЕНИ РАЗВОЈ	UR.1
Prof. dr Goran Devedžić WAITING FOR THE FOUNDATION (WORLD) MODEL FOR CA _x TECHNOLOGIES	UR.29
Prof. dr Stevo Borojević, Prof. dr Branislav Sredanović, Prof. dr Đorđe Čiča, Milan Gajanin, Jelena Marković, Dejan Vujasin IZVOD IZ ISTRAŽIVANJA U OBLASTI PROJEKTOVANJA TEHNOLOŠKIH PROCESA, SAVREMENIH METODA OBRADE I MAŠINA ALATKI	UR.38



Миљковић, З.¹⁾**Емпиријско управљање аутономних роботских система - сопствени развој²⁾****Резиме**

Нова достигнућа науке, посебно технологије базиране на развоју вештачке интелигенције током 21. века, допринели су унапређењу роботских система који, поред извршавања сложених технолошких задатака, могу да уче и да се прилагођавају реалним условима на основу прикупљених информација из сопственог окружења. Током тридесетогодишњих интензивних научних истраживања развијана је емпиријска управљачка стратегија, посебно за аутономну навигацију, препознавање објеката машинским гледањем и њихову манипулацију, односно интелигентно управљање аутономних роботских система сопственог развоја. Та истраживања, развој и примена, у домену унапређивања интелигентних технолошких система, усмерени су ка пројектовању и реализацији емпиријског управљања које има могућност разумевања процеса у дигитално оријентисаном окружењу, односно способност одабира најбољег могућег понашања аутономног робота. У овом уводном раду, дат је преглед неких од кључних научних резултата утемељених кроз више одбрањених докторских дисертација на Машинском факултету у Београду, као и остварених на пројектима MISSION4.0 и MCSecurity, посебно се ослањајући на оне публиковане у престижним међународним и националним научним часописима, те преко објављених поглавља у научним монографијама светски признатих издавача, саопштених и потом одштампаних научних радова у зборницима истакнутих конференција одржаних у иностранству и региону, затим у оквиру верификованих техничких решења и веома битних скупова података са отвореним приступом. Ти остварени резултати подразумевали су развој и континуирану еволуцију оригиналног, емпиријски оријентисаног управљачког система на бази генерисаног сопственог искуства аутономног робота, стеченог кроз интензивну интеракцију са окружењем, уз примену напредних концепата дубоког учења и *computational intelligence* техника вештачке интелигенције.

Кључне речи: емпиријско управљање аутономних роботских система, машинско гледање, вештачке неуронске мреже, дубоко учење ојачавањем, машинско учење демонстрацијом.

1. УВОД

Аутономни роботи, који „разумеју“ технолошки задатак и радно окружење, данас се интензивно развијају, и то посебно у домену данашњих потреба њихове имплементације унутар *интелигентних технолошких система*, сходно концепту *Индустрија 4.0*. Ово се превасходно односи на развој нове генерације управљачких и сензорских система, посебно система препознавања базираних на машинском гледању и алгоритмима учења, што је и предмет овог уводног рада. Развојем оваквих система аутономних робота омогућава се њихова ефикаснија примена на постојећим задацима, али и проширују области примене на комплексне технолошке задатке који до сада нису, или не у потпуности, били погодни за роботизацију. Посебно треба нагласити да су предвиђања компетентних истраживачких центара у Јапану (*Nagoya University*) [1], Европи (*Swiss Federal Institute of Technology*) и САД (*MIT*) [2], указала на то да ће током првих декада 21. века бити успостављена широка аутономност роботских система, а који имају остварену способност расуђивања, током перманентног

¹⁾ Проф. др Зоран Миљковић, Универзитет у Београду - Машински факултет, (zmiljkovic@mas.bg.ac.rs).

²⁾ У овом уводном раду, даје се преглед кључних резултата истраживања, остварених у оквиру научних пројеката, и то актуелног MCSecurity, ев. број: 17801 и раније реализованог MISSION4.0, ев. број: 6523109, који су финансијски подржани од Фонда за науку Републике Србије, као и пројекта Машинског факултета у Београду, ев. број: 451-03-34/2026-03/200105, финансијски подржаног од Министарства науке, технолошког развоја и иновација Владе Републике Србије од 5.2.2026.

дубоког учења. Посебно, истраживачки резултати водећих лабораторија MIT-а, недвосмислено потврђују да је данашњи степен развоја вештачке интелигенције у роботизи запањујући, при чему се у том контексту посебно истиче *аутономност* робота, као кључна особина која се остварује кроз нове, емпиријски оријентисане, управљачке стратегије.

Аутономни роботски системи, одликују се способношћу да успешно раде у неструктурисаном радном окружењу које увек има одговарајући степен неодређености. Наиме, *интелигентни технолошки системи* данас подразумевају учешће аутономних индустријских робота и машина алатки који извршавају технолошке задатке у интерактивном и кооперативном раду (концепт *Индустрија 4.0*), као и колаборативном, у интеракцији са човеком (концепт *Индустрија 5.0*), како би се, у инжењерски недовољно одређеном окружењу, успешно реализовао постављени технолошки задатак. Сопствена истраживања, чији се преглед кључних резултата даје у овом уводном раду, имала су за циљ развој и примену емпиријски оријентисаних управљачких алгоритама за аутономне роботе на бази коришћења система препознавања и машинског учења, посебно за најчешћи технолошки задатак манипулације препознатим и идентификованим објектима у радном окружењу, као и за успешну навигацију током избегавања препрека приликом кретања мобилног робота, уз изградњу мапе тог окружења, коришћењем стерео визуелног управљања.

Свеобухватна истраживања, у домену развоја *емпиријског управљања аутономних роботских система*, током тридесетогодишњег научноистраживачког рада [3,4,5,6,7,8,9,10,11,12,13,14,15,16,40] у Лабораторији за индустријску роботiku и вештачку интелигенцију (ROBOTICS & AI) Катедре за производно машинство (КаПроМ) Универзитета у Београду - Машинског факултета, обухватила су:

- Идентификацију и класификацију технолошких задатака за аутономне роботске системе са становишта избора коришћеног система препознавања;
- Анализу развијених модела машинског учења, посебно вештачких неуронских мрежа;
- Континуирани, еволутивни развој емпиријског управљачког алгорита за аутономне роботе, применом система препознавања - машинског гледања, као и машинског учења;
- Развој система препознавања за идентификацију снимљених објеката камером и издвајање значајних карактеристика тих објеката, као и експерименталног система за сензорско-моторну координацију робота вертикалне зглобне конфигурације, превасходно при манипулацији препознатих објеката на бази машинског учења помоћу система вештачких неуронских мрежа;
- Развој дубоког учења ојачавањем и машинског учења демонстрацијом за аутономне мобилне роботе који се користе у унутрашњем транспорту радних предмета.

Емпиријско управљање [17,18,19] представља наградњу основно устројеног управљачког система, због неспецифичног понашања робота у условима увек присутне неодређености. Аутономни роботи, у стању су да остваре адаптивно понашање које је базирано на искуству, и то у датом радном окружењу, без учешћа човека-оператора (концепт *Индустрија 4.0*), пре свега захваљујући *емпиријском управљању*. Ова констатација мора бити и верификована, што је представљало изузетан научни изазов. Наиме, научни циљ је везан за остваривање способности „емпиријске машине“ – аутономног робота да учи, односно да буде у стању да развије адекватно понашање на основу сопственог искуства и да одлучује о наредним акцијама. Способност „емпиријске машине“ се огледа у томе да може да генерише такво *аутономно понашање* које допушта откривање посебних односа у сопственом дигитално оријентисаном окружењу, као и њихову уградњу у своје будуће понашање на бази сопственог искуства.

Оно што се одмах намеће у овом „смелом“ приступу, везано је за то да су фундаментални захтеви аутономног роботског система, уствари постављени услови остваривања најбољег могућег понашања у датом радном окружењу. На први поглед, рекло би се одмах, да се ради о филозофском поимању *производног инжењерства у 21. веку*, међутим објективно посматрано, емпиријски управљачки систем, да би остварио оптимално понашање треба да реализује следећа три корака, коришћењем сопствене способности машинског учења, а то су:

1. Генерисање оптималног понашања под одређеним условима.
2. Утврђивање (експерименталним тестирањем) да ли је то понашање остварено.
3. Остваривање понашања које има највећу вероватноћу успешне реализације технолошког задатка, сходно датим условима у радном окружењу.

Поставке емпиријског управљања инспирисане су *природом* и њеним сложеним процесима. Сходно понашању бића у природи која интуитивно предосећају које је понашање најповољније, према датим условима и постављеном задатку, *теорија емпиријског управљања* указује на неопходност адаптивног понашања машина-робота у раду. Самоодрживо понашање *интелигентних агената* (машина алатка, робот, управљање, софтвер, итд.), који свој будући рад заснивају на оцени успешности извршавања прописаног задатка у прошлости, крајњи је циљ емпиријског концепта пројектовања

оваквог управљачког система. Дакле, *системи управљања инспирисани природним процесима* се називају „природни“ управљачки системи (енгл. *Natural control systems*) [17], а у раду ће се представити еволутивни развој емпиријског управљања аутономних роботских система, наслоњен на управо природне постулате, а који је реализован у Лабораторији за индустријску роботiku и вештачку интелигенцију, и то кроз примену дубоког учења ојачавањем и машинског учења демонстрацијом.

2. ЕМПИРИЈСКО УПРАВЉАЊЕ - ТЕОРИЈСКЕ ОСНОВЕ

„Природни“ управљачки систем сваког интелигентног агента подразумева тзв. емпиријски процес који обухвата: истраживање окружења, запажање односно „разумевање“ догађаја у њему, као и примену успешног понашања у домену изабраног технолошког задатка [17,18]. Постоји пет фаза емпиријског процеса које представљају основ за генерисање алгорита управљања базираног на прикупљеним сензорским информацијама из технолошког окружења. Ови основни елементи (фазе) емпиријског процеса, приказани су у табели 1.

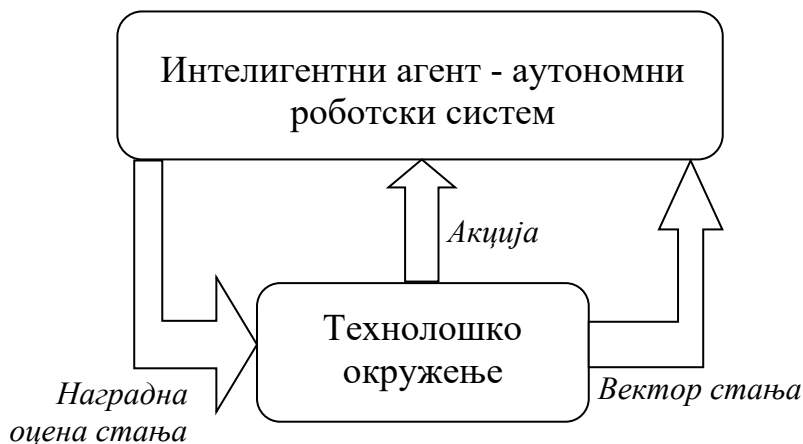
Табела 1. Елементи емпиријског процеса	
1.	Предвиђање излаза на основу нивоа поузданости за сваки улазно-излазни уређени пар
2.	Покушај остваривања изабраног излаза
3.	Поређење оствареног излаза са предвиђеним излазом
4.	У случају подударања оствареног и предвиђеног излаза, меморисаће се скуп управљачких акција и модел предвиђања
5.	Уколико не може да се оствари конзистентност у спровођењу управљачких акција при извршавању задатка, треба одбацити тренутни модел предвиђања, а затим усвојити нови који је боље успостављен

Једна од најбитнијих поставки везаних за *емпиријско управљање* је та да емпиријски систем мора да учи самостално током процеса истраживања окружења. У том процесу је неопходно да систем „испроба“ неку од могућих акција, као што то чине системи у природи. Целокупан процес учења се може посматрати као *скуп позитивних наградних оцена стања* у смислу да управљачки систем мора да научи погодно деловање, односно акцију, уз истовремено одбацавање оне акције која се не може извести или која није најповољнија за тренутно стање система. Елементи емпиријског система неопходни за остваривање овако дефинисаног процеса учења представљени су у табели 2.

Табела 2. Елементи емпиријског система	
1.	Сензорска информација из окружења
2.	Актуатори у функцији остваривања промене стања
3.	Поступак мерења извршене акције
4.	Управљачки систем за праћење и меморисање успешно изведених одлука
5.	Технолошко окружење које садржи информације неопходне за правилно функционисање управљачког система

Основни концепт *теорије емпиријског управљања* може да се представи сликом 1. која приказује односе између интелигентног агента (система управљања), технолошког окружења и наградних оцена

стања које указују на степен успешности изведених акција (степен подударности тренутног и оптималног понашања роботског система, сходно постављеном задатку).



Слика 1. Основни концепт теорије емпиријског управљања.

Произвољни скуп задатака који се поставља пред систем емпиријског управљања може бити извршен у случају да су решена три основна проблема наведена у табели 3. У литератури [3,9,17] је представљена емпиријска управљачка стратегија за индустријски робот који учи, а која се базира на овим „једноставним“ принципима.

Табела 3. Принципи емпиријске управљачке стратегије	
1.	Покушај остваривања захтеваног понашања према датим условима
2.	Одређивање успешности остварене акције
3.	Остваривање понашања које има највећу вероватноћу успешности приликом извршавања постављеног задатка у реалним условима

Као што је речено у уводу, аутономни индустријски роботи који „разумеју“ технолошки задатак и окружење у коме се налазе, саставни су део интелигентних технолошких система [3,9]. У циљу реализације оваког понашања роботских система, као и реализације могућности прилагођавања променама стохастичке природе у дефинисаном радном простору, управљачки алгоритми се морају базирати на прикупљеним сензорским информацијама. Сходно овоме, сви алгоритми представљени у овом раду подлежу *правилима емпиријског управљања* [3,17], и то у циљу реализације аутономности робота при извршавању постављених задатака у реалном времену. У наставку су дате основе за два приступа пројектовању *емпиријских управљачких система* која се данас најчешће користе у пракси, а иначе су била предмет и циљ вишедеценијски спроведених сопствених истраживања [3,4,5,6,7].

2.1 Машинско учење ојачавањем - основни појмови и математичка формулација

Машинско учење ојачавањем (енгл. *Reinforcement Learning*) представља вид машинског учења у коме интелигентни агент учи на основу емпиријских података добијених из технолошког окружења у коме се налази [20]. Интеракцијом са окружењем, односно процесирањем сензорских информација из окружења, мобилни робот врши естимацију (оцену) тренутног стања у коме се налази и генерише оптималан одговор у односу на дефинисани задатак у складу са тренутним нивоом обучености. За разлику од значајног броја техника машинског учења [21,22,23,24,25], у овом случају жељено понашање није експлицитно задато, већ интелигентни агент самостално, кроз стечено искуство, одређује која је акција најповољнија према стању у коме се налази и циљу који је постављен. Другим речима, мобилни робот представља и „ученика“ и „учитеља“ у процесу машинског учења [26].

Аутономност било ког роботског система може се дефинисати као способност самосталног доношења одлука при реализацији одређеног задатка, и то коришћењем техника вештачке интелигенције и информација добијених од спољашњих сензора, попут камере. Како би ово било

могуће, у оквиру *емпиријске теорије* [17] уведен је појам наградне оцено стања (енгл. *Reward*) која указује на степен успешности одабране акције у претходном стању система. Наградна оцена стања је, у општем случају, нумеричка вредност (позитивна или негативна) која је повезана са условима успешног или неуспешног понашања робота. У циљу развоја аутономног понашања, неопходно је да интелигентни агент, кроз интеракцију са окружењем, спозна стање у коме се налази, акцију која га је довела у то стање и да, према њиховој повољности, одреди наградну оцену. За разлику од конвенционалних типова учења, у случају *емпиријске теорије* и *машинског учења ојачавањем*, мобилни робот мора да научи услове (енгл. *Conditionals*) за доделу оцено стања, односно мора да изврши естимацију погодности свих претходних акција које су одредиле његово тренутно стање. Сходно наведеном, ова техника вештачке интелигенције не спада под категорију надгледаног (супервизорског), као ни ненадгледаног (несупервизорског) машинског учења, већ запажен број истраживача сврстава *машинско учење ојачавањем* у независну групу.

Машинско учење ојачавањем засновано је на Марковљевом процесу одлучивања (енгл. *Markov Decision Processes - MDP*). У сваком дискретизованом временском интервалу (или итерацији управљања), интелигентни агент треба да изврши естимацију тренутног стања и одабере одговарајућу акцију. Овај процес подразумева Марковљево правило (енгл. *Markov property*) [20], а које имплицира да наредно стање зависи само од тренутног стања и примењене акције у том стању:

$$S_{t+1} = f(S_t, a_t, \dots, z_t), \quad (1)$$

где је: S_{t+1} - наредно стање, S_t - тренутно стање, a_t - акција мобилног робота, z_t - произвољни поремећај. Стохастичке промене разматраног система описане су помоћу параметра z_t , који у општем случају може бити представљен и помоћу вероватноће прелаза стања у облику $p(\mathbf{x}_{t+1} | \mathbf{x}_t, a_t)$. У општем случају, када спознаје интелигентног агента не обухватају у потпуности све информације о процесу/стању технолошког окружења, наредно стање не подлеже *Марковљевом процесу одлучивања* (енгл. *Non-Markovian process*), те се исто може описати преко:

$$S_{t+1} = f(S_t, S_{t-1}, S_{t-2}, \dots, a_t, a_{t-1}, a_{t-2}, \dots, z_t). \quad (2)$$

Показано је да је значајан број процеса могуће приближно трансформисати тако да делимично подлеже Марковљевом правилу (енгл. *Near optimal solution of MDP*) [27,28]. У оквиру сопствено спроведених истраживања, сви разматрани системи управљања, процеси и развијени алгоритми подлежу овом Марковљевом правилу.

Познато је да је модификацијом појединих метода итерације стратегије, уз познат модел процеса и/или окружења, могуће одредити стратегију доношења акција π_t путем машинског учења [20]. У пракси је међутим, у великом броју случајева, немогуће одредити адекватан математички модел разматраног сложеност процеса или технолошког окружења. Из тог разлога, алгоритам машинског Q -учења ојачавањем, од изузетног је значаја при решавању проблема управљања у реалном времену, зато што за његову имплементацију није потребно познавати математички модел објекта управљања. У овом случају, стратегија доношења акција π_t се одређује директном интеракцијом интелигентног агента са окружењем, уз анализу емпиријски добијених информација од спољашњих сензора.

Алгоритам Q -учења први је увео Ваткинс (енгл. *Watkins*) 1989. године [29,30], као јединствени метод за одређивање стратегије доношења одлука код аутономних агената. Заснива се на одређивању вредносне функције пара стање-акција $Q(\mathbf{x}_t, a_t)$ (енгл. *Action-Value function*), дефинисане преко релације:

$$\begin{aligned} Q(\mathbf{x}_t, a_t) &= E[r(\mathbf{x}_t, a_t) + \gamma V^*(\mathbf{x}_{t+1})] = \\ &= r(\mathbf{x}_t, a_t) + \gamma \sum_{\mathbf{x}_{t+1} \in X} P(\mathbf{x}_{t+1} | \mathbf{x}_t, a_t) V^*(\mathbf{x}_{t+1}), \end{aligned} \quad (3)$$

која представља очекивану вредносну оцену стања након одабира акције a_t у стању $\mathbf{x}_t$, уз примену оптималне стратегије доношења акција. Из ове једначине, коришћењем Белманове једначине оптималности, добијамо:

$$\begin{aligned} V^*(\mathbf{x}_t) &= TV^*(\mathbf{x}_t) = \\ &= \min_a \left[r(\mathbf{x}_t, a) + \gamma \sum_{x_{t+1} \in X} P(x_{t+1} | \mathbf{x}_t, a) V^*(\mathbf{x}_{t+1}) \right] = \\ &= \min_a Q(\mathbf{x}_t, a), \end{aligned} \quad (4)$$

тако да је једначина (3) сада облика:

$$Q(\mathbf{x}_t, a_t) = r(\mathbf{x}_t, a_t) + \gamma \sum_{x_{t+1} \in X} P(x_{t+1} | \mathbf{x}_t, a_t) \min_a Q(\mathbf{x}_{t+1}, a), \quad (5)$$

односно:

$$Q(\mathbf{x}_t, a_t) = (T^q Q)(\mathbf{x}_t, a_t), \quad (6)$$

где је T^q оператор функције стање-акција, дефинисан преко

$$(T^q Q)(\mathbf{x}, a) = r(\mathbf{x}, a) + \gamma \sum_{y \in X} P(y | \mathbf{x}, a) \min_a Q(y, a). \quad (7)$$

Иначе, вредносна функција пара стање-акција је независна од тренутне стратегије одабира акција. Другим речима, вредносна функција $Q(\mathbf{x}_t, a_t)$ одређује оптималну акцију a_t за тренутно стање $\mathbf{x}_t$ преко израза $\pi^*(\mathbf{x}) = \arg \min [Q(\mathbf{x}, a)]$. Псеудокод алгоритма машинског Q -учења ојачавањем дат је у табели 4. [20].

Табела 4. Псеудокод машинског Q -учења ојачавањем	
1.	Иницијализација вредносне функције пара стање-акција Q
2.	Одређивање тренутног стања $\mathbf{x}_t$ и одабир одговајуће акције a_t
3.	Додељивање наградне оцене стања r_t и одређивање наредног стања $\mathbf{x}_{t+1}$
4.	Ажурирати вредносне функције пара стање-акција $Q(\mathbf{x}_t, a_t)$ према: $Q_{t+1}(\mathbf{x}_t, a_t) = Q_t(\mathbf{x}_t, a_t) + a_t \left[r(\mathbf{x}_t, a_t) + \gamma \hat{V}_t(\mathbf{x}_{t+1}) - Q_t(\mathbf{x}_t, a_t) \right],$ где је $\hat{V}_t(\mathbf{x}_{t+1}) = \min_a [Q(\mathbf{x}_{t+1}, a)]$
5.	Поновити процедуру све док не буду задовољени постављени критеријуми заустављања

Треба истаћи да значајан број примера примене овог алгоритма подразумева тзв. табеларну репрезентацију $Q(\mathbf{x}_t, a_t)$, у којој врсте табеле представљају сва могућа стања разматраног система, а колоне скуп свих дефинисаних акција. Вредности табеле у почетном тренутку су једнаке нули или представљају насумично изабране бројеве, сходно проблему који се третира. Ставке 2 и 3 из табеле 4. чине уређену четворку $\langle \mathbf{x}_t, a_t, \mathbf{x}_{t+1}, r(\mathbf{x}_t, a_t) \rangle$, а која се назива уређена четворка на бази искуства.

Један од проблема који настаје приликом развоја и имплементације нових алгоритама управљања базираних на машинском учењу ојачавањем је немогућност прецизног дефинисања односа „истраживање окружења - експлоатација стеченог знања“ (енгл. *Exploration-Exploitation problem*). Превелики број стања посећених током прве фазе (истраживање окружења) има утицај на брзину учења, док рана експлоатација недовољно велике базе знања проузрокује спорију конвергенцију ка оптималном решењу. Избором одговарајуће технике одабира акција (енгл. *Action Selection technique*), овај проблем нема превелики утицај на анализу перформанси развијене методе учења. Опште су прихваћене и подједнако заступљене две технике одабира акција:

1. „Похлепна“ стратегија одабира (енгл. *Greedy action selection strategy*), код које је вероватноћа одабира акције зависи од константе ϵ , и дефинисана је изразом:

$$\pi(\mathbf{x}) = \begin{cases} \text{случајан број (енгл. } random) \text{ из скупа } A(\mathbf{x}), & \text{ако је } \xi < \epsilon \\ \arg \max_{a \in A(\mathbf{x})} Q(\mathbf{x}, a), & \text{ако је } \xi \geq \epsilon \end{cases}, \quad (8)$$

где ξ означава случајан број који подлеже униформној расподели.

2. *Softmax* техника на бази Болцманове расподеле:

$$\pi(a | \mathbf{x}) = P(a_t = a | \mathbf{x}_t = \mathbf{x}) = \frac{e^{Q(\mathbf{x}, a_t)/T}}{\sum_{j=1}^n e^{Q(\mathbf{x}, a_j)/T}}, \quad i = 1, \dots, n, \quad n = \max_a(A), \quad (9)$$

где T представља параметар регулације вероватноће одабира. Овај параметер најчешће није константан, већ се прогресивно смањује током истраживања окружења.

У докторату [29], Ваткинс је показао да под одређеним условима вредносна функција уређеног пара стање-акција $Q(\mathbf{x}_t, a_t)$ увек конвергира ка њеној оптималној вредности, и то са вероватноћом 1, односно стопроцентно. За случај да простор акција и/или простор стања није дискретан, развијено је неколико додатака основном алгоритму Q -учења, од којих се посебно издваја понуђено у [31].

Битно је истаћи и неколико предности које чине алгоритам Q -учења изузетно популарним у светској научној заједници:

- За разлику од осталих метода машинског учења ојачавањем, доказана је конвергенција алгоритма Q -учења за реалне проблеме управљања [29,32];
- Алгоритам представља екстензију познатог концепта оптималног управљања у смислу да одређује оптималну стратегију доношења акција, односно оптимално понашање интелигентног агента кроз интеракцију са спољашњим окружењем;
- Није неопходно познавати математички модел система/процеса/окружења за успешну имплементацију овог алгоритма.

Показано је да, у општем случају, овај алгоритам обезбеђује боље перформансе управљачког система [33], а сопствена истраживања су то експериментално потврдила [3,4,5,6,7,8,9,10,11,12,13,14,15,16,40].

2.2 Машинско учење демонстрацијом - теоријске основе

Током првих декада 21. века, *машинско учење демонстрацијом* представља изузетно популаран приступ у пројектовању *интелигентних управљачких система*. Један од главних разлога ове тврдње лежи у једноставности целог концепта, јер робот или било који објекат управљања може да се обучи уз помоћ демонстрација оптималног понашања које дефинише човек и/или роботски систем [34,35,36]. Појавом овог приступа, омогућена је непосредна интеракција човека и робота (односно човека и машине), чиме је значајно поједностављено управљање са аспекта стохастичких утицаја из спољашњег технолошког окружења. Одличан преглед стања, у овој области истраживања, дат је у [36].

Постоје две основне поделе у оквиру ове парадигме машинског учења, дате на бази искуства:

1) *Машинско учење имитацијом*. У овом случају објекат управљања (роботски систем) не припада простору учења, већ се он односи на човека као демонстратора. Другим речима, робот који је објекат управљања, својим спољашњим сензорима (камером, ултразвучним и/или инфрацрвеним

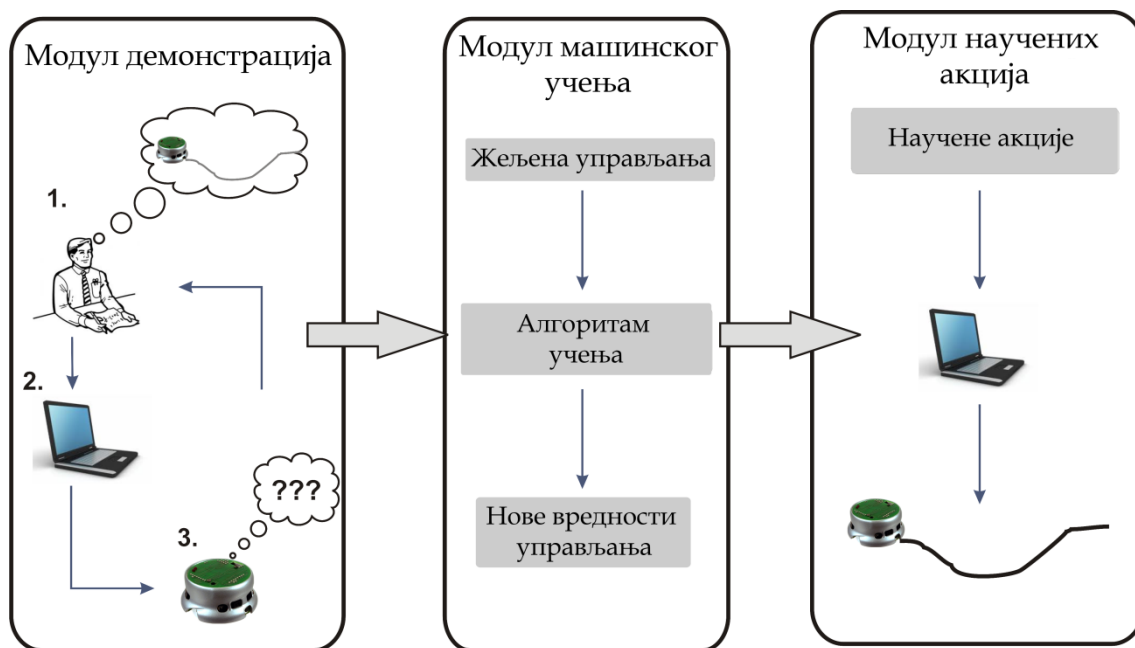
сензорима, и сл.) добија информације од демонстратора („учитеља“), такође опремљеног различитим сензорским системима.

2) *Машинско учење демонстрацијом*, базира се на другој идеји. Човек демонстратор указује на оптималне акције директним управљањем робота, док (истовремено) сензори роботског система бележе информације из окружења. За разлику од учења имитацијом, код овог приступа демонстратор користи робот како би прикупио све неопходне сензорске информације за успешно моделирање жељеног понашања. Издвајају се два приступа прикупљања информација из окружења: (i) метода телеоперације (енгл. *Teleoperating*), код које демонстратор директно управља роботом (задавајући жељене команде актуаторима), истовремено меморишући сензорске информације, и (ii) метода опонашања (енгл. *Shadowing*), где робот бележи оптималне акције осматрајући понашање другог роботског система којим управља демонстратор [2]. Прва метода је највише заступљена у научној и стручној литератури [37,38].

Концепт учења демонстрацијом сличан је представљеном учењу ојачавањем, с обзиром на то да се базирају на емпиријским подацима прикупљеним из окружења. Међутим, битна разлика лежи у начину прикупљања ових информација. У случају машинског учења ојачавањем, робот налази оптималну стратегију одабира акција путем директне интеракције са технолошким окружењем, за разлику од учења демонстрацијом које унапред дефинише стратегију одабира акција демонстрацијама оптималног понашања. У литератури је заступљен и одређени број приступа који комбинују ове две парадигме учења: на пример, истраживање представљено у [37] третира заједничке особине учења ојачавањем и учења демонстрацијом.

У оквиру машинског учења демонстрацијом, разликују се и две схеме управљања сходно структури демонстратора и „ученика“. Хетерогена схема подразумева различите структуре робота демонстратора и робота као објекта управљања, што омогућава тестирање развијених емпиријских алгоритама на различитим роботским системима. У највећем броју случајева, у пракси је заступљена хомогена схема у којој иста структура робота егзистира током целог процеса учења. У оквиру спроведених сопствених истраживања [3,4,5,6,7,8,9,10,11,12,13,14,15,16,40], ова схема је коришћена у свим новоразвијеним и тестираним алгоритмима управљања.

У општем случају постоје три модула машинског учења демонстрацијом (слика 2.): модул демонстрација (енгл. *Demonstration Module*), модул машинског учења (енгл. *Learning Module*) и модул научених акција (енгл. *Drive Modul*) било ког интелигентног агента.



Слика 2. Основни модули машинског учења демонстрацијом.

Модул демонстрација подразумева запис сензорских вредности неопходних за успешно извршавање постављеног задатка. Роботски систем, у овом модулу, меморише сензорске информације и извршава њима одговарајуће унапред дефинисане акције. Овакви подаци се затим „смештају“ у базу уређених парова (пар сензорска информација-команда актуатору) која се користи за обучавање роботског система у следећој фази.

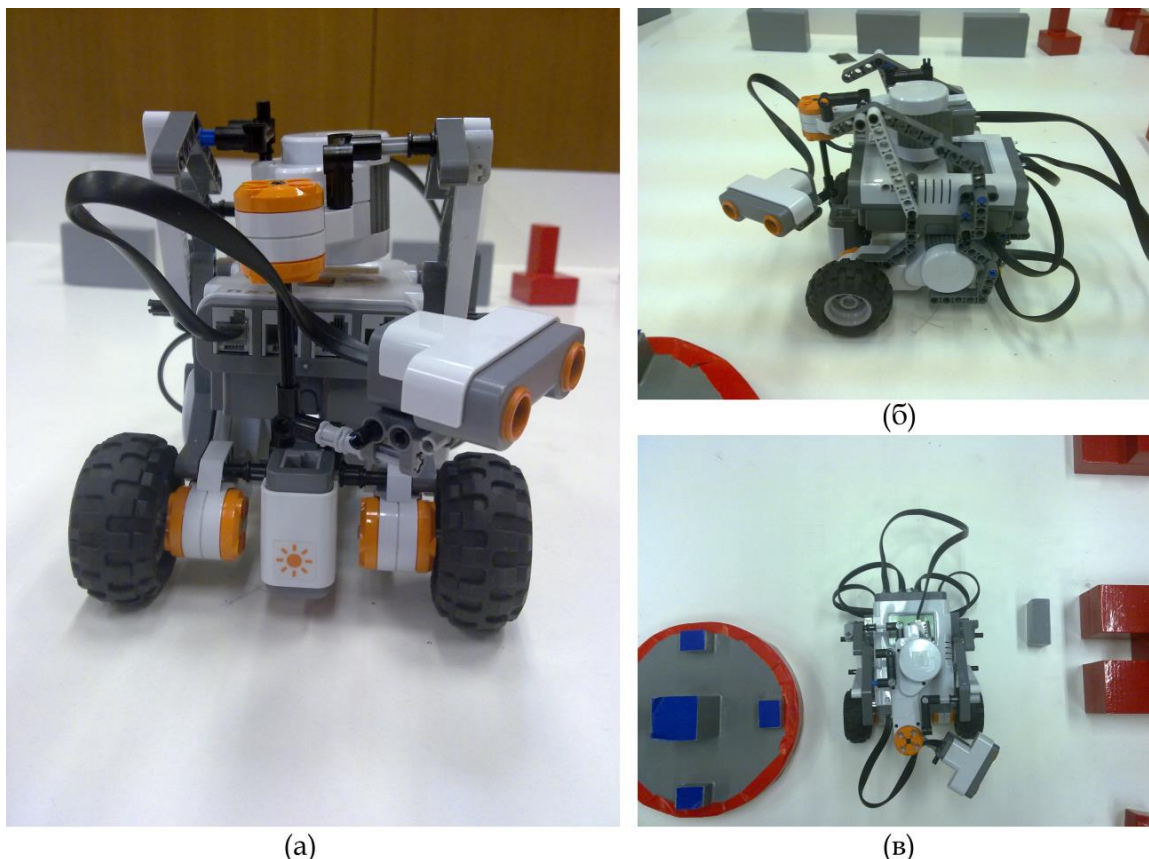
Модул машинског учења подразумева процес нелинеарног пресликавања скупа улазних података (очитавања са спољашњих и/или унутрашњих сензора) у скуп излазних података (скуп акција робота). Машинско учење у овом модулу може обухватати различите парадигме вештачке интелигенције као што су на пример вештачке неуронске мреже, генетички алгоритми, *fuzzy* логика, или напредне статистичке методе као што су скривени Марковљеви модели, методе Гаусове регресије, итд. Избор одговарајућег алгоритма учења зависи пре свега од проблема који се третира, односно од информација у претходно генерисаној бази.

Модул научених акција обухвата извршавање сваке од акција и израчунавање грешке са аспекта успешности нелинеарног пресликавања скупова дефинисаних у претходном модулу. Скуп акција са најмањом израчунавом грешком представљаће скуп оптималних акција при извршавању разматраног задатка. Детаљан опис ових модула машинског учења демонстрацијом дат је у [34,36].

3. АЛГОРИТМИ ЕМПИРИЈСКОГ УПРАВЉАЊА - СОПСТВЕНИ ЕВОЛУТИВНИ РАЗВОЈ

Код аутономних робота, са аспекта развоја управљачких алгоритама, веома важне особине система представљају *сензитивност* и *расуђивање* [3,4,5,6,7]. Способности анализе прикупљених сензорских информација у реалном времену и машинског учења на бази емпиријских података из окружења представљају основне кораке ка остваривању аутономног понашања мобилног робота, а уједно су и основни захтеви при реализацији интелигентног унутрашњег транспорта. У циљу реализације роботизованог интелигентног унутрашњег транспорта сировина, полуфабриката, материјала и готових делова, потребно је развити управљачке системе који „разумеју“ технолошке задатке и реагују на непредвидиве промене у радном окружењу.

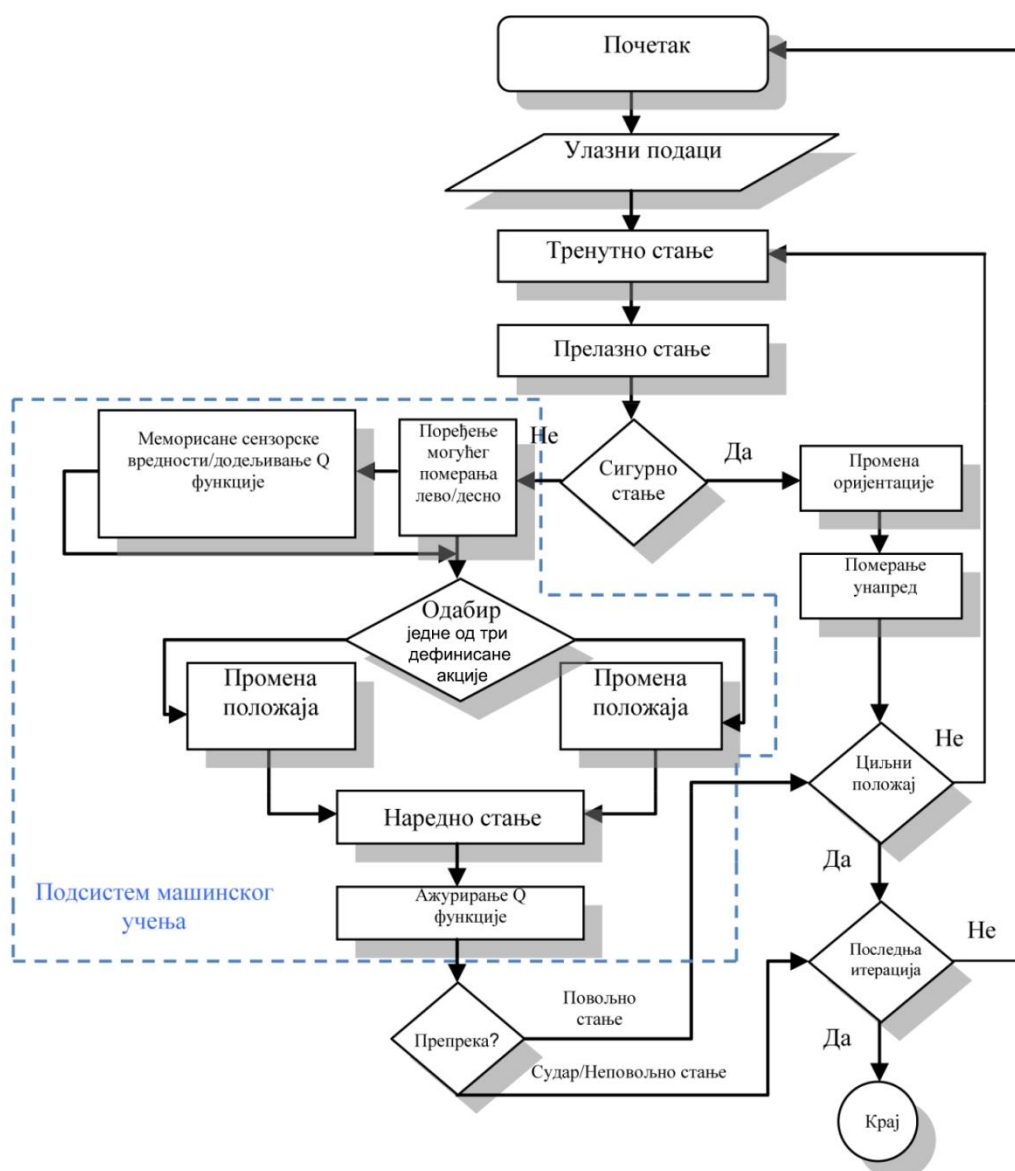
Верификација развијеног система емпиријског управљања на бази машинског учења ојачавањем, извршена је у лабораторијском моделу технолошког окружења (слика 3.) применом *LEGO Mindstorms NXT* мобилног робота, и то у домену задатака навигације, локализације, избегавања препрека и изградње мапе радног окружења. Конфигурација мобилног робота се састоји од два погонска точка, једног помоћног (задњег) точка, ултразвучног сензора који је постављен у хоризонталној равни, као и управљачке јединице која директно управља елементима система мобилног робота (слика 3.). Основне, техничке карактеристике мобилног робота приказане су у табели 5.



Слика 3. Конфигурација мобилног робота *LEGO Mindstorms NXT*:
(a) поглед спреда, (б) поглед са стране, (в) поглед одозго.

Табела 5. Техничке карактеристике - <i>LEGO Mindstorms NXT</i> мобилног робота	
Назив	Опис
Управљачки систем	32-битна управљачка јединица
Погонски систем	Три DC мотора са инкременталним енкодерима
Основни сензори	Ултразвучни сензор. Домет ~ 2,5 m
Улаз/излаз (I/O)	USB 2.0 и Bluetooth interface
Напајање	Батерије AA
Управљачки софтвер	<i>MATLAB</i> [®] преко <i>USB</i> порта

На основу искуства стеченог развојем модела кретања робота, основна верзија управљачког алгоритма је модификовна у циљу бржег учења у реалном времену (слика 4.).

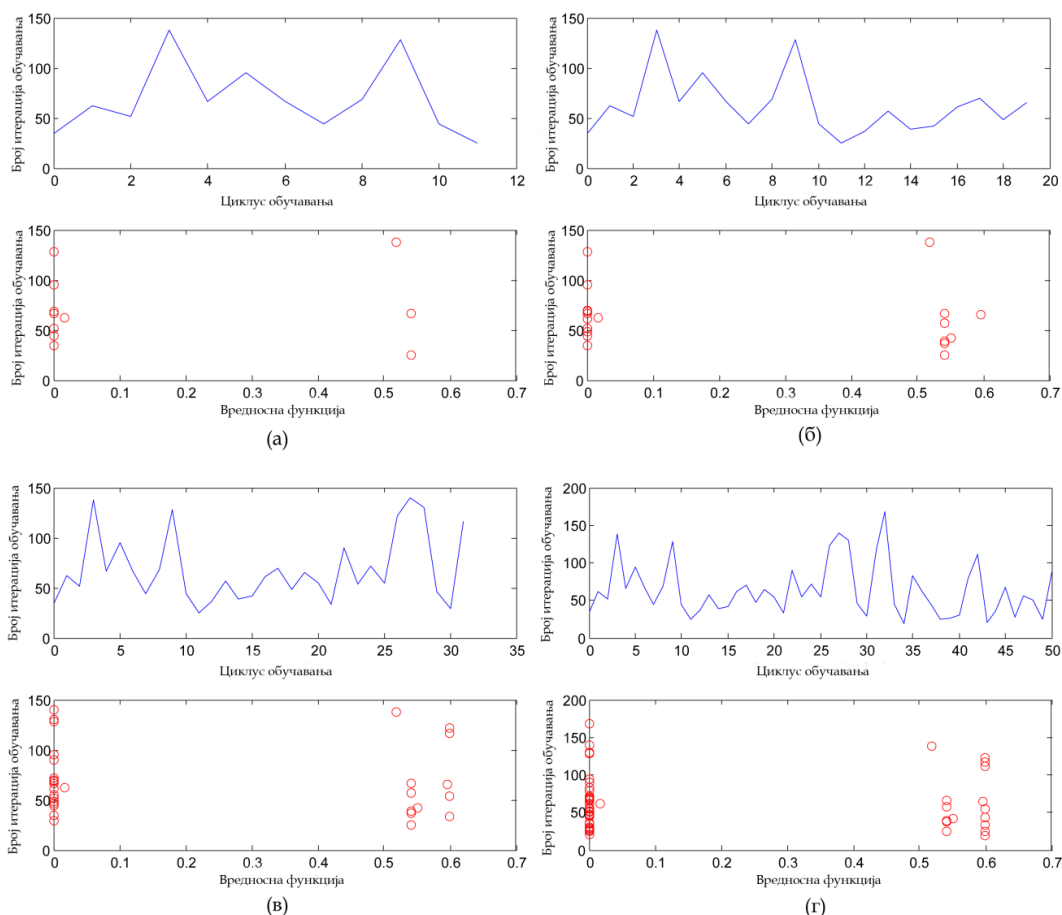


Слика 4. Нови алгоритам емпиријског управљања мобилног робота.

Подсистем за машинско учење обухвата блок одабира акција који врши селекцију на основу информација о могућим померањима и претходно меморисаним вредностима Q функције. Другим речима, акција мобилног робота се бира на основу два типа критеријума: тренутних информација о непознатом окружењу (вредност оцене пара стање-акција) и могућности остваривања селектованог понашања. Овакво решење директно утиче на смањење броја колизија робота са различитим типовима препрека приликом фазе истраживања радног окружења.

Ултразвучни сензор врши читавања до 180° на сваких 45° у математички позитивном смеру, укључујући и почетни положај (угао од нула степени гледано одозго). Добијена мерења учествују у формирању матрице на основу које ће се касније вршити верификација избора одабране одлуке према меморисаним вредностима у одговарајућој табели Q функције. Ова табела је на почетку експеримента дефинисана као празан скуп, у коју ће се, као што је већ напоменуто у претходном поглављу, уписивати вредности добијене према изразу (5). Наградна оцена стања, усвојена је према изразу (10). Са J је означен вектор од пет добијених вредности растојања мерених помоћу ултразвучног сензора у претходном стању, док је са U означен вектор растојања у тренутном стању. Техника одабира акција, у овом случају је „похлепни“ алгоритам који према параметру ε бира акцију са највећом вредношћу Q функције у том тренутку. За мобилни робот, изабране и усвојене конфигурације (слика 3.), одабране су произвољне вредности стартне и циљне позиције. Усвојени скуп акција је дефинисан на следећи начин: робот се може кретати право, скретати улево под углом од 45 степени, као и скретати у десну страну под истим углом. Експериментални резултати обучавања *LEGO Mindstorms NXT* мобилног робота у лабораторијском моделу технолошког окружења дати су у наставку (слика 5.).

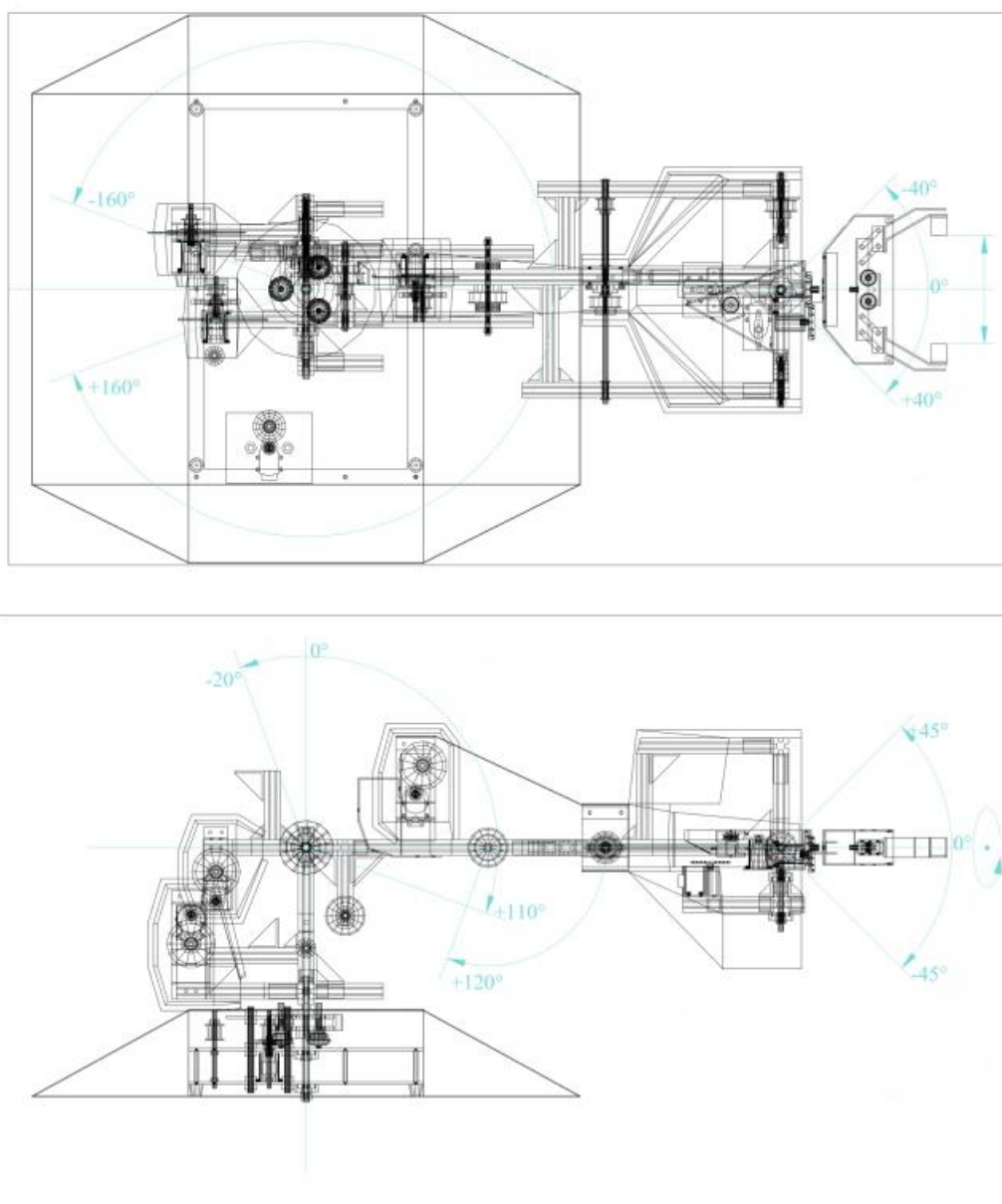
$$r = \begin{cases} 0, \min J = \min U \\ 1, \min J < \min U \\ -1, \min U < \min J \\ 2, x_{\text{тренутно}} = x_{\text{циљ}} \end{cases} \quad (10)$$



Слика 5. Експериментални резултати обучавања *LEGO Mindstorms NXT* мобилног робота у лабораторијском моделу технолошког окружења.

Управљање роботског система, на бази информација добијених од камере у реалном времену, подразумева решавање различитих проблема као што су: проблем видног поља камере, постојања шума у аквизицији сигнала, нетачних параметара калибрационе матрице и прилагођавања условима нагледне промене осветљења у окружењу. Како би се ови проблеми успешно превазишли, развијен је *хибридни емпијски управљачки систем на бази машинског учења ојачавањем* и управљања на основу грешке у параметрима слике. Нови интелигентни систем се састоји од две независне фазе управљања. У првој фази, оријентација робота је коригована како би изабрани карактеристични објект припадао жељеној области у равни слике, док друга фаза обухвата приступање тродимензионалном објекту уз кориговање оријентације у зависности од положаја карактеристичних објеката на слици. Дефинисањем жељених области карактеристичних објеката у равни слике омогућава се кретање робота оптималном путањом.

Како би испитали понашање система у условима поремећаја као што су изненадна промена осветљења у окружењу, изражено постојање шума у сигналу и погрешна естимација растојања до 3D објекта, нови управљачки систем на основу грешке у параметрима слике имплементиран је на антропоморфном роботу вертикалне зглобне конфигурације *NeuroArm Manipulator System* [5]. Техничке карактеристике овог роботског система дате су у табели 6. Максимални углови ротације (дефинишу достиживи радни простор робота) овог манипулатора, приказани су на слици 6. Компатибилна камера *CMUCam3* инсталирана је и постављена изнад енд-ефектора робота у „eye-in-hand“ схеми управљања.



Слика 6. Максимални углови ротације у зглобовима робота *NeuroArm Manipulator System* (слика преузета са сајта www.neurorobotics.co.uk).

Табела 6. Техничке карактеристике – основна конфигурација <i>NeuroArm Manipulator System</i> роботског система	
Назив	Опис
Процесор	Atmel® Atmega 128 (<i>Low-level</i>) и Intel® XScale 400 MHz Pxa255 (<i>High-level</i>)
Погонски систем	7 DC серво мотора са уграђеним инкременталним енкодерима (6 зглобова + хватач)
Комуникацијски протокол	RS232 & I2C порт
Основни сензори	Сензори за мерење позиције, брзине и сила у зглобовима
Носивост	До 3.5 kg
Напајање	Адаптер за директно напајање из мреже
Управљачки софтвер	MATLAB® преко RS232 порта

Првенствено је било неопходно успоставити везу између координата карактеристичних објеката у равни слике и углова ротације у зглобовима роботског система. У општем случају, решење директног кинематичког проблема за последњи сегмент (енд-ефектор) дато је у следећој форми [39,40]:

$${}^0\mathbf{T}_n = {}^0\mathbf{A}_1 {}^1\mathbf{A}_2 \dots {}^{n-1}\mathbf{A}_n = \kappa(\underline{q}), \quad (11)$$

где је: ${}^{n-1}\mathbf{A}_n$ хомогена матрица трансформација n -тог сегмента у односу на локални координатни систем $(n-1)$ -тог сегмента, ${}^0\mathbf{T}_n$ хомогена матрица трансформације последњег сегмента (енд-ефектора) у односу на референтни координатни систем који је круто везан за базу робота, и $\underline{q}$ унутрашња или генералисана координата која одређује релативна померања сегмената. Познато је да је решење инверзног кинематичког проблема облика:

$$\underline{q} = \kappa^{-1}(\mathbf{T}). \quad (12)$$

Битно је нагласити да решење инверзног кинематичког проблема не мора бити и јединствено, што узрокује појаву сингуларитета у појединим положајима унутар радног простора робота. Брзина енд-ефектора може да се одреди на основу промене генералисаних координата у времену и познате Јакобијан матрице

$${}^0\dot{\underline{x}}_n = {}^0\mathbf{J}_\theta \dot{\underline{q}}. \quad (13)$$

У случају познавања брзине у спољашњим координатама, лако се може одредити брзина у зглобовима:

$$\dot{\underline{q}} = {}^0\mathbf{J}_\theta^{-1} {}^0\dot{\underline{x}}_n. \quad (14)$$

У случају кинематичког сингуларитета, није могуће одредити инверзију Јакобијан матрице што утиче на управљивост система. Сходно изразу (13), једначина која успоставља везу између два координатна система сегмената је:

$${}^B\dot{\underline{x}} = {}^B\mathbf{J}_A {}^A\dot{\underline{x}}. \quad (15)$$

Координате карактеристичних објеката у равни слике зависе од унутрашњих параметара камере и положаја 3D објекта у односу на камеру, а у општем случају су одређене према:

$$\underline{f} = f\left({}^c \underline{x}_t\right), \quad (16)$$

односно у линеаризованом облику

$$\delta \underline{f} = {}^f \mathbf{J}_c \left({}^c \underline{x}_t\right) \cdot \delta {}^c \underline{x}_t. \quad (17)$$

Величина ${}^f \mathbf{J}_c \left({}^c \underline{x}_t\right)$ у једначини (17) означава Јакобијан матрицу слике односно матрицу интеракције. Увођењем овог параметра дефинисан је однос између промене положаја камере и промене положаја објекта у равни слике. Под претпоставком да постоји инверзија Јакобијан матрице, израз (17) се може трансформисати у:

$${}^c \dot{\underline{x}}_t = {}^f \mathbf{J}_c^{-1} \left({}^c \underline{x}_t\right) \dot{\underline{f}}. \quad (18)$$

Користећи једначину (15) могуће је успоставити релацију између промене положаја 3D објекта у односу на камеру и промене положаја истог објекта у односу на енд-ефектор:

$${}^{t6} \dot{\underline{x}}_t = {}^{t6} \mathbf{J}_c^f {}^c \dot{\underline{x}}_t. \quad (19)$$

С обзиром на то да се брзина из претходног израза може трансформисати у брзину зглобова робота преко:

$$\dot{\underline{\theta}} = {}^{t6} \mathbf{J}_\theta^{-1}(\underline{\theta}) {}^{t6} \dot{\underline{x}}_t, \quad (20)$$

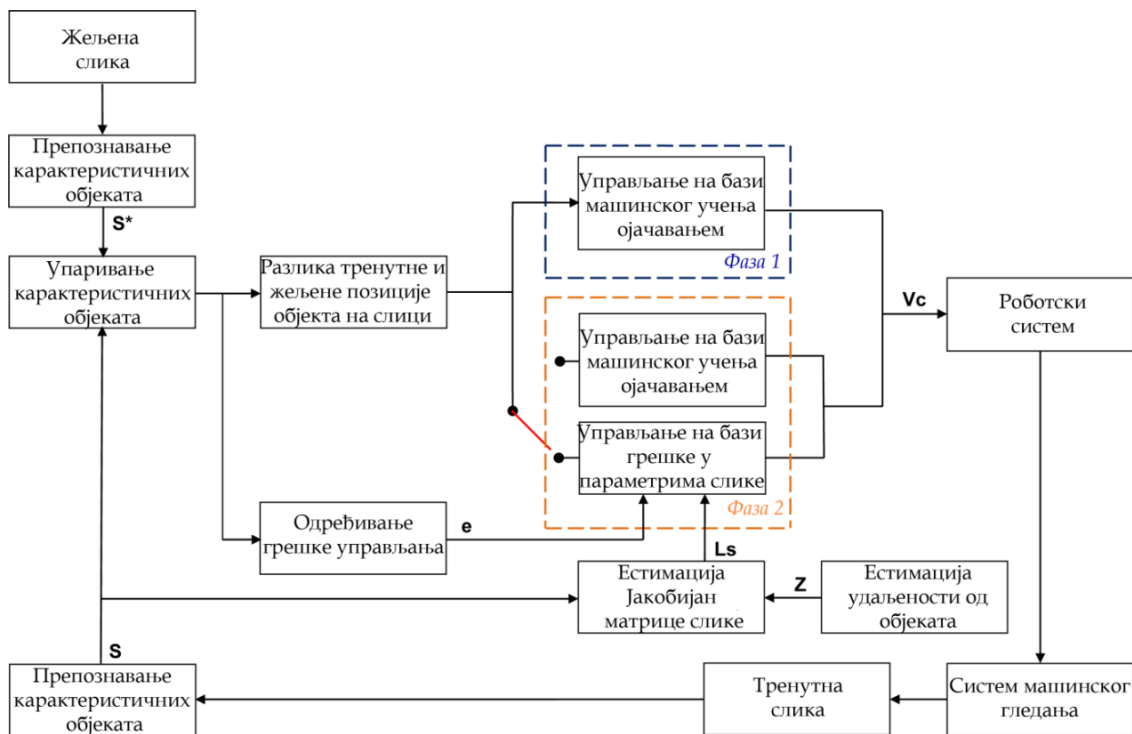
коначна једначина која повезује брзину зглобова робота и промену координата карактеристичних објекта у равни слике је:

$$\dot{\underline{\theta}} = {}^{t6} \mathbf{J}_\theta^{-1}(\underline{\theta}) \cdot {}^{t6} \mathbf{J}_c^f \cdot {}^f \mathbf{J}_c^{-1} \left({}^c \underline{x}_t\right) \cdot \dot{\underline{f}}. \quad (21)$$

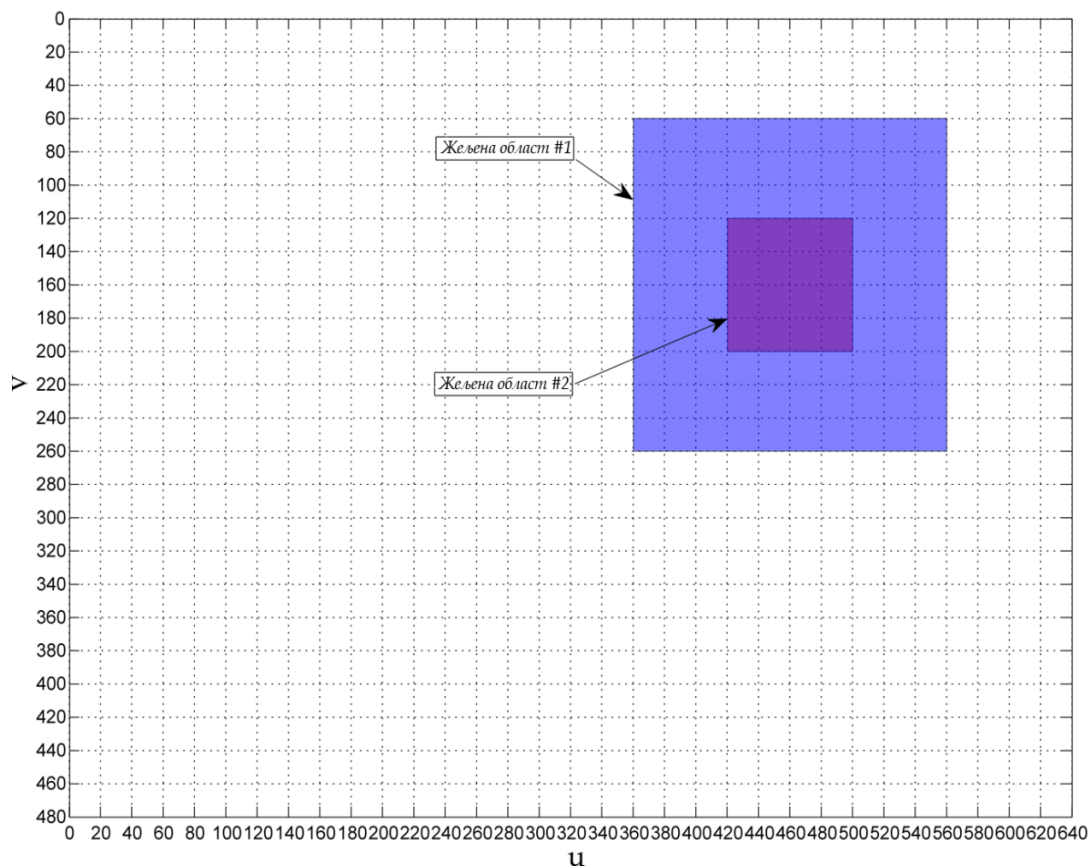
Након успостављања претходне релације, могуће је приступити реализацији експеримента у реалном времену, и то у условима нагле промене амбијенталног осветљења. Како би се делимично или у потпуности отклонили погрешно упарени објекти, имплементиран је *RANSAC* (енгл. *Random Sample Consensus*) алгоритам [5,41]. У циљу оптимизације времена неопходног за процесирање слике и прорачун управљачких команди коришћен је шаблон са три издвојена региона. Резолуција камере износи 177x144 пиксела.

Нови хибридни управљачки систем обухвата две независне фазе управљања [13]. Прва фаза се заснива на имплементацији алгоритма машинског учења ојачавањем на бази информација добијених од камере и вештачких неуронских мрежа, а односи се на задатак промене оријентације камере (која је круто везана за структуру робота). Друга фаза решава проблем позиционирања енд-ефектора у односу на 3D објекат дефинисан у равни слике, уз кориговање положаја у случају нагле промене осветљења, односно великог броја погрешно упарених карактеристичних објекта. Структурни блок дијаграм система управљања дат је на слици 7.

Тренутна фаза управљања одређује се уз помоћ жељених области изабраног карактеристичног објекта у равни циљне слике (енгл. *Desired feature area*). На почетку управљања, карактеристични објекат се налази на одређеној удаљености од његовог жељеног положаја. Основна идеја је да, кориговањем оријентације робота, положај карактеристичног објекта буде у жељеној области на крају прве фазе управљања. Дефинисањем друге жељене области за карактеристични објекат омогућен је почетак друге фазе управљања у којој тренутни положај изабраног објекта одређује који је од два подсистема управљања активан (видети слику 8). Жељене области објекта за обе фазе подељене су на јединице величине 20 пиксела (слика 8.). Како би се омогућио што ранији почетак друге фазе управљања (фазе позиционирања), жељена област у првој фази заузима шири простор у односу на област дефинисану у фази позиционирања робота.

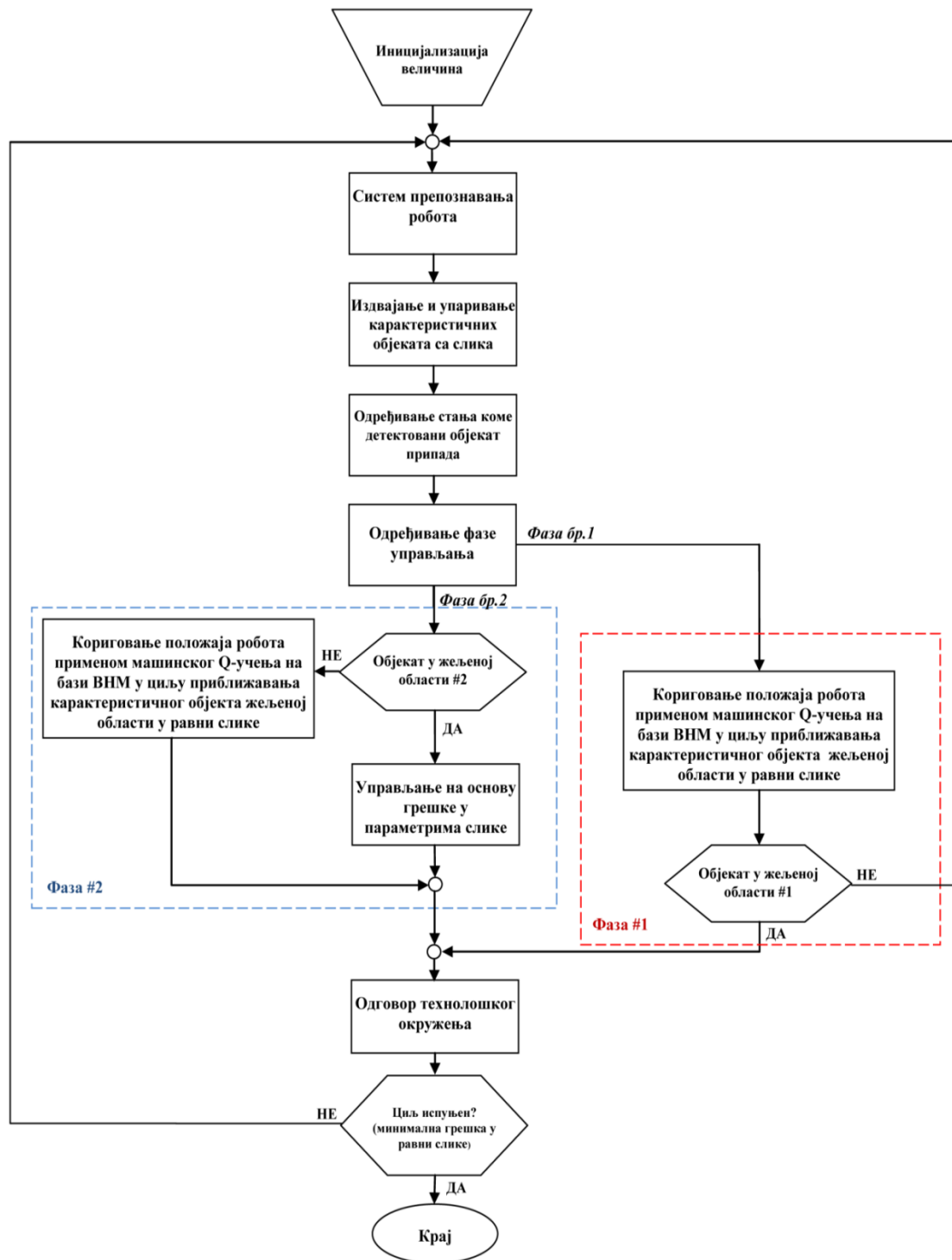


Слика 7. Интелигентни систем управљања на бази информација добијених од камере и машинског учења ојачавањем.



Слика 8. Жељене области изабраног карактеристичног објекта у равни слике.

Концепцијско решење новог хибридног емпиријског управљачког система представљено је на слици 9. Дати оригинални алгоритам приказује ток информација унутар главног управљачког кода реализованог у *MATLAB*[®] програмском окружењу.



Слика 9. Алгоритам новог хибридног емпиријског управљачког система робота вертикалне зглобне конфигурације.

Прво је потребно иницијализовати све величине које фигуришу у претходно датим подсистемима управљања. Неопходно је подесити почетне вредности тежинских односа вештачке неуронске мреже на мале случајно изабране величине, одабрати одговарајуће вредности за параметре γ и α из једначине грешке учења (22), одредити параметар активације фазе управљања, а затим дефинисати жељене области у равни слике за обе фазе управљања.

$$Q(\mathbf{x}_t, a_t) = r_t + \gamma \max_{a_{t+1}} Q(\mathbf{x}_{t+1}, a_{t+1}). \quad (22)$$

Такође, потребна је и аквизиција жељене слике (односно слике која треба да се подудари са сликом у крајњем стању роботског система), како би одређивање карактеристичних објеката и грешке између параметара две слике (тренутне и циљне) било могуће. Подразумева се да је камера претходно калибрисана, односно да су познате величине фокалне (жигне) даљине f , резолуције слике $U \times V$, димензије пиксела β , као и положаја тачке у пресеку главне оптичке осе и равни слике (u_0, v_0) .

Као што је напоменуто, прва фаза управљања базираног на информацијама добијеним од камере се односи на кориговање грешке оријентације камере у простору, односно на одређивање величина w_x, w_y и w_z . Корекција положаја роботског система, остварена је применом машинског учења ојачавањем (Q -учења), коришћењем вештачких неуронских мрежа са дефинисаном базом карактеристичних вредности за обучавање [13]. Крајњи циљ ове фазе је такав положај робота у коме карактеристични објекат у тренутној слици припада првој жељеној области. Уколико је тај услов испуњен, параметар активације узима нову вредност и аутоматски се прелази на другу фазу управљања.

У другој фази управљања разликују се два подсистема управљања у зависности од положаја карактеристичног објекта. У случају да се карактеристични објекат налази у другој жељеној области врши се корекција позиције камере у простору, односно одређују се величине v_x, v_y и v_z применом класичног управљања на основу грешке у параметрима слике. Уколико се у неком тренутку карактеристични објекат нађе изван ове области, било услед поремећаја у параметрима слике или произвољних спољашњих поремећаја, корекција оријентације камере се остварује уз примену машинског Q -учења на бази вештачких неуронских мрежа. Ова фаза (као и комплетан циклус) се завршава када је разлика између тренутне и циљне слике минимална, односно када је разлика у положајима карактеристичних објеката у равни обе слике испод претходно прописане границе. У табели 7. дат је псеудокод машинског Q -учења на бази вештачких неуронских мрежа у случају управљања робота на основу информација добијених од камере.

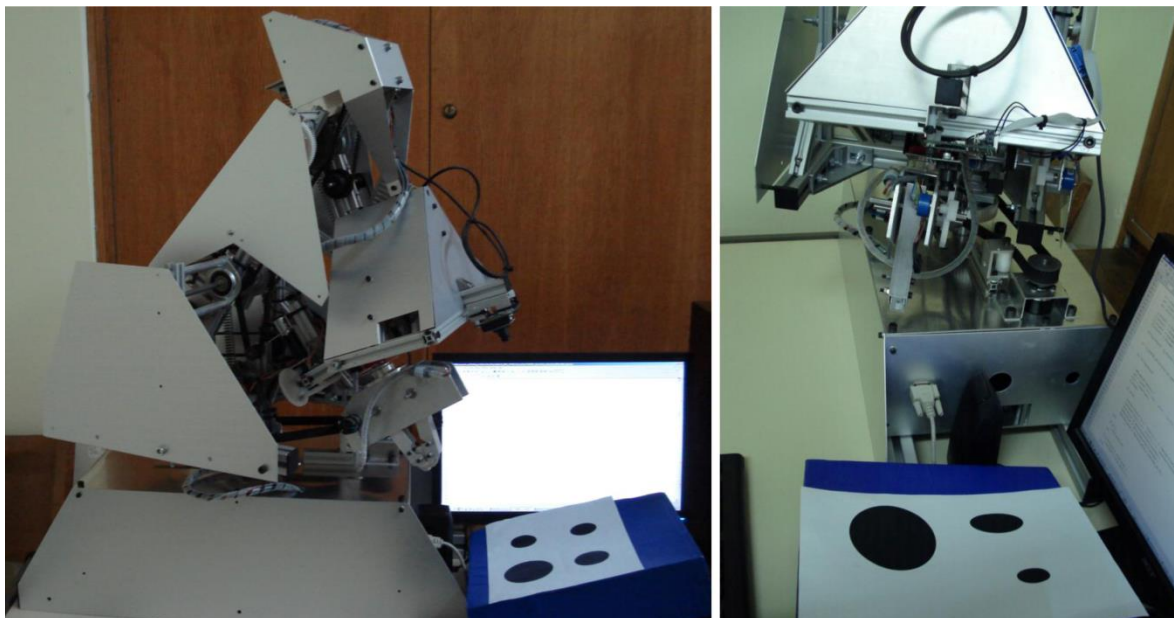
Табела 7. Псеудокод алгорита машинског Q-учења ојачавањем на бази вештачких неуронских мрежа у случају управљања на основу информација добијених од камере	
1.	Иницијализација тежинских односа вештачке неуронске мреже. Дефинисање наградне оцене стања и параметара $\alpha, \gamma, \varepsilon, r_p$
2.	Одредити стање $\mathbf{x}_t$ и одговарајућу акцију a_t према претпостављеној стратегији доношења акција $\mathcal{L}_t$
3.	Одредити наредно стање $\mathbf{x}_{t+1}$ сходно положају изабраног карактеристичног објекта у равни слике и доделити наградну оцену стања r_t
4.	Извршити простирање сигнала унапред и ажурирати вредносну функцију пара стање-акција према једначини (5)
5.	Допунити базу података новим обучавајућим паром сходно једначини (23)
6.	Обучити вештачку неуронску мрежу користећи све податке из базе применом генерализаног делта правила
7.	Поновити процедуру од 2-6 све док не буду задовољени постављени критеријуми заустављања

$$r_p = \sqrt{w_{e_1} (s_n - s_o)^2 + w_{e_2} (a_n - a_o)^2 + w_{e_3} (r_n - r_o)^2}, \quad (23)$$

где је: s_n и s_o будуће и тренутно стање респективно, a_n и a_o будућа и тренутна акција, r_n и r_o будућа и тренутна наградна оцена стања, а w_{e_1} , w_{e_2} , и w_{e_3} одговарајаћа појачања.

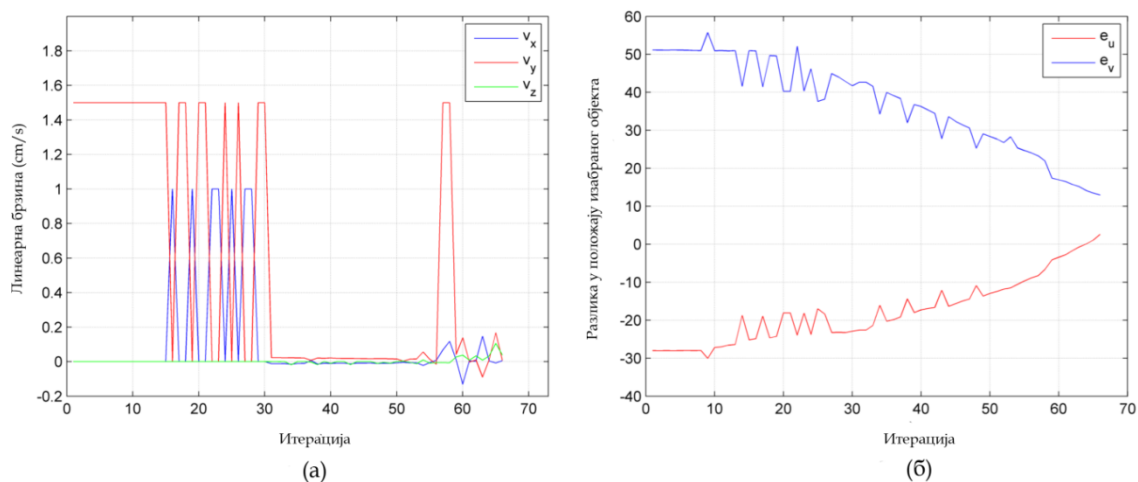
Дакле, роботски систем за експерименталну верификацију чинили су: робот вертикалне зглобне структуре *NeuroArm Manipulator System* [5], компатибилна камера *CMUCam3* и десктоп рачунар (процесор *AMD Athlon II X4 630 2.8 GHz* и *3GB RAM*) са *WindowsXP* оперативним системом (слика 10.). У циљу бржег прорачуна управљачких команди, коришћен је шаблон са издвојеним регионима црне

боје (два шаблона са 3 и 4 региона). *SURF* алгоритам [5], коришћен је за препознавање и упаривање карактеристичних објеката. *RANSAC* алгоритмом [5] елиминисани су погрешно упарени објекти у равни тренутне и циљне слике.



Слика 10. Роботски систем *NeuroArm Manipulator System* са интегрисаном компатибилном камером *CMUCam3*.

Резултати првог експеримента приказани су на слици 11. и слици 12. Промена транслаторне (линеарне) брзине указује на то да прва фаза траје 14 итерација управљања [5], као и да је интелигентни подсистем активан на почетку друге фазе управљања. Очигледно је да постоји корекција позиције у x и y правцу, са циљем одржавања оптималне путање роботског система. Слика 11.(б) показује промену грешке у координатама изабраног објекта (u, v) . Уочљиво је да итерације у којима је активан интелигентни подсистем (слика 11.(а)) одговарају тренуцима израженије промене у разлици координата изабраног објекта у равни тренутне и циљне слике (слика 12.(б)). Поредећи слику 12.(б) и слику 12.(в) закључује се да постоји незнатна разлика у жељеном и крајњем положају робота.



Слика 11. Резултати првог експеримента: (а) промена линеарне (транслаторне) брзине камере, (б) разлика у положају изабраног објекта у равни тренутне и циљне слике [5].

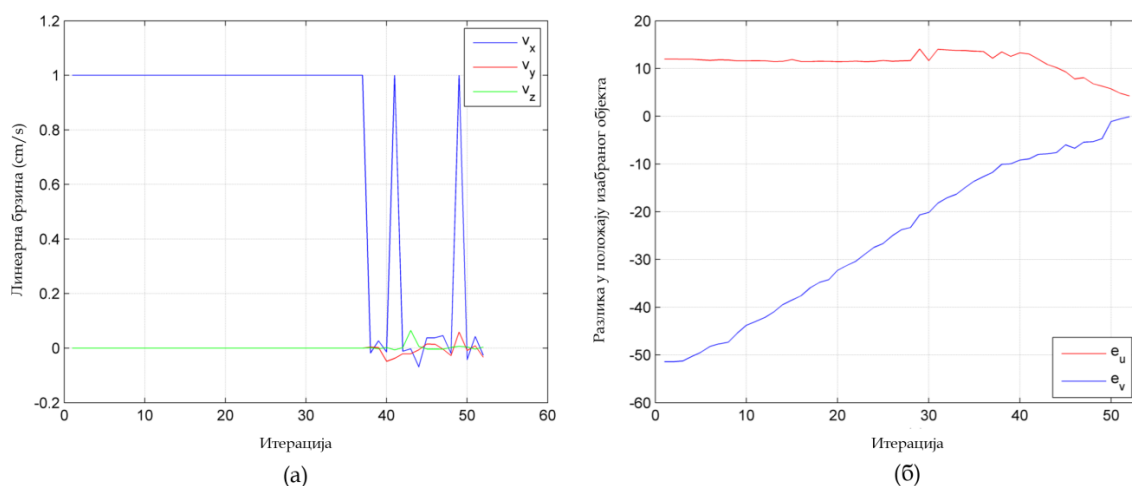
Резултати другог експеримента приказани су на слици 13., слици 14 и слици 15. С обзиром да су у почетном положају камере познате координате два објекта у равни слике, прва фаза управљања траје дуже како би се изабрани карактеристични објекат налазио у оквиру жељене области у равни слике.



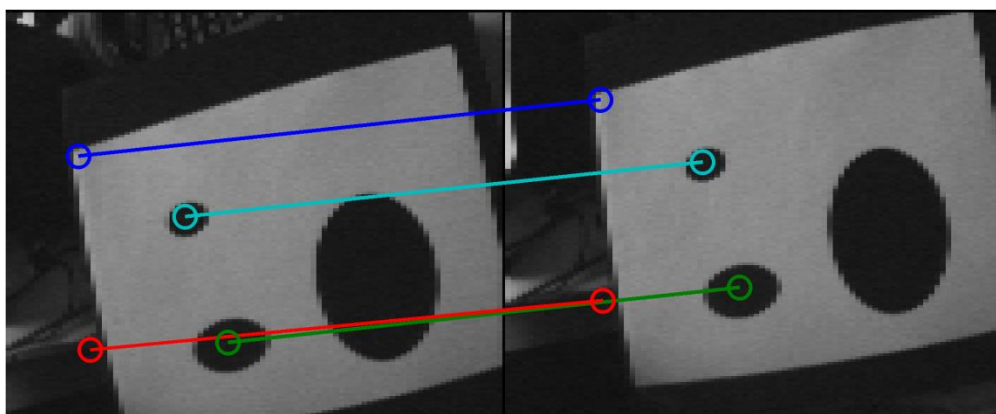
(a) (б) (в)

Слика 12.. Резултати првог експеримента: (а) почетна слика, (б) циљна (жељена) слика, (в) слика на крају експеримента [5].

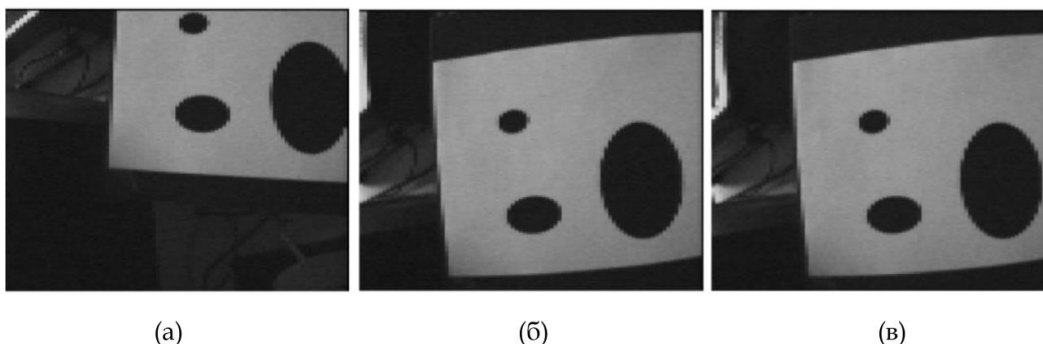
Након почетка друге фазе управљања (38 итерација управљања [5]), интелигентни подсистем је активан још два пута. Слика 13.(б) указује на то да је положај робота на крају експеримента задовољавајући, пошто је разлика у координатама изабраног карактеристичног објекта у жељеној и крајњој слици, у последњој итерацији управљања, мања од 5 пиксела. Слика 14. приказује упарене карактеристичне објекте тренутне и циљне слике у току експеримента. Оптимално кретање камере је евидентно приказано на слици 15., јер се поређењем слике 15.(б) и слике 15.(в) закључује да положај робота на крају експеримента приближно одговара његовом жељеном положају.



Слика 13. Резултати другог експеримента: (а) промена линеарне (транслаторне) брзине камере, (б) разлика у положају изабраног објекта у равни тренутне и циљне слике [5].



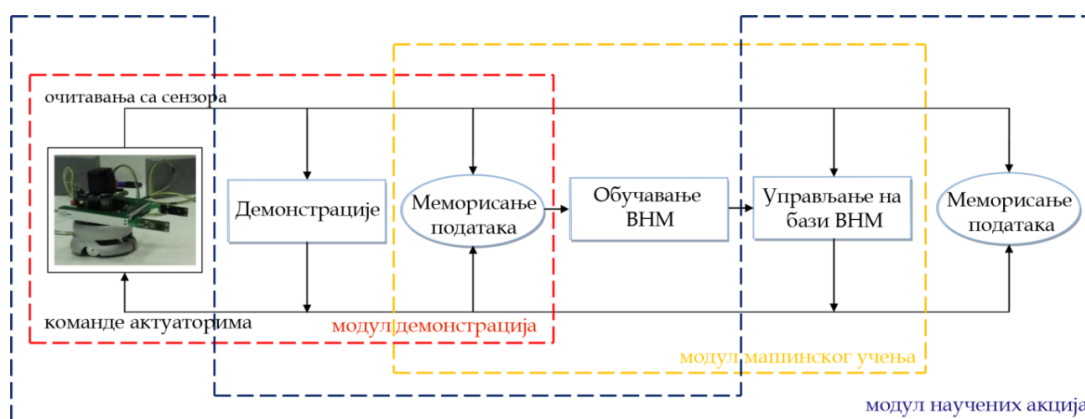
Слика 14. Резултати другог експеримента: препознати и упарени карактеристични објекти у тренутној и циљној слици након примене SURF алгоритма [5].



Слика 15. Резултати другог експеримента: (а) почетна слика, (б) циљна (жељена) слика, (в) слика на крају експеримента [5].

4. ЕМПИРИЈСКО УПРАВЉАЊЕ НА БАЗИ МАШИНСКОГ УЧЕЊА ДЕМОНСТРАЦИЈОМ

Машинско учење демонстрацијом, детаљно је објашњено у потпоглављу 2.2, тако да је овде представљена схема која је имплементирана у *оригинални хибридни управљачки алгоритам* (слика 16.). Као што се види са слике 16., дефинисана су три независна модула машинског учења демонстрацијом: (i) *модул демонстрација*, (ii) *модул машинског учења*, и (iii) *модул научених акција*.

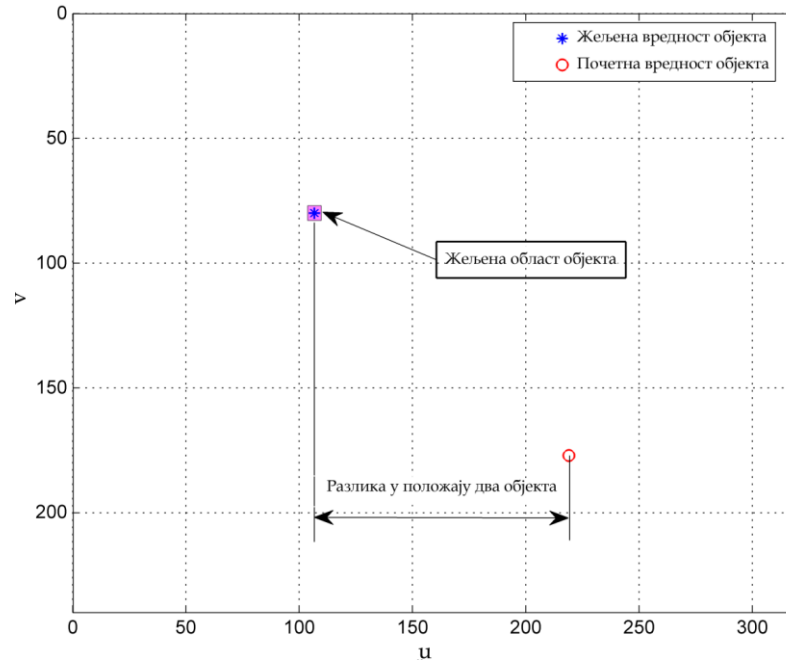


Слика 16. Три основна модула машинског учења демонстрацијом новоразвијеног хибридног емпиријског алгоритма управљања.

Модул демонстрација обухвата процес меморисања акција мобилног робота (односно команди актуаторима) и одговарајућих сензорских информација добијених из технолошког окружења. Усвојене акције односе се на корекцију оријентације (ротацију робота у достигнутој позицији), док информације из окружења представљају разлику између положаја изабраног карактеристичног објекта у равни жељене и тренутне слике. Дефинисањем жељене области карактеристичног објекта (на начин описан у поглављу 3, слика 8.), омогућава се одређивање периода активације сваког од подсистема управљања. Као и у случају хибридног система на бази машинског учења ојачавањем, основна идеја имплементације жељене области објекта је да се кориговањем оријентације мобилног робота омогући да положај карактеристичног објекта буде у жељеној области. На слици 17. приказан је пример разлике у положају између истог карактеристичног објекта у равни тренутне и жељене слике.

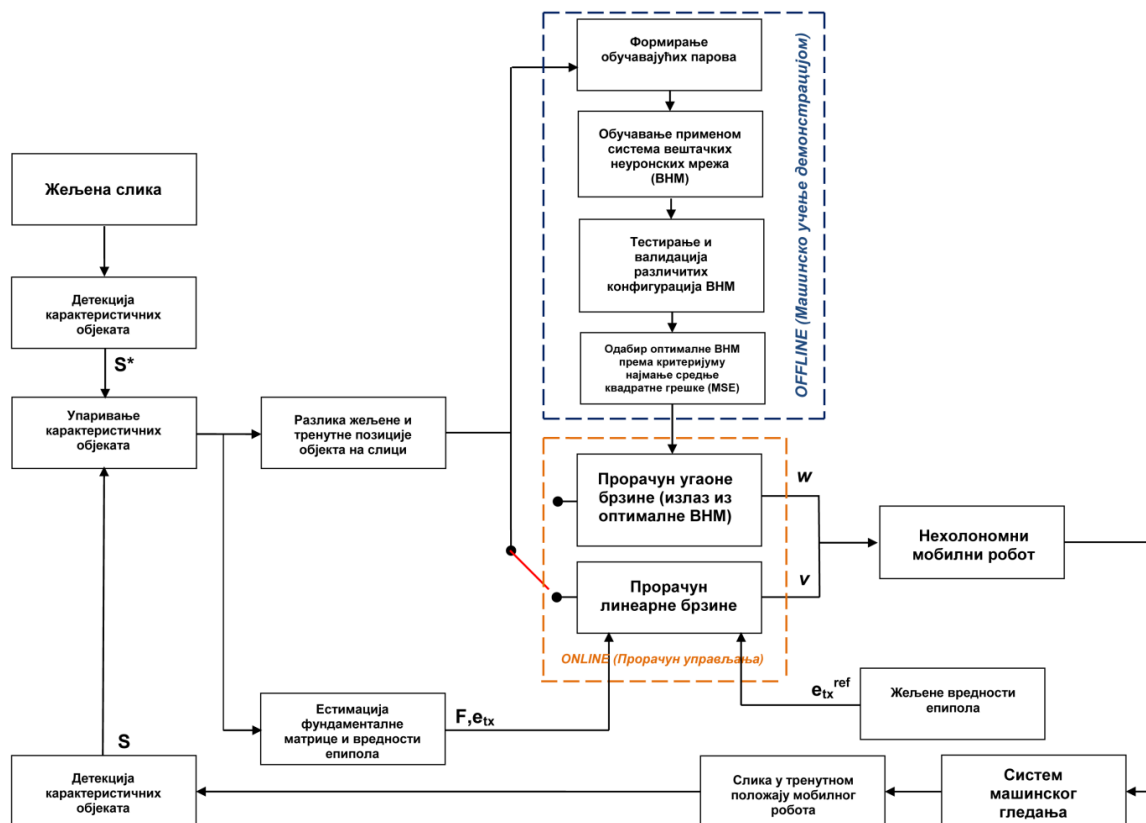
Модул машинског учења обухвата процес нелинеарног пресликавања скупа улазних у скуп излазних података. Базира се на систему вештачких неуронских мрежа са простирањем сигнала унапред (енгл. *Feedforward neural networks*). Меморисани подаци из претходног модула се овде користе за обучавање вештачке неуронске мреже (улаз су разлике у положају карактеристичног објекта, а излаз из мреже представљају управљачке команде робота). Оптимална вештачка неуронска мрежа (оптимална по питању броја неурона у слојевима, броју слојева) одређена је према критеријуму средње квадратне грешке на тест скупу података (садржи 25% од укупног скупа података). Изабрана мрежа се затим имплементира у *модул научених акција*, у коме се тестирају њене перформансе у реалном времену. За произвољно изабрани број улазних података (разлике у положају карактеристичног објекта) тестира се понашање мобилног робота у технолошком окружењу. Интелигентни управљачки подсистем који се показао најбољим са аспекта корекције оријентације мобилног робота,

имплементиран је у новоразвијени хибридни управљачки систем у циљу успешног навођења у оквиру лабораторијског модела технолошког система.



Слика 17. Дефинисање жељене области изабраног карактеристичног објекта.

Концепцијско решење разматраног новог алгоритма емпиријског управљања представљено је на слици 18. Дати алгоритам приказује ток информација у емпиријском управљачком систему, реализованог у *MATLAB*[®] софтверском окружењу.



Слика 18. Концепцијско решење новог алгоритма емпиријског управљања.

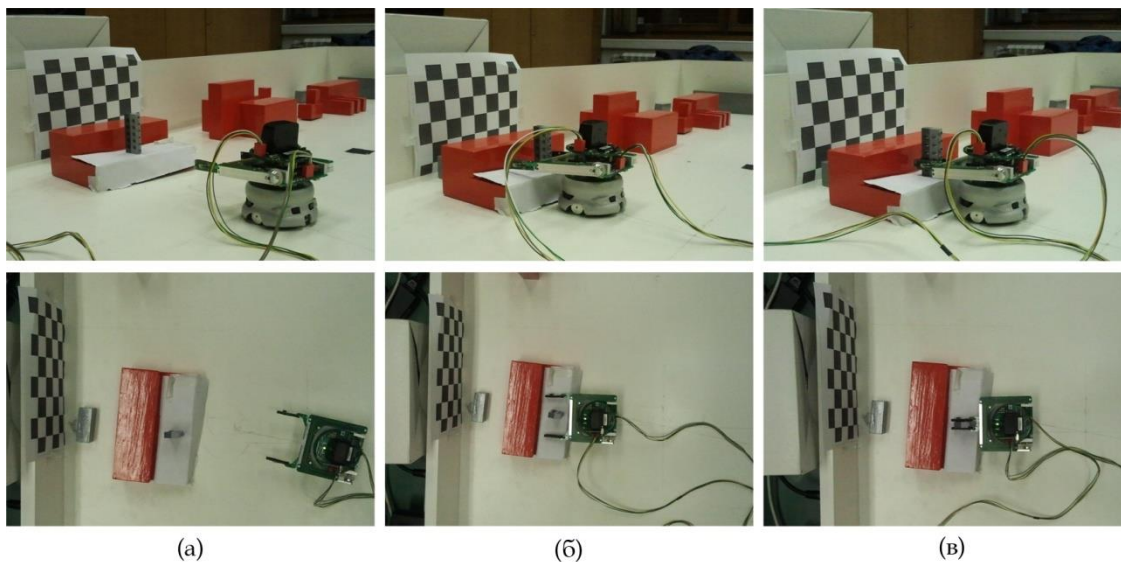
Роботски систем за експерименталну верификацију чине [4,5,6]: мобилни робот *KheperaII*, хватач *KheGrip*, компатибилна камера *CMUcam VISION TURRET-KheCMUCam* и десктоп рачунар (3 GHz,

4GB RAM), са *Windows 7* оперативним системом. Техничке карактеристике камере дате су у табели 8.. Комуникација између рачунара и робота, као и између рачунара и камере, извршава се путем *RS232* порта. Као и у свим претходним експериментима визуелног навођења, коришћен је шаблон изгледа шаховске табле. Димензије црно беле слике су 320x240 пиксела.

Табела 8. Спецификација камере <i>KheCMUCam</i>	
Сензор	CMOS, Omnivision OV7620
Резолуција слике	176x255, RGB, 24 бита по пикселу
Чип за претпроцесирање слике	SX53
Комуникација са рачунаром	Серијска веза
Пречник	55 mm
Висина	40 mm

За разлику од осталих приступа [42,43], нови хибридни интелигентни систем управљања (слика 18.) омогућава корекцију оријентације мобилног робота након достизања жељене позиције, чиме је елиминисан утицај поремећаја из реалног окружења [4,5,6].

У складу са оствареним резултатима, слика 19. приказује положај *KheperaII* мобилног робота на почетку и на крају експеримента. Евидентно је да је мобилни робот достигао жељени положај, као и да је успешно извршио хватање дела у последњој итерацији управљања. Најзначајнија предност новоразвијеног интелигентног управљања мобилног робота лежи у његовој адаптабилности: емпиријски алгоритам се може применити секвенцијално, и то на више технолошких задатака. Такође, изузетну предност представља и чињеница да није неопходно обезбедити додатну инфраструктуру или вештачки постављене маркере и објекте у технолошком окружењу ради успешне реализације визуелног навођења мобилног робота.



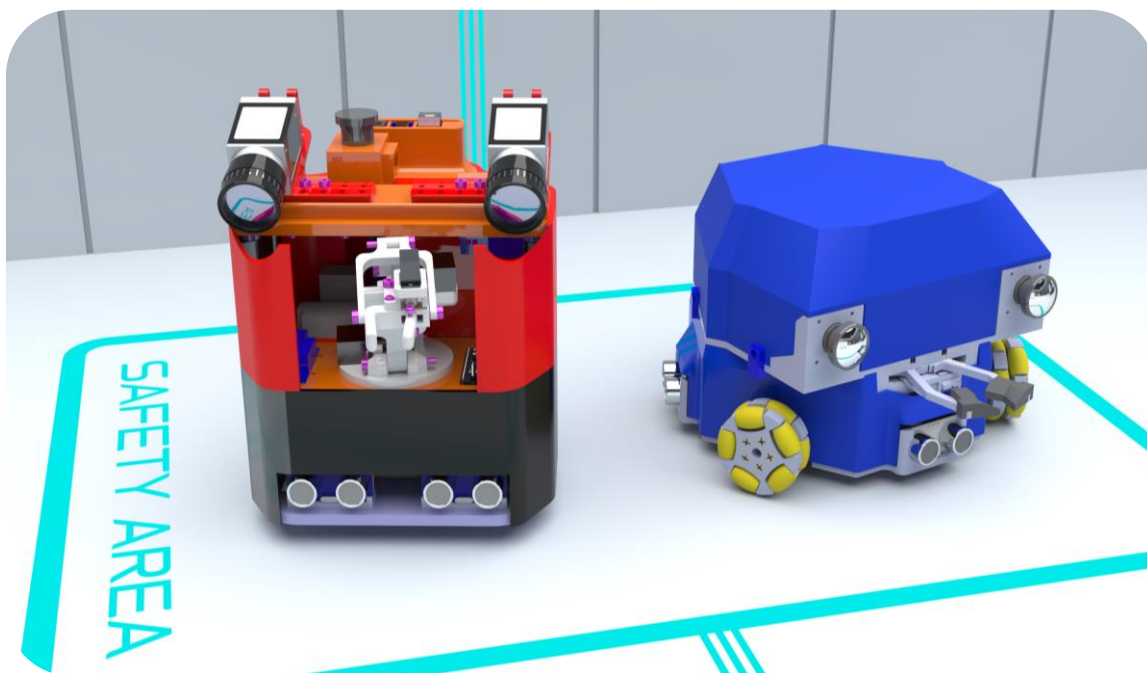
Слика 19. Експеримент визуелног навођења мобилног робота *KheperaII*:
 (а) робот у почетном положају, (б) робот у коначном положају,
 (в) хватање радног предмета у последњој итерацији управљања робота.

Позиција и оријентација мобилног робота *KheperaII* на крају експеримента, потврђују робустност развијеног система управљања и указују на неопходност имплементације техника вештачке интелигенције у циљу превазилажења бројних проблема везаних за *a priori* присутну неодређеност.

5. ДИСКУСИЈА

Пројекат Универзитета у Београду – Машинског факултета (<http://mission4-0.mas.bg.ac.rs/>) под називом „*Deep Machine Learning and Swarm Intelligence-based Optimization Algorithms for Control and Scheduling of Cyber-Physical Systems in Industry 4.0*“ (акроним MISSION4.0, ев. број: 6523109), реализован у периоду од 1. септембра 2020. до 31. децембра 2022. године, био је први научноистраживачки пројекат у домену машинства који је финансиран од стране Фонда за науку Републике Србије, и то у оквиру Програма за развој и примену вештачке интелигенције.

Основи циљ престижног пројекта MISSION4.0, односио се на развој иновативних алгоритама базираних на техникама вештачке интелигенције, између осталог и за емпиријско управљање, терминирање и остваривање сајбер безбедности кибернетско-физичких производних система Индустије 4.0 [44,45,46,47,48,49,50]. Стратегије стерео визуелног управљања мобилних робота сопственог развоја, под акронимима RAICO и DOMINO (слика 20.), на основу развоја и примене техника дубоког учења, представљао је основни циљ радног пакета три. У оквиру радног пакета четири, циљ је био развој система за емпиријско управљање мобилних роботских система на бази фази логике и дубоког учења ојачавањем.

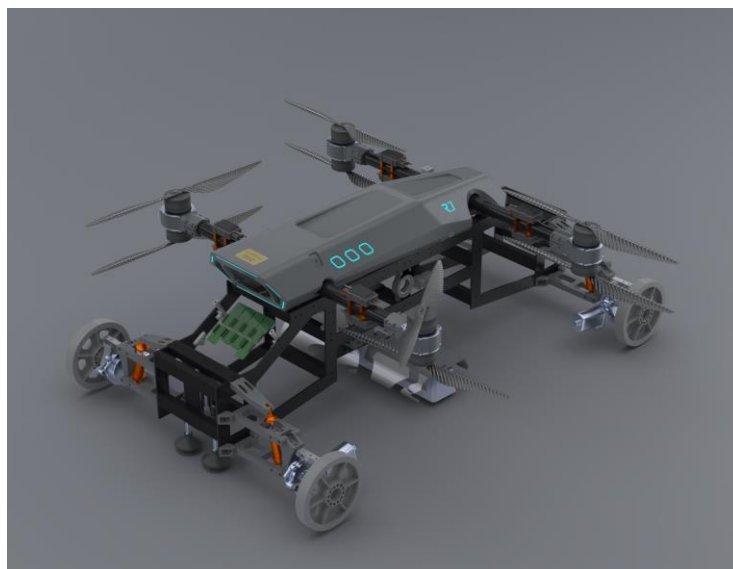


Слика 20. Прототипови мобилних робота RAICO и DOMINO.

Данас, актуелна истраживања емпиријског управљања аутономних роботских система [51,52,53] везана су за пројекат „*Cybersecurity of Motion Control System in Industry 4.0*“ (акроним MCSecurity, ев. број: 17801), Универзитета у Београду - Машинског факултета (<https://mcsecurity.mas.bg.ac.rs/>), и то у оквиру Програма ДИЈАСПОРА 2023 - заједнички програм истраживања, који такође финансира Фонд за науку Републике Србије, у периоду од 1. јануара 2025. до 31. децембра 2026. године.

У оквиру научноистраживачких активности пројекта MCSecurity, тренутно се развија дубоко машинско учење ојачавањем за мобилни робот-летелицу (слика 21.), као и алгоритам за откривање сајбер напада на систем перцепције мобилног робота PAL TIAGo (слика 22.), а који се успешно користи у оквиру Лабораторије за индустријску роботiku и вештачку интелигенцију (ROBOTICS & AI) Катедре за производно машинство (КаПроМ) Универзитета у Београду - Машинског факултета.

Дакле, кроз научне пројекте, неопходно је спроводити сталну анализу карактеристика нових алгоритама емпиријског управљања који подразумевају даљи развој, увођење и примену сродних истраживачких праваца у роботизи, заснованих на вештачкој интелигенцији. Визуелно навођење мобилног робота, базирано на хибридном алгоритму управљања коришћењем информација добијених од камере, уз имплементацију дубоког учења ојачавањем, у циљу апроксимације функције пара стање-акција $Q(\mathbf{x}_t, \mathbf{a}_t)$, те подстицање коришћења машинског учења демонстрацијом, ради унапређења перформанси аутономног мобилног робота у реалном времену, перманентно је у фокусу Лабораторије.



Слика 21. Концепцијско решење мобилног робота - летелице.



Слика 22. Експериментална поставка и прикупљање података. Мобилни робот PAL TIAGo реализује визуелно серво управљање у затвореном лабораторијском окружењу.

6. ЗАКЉУЧАК

У оквиру дате структуре рада, систематично су представљени кључни научноистраживачки резултати проистекли из тридесетогодишњег развоја емпиријског управљања аутономних роботских система, а који се и даље еволутивно користе у оквиру два пројекта, раније „*Deep Machine Learning and Swarm Intelligence-based Optimization Algorithms for Control and Scheduling of Cyber-Physical Systems in Industry 4.0*“ - *MISSION4.0*, те данас актуелног „*Cybersecurity of Motion Control System in Industry 4.0*“ - *MCSecurity*. Остварени и приказани резултати, груписани су у две категорије: дубоко учење ојачавањем и машинско учење демонстрацијом. Визуелно оријентисано, на бази коришћења камера, емпиријско управљање аутономних роботских система и примена богатог спектра техника вештачке интелигенције у остваривању концепта *Индустрија 4.0*, детаљно су разрађени у раду, и то кроз еволутивни развој

алгоритама емпиријског управљања. У погледу актуелних истраживања, везаних за стерео визуелно управљање мобилних роботских система, она су фокусирана на примену дубоког машинског учења, односно на примену конволуционих неуронских мрежа (енгл. *Convolutional Neural Networks* - CNN), ради обезбеђивања адекватних информација систему перцепције аутономних мобилних робота, посебно кроз учење робота RAICO и DOMINO. Приоритетни фокус Лабораторије за индустријску роботiku и вештачку интелигенцију (ROBOTICS & AI) Катедре за производно машинство (КаПроМ) Универзитета у Београду - Машинског факултета, оријентисан је на развој CNN модела за успешно спровођење комплексних процедура машинског гледања мобилног робота, те на даљи развој и примену дубоког учења ојачавањем за сложене задатке аутономне навигације, попут оних за мобилни робот-летелицу, као и на примену метахеуристичких алгоритама биолошки оријентисане оптимизације у процесима селекције и класификације великих скупова података прикупљених аквизицијом са сензора, преко потребним за стерео визуелно емпиријско управљање у реалном окружењу.

7. ЛИТЕРАТУРА

- [1] Fukuda,T.,Arai,F.,Ishihara,H.,Sekiyama,K.,Matsuura,H., *Laboratory of Micro System Control – Research Activities in 1997*, Nagoya University, Japan, 1998.
- [2] Brooks,R.A., *Artificial Life and Real Robots*, MIT Press, pp.3-10, 1992.
- [3] Миљковић,З., *Развој управљачких алгоритама за аутономне индустријске роботе на бази система препознавања и учења*, докторска дисертација, Универзитет у Београду – Машински факултет, 31.08.2000.
- [4] Вуковић,Н., *Развој машинског учења интелигентног мобилног робота базиран на систему вештачких неуронских мрежа*, докторска дисертација (DOI: 10.2298/BG20120928VUKOVIC), Универзитет у Београду – Машински факултет, 28.09.2012.
- [5] Митић,М., *Емпиријско управљање интелигентног мобилног робота базирано на машинском учењу*, докторска дисертација, Универзитет у Београду – Машински факултет, 15.04.2014.
- [6] Петровић,М.М., *Вештачка интелигенција у пројектовању интелигентних технолошких система*, докторска дисертација, Универзитет у Београду - Машински факултет, 24.06.2016.
- [7] Јокић,А., *Визуелно управљање мобилног робота базирано на биолошки инспирисаним техникама вештачке интелигенције*, докторска дисертација, Универзитет у Београду - Машински факултет, 16.10.2025.
- [8] Miljković,Z., Lazarević,I., *Control Strategy for Learning Industrial Robot Based on Artificial Neural Network System*, Scientific paper printed in „Advances in Systems, Signals, Control and Computers – SSCC’98“, Vol. III, Edited by Vladimir B. Bajić, Published by IAAMSAD and the South African Branch of the Academy on Nonlinear Sciences (ISBN 0-620-23136-X), pp.124-128, Durban-South Africa, 1998.
- [9] Miljković,Z., Babić,B., *Empirical Control Strategy for Learning Industrial Robot*, FME Transactions (ISSN 1451-2092), New Series, Vol.35 No.1, pp.1-8, University of Belgrade – Faculty of Mechanical Engineering, 2007.
- [10] Vuković,N., Miljković,Z., *New Hybrid Control Architecture for Intelligent Mobile Robot Navigation in a Manufacturing Environment*, FME Transactions (ISSN 1451-2092), New Series, Vol.37 No.1, pp.9-18, University of Belgrade – Faculty of Mechanical Engineering, 2009.
- [11] Mitić,M., Miljković,Z., Babić,B., *Empirical Control System Development for Intelligent Mobile Robot Based on the Elements of the Reinforcement Machine Learning and Axiomatic Design Theory*, FME Transactions (ISSN 1451-2092), New Series, Vol.39 No.1, pp.1-8, University of Belgrade – Faculty of Mechanical Engineering, 2011.
- [12] Miljković,Z., Vuković,N., Mitić,M., Babić,B., *New Hybrid Vision-Based Control Approach for Automated Guided Vehicles*, The International Journal of Advanced Manufacturing Technology (Print ISSN 0268-3768), Vol. 66 Issues: 1-4, pp.231-249, Springer-Verlag London Ltd., April 2013. <http://link.springer.com/article/10.1007%2Fs00170-012-4321-y> (Online ISSN 1433-3015_Available online: 6 July 2012_First™ Articles; DOI: 10.1007/s00170-012-4321-y).
- [13] Miljković,Z., Mitić,M., Lazarević,M., Babić,B., *Neural Network Reinforcement Learning for Visual Control of Robot Manipulators*, Journal Expert Systems with Applications (ISSN 0957-4174), Vol. 40 Issue: 5, pp.1721-1736, Elsevier, April 2013. (Available online: 9 October 2012; DOI: 10.1016/j.eswa.2012.09.010), <http://www.sciencedirect.com/science/article/pii/S0957417412010640>
- [14] Mitić,M., Miljković,Z., *Neural Network Learning from Demonstration and Epipolar Geometry for Visual Control of a Nonholonomic Mobile Robot*, Journal Soft Computing (ISSN 1432-7643), Vol. 18

- Issue: 5, pp.1011-1025, Springer Berlin Heidelberg-Germany, May 2014. (Available online: 12 Sept. 2013; DOI: 10.1007/s00500-013-1121-8), <http://rd.springer.com/article/10.1007/s00500-013-1121-8>
- [15] Vuković,N., Mitić,M., Miljković,Z., *Trajectory learning and reproduction for differential drive mobile robots based on GMM/HMM and Dynamic Time Warping using Learning from Demonstration framework*, Journal *Engineering Applications of Artificial Intelligence* (ISSN 0952-1976), Vol.45, pp.388-404, <http://www.sciencedirect.com/science/article/pii/S0952197615001499> Elsevier, 1 October 2015 (DOI: [10.1016/j.engappai.2015.07.002](https://doi.org/10.1016/j.engappai.2015.07.002)).
 - [16] Jokić,A., Petrović,M.M., Miljković,Z., *Semantic Segmentation Based Stereo Visual Servoing of Nonholonomic Mobile Robot in Intelligent Manufacturing Environment*, Journal *Expert Systems with Applications* (ISSN 0957-4174), Volume 190, p.116203, 15 March 2022, Elsevier, <https://www.sciencedirect.com/science/article/abs/pii/S0957417421015189> (Available online: 15 November 2021; DOI: 10.1016/j.eswa.2021.116203).
 - [17] Brown,R.A., *Machines that learn – based on the principles of empirical control*, Oxford University Press, 1994.
 - [18] Brown,R.A., *Consciousness in a self-learning, memory-controlled, compound machine*, Journal *Neural Networks*, Vol. 10, No. 7, pp.1333-1343, 1997.
 - [19] Becerra,H.M., López-Nicolás,G., Sagues,C., *A sliding-mode-control law for mobile robots based on epipolar visual servoing from three views*, IEEE *Transac. on Robotics*, Vol. 27, No. 1, pp.175-183, 2011.
 - [20] Sutton,R.S., Barto,A.G., *Reinforcement Learning: An Introduction*, MIT Press, 1998.
 - [21] Alpaydin, E., *Introduction to Machine Learning*, 2nd Edition, MIT Press, 2010.
 - [22] Nolfi,S., Floreano,D., *Evolutionary robotics: The biology, intelligence, and technology of self-organizing machines*, The MIT press, 2000.
 - [23] Pfeifer,R., *Understanding intelligence*, The MIT Press, 2001.
 - [24] Russell,S., *Rationality and Intelligence*, Journal *Artificial Intelligence*, Vol. 94, No. 1-2, pp.57-77, 1997.
 - [25] Russel,S., Norvig,P., (Editors), *Artificial intelligence: a modern approach*, 4th Edition, Pearson Series in Artificial Intelligence, 2022.
 - [26] Kaelbling,L.P., Littman,M.L., Moore,A.W., *Reinforcement learning: a survey*, Journal of *Artificial Intelligence Research*, Vol. 4, pp.237–285, 1996.
 - [27] Bertsekas,P.D., *Dynamic programming and optimal control*, Vol. I, 3rd Edition, 2005.
 - [28] Bertsekas,P.D., *Dynamic programming and optimal control*, Vol. II, 4th Edition: Approximate Dynamic Programming, 2012.
 - [29] Watkins,C., *Learning from delayed rewards*, PhD dissertation, King’s College, London, 1989.
 - [30] Watkins,C., Daya,P., *Technical note: Q-learning*, Journal *Machine Learning*, Vol. 8, No. 3–4, pp.279–292, 1992.
 - [31] Gaskett,C., Fletcher,L., Zelinsky,A., *Reinforcement learning for visual servoing of a mobile robot*, In: Proc. of the Australian Conference on Robotics and Automation (ACRA2000), pp.149, 2000.
 - [32] Jaakkola,T., Jordan,M.I., Singh,S.P., *On the convergence of stochastic iterative dynamic programming algorithms*, Journal *Neural Computation*, Vol. 6, No. 6, pp.1185–1201, 1994.
 - [33] Lin,L.-J., *Self-improving reactive agents based on reinforcement learning, planning and teaching*, Journal *Machine Learning*, Vol. 8, pp.293–321, 1992.
 - [34] Argall,B.D., *Learning mobile robot motion control from demonstration and corrective feedback*, PhD dissertation, Robotics Institute - Carnegie Mellon University, Pittsburgh, 2009.
 - [35] Argall,B., Chernova,S., Veloso,M., Browning,B., *A survey of robot learning from demonstration*, Journal *Robotics and Autonomous Systems*, Vol. 57, pp.469–483, 2009.
 - [36] Argall,B., Browning,B., Veloso,M., *Teacher feedback to scaffold and refine demonstrated motion primitives on a mobile robot*, Journal *Robotics and Autonomous Systems*, Vol. 59, pp.243–255, 2011.
 - [37] Abbeel,P., Coates,A., Ng,A.Y., *Autonomous Helicopter Aerobatics through Apprenticeship Learning*, *International Journal of Robotics Research*, Vol. 29, pp.1608-1639, 2010.
 - [38] Pomerleau,D., *Neural network vision for robot driving*, In: M. Arbib, (Ed.) *The Handbook of Brain Theory and Neural Networks* (1st Edition), 1998.
 - [39] Corke,P., *Visual control of robots: high-performance visual servoing*, John Wiley & Sons, 1996.
 - [40] Mitić,M., Miljković,Z., Lazarević,M., Babić,B., Lazarević,I., *Image-based visual servo control of robot manipulator under parameter uncertainties*, Proceedings of the 29th Danubia-Adria Symposium on Advances in Experimental Mechanics, pp.174-177, Belgrade, 26-29 September, 2012.
 - [41] Fischler,M.A., Bolles,R.C., *Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography*, Journal *Communications of the ACM*, Vol. 24, No. 6, pp.381–395, 1981.

- [42] López-Nicolás, G., Sagüés, C., Guerrero, J.J., Kragic, D., Jensfelt, P. *Switching visual control based on epipoles for mobile robots*, *Journal Robotics and Autonomous Systems*, Vol. 56, No. 7, pp.592-603, 2008.
- [43] López-Nicolás, G., Guerrero, J.J., Sagüés, C., *Visual control of vehicles using two-view geometry*, *Journal Mechatronics*, Vol. 20, No. 2, pp.315-325, 2010.
- [44] Јокић, А., Петровић, М.М., Миљковић, З., Бабић, Б., *Визуелно управљање мобилног робота у технолошком окружењу на бази информација добијених од камере* (техничко решење - *нова метода*: односи се на проблем визуелног емпиријског управљања мобилног робота (<https://www.indeks.rs/srbin-rajko-mali-inteligentni-mobilni-robot/>) базираног на грешкама у параметрима слике препознатих објеката унутар радног окружења (енгл. *Image-based visual servoing - IBVS* (<https://www.rts.rs/magazin/tehnologija/5006598/robotika-institut-mihajlo-pupin.html>) и одликује се робусношћу у домену тачности позиционирња мобилног робота и конвергенције грешке учења (<https://www.rts.rs/lat/magazin/tehnologija/4581215/kognitivni-mobilni-robot-rajko-masinski-fakultet.html>), а развијана је у пројекту ТР-35004 МПНПТР Владе Реп. Србије и у пројекту MISSION4.0, ев. број: 6523109 Фонда за науку Републике Србије), 2018.
- [45] Miljković, Z., Petrović, M.M., *Single Mobile Robot Scheduling Problem: A Survey of Current Biologically Inspired Algorithms, Research Challenges and Real-World Applications*, *Scientific keynote invited paper* printed in: Karabegović I. (editor) *New Technologies, Development and Application II. NT 2019. Lecture Notes in Networks and Systems*, Vol 76. Springer, 6330 Cham, Switzerland, (Series ISBN 978-3-030-18071-3) and First Online: 24 April 2019 (Online ISBN 978-3-030-18072-0), https://doi.org/10.1007/978-3-030-18072-0_4, pp.33-41, 2020.
- [46] Jokić, A., Petrović, M.M., Miljković, Z., *Dataset for semantic segmentation of the laboratory model of manufacturing environment* (Version 0.1.0), [Data set], Zenodo, <https://zenodo.org/record/4138944> (2020).
- [47] Јокић, А., Петровић, М.М., Миљковић, З., Бабић, Б., *Стерео визуелни систем перцепције мобилног робота базиран на дубоком машинском учењу* (техничко решење - *нова метода*: односи се на проблем стерео визуелног система перцепције мобилног робота (<https://startit.rs/rajko-i-domino-roboti-iz-srbije-koji-ce-automatizovati-industriju/>) и одликује се робусношћу у домену дубоког машинског учења применом конволуционе вештачке неуронске мреже (<https://ms-my.facebook.com/tvRTSNauka/videos/kako-bi-podstakao-razvoj-projekata-iz-oblasti-ve%C5%A1ta%C4%8Dke-inteligencije-fond-za-nau/254937446518166/>), са детекцијом објеката у реалном времену и тачношћу од 91,85%, а развијана је у пројекту MISSION4.0 (<https://www.rts.rs/magazin/tehnologija/4581215/kognitivni-mobilni-robot-rajko-masinski-fakultet.html>), у оквиру Програма за развој пројеката из области вештачке интелигенције (<https://www.youtube.com/watch?v=15UM9qdXg2o>), који је финансирао Фонд за науку Републике Србије, евиденц. број 6523109), <https://enauka.gov.rs/handle/123456789/581501>, 2020.
- [48] Miljković, Z., Jokić, A., Petrović, M.M., *Image Registration Algorithm for Deep Learning-Based Stereo Visual Control of Mobile Robots*, Chapter 13 printed in the scientific monograph book: *Deep Learning for Unmanned Systems*, Edited by Anis Koubaa & Ahmad Taher Azar, Series: *Studies in Computational Intelligence* (Volume 984), (ISBN 978-3-030-77938-2 & eBook ISBN 978-3-030-77939-9). printed by Springer Cham, https://doi.org/10.1007/978-3-030-77939-9_13, pp.447-479, October 2, 2021.
- [49] Miljković, Z., Jevtić, Đ., *Object Detection and Reinforcement Learning Approach for Intelligent Control of UAV*, *Scientific paper* printed in: Karabegović I., Kovačević, A., Mandžuka, S. (eds) *New Technologies, Development and Application V. NT 2022. Lecture Notes in Networks and Systems*, Vol. 472. Springer, 6330 Cham, Switzerland, (Series ISBN 978-3-031-05229-3) and First Online: 26 May 2022. (Online ISBN 978-3-031-05230-9), printed by Springer Cham, Switzerland, https://link.springer.com/chapter/10.1007/978-3-031-05230-9_79, pp.659-669, 2022.
- [50] Miljković, Z., Đokić, L., Petrović, M.M., *Application of convolutional neural networks for visual control of intelligent robotic systems*, Chapter 3 printed in the scientific monograph book: *Soft Computing in Smart Manufacturing - Solutions toward Industry 5.0*, Edited by Tatjana Šibaliја & J. Paulo Davim, Series: *Advanced Mechanical Engineering* (Volume 7), (ISBN 978-3-11-069317-1 & eBook ISBN 978-3-11-069322-5), pp.83-112, © 2022 Walter de Gruyter GmbH, Berlin/Boston, Available online: December 6, 2021. (<https://doi.org/10.1515/9783110693225-003>), Published by De Gruyter, 2022.
- [51] Jokić, A., Jevtić, Đ., Brenjo, K., Petrović, M.M., Miljković, Z., *Deep Learning-based Visual Servoing Algorithm for Wheeled Mobile Robot Control*, *Proceedings of the 15th International Scientific Conference MMA 2024 – Flexible Technologies* (ISBN 978-86-6022-681-7), pp.71-74 (DOI: [10.24867/MMA-2024-02-011](https://doi.org/10.24867/MMA-2024-02-011)), Novi Sad, Serbia, 24-26 September, 2024.

- [52] Miljković,Z., Jokić,A., Jevtić,Đ., *Cognitive Mobile Robotics Based on Intelligent Mechanisms of Learning*, Plenary Session, Invited paper - Keynote lecture, Proceedings of the 40th International Conference on Production Engineering of Serbia - ICPES 2025, <http://spms.fink.rs/doc/2025/Conference%20proceedings%20SPMS2025%20v25%20SR%20RT%20VM%201.pdf> (ISBN 978-86-6055-197-1, doi: <https://doi.org/10.46793/ICPES25.002M>), pp.2-8, Niš, Serbia, 18-19 September, 2025.
- [53] Jokić,A., Jevtić,Đ., Petrović,M.M., Miljković,Z., *Image dataset for industrial cyber-security research within mobile robotic domain* (1.0.0), [Data set], Zenodo, <https://doi.org/10.5281/zenodo.17234795> (2025).

Miljković,Z.

Empirical control of autonomous robotic systems - own development

Abstract: *New scientific achievements, especially technologies based on development of artificial intelligence during the 21st century, have contributed to the improvement of robotic systems to performing complex manufacturing tasks, and they can learn as well as adapt to real-world conditions based on data/information collected from their own environment. During thirty years of intensive scientific research, an empirical control strategy has been developed, especially for autonomous navigation, object recognition based on machine vision and their manipulation, i.e. intelligent control of autonomous robotic systems of our own development. This research, development and application, in domain of improving intelligent manufacturing systems, are aimed at designing and implementing empirical control that has the ability to understand processes in digitally oriented environment, i.e. the ability to select the best possible behaviour of an autonomous robot. This introductory paper provides an overview of some of the key scientific results established through several defended doctoral dissertations at the Faculty of Mechanical Engineering in Belgrade, as well as those achieved in the MISSION4.0 and MCSecurity projects, particularly relying on those published in prestigious international and national scientific journals, and through published chapters in scientific monographs by the world-renowned publishers, announced and subsequently printed scientific papers in proceedings of prominent conferences held abroad and in the region, as well as within the framework of verified technical solutions and very important open access datasets. These achieved results implied the development and continuous evolution of an original, empirically oriented control system based on the autonomous robot's own generated experience, acquired through the intensive interaction with the environment, considering the application of advanced concepts of deep learning and computational intelligence techniques.*

Key words: *empirical control of autonomous robotic systems, machine vision, artificial neural networks, deep reinforcement learning, learning from demonstration.*