



Department of Computer Science and Engineering
College of Engineering
Qatar University

Senior Project Report

Intelligent Mobile Target Visitation of a UAV using DRL:
A Practical Implementation of the Work by Hendawy *et al.*

Project Group Members

Abdelrahman Soliman (201701600)
Mohamad Mohamad Ali Bahri (201806966)
Mohamed Daniel Bin Mohamed Izham (201802738)

Supervisor

Dr. Amr Mahmoud Salem Mohamed

2021

This project report is submitted to the Department of Computer Science and Engineering of
Qatar University in partial fulfillment of the requirements of the Senior Project course.

Declaration

This report has not been submitted for any other degree at this or any other University. It is solely the work of us except where cited in the text or the Acknowledgements page. It describes work carried out by us for the capstone design project. We are aware of the university's policy on plagiarism and the associated penalties and we declare that this report is the product of our own work.

Student:

Date:

Signature:

Student:

Date:

Signature:

Student:

Date:

Signature:

29 **Abstract**

30 Hello, here is some text without a meaning. This text should show what a printed text will
31 look like at this place. If you read this text, you will get no information. Really? Is there no
32 information? Is there a difference between this text and some nonsense like “Huardest gefburn”?
33 Kjift – not at all! A blind text like this gives you information about the selected font, how
34 the letters are written and an impression of the look. This text should contain all letters of the
35 alphabet and it should be written in of the original language. There is no need for special content,
36 but the length of words should match the language.

37 **Acknowledgment**

38 Hello, here is some text without a meaning. This text should show what a printed text will
39 look like at this place. If you read this text, you will get no information. Really? Is there no
40 information? Is there a difference between this text and some nonsense like “Huardest gefburn”?
41 Kjift – not at all! A blind text like this gives you information about the selected font, how
42 the letters are written and an impression of the look. This text should contain all letters of the
43 alphabet and it should be written in of the original language. There is no need for special content,
44 but the length of words should match the language.

45	Table of Contents	
46	Declaration	i
47	Abstract	iii
48	Acknowledgment	iv
49	List of Figures	vii
50	List of Tables	vii
51	1 Introduction and Motivation	1
52	1.1 Problem statement	1
53	1.2 Project significance	1
54	1.3 Project objectives	1
55	2 Background and Related Work	1
56	2.1 Background	1
57	2.2 Related work	2
58	3 Requirements Analysis	4
59	3.1 Functional requirements	4
60	3.2 Design constraints	4
61	3.3 Design standards	4
62	3.4 Professional code of ethics	5
63	3.5 Assumptions	5
64	4 Proposed Solution	6
65	4.1 Solution overview	6
66	4.2 High level architecture	6
67	4.3 Hardware/software to be used	6
68	5 Proof of Concept	6
69	6 Market Research and Business Viability	6
70	7 Project Plan	7
71	7.1 Project milestones	7
72	7.2 Project timeline	7
73	7.3 Anticipated risks	7
74	8 Short Guide	7
75	8.1 Abbreviations	8
76	8.2 Figure	8
77	8.3 Equations	9
78	8.4 Simple table	9

79	8.5	Table from a csv file	10
80	8.6	Graph from a csv file	11
81	8.7	Citations	11
82	8.8	Cross-references	11
83	Appendix		14

84 **List of Figures**

85	1	High-Level Architecture	6
86	2	The arch linux logo	8
87	3	The relationship between potential and kinetic energies.	11

88 **List of Tables**

89	1	Slope, intercept and their uncertainties	9
90	2	Translational and rotational energies.	10

1 Introduction and Motivation

1.1 Problem statement

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

1.2 Project significance

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

1.3 Project objectives

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

2 Background and Related Work

2.1 Background

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

2.2 Related work

The second important concept of the project is the computer simulation. Simulation is a cost-effective, time-saving, flexible and safe way to experiment with a drone at the expense of a reduction in accuracy compared to the real world. The research realm has shown that the use of simulation is highly attractive in deep reinforcement learning (DRL) studies with drones. This is because DRL involves making gradual improvements to a model based on repeated cycles of experience, and computer simulation allows these iterations to be carried out cheaply.

According to the literature review, the combination of Gazebo, the Robot Operating System (ROS) and PX4 are the most widely used software stack for the software-in-the-loop (SITL) development because they are open-source. In contrast, Sphinx and Olympe, which are used in this project, are closed-source. Although open-source software allows for full control and the flexibility to tinker since the source code is freely available, it is less stable and time-consuming to debug when there is an error in the source code. Another conclusion from the literature is that the transfer from simulation to the real world, which this project also aims to accomplish, is lacking in the studies reviewed. As a result, it is not possible to comment on how the simulation findings effectively translate to the physical environment.

The use of simulation makes rapid experiments in realistic settings and iterative unmanned aerial vehicle (UAV) designs possible, both of which are important in artificial intelligence (AI) training. Zhou, Li, Ding, *et al.* demonstrates this by using a combination of the open-source 3D dynamic simulator Gazebo and the autopilot system PX4. Through this, they avoided the time-consuming steps of carrying out physical experiments and adjusting parameters according to the environmental settings [1]. Thanks to the simulation, the authors were also able to propose a generic framework to integrate the Deep Q-Network (DQN) algorithm into the simulated UAV environment [1]. In our work, the same Gazebo physics engine simulation software is used, and DRL is similarly trained for the high-level control of the UAV. However, the authors used the ROS-PX4 as the controller while in our work, the Olympe program is used. The main criticism against this paper is that the operating system, which was Ubuntu 16.04, and the ROS version, which was Kinetic Kame, were old and no longer supported even though the paper was written in 2020. Nevertheless, the explanation and the flowchart illustrating the Q-learning in the context of drone control are instructive for our work going forward.

The time-saving benefit and the ease of experimentation afforded by the use of computer simulation are further emphasized when studying uncertain environments. Dealing with an unknown environment for search and navigation applications, Walker, Vanegas, Gonzalez, *et al.* used simulation to train a UAV to solve a local planning problem by framing the problem as a partially observable Markov decision process (POMDP) using continuous action spaces [2]. Similar to the previous paper, the ROS-PX4 stack and the Gazebo simulation software were used compared to Olympe and Sphinx in this project [2]. In addition, both the paper and this project study path planning with DRL but our work uses it for target visitation in an obstacle-free environment while the authors used it for searching and navigation in both obstacle-free and non-obstacle-free environments. However, the use of a UAV indoor by the authors as an application does not leverage the unique features of UAVs, but it is a good starting point and easier to implement in the real world when the outdoor flight is restricted. A useful lesson that this paper provides for our project is the use of the open AI gym in creating the UAV environment

168 resulting in clearer abstraction in the codebase for the training process.

169 Yet another UAV research-related applications that profits from the use of computer simulation
170 is the testing of new sensors on the UAV. García and Molina argues that the future of UAVs relies
171 on the use of advanced sensors and the ease of analyzing their functions in real operational
172 conditions [3]. To demonstrate such viability, they connected a LIDAR sensor to a PixHawk
173 flight controller and tested the improvement that the new sensor provided in the application
174 of navigation and obstacle avoidance. Importantly, prior to that, they used QGroundControl
175 and the PX4 platforms to analyze the addition of a LIDAR sensor on a simulated 3DR Iris
176 UAV. Unlike our work, the authors' focus for using the simulation was on sensor integration
177 and not DRL which did not feature in the paper. The main criticism of this work is that the
178 sensor is simulated without noise when in the real world, the data captured by the sensor is
179 invariably noisy. Although the objectives between the authors' and our work are different, it is
180 still very helpful to learn from the extensive software stack and architecture guide presented by
181 the authors.

182 The third and final concept of the project is hardware realization for drone visits. The hard-
183 ware part is essential in the implementation in the real world, where the simulation sometimes
184 strays from the truth. There is a lack of hardware implementations in the field of research
185 regarding drones with DRL, and most of the research papers focus on the simulations. According
186 to the literature review, the convolutional neural network (CNN) models were used in the majority
187 of the papers for object detection. Also, the controller boards and custom drone kits were used
188 instead of commercial drones. Those kits give the researcher and user more flexibility since the
189 drone is customizable in hardware and software. But in our design, we will use a commercial
190 drone so that we focus on the DRL, not the actual drone build process.

191 The use of a microcomputer with Quadcopter UAVs and autopilot software will help in the
192 hardware implementation part. Khan, Tufail, Khan, *et al.* used the drone in the agriculture field
193 to spray pesticides and monitor the crops. Unlike our work, the drone was limited to specific
194 boundaries and fixed targets such as crops. They used a Raspberry Pi microcomputer board
195 attached to the drone, which will handle two different operations. Firstly, it will control the
196 drone using an open-source software called Arducopter autopilot which will handle the trip of
197 the drone and autonomous flight option. The second operation is to deal with the Intel neural
198 computer stick 2, which will deploy the CNN model and deal with the computation part [4].
199 Although this work is close to ours, there are some differences, one of them is using a custom
200 drone which is not considered since we are limited in time. Since we will use the Anafi drone,
201 the Olympe program will take control of the drone, which will be installed on the Raspberry Pi.
202 Finally, using CNN only is not enough without DRL which makes the drone more intelligent and
203 accurate.

204 A helpful example that uses a commercial drone with an onboard computer and uses SDK
205 with image processing techniques. The hardware architecture in Wang, Gu, Huang, *et al.* work
206 for this paper includes a DJI commercial drone and an onboard computer called manifold, which
207 is from the same manufacturer. Also, onboard sensors like camera, GPS and inertial sensor are
208 included. Finally, external battery for the manifold computer and Wi-Fi adapter that is used
209 for connection between the drone and the onboard computer. This hardware architecture is
210 inspirational, and our design is somehow close to it with minor changes in the onboard computer
211 and without the existence of the sensors. Image and video processing techniques were used, such

as segmentation to keep detecting moving targets was presented in [5]. For the navigation part, they used predetermined waypoints related to historical path cost. However, in our work, probability and mobility patterns will be used to guess the target's location.

An embedded system connected and attached to the UAV and uses mobility pattern recognition, which shortens response time and saves transmission bandwidth. Wang, Zhao, Yang, *et al.* work used a quadrotor UAV supported with GPS module and a Pix Hawk flight controller. The power sources in the architecture were two lithium batteries, one for the drone and one for the embedded system. The system uses NVIDIA Jetson development kits which give enough computing power for the processing and communication between the flight controller and the system. The Jetson board is connected to the flight controller using serial communication while connected to the ground controller using Wi-Fi. Communication tools and protocols used in Wang, Zhao, Yang, *et al.* work will help us to determine the best way to communicate between the development board and the drone without any delay or interference [6].

3 Requirements Analysis

3.1 Functional requirements

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

3.2 Design constraints

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

3.3 Design standards

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the

248 alphabet and it should be written in of the original language. There is no need for special content,
249 but the length of words should match the language.

250 **3.4 Professional code of ethics**

251 Hello, here is some text without a meaning. This text should show what a printed text will
252 look like at this place. If you read this text, you will get no information. Really? Is there no
253 information? Is there a difference between this text and some nonsense like “Huardest gefburn”?
254 Kjift – not at all! A blind text like this gives you information about the selected font, how
255 the letters are written and an impression of the look. This text should contain all letters of the
256 alphabet and it should be written in of the original language. There is no need for special content,
257 but the length of words should match the language.

258 **3.5 Assumptions**

259 Hello, here is some text without a meaning. This text should show what a printed text will
260 look like at this place. If you read this text, you will get no information. Really? Is there no
261 information? Is there a difference between this text and some nonsense like “Huardest gefburn”?
262 Kjift – not at all! A blind text like this gives you information about the selected font, how
263 the letters are written and an impression of the look. This text should contain all letters of the
264 alphabet and it should be written in of the original language. There is no need for special content,
265 but the length of words should match the language.

4 Proposed Solution

4.1 Solution overview

4.2 High level architecture

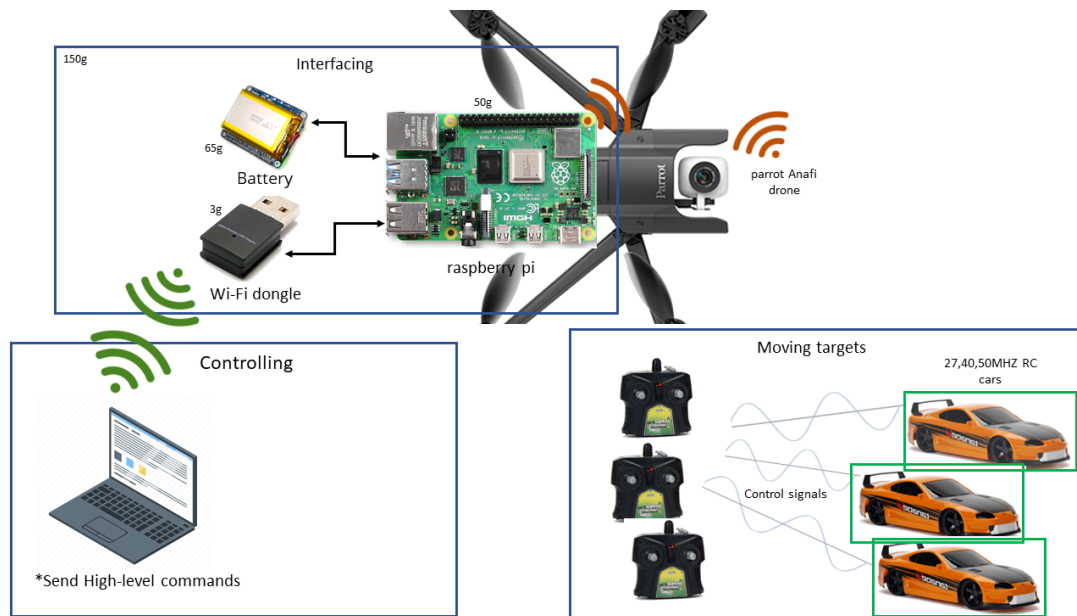


Figure 1: High-Level Architecture

4.3 Hardware/software to be used

5 Proof of Concept

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

6 Market Research and Business Viability

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the

284 alphabet and it should be written in of the original language. There is no need for special content,
285 but the length of words should match the language.

286 **7 Project Plan**

287 **7.1 Project milestones**

288 Hello, here is some text without a meaning. This text should show what a printed text will
289 look like at this place. If you read this text, you will get no information. Really? Is there no
290 information? Is there a difference between this text and some nonsense like “Huardest gefburn”?
291 Kjift – not at all! A blind text like this gives you information about the selected font, how
292 the letters are written and an impression of the look. This text should contain all letters of the
293 alphabet and it should be written in of the original language. There is no need for special content,
294 but the length of words should match the language.

295 **7.2 Project timeline**

296 Hello, here is some text without a meaning. This text should show what a printed text will
297 look like at this place. If you read this text, you will get no information. Really? Is there no
298 information? Is there a difference between this text and some nonsense like “Huardest gefburn”?
299 Kjift – not at all! A blind text like this gives you information about the selected font, how
300 the letters are written and an impression of the look. This text should contain all letters of the
301 alphabet and it should be written in of the original language. There is no need for special content,
302 but the length of words should match the language.

303 **7.3 Anticipated risks**

304 Hello, here is some text without a meaning. This text should show what a printed text will
305 look like at this place. If you read this text, you will get no information. Really? Is there no
306 information? Is there a difference between this text and some nonsense like “Huardest gefburn”?
307 Kjift – not at all! A blind text like this gives you information about the selected font, how
308 the letters are written and an impression of the look. This text should contain all letters of the
309 alphabet and it should be written in of the original language. There is no need for special content,
310 but the length of words should match the language.

311 **8 Short Guide**

312 Please read the guides available online about the right way to write L^AT_EX such as how to include
313 a math symbol in text (e.g. x not x) and a proper noun with all capitals (e.g. SQL not SQL).

314 Below are examples of different constructs in a report. You can copy-paste and change the
315 content. For more information, refer to the relevant package manual in CTAN.

8.1 Abbreviations

To add an abbreviation (e.g. UAV), append the following line in the list of abbreviations portion in `main.tex`:

```
\newacronym{uav}{\textsc{uav}}{unmanned aerial vehicle}
```

To use the abbreviation, there are 3 ways to do so:

1. In a normal case: `\gls{uav}`
2. For its plural form: `\glspl{uav}`
3. In the beginning of a sentence: `\Gls{uav}`
4. A combination of cases 2 and 3: `\Glspl{uav}`

For example:

An UAV has many unique features. UAVs have been used in many different applications.

8.2 Figure



Figure 2: The arch linux logo

329 8.3 Equations

$$E_p = mgh = mg(x_f - x_i) \quad (1)$$

$$E_k = E_t + E_r$$

$$E_t = \frac{1}{2}mv^2 \quad (2)$$

$$E_r = \frac{1}{2}I\omega^2 \quad (3)$$

$$I = \frac{1}{2}MR^2 \quad (4)$$

$$\omega = \frac{v}{r}$$

$$E_k = \frac{1}{2}mv^2 + \frac{1}{2}I\left(\frac{v}{r}\right)^2 \quad (5)$$

330 where E_p is the potential energy, E_k the kinetic energy, E_t the translational energy and E_r the
331 rotational energy.

$$\begin{aligned} \frac{\partial E_p}{\partial m} &= \frac{\partial}{\partial m}(mgh) \\ &= gh \end{aligned}$$

$$\begin{aligned} \frac{\partial E_p}{\partial g} &= \frac{\partial}{\partial g}(mgh) \\ &= mh \end{aligned}$$

$$\begin{aligned} \frac{\partial E_p}{\partial h} &= \frac{\partial}{\partial h}(mgh) \\ &= mg \end{aligned}$$

332 8.4 Simple table

Table 1: Slope, intercept and their uncertainties

Slope		Intercept (J)	
Value	Error	Value	Error
1.0933	0.0300	0.0148	0.0157

333 **8.5 Table from a csv file**

Table 2: Translational and rotational energies.

m kg	v_m m s^{-1}	E_t J	δE_t J	E_r J	δE_r J
0.055	0.17	0.000 79	0.000 01	0.280	0.007
0.075	0.20	0.001 50	0.000 02	0.387	0.010
0.095	0.23	0.002 51	0.000 03	0.512	0.013
0.115	0.25	0.003 59	0.000 03	0.605	0.015
0.135	0.27	0.004 92	0.000 04	0.706	0.018

334 8.6 Graph from a csv file

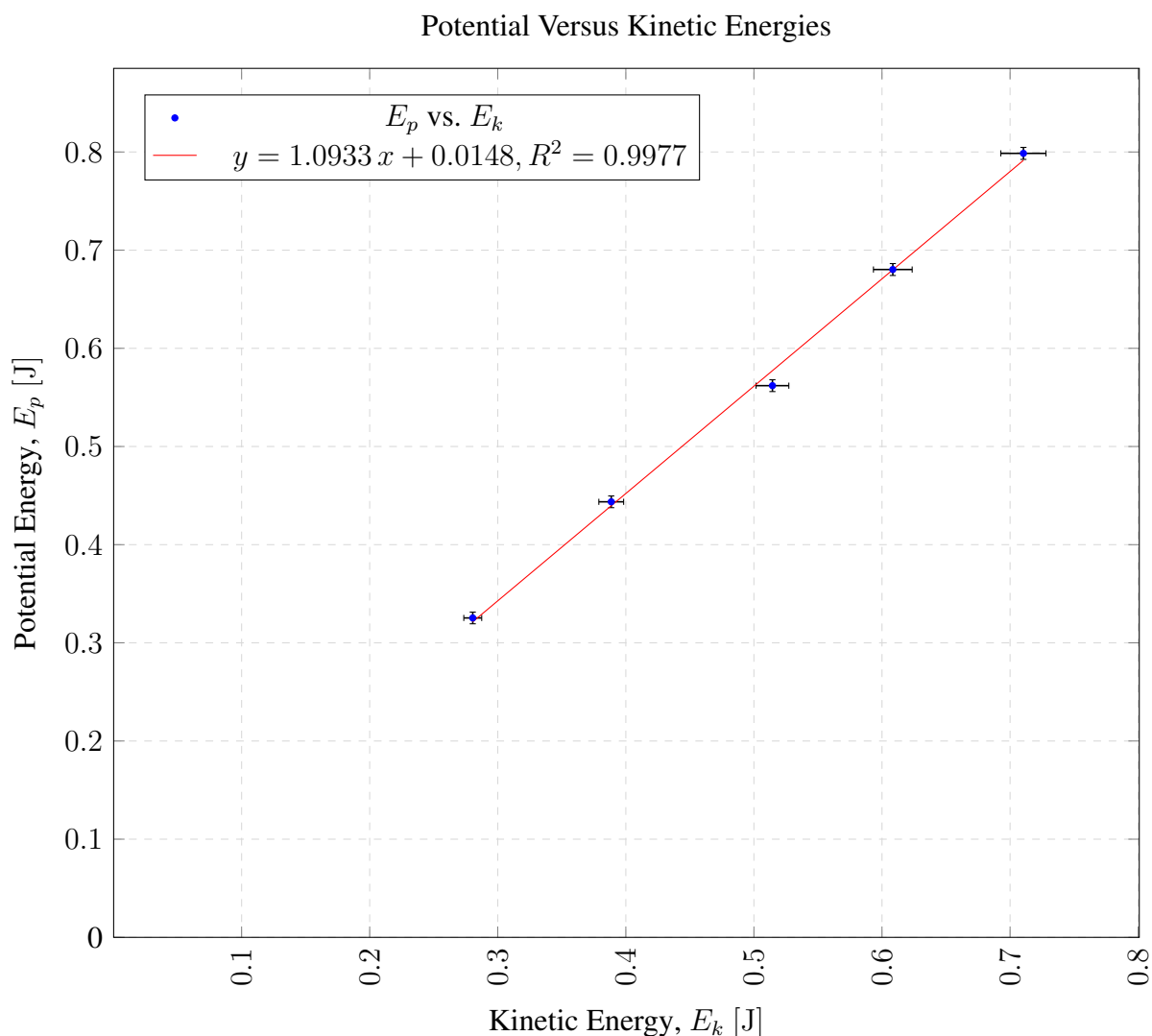


Figure 3: The relationship between potential and kinetic energies.

335 8.7 Citations

- 336 • **in-text citation:** use `\cite{dirac}` to produce [7] or `\textcite{dirac}` to pro-
 337 duce Dirac [7]
- 338 • **citation in parentheses:** `\parencite{knuthwebsite}` produces [8] (for IEEE, this
 339 has no difference to the `\cite{}` command above.)

340 8.8 Cross-references

341 Label using suitable names with the following format: figure `\label{fig:<name>}`, tables
 342 `\label{tab:<name>}`, sections `\label{sec:<name>}` and equations

343 `\label{eq:<name>}`.
344 **Then when cross-referencing, use `\cref{<type>:<name>}`**
345 **(or `\Cref{<type>:<name>}` when used at the beginning of a sentence)**

References

- [1] S. Zhou, B. Li, C. Ding, L. Lu, and C. Ding, “An efficient deep reinforcement learning framework for UAVs,” in *2020 21st International Symposium on Quality Electronic Design (ISQED)*, IEEE, Mar. 2020, pp. 323–328. DOI: 10.1109/isqed48828.2020.9136980.
- [2] O. Walker, F. Vanegas, F. Gonzalez, and S. Koenig, “A deep reinforcement learning framework for UAV navigation in indoor environments,” in *2019 IEEE Aerospace Conference*, IEEE, Mar. 2019, pp. 1–14. DOI: 10.1109/aero.2019.8742226.
- [3] J. García and J. M. Molina, “Simulation in real conditions of navigation and obstacle avoidance with PX4/Gazebo platform,” *Personal and Ubiquitous Computing*, Jan. 2020. DOI: 10.1007/s00779-019-01356-4.
- [4] S. Khan, M. Tufail, M. Khan, Z. Khan, J. Iqbal, and A. Wasim, “Real-time recognition of spraying area for uav sprayers using a deep learning approach,” Apr. 2021. DOI: 10.1371/journal.pone.0249436.
- [5] Q. Wang, J. Gu, H. Huang, and Y. Zhao, “Online drone-based moving target detection system in dense-obstructer environment,” IEEE, Dec. 2018. DOI: 10.1109/PADSW.2018.8644908.
- [6] C. Wang, R. Zhao, X. Yang, and Q. Wu, “Research of uav target detection and flight control based on deep learning,” IEEE, May 2018. DOI: 10.1109/ICAIBD.2018.8396188.
- [7] P. A. M. Dirac, *The Principles of Quantum Mechanics*, ser. International series of monographs on physics. Clarendon Press, 1981, ISBN: 9780198520115.
- [8] D. Knuth. “Knuth: Computers and typesetting.” (), [Online]. Available: <http://www-cs-faculty.stanford.edu/~uno/abcde.html>.

369 Appendix

370 Hello, here is some text without a meaning. This text should show what a printed text will
371 look like at this place. If you read this text, you will get no information. Really? Is there no
372 information? Is there a difference between this text and some nonsense like “Huardest gefburn”?
373 Kjift – not at all! A blind text like this gives you information about the selected font, how
374 the letters are written and an impression of the look. This text should contain all letters of the
375 alphabet and it should be written in of the original language. There is no need for special content,
376 but the length of words should match the language.