



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

79	8.5	Table from a csv file . . . . .	9
80	8.6	Graph from a csv file . . . . .	10
81	8.7	Citations . . . . .	10
82	8.8	Cross-references . . . . .	10
83	<b>Appendix</b>		<b>12</b>

84 **List of Figures**

85	1	High-Level Architecture . . . . .	5
86	2	The arch linux logo . . . . .	8
87	3	The relationship between potential and kinetic energies. . . . .	10

88 **List of Tables**

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

# 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. **Zho20** 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 **Zho20**. 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 **Zho20**. 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, **Wal19** 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 **Wal19**. Similar to the previous paper, the ROS-PX4 stack and the Gazebo simulation software were used compared to Olympe and Sphinx in this project **Wal19**. 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 resulting in clearer

168 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. **Gar20** argues that the future of UAVs relies on the use of  
171 advanced sensors and the ease of analyzing their functions in real operational conditions **Gar20**.  
172 To demonstrate such viability, they connected a LIDAR sensor to a PixHawk flight controller  
173 and tested the improvement that the new sensor provided in the application of navigation and  
174 obstacle avoidance. Importantly, prior to that, they used QGroundControl and the PX4 platforms  
175 to analyze the addition of a LIDAR sensor on a simulated 3DR Iris UAV. Unlike our work, the  
176 authors' focus for using the simulation was on sensor integration and not DRL which did not  
177 feature in the paper. The main criticism of this work is that the sensor is simulated without  
178 noise when in the real world, the data captured by the sensor is invariably noisy. Although the  
179 objectives between the authors' and our work are different, it is still very helpful to learn from  
180 the extensive software stack and architecture guide presented by the authors.

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

190 The use of a microcomputer with Quadcopter UAVs and autopilot software will help in the  
191 hardware implementation part. **Khan21** used the drone in the agriculture field to spray pesticides  
192 and monitor the crops. Unlike our work, the drone was limited to specific boundaries and fixed  
193 targets such as crops. They used a Raspberry Pi microcomputer board attached to the drone ,  
194 which will handle two different operations. Firstly, it will control the drone using an open-source  
195 software called Arducopter autopilot which will handle the trip of the drone and autonomous  
196 flight option. The second operation is to deal with the Intel neural computer stick 2, which will  
197 deploy the CNN model and deal with the computation part **Khan21**. Although this work is close  
198 to ours, there are some differences, one of them is using a custom drone which is not considered  
199 since we are limited in time. Since we will use the Anafi drone, the Olympe program will take  
200 control of the drone, which will be installed on the Raspberry Pi. Finally, using CNN only is not  
201 enough without DRL which makes the drone more intelligent and accurate.

202 A helpful example that uses a commercial drone with an onboard computer and uses SDK  
203 with image processing techniques. The hardware architecture in **Wang18** work for this paper  
204 includes a DJI commercial drone and an onboard computer called manifold, which is from the  
205 same manufacturer. Also, onboard sensors like camera, GPS and inertial sensor are included.  
206 Finally, external battery for the manifold computer and Wi-Fi adapter that is used for connection  
207 between the drone and the onboard computer. This hardware architecture is inspirational , and  
208 our design is somehow close to it with minor changes in the onboard computer and without the  
209 existence of the sensors. Image and video processing techniques were used, such as segmentation  
210 to keep detecting moving targets was presented in **Wang18**. For the navigation part, they used  
211 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. **Zhao18** 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 **Zhao18** work will help us to determine the best way to communicate between the development board and the drone without any delay or interference **Zhao18**.

## 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 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.

### 3.4 Professional code of ethics

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.

### 3.5 Assumptions

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.

## 4 Proposed Solution

### 4.1 Solution overview

### 4.2 High level architecture

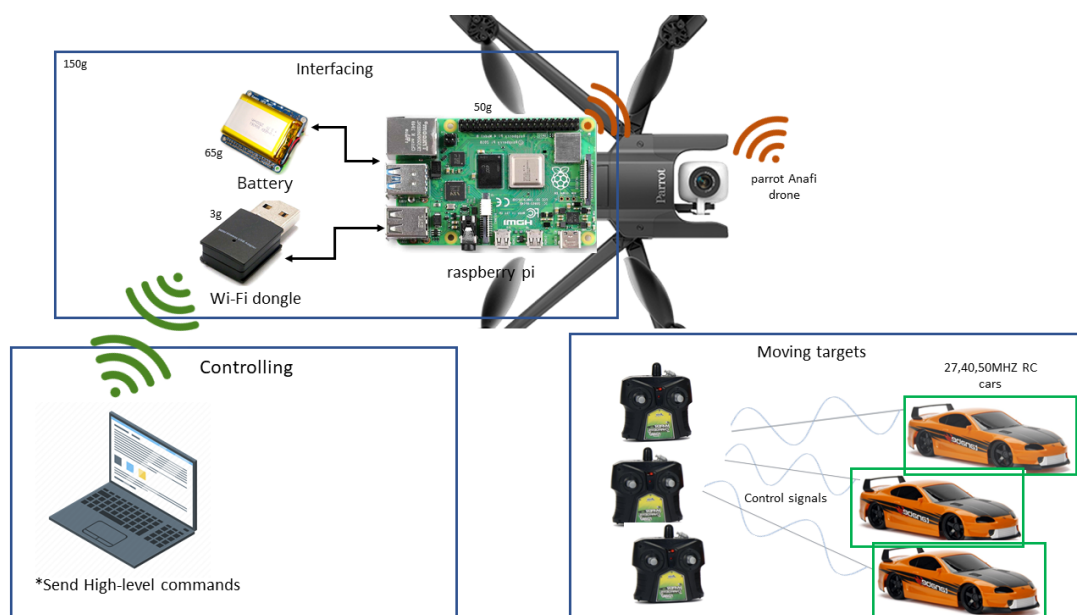


Figure 1: High-Level Architecture

### 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 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.

## 7 Project Plan

### 7.1 Project milestones

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.

### 7.2 Project timeline

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.

### 7.3 Anticipated risks

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.

## 8 Short Guide

Please read the guides available online about the right way to write L<sup>A</sup>T<sub>E</sub>X such as how to include a math symbol in text (e.g.  $x$  not x) and a proper noun with all capitals (e.g. SQL not SQL).

Below are examples of different constructs in a report. You can copy-paste and change the 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.



Figure 2: The arch linux logo

## 326 8.2 Figure

## 327 8.3 Equations

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

$$E_k = E_t + E_r$$

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

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

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

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

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

328 where  $E_p$  is the potential energy,  $E_k$  the kinetic energy,  $E_t$  the translational energy and  $E_r$  the  
329 rotational energy.

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

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

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

## 330 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

## 331 8.5 Table from a csv file

Table 2: Translational and rotational energies.

$m$	$v_m$	$E_t$	$\delta E_t$	$E_r$	$\delta E_r$
kg	$\text{m s}^{-1}$	J	J	J	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



## 332 8.6 Graph from a csv file

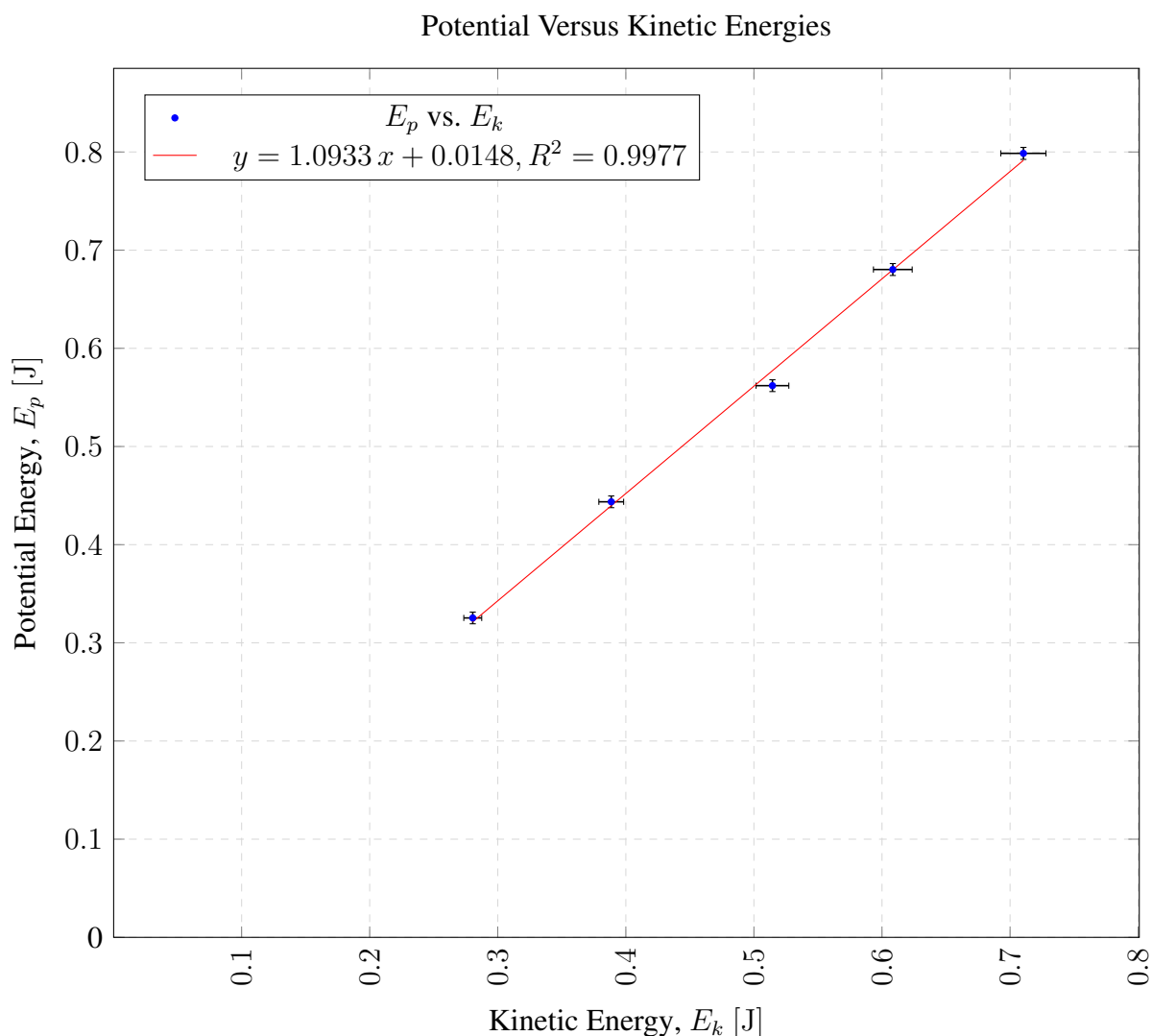


Figure 3: The relationship between potential and kinetic energies.

## 333 8.7 Citations

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

## 338 8.8 Cross-references

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

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

## 344 **Appendix**

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