**Dataset of drone esncounters**

A group of people walking down a road

Description automatically generated with medium confidence

**Academic Guidance:** Dr Jessica Cauchard

**Date: 01.07.21**

**Project members:**

Jonathan Lahat,

Alon Karmona,

Roy Yogev

**Table of Contents**

Abstract ……………………………………………………………………….…….…. 2 – 3

List of tables, List of Figure, List of Terms and Abbreviations..………...…….………. 4

Introduction……………………...…………………………….….………..................... 5 - 6

Literature review ……………...…….……………………….……………….….……. 7 - 11

Protocol ………..……………….……………………………..……………..………. 12 - 17

Systems used………………..…………………………...…………….……...….………. 18

Experiment process.………..…………………………...……….……...….………. 19 - 21

Analysis …………………………………………………...………….….….………. 22 – 23

Our Dataset …………………………...…………….….………………….……………... 24

Discussion …………………………………….…………...…………….…..……………. 26

Conclusions ..……………………………………………...…………….…..……………. 27

Summary and future work………………………………...…………….…..……………. 27

References....……………………………………………...…………….…..……………. 27

Appendix…………………………………………………...…………….…..……………. 27

**Abstract**

The expansion in drones' use in recent years is a direct outcome of the research and developments made in the field. The aerodynamic abilities, lightweight and relatively low price – helped the Unmanned Aerial Vehicles become a key component in different fields in the different sectors, from its use in classified military missions – to a simple Pizza delivery.

Another effect of the R&D in drone technology is the level of autonomy that the drone has. The levels vary between strictly manual ones and drones that are mission-oriented when they are sent to a mission and returned upon completion.

The fast growth in their presence in populated areas has exposed another tier in the sub-field of Human-Robot Interaction (HRI) - Human Drones Interaction (HDI). That is the study of the interaction between drones with by-passers – who are not the main subjects of interest in the drone's mission (not the user or the target). An autonomous drone that is sent to provide an essential drug to a patient (e.g.)  – needs to know how its environment reacts to his presence and act proactively.

In our project, we have created a dataset of annotated footage of the interaction between drones and bystanders – from the point of view of the drone. After studying the different datasets of videos published in the HDI field and analyzing their specifications, we have decided to use a deception method in which we have created a situation that simulates an encounter between an autonomous drone and by-passers. The footage was collected in public areas and with people that were not aware of the experiment.

The deception has helped us collect spontaneous reactions of participants, the participants were only notified about the experiment after the footage was already taken. This was followed by asking for their permission to use the videos for academic reasons. After granting permission, the users were also asked a few questions about their experience of the interaction. After editing and keeping the approved videos, we have done a basic Positive/Negative/Neutral annotation to the video frames. In addition, we have also pointed out the frame in which the participant notices the drone.

The datasets will be used as the basis on which designed models will perform a process of Supervised Learning. Using this Machine Learning process, we will enable the drone to analyze its surroundings in real-time, correctly classify the nature of the interaction with by-passers and make decisions accordingly.

Keywords: Human-Drone Interaction, Dataset, drone, videos, by-passers, Annotation, Machine-Learning

**List of Figure**

1.1) Experiment process chart……………………………………………………..15

2.1) DJI MAVIC Air 2 image…………………………………………………………18

2.2) Image Annotation Lab Symbol………………………………………………..18

2.3) Windows 10 Symbol……………………………………………………………18

2.4) Excel software symbol…………………………………………………………18

3.1) Illustration of Team's Layout………………………………………………….19

4.1) Frame's example image………………………………………………………21

4.2) Labeling1 image……………………………………………………………….21

4.3) Labeling2 image……………………………………………………………….21

5.1) Age distribution bar chart……………………………………………………..22

5.2) People reaction distribution bar chart……………………………………….23

5.3) Labeling distribution bar chart….……………………………………………24

6.1) Dataset example image………………………………………………………24

**List of Terms and Abbreviations**

JSON - JavaScript Object Notation

AI - Artificial Intelligence

ML - Machine Learning

HDI –  Human-Drone Interaction

HRI - Human-Robot Interaction

UAV - Unmanned Aerial Vehicles

R&D - Research and Development

**Introduction**

Drones, also known as Unmanned Aerial Vehicles (UAVs), are pilotless aircraft systems that are used for a wide range of tasks, from battlefield surveillance, borders patrol and air ambulance to agriculture monitoring, photography, package delivery and others [34,40].

The expansion in their use in recent years is a result of the research and developments made in the field. Governmental, military and private sections are taking advantage of the reduced cost, time and risk taken in the performance of a vast variety of applications.  
Furthermore, their growing popularity has resulted a convenient price and a fairly low maintenance expenses that make Drones a lucrative tool [6,15,24,26,29,31,35,36,37].

Many of the missions that have been allocated to drones involve different levels of interactions with humans. In missions such as medical aid delivery and Search and Rescue actions- unplanned complications can lead to a failure that has dire consequences.  
Nonetheless, we are at the very early stage of understanding how people will interact with them in real time.  
The research regarding their effect on humans, from different ages and backgrounds, hasn't met yet the fast growth in their presence in populated areas [20].

With that, technological developments are contributing to the UAVs capability to make decisions and raises their level of autonomy. In our research we want to focus on studying the reactions of passers-by to the flying drones.  
The prognosis of how a positive interaction can contribute to the successful completion of crucial tasks, and vice versa, is our motivation to study those encounters.  
  
Based on that knowledge, autonomous drones will be able to correctly classify the nature of the interaction with by -passers and make decisions accordingly.  
For instance, a drone has been missioned to deliver an essential drug. On the way to the clinic, it encounters a man that does not favor his presence around his house and starts to throw objects on the drone. We would want the drone to correctly identify the situation as a negative encounter that could risk its mission- and decide to change its path.

In this project, we will be building a dataset of videos taken by drones that documents different encounters with by-passers and classify them. The datasets will be used as the basis on which designed models will perform a process of Supervised Learning.

**Literature review**

Human- Drone Interaction

Human–Robot Interaction (HRI) is a field of study dedicated for understanding, designing, and evaluating robotic systems for use by or with humans [20]. The investment in research and development in robotic systems, over the years, has led to a continuous technological growth in the field [41]. The HRI field of study has been sensitive to the changes and followed the developments accruing- with research.  
The fast development of drone technology in recent years - opens a new sub-field of HRI, Human-Drone interaction (HDI). The flying capabilities and the fact that drones technology has allowed availability in many sizes and designs- brings new research challenges to this field.

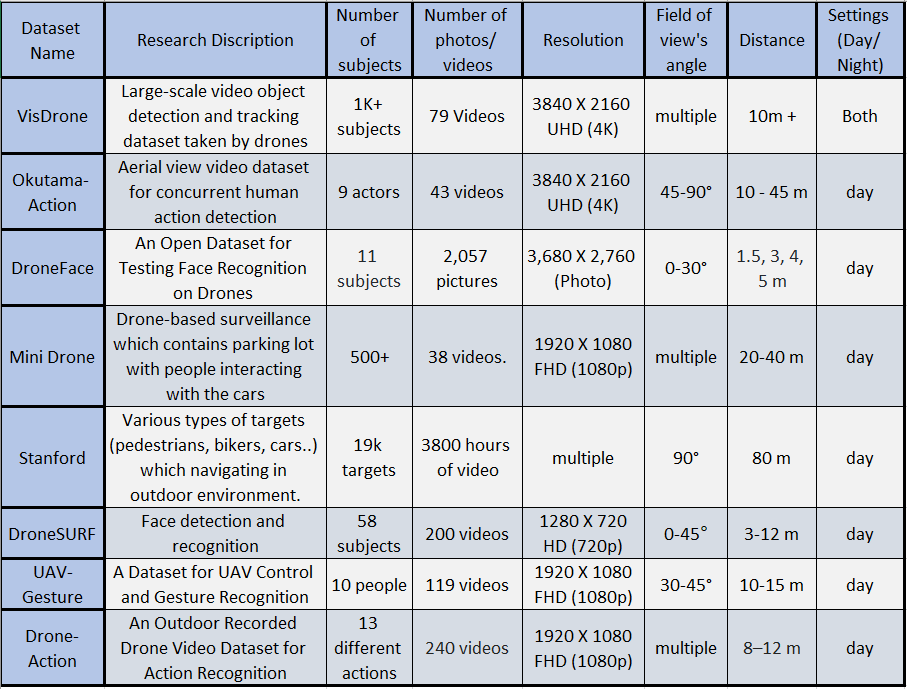
The most basic interaction of humans and drones is the control of the drone’s position.  
In the traditional form, the user directly controls the flight path using an interface. In a more advanced form- the drone is task-oriented controlled, in that case the drone is given a task and it is sent to accomplish it. The feedback and control of the operator, in the different forms, varies [17].

Although HDI is a fairly new discipline, various studies have been conducted on human-drone interactions [42].  Prior work shows how people naturally interact with drones by executing a study with participants performing a range of tasks with the UAVs. This study found a combination of gestures, voice and multimodal interactions that felt intuitive to participants [38].  
Another study shows that the appearance of the drone and the knowledge of its safety - affects how people naturally interact with it [11]. The study found that most users naturally touched the drone when they knew it was a safe-to-touch Drone. Similarly, studies created a prototype of a social drone with a social shape, face and voice for human interaction has made the drone more approachable [13,14].  
In a study that investigated the preferences of drones’ users, researchers have shown that bystanders favours a distance of at least 2 meters between them and the drones [22].

As our study focuses on accurate interaction classifications, we have explored the visual methods of correctly identifying a person’s reaction. Both full-body movement features and face expressions recognition has been proved to be a useful tool to do so [3,4].  
Creating contrived situations and acting them out would not reflect the way features are distributed in real spontaneous actions [33]. accordingly, and in order to create a high-quality HDI dataset, the data will be collected from natural encounters with bystanders.  
The Covid-19 pandemic has created a new condition in which people wear a mask in public areas. Due to the pandemic and in respect to people's aerial space, we have chosen to focus the dataset on the recognition of body gestures.

Existing Drone Datasets

To understand the specific requirements we need to apply in the dataset we wish to create; and in order for it to be as informative as possible for further usage - we have analyzed other databases with a similar purposes [9,10,12,16,25,28,32]. After extracting their specifications, we prepared a comparison table based on their different parameters. Table 1 shows the features and settings which has been used to create similar datasets.



As featured in the table, the researchers who built the various databases had different goals. Some of whom wanted to identify objects or targets in the photographed space, some of the studies involved face recognition - and the rest involved identifying body gestures.  
The datasets UAV-Gesture and Drone-Action aimed for body gesture recognition too, but the participants were previously instructed to perform certain movements.  
Mini-drone video dataset documents events occurring in parking lots, the interactions are either with cars or with other people - but not with the drone itself.  
Therefore, it is not possible to rely on these databases to identify a spontaneous encounter between a person and a UAV.

The dataset we will build will contain videos only and as we want our subjects to be unfamiliar with the study - they will only participate once.   
In the data collection stage, the different datasets have used different camera resolution – although maintaining the common Aspect Ratio (16:9) [43]. The lowest resolution used was by DroneSURF with HD video camera, while VisDrone and Okutama-Action Action used 4K resolution.

Except for Okutama-Action, all datasets that have been seeking face recognition or body gesture recognition did not exceed 45º angle. Okutama-Action researches mentioned that as the angle got higher (to the extent of 90º), a severe distortion in perspective and self-occlusions in videos was caused [11].  
When it comes to choosing the range of distances, one must consider the dependence that exists on the resolution of the camera. In addition, we would like to consider that since the participants do not know the nature of the experiment before the drone will initiate an interaction- we must consider keeping a considerable distance. When analyzing the distances in the various datasets, it seems that the accepted range is between 3-15 meters.  
  
Video processing and uses of dataset

A dataset is a collection of [data](https://en.wikipedia.org/wiki/Data), the term "data" refers to a collection of facts (numbers, words, video, etc) that have been brought together in a form that computers can utilize. We can generally say that, in computing fields, data refers to information that is machine-readable rather than human-readable [45] .Having collected this information and made it machine-readable, we can analyze it and make models based on it to improve services, design public policies, boost businesses, advance science, make better informed decisions, and a whole lot more [8,27].

Particularly video and image datasets are used to build systems of human activity and action recognition. Several applications such as visual surveillance, video retrieval and human-computer interaction, among others, depend on the ability to recognize actions. The detection of human activity relies on a series of previously accomplished tasks- such as image capture, segmentation, tracking, identification, and classification. [2,19]

Today we are even able to build models of machine learning [21], deep learning [23], and computer vision from these datasets [30].  Machine learning is a branch of artificial intelligence (AI) focusing on developing programs that learn from data and improve their accuracy over time without having to be programmed to do so. The [algorithms](https://whatis.techtarget.com/definition/algorithm) based on machine learning use historical data as input to predict new output values. These algorithms power many of the services we use today, recommendation systems like Netflix, YouTube, and Spotify, search engines like Google and Baidu, social-media feeds such as Facebook and Twitter, voice-assistants like Siri and Alexa and more [39,46].

Deep learning is a subset of machine learning that constructs artificial [neural networks](https://deepai.org/machine-learning-glossary-and-terms/neural-network) to mimic the structure and function of the human brain. Both machine learning and deep learning algorithms have transformed the field of computer vision in the last years.   
The field of computer vision focuses on the design of computer systems that can recognize, understand and interpret image and video data. For example it enables us to use facial recognition technology, to build Autonomous vehicles in the future, diagnose an X-ray or mammography and the list goes on [31,44,47] . As technology in this field becomes more advanced and as drones become more advanced, the combination of these developments can lead to some great things. It is not hard to imagine that a drone could accurately detect faces and use them for security, recognize human behaviour and act accordingly and more. All of this requires good datasets.   
  
We believe that our dataset enables the first step in detecting human activity and identifying actions from people who encounter a drone. In addition, it can be used to study bystander drone interactions and design models based on them.

**Data collection – Protocol (Planning)**

Introduction

In our project, we will be building a dataset of videos, taken by drones, that documents different encounters with by-passers. In the data collection stage, the drone will initiate an interaction with a participant in a public area.   
The purpose of the dataset creation is to be the first step in a Supervised Machine Learning process, that will eventually improve humans and autonomous-drones’ interactions.

Thus, we made this Protocol for better understanding of our experiment procedure and ensuring high-quality deliverables.

Experiment method

We have researched for experiment methods that will help us document a sincere and spontaneous reaction to a drone. In our research, we have found the “Ghost Driver” experiment.  
The experiment, that was conducted by Stanford University researches, has used deception to conceal a driver of the car -so it will appear to be a full autonomous system.

The research documented the reactions of pedestrians that encountered the car (which took place at a crossroad). The Data included video recordings and participant responses to post-interaction questionnaires. This method is beneficial for safely acquiring empirical data.

We have decided to follow the same protocol.

Data collection Goals

1. Building a dataset of drone videos showing encounters between people and drones – from the drone point of view.
2. After collecting the data and getting the participant’s approval – labeling the encounters.
3. Determine the optimal settings (Angle, Distance, etc.) for encounter documentation and classification.

**Participants**: Non-bias by-passers in public areas.

**Requirements:** To understand the specific requirements we need to apply in the dataset, we have analyzed other databases with similar purposes. After extracting their specifications, we concluded the appropriate requirements for our experimen

1. Distance: 1-5m height, 2-15m horizontal range.
2. Angle: 0-45 degrees.
3. Lighting: daylight, sunny day.
4. Shooting mode: stable and horizontal filming.
5. Type of documentation: video.
6. camera resolution: Aspect Ratio (16:9), resolution 720p (HD).

**The product:** DJI MAVIC AIR 2 - Off the shelf product.   
Technical details:

1. Weight: Approximately 570 g (Propellers and Battery Included)
2. Dimensions: Folded: 180×97×74 mm Unfolded: 183×253×77 mm
3. Max Flight Distance: 18.5 Km
4. Max Speed: 19m/s
5. Max Flight Height: 5Km
6. Video resolution: 4K/60fps

**Procedures:** we will follow the *ISRAEL DRONE REGULATIONS* to ensure safety process and law obedience. Therefore, our procedure will follow:

1. The drone operator will maintain a direct eyesight channel with the drone while flying.
2. The drone will not fly above 50 meters (164 feet) from the ground.
3. The drone will not fly within 2 kilometers (1.2 miles) of any airport or airfield.
4. The drone will not fly in No-Fly Zones.
5. Drones may not fly within 250 meters (820 feet) of Populated places and buildings\*.
6. The drone will not fly during night-time (allowed during daytime only).

\*Most of the regulation apply to large drones that are over 250g. We have chosen to use a relatively small drone that, as mentioned, weighs only 80g.  
In addition, the drone will only be flown after being equipped with propeller shields.  
Both contributes to the safety of by-passers EX GRATIA.

**Experiment process**

Before arriving

1. Preparation of the required equipment-

A. Verify a proper and complete drone, with a 4-blade propeller and 4 propeller shields.

B. Fully charge the 3 batteries.

C. Print 50 participation consent forms (Appears in.

D. Print 50 participation questionnaire forms.

2. Choosing a location, based on -

A. Average number of passers-by at the spot.

B. Properly area sterility (according the law).

C. Ensuring that the location gets enough natural lighting (for the quality of the documentation).

D. Avoiding trees or other obstacles that can limit the drone’s movement.

3. Check the weather and the forecast – focus on relevant attributes as rain, cloudiness, wind, etc.

4. Assign 2 staff members to carry out the task.

Post arrival

1. Division into positions -

A. Team member 1 - The drone pilot. Located on the side of the street in a hidden place – yet remaining a direct eyesight with the drone. Responsible for safety flying and filming.

B. Team member 2 – Experiment leader. Located close to the designated filming area. Responsible for the safety of the event, gives the order to start and end the flight, and interact with the participant afterwards.

2. Beginning of flight + filming. For each participant the actions will be:

A. The drone takes off at about 30 meters from the participant.

B. Beginning of recording.

C. Locate the drone in pre-agreed height.

D. Start Flying the drone towards the targeted participant, the speed will not exceed 5 km/h.

E. Flying the drone over the participant. Team member 2 is located perpendicular to the drone and participant meeting place and ensures that there is no safety hazard to the participant or the drone. If he detects a danger, he intervenes the event and reports to crew member 1 to move the drone away immediately.

F. Completion of the recording. Landing of the drone slowly in an open space away from the participant.

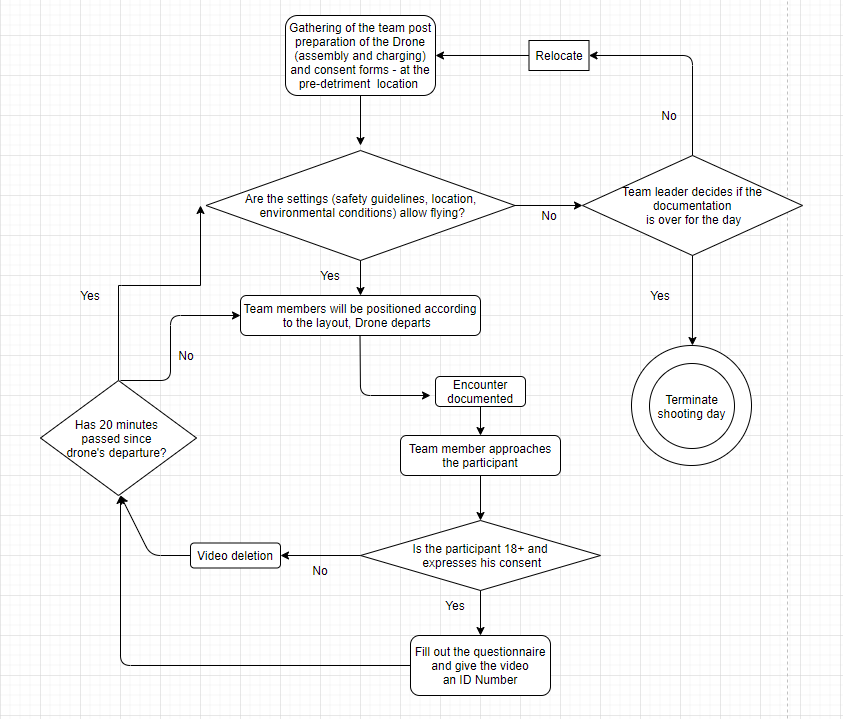
G. Team member 2 interacts with the participant. He provides an explanation of the research and asks for his consent to use the video for an academic research.

G.1. If the participant **refuses** - deletes the video in front of the participant.

G.2. If the participant **confirms** – The encounter gets an ID number. Filling the form with the participant details, ask him to sign the form, offering him to answer a short experience questionnaire.

H. End of session. Bringing the drone to the starting point, switching battery if needed, and performing another session.

3. At the end of each day, Team member 1 uploads the videos to the designated folder on Google Drive. The video will be given the title that will associate it with the encounter ID number – for future process.



*1.1) Experiment process chart*

**Safety guidelines**

In our project we will use DJI MAVIC Air 2 drone to collect data (videos) for the goal of creating a dataset. We are fully aware of the complexity of the drone flying action; therefore, we have determined precaution and safety guidelines.  
In order to minimize the existing risk, we will take into account environmental settings, population density and rules that we will enforce during flight and filming.  
We will present below the safety guidelines that we will follow in the dataset collection, the guidelines were gathered from websites discussing flying drones and in accordance to our experiment requirements:

* The drone will be flown only for the research purposes.
* We will not fly the drone above any restricted areas or in airports area.
* We will not fly our drone in an area where there is another drone being used.
* We will not fly the drone above groups of people – but in sterile area with a small presence of by-passers.
* We will not fly the drone next to emergency incident.
* The drone will be used only at day light where we could keep a constant eyesight.
* The drone will not exceed 8 meters height.
* The vertical distance from the participant will be in the range 3-30 meters.
* The drone will not be closer than 3 meters to the participant at any time – and will not fly above him/her. Maximum flight speed of the drone will be 5 km per hour.
* Before each flight, we will perform drone- systems operation, which includes verifying correct parts, and an initial flight to check the correctness of the drone before starting to collect the data.
* During the entire flight time, we will be at least 2 operators in place, with Operator 1 being the pilot and Operator 2 being an observer near the filmed participant who will secure the event closely. Both operators will be in direct eyesight with the drone and in any case of concern / danger, the pilot will move the drone away and the observer will cling to the participant to minimize the risk.
* When we will be a team of 3 - one person will secure one person will fly and one person will approach with the letter of consent.

Formal chronological steps:

A. The drone takes off at about 30 meters from the participant.

B. Beginning of recording.

C. Locate the drone in pre-agreed height.

D. Start Flying the drone towards the targeted participant, the speed will not exceed 5 km/h.

E. Flying the drone over the participant. Team member 2 is located perpendicular to the drone and participant meeting place and ensures that there is no safety hazard to the participant or the drone. If he detects a danger, he intervenes the event and reports to crew member 1 to move the drone away immediately.

F. Completion of the recording. Landing of the drone slowly in an open space away from the participant.

**Systems used:**

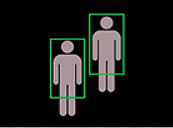
For our project we used:

* Drone –DJI MAVIC Air 2



*2.1 Figure 1 - DJI MAVIC Air 2*

* Image Annotation Lab 2.2.0



*2.2 Figure 2 – Image Annotation Lab Symbol*

* Windows 10 -

*2.3 Figure 3 – Windows 10 Symbol*

* Microsoft Excel -



*2.4 Figure 4 – Excel Symbol*

**Experiment process**

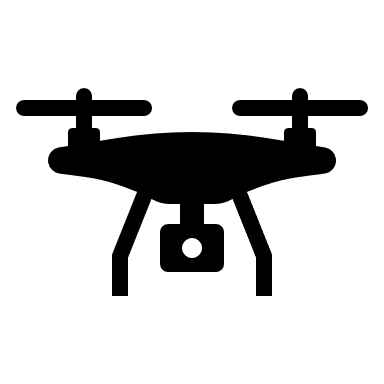
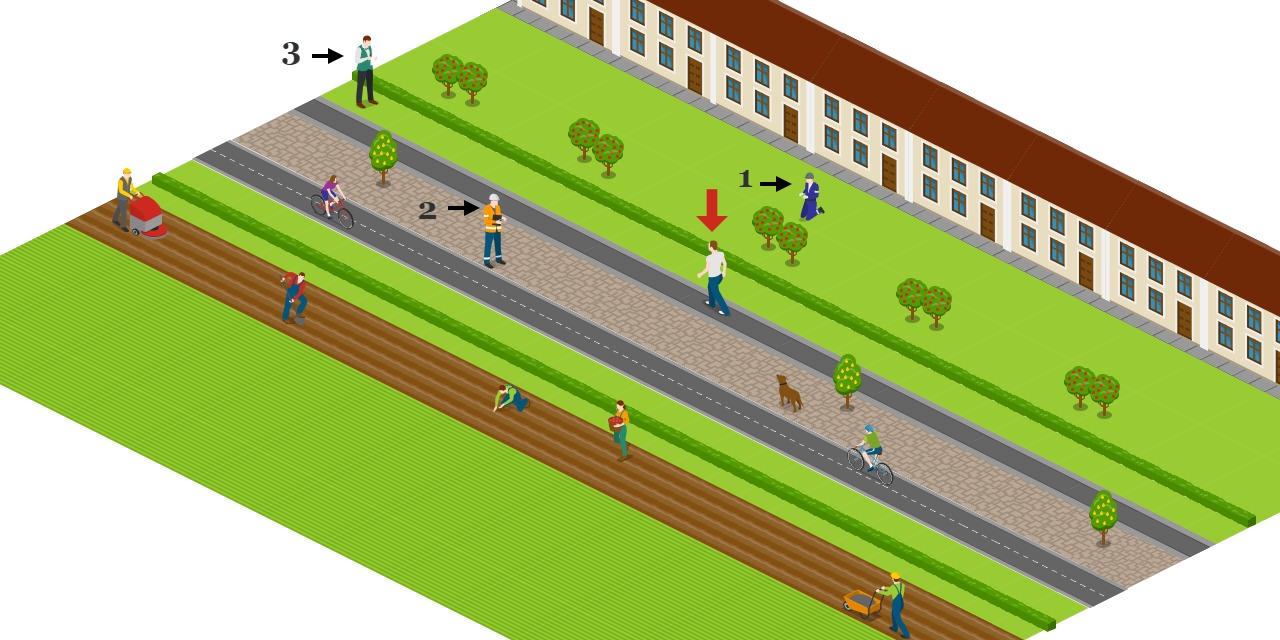
Before arriving, we had to decide of a good location for us. The decision was based on few conditions like the predict average number of passersby at the spot, the area sterility, a location with good enough lighting and avoiding trees or other obstacles that can limit the drone’s movement or sight, a good weather and forecast, and at least 2 staff members to carry out the task.

After it, we verify the needed equipment which includes a proper and complete drone with fully charged 3 batteries, and enough participation consent forms and participation questionnaire forms.

Once we arrived, we divided into 3 positions: (examples in our illustration below)

**1.** The drone pilot - Located on the side of the street in a hidden place (remaining a direct eyesight with the drone)  
**2.**  safety supervisor - Located close to the designated filming area – ready to act in case of safety issue.

**3.** The questioner - Reaches the person after the encounter has ended, asks for the person's consent to use the video, and if agrees he asks to answer a short questionnaire.



*3.1 Illustration of Team's Layout*

As written in the protocol, the drone takes off at about 30 meters from the participant. The drone pilot locates the drone in pre-agreed height and starts recording. The passersby walking towards, and team member 2 is located perpendicular to the drone and participant meeting place and ensures that there is no safety hazard to the participant or the drone. If he detects a danger, he intervenes in the event and reports to crew member 1 to move the drone away immediately. When the passersby approach 2 meters from the drone, the drone pilot flies it away, stops the recording and Lands the drone slowly in an open space, away from the participant.

After that, we turned to the subject, and first we started each of the interviews by providing an explanation of the research and asking for the participants consent to be a part of our academic research, and if not, the video would be deleted in front of him immediately. To our delight, most of the participants agreed that we would use the videos, and we asked them to fill the form with the participant details, sign the form and fill a short questionnaire about their experience. By that we complete the session.

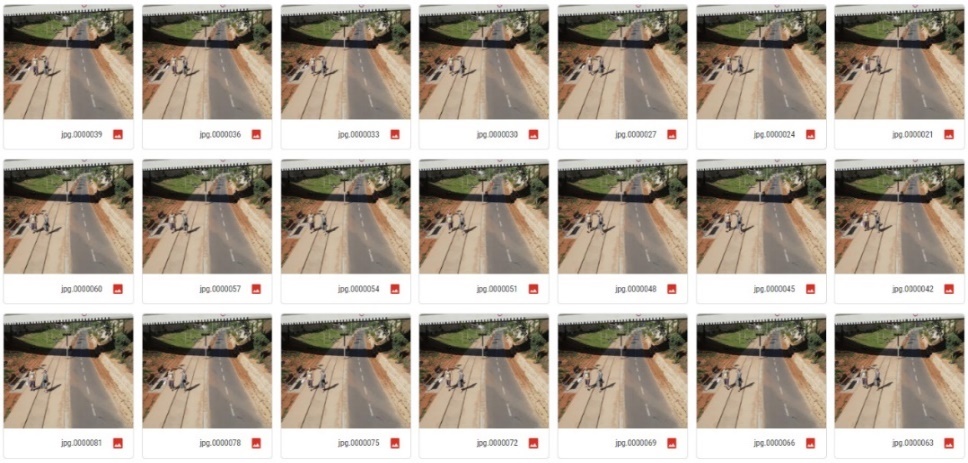
After that, we bring the drone to the starting point, switching batteries if needed, and performing another session. At the end of each day, we uploaded the videos to the designated folder on Google Drive.

Annotation process

The Annotation process was performed after we had filmed the bystanders and had a data base which contained videos of interaction between bystanders and the Drone. Image Annotation Lab was used to perform the annotation process, which was divided into three steps:

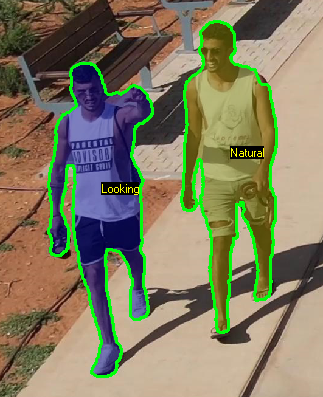
**Edit the videos** - The first step in the annotation process was to edit the video so that it contained only the times when the subject was close enough to the drone and when he or she was being filmed. Based on 67 videos, which includes the whole interaction, we edited the videos in such a way that the videos were between 3-20 seconds long, with an average of 7.6 seconds per video. Microsoft Films & TV software was used for this process.

**Divide the videos into frames**   - Each frame of video must be as precisely annotated as an individual image. So, after we had edited the videos and had a high-



*4.1) Frame's example image*

quality dataset of edited videos, the first step in the Image Annotation Lab software was to divide the edited videos into frames. Each video contained between 30- 250 frames that will be marked and labeled in the next stages.



*4.2 ,4.3 - Example of a video after divided into frames*.

**Mark the relevant object and label-** For identifying the relevant object in the frame there are several techniques. Because of the complexity of the human body and our interest in finding out every little change in the subject body, we decided to mark the object with polygon technique.  In this technique, objects are marked by plotting points on each vertex of the target object. It allows all the object's exact edges to be annotated, regardless of its shape. The process was conducted by AI technique and had to be done for each frame. After all the relevant objects in the frame had been marked in a polygon technique the final stage in the annotation process is to label each polygon in each frame with one of 4 labeling types, positive, negative, neutral, and looking.

**Analysis**

**Questionnaire** **Analysis**

After documenting the interactions and sifting through the non-approved encounters, we linked the videos to the forms using the tracking table. To ensure backup and to ease the numeric analysis, we transferred the data from the questionnaires to a tabular document in which we added the data according to the video ID.

We added comments in videos that would seem relevant to people who would like to work with the database in the future - but we made sure to document the rest of the data accurately so we will not create bias.

**Data Analysis**

The data we analyzed was taken by two main sources:  
1. The data that was collected by the questions answered in the questionnaires. It contains both personal details of the subject and information about his experience.   
2. The second source of data comes from the annotation process. After we edited and labeled the frames in the videos, we were able to extract data from them.

From the questionnaire data, we examined our gender and age. 48% of our participants has identified themselves as Males and 52% as Females.   
The age distribution seems like a normal distribution with a right tail, most of the participants were between 20 - 40 years old.

*5.1) Age distribution bar chart.* In addition, to better understand the people's reaction, we have asked them a few questions about the feelings they have had whilst the encounter.   
To the questions if they ever encountered a drone before and if they noticed the drone during the interaction, the answers were that 94% and 71% - respectively.

In addition, we asked the participants what they were feeling when they encountered the drone. They have had 6 possibilities of emotions to choose from, the options were based on the psychologist Paul Eckman - six basic emotions: fear, anger, joy, sadness, disgust, and surprise. Although most of the people stated that they have encountered a drone before, "surprise" was the most common feeling during the interactions. After "surprise", the other common feelings were "anger" and "joy".

*5.2) People's reaction distribution bar chart*

From the data we have extracted from the videos we have found that in most of the encounter the passersby do not interact with the drone. Most of the labels were classified as “neutral” and the eye contact is not long.  
Another finding was that most people did not created gestures to the presence of the drones. People that stated that they have felt “anger” during the encounter (for an example) – did not express that not with words or hand gestures and not with other violent/extreme measure.

From the annotation process, we extracted all the data from the JSON files and we examined the distribution of the labeling. As we experienced during the experiment, neutral was the most common reaction to the drone. In the labels, 62% of the frames were tagged as "Neutral", followed by 21% of "Looking", "Positive" with 17% and negative with only 0.4% of the frames.

*5.3) Labeling distribution bar chart*

**Our Dataset**

After we had filmed all the videos and completed the annotation process, we organized the dataset. The dataset includes 5 different items:

**The questionnaires** - we have 67 questionnaires containing raw data of the responses received from the study participants.

**The videos of the human drone interaction** - 67 full-length videos that document the entire interaction between passers-by and the drone.

**The videos after we edited them** - 67 videos contained only the times when the subject was close enough to the drone and when he or she was being filmed. Video lengths ranged from 3 to 20 seconds, with an average of 7.6 seconds.

**Folder for each video which contains the frames** - 67 folders in which each folder contains the entire frames of that video. every folder contains between 30 to 250 frames.

**Json files** - JavaScript Object Notation (JSON) is a standard text-based format for representing structured data based on JavaScript object syntax. It is a simple code that is easily translated to human-readable text, containing pairs of attribute-values and arrays. For each Video, we created a JSON file that included the path to each frame, the coordinate of the polygon, and the labeling. In total we have in our dataset 67 Json files.

*6.1) Dataset example image*

**Discussion**

Possible limitations of our dataset:

Although our dataset is based on random participants that were passing by in the various locations – there are some “blind spots” that we have identified:

**Location’s demographic** – The experiment was conducted in Tel Aviv, Israel. It is a large city in the heart of the metropolitan area of ​​Israel's coastal plain. The city is considered progressive and liberal and may represent a population that is more familiar and friendly to new technology and its integration in daily life.

**Age Restriction** - Although there are several videos taken of minors, alongside a parent who has expressed his consent to use the video - we refrained from documenting passers-by who seemed to us younger than 18 (the age of consent in Israel). This population has a unique behavior that could be expressed in its reaction to the drones, not having such documentation is undoubtedly a blind spot in our database.

**Number of people in video** – To assist with the annotation stage of our dataset creation process, we have chosen locations that would have a fair number of pedestrians – but not too busy. Eventually, drones are and will be active in some locations that could be busier. We assume that people also may react different when alone - than when in a crowded place.

**Pets** – When planning the experiment’s methodology and process – we have not taken into consideration the presence of pets in the encounters. Many of the videos contain pets (especially dogs) that react to the drone and affect the passerby’s reaction to the drone.

**Conclusions:**

When studying the field and preparing our experiment, we have learnt about the different regulations regarding flying drones in populated areas. The main goal of the regulations and rules is to enable the use of the technology – while trying to limit safety and privacy issues that might occur.

The integration process of drones was noticed in the data we have collected. Although most of the participants have stated that they have encountered a drone in the past (94%) – the leading emotion reaction to our experiment was “Surprise” (over half of the participants). In addition, people who said they felt “Anger” during the experiment, have not acted in an extreme/violent way towards the drone.

Our project’s goal was to assist the integration process. With the creation of the dataset of annotated videos and the questionnaires, we have built the basis on which the technology could examine the passersby behavior while using Machine Learning algorithms. In the future, on top of the general regulations, autonomous drones will be able to be sensitive to the unique encounter and the passersby reaction.

**Summary and future work:**

We have successfully created a data set of videos documenting spontaneous encounters between passersby and a drone – from the drone’s point of view.

The videos are edited and labeled, and questionnaires were added to thicken the data regarding the encounters. The dataset contains a video documentation of over 60 sessions – in which every frame is annotated.

In the future, additional researchers could use the method we have created to collect similar videos. Furthermore, researchers can create unique models based on Machine Learning algorithms and apply them on the dataset.

***References***

**Literature review references**

[1]Huang, H. M., Pavek, K., Novak, B., Albus, J., & Messin, E. (2005). A framework for autonomy levels for unmanned systems (ALFUS). *Proceedings of the AUVSI’s Unmanned Systems North America*, 849-863.‏

[2]Chaquet, J. M., Carmona, E. J., & Fernández-Caballero, A. (2013). A survey of video datasets for human action and activity recognition. *Computer Vision and Image Understanding*, *117*(6), 633-659.‏

[3] Piana, S., Stagliano, A., Camurri, A., & Odone, F. (2013). A set of full-body movement features for emotion recognition to help children affected by autism spectrum condition. In *IDGEI International Workshop*.‏

[4] Shan, C., Gong, S., & McOwan, P. W. (2007, September). Beyond Facial Expressions: Learning Human Emotion from Body Gestures. In *BMVC* (pp. 1-10).‏

[5]Cowie, R. (2008, September). Building the databases needed to understand rich, spontaneous human behaviour. In *2008 8th IEEE International Conference on Automatic Face & Gesture Recognition* (pp. 1-6). IEEE.‏

[6]Cohn, P., Green, A., Langstaff, M., & Roller, M. (2017). Commercial drones are here: The future of unmanned aerial systems. *McKinsey & Company*.‏

[7] Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep learning for computer vision: A brief review. *Computational intelligence and neuroscience*, *2018*.‏

[8]Aminanto, E., & Kim, K. (2016). Deep learning in intrusion detection system: An overview. In *2016 International Research Conference on Engineering and Technology (2016 IRCET)*. Higher Education Forum.‏

[9] Hsu, H. J., & Chen, K. T. (2017, June). DroneFace: an open dataset for drone research. In *Proceedings of the 8th ACM on multimedia systems conference* (pp. 187-192).‏

[10]Cauchard, J. R., E, J. L., Zhai, K. Y., & Landay, J. A. (2015, September). Drone & me: an exploration into natural human-drone interaction. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing* (pp. 361-365).‏

[11] Abtahi, P., Zhao, D. Y., E, J. L., & Landay, J. A. (2017). Drone near me: Exploring touch-based human-drone interaction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *1*(3), 1-8.‏

[12]kalra, I., Singh, M., Nagpal, S., Singh, R., Vatsa, M., & Sujit, P. B. (2019, May). Dronesurf: Benchmark dataset for drone-based face recognition. In *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)* (pp. 1-7). IEEE.‏

[13] Cauchard, J. R., Zhai, K. Y., Spadafora, M., & Landay, J. A. (2016, March). Emotion encoding in human-drone interaction. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 263-270). IEEE.‏

[14] Yeh, A., Ratsamee, P., Kiyokawa, K., Uranishi, Y., Mashita, T., Takemura, H., ... & Obaid, M. (2017, October). Exploring proxemics for human-drone interaction. In *Proceedings of the 5th international conference on human agent interaction* (pp. 81-88).‏

[15]Bassi, E. (2020). From Here to 2023: Civil Drones Operations and the Setting of New Legal Rules for the European Single Sky. *Journal of Intelligent & Robotic Systems*, 1-11.‏

[16]Giones, F., & Brem, A. (2017). From toys to tools: The co-evolution of technological and entrepreneurial developments in the drone industry. *Business Horizons*, *60*(6), 875-884.‏

[17]Funk, M. (2018). Human-drone interaction: let's get ready for flying user interfaces!. Interactions, 25(3), 78-81.‏ HUMAN-DRONE INTERACTION: LET’S GET READY FOR FLYING USER INTERFACES

[18]Mirri, S., Prandi, C., & Salomoni, P. (2019, August). Human-Drone Interaction: state of the art, open issues and challenges. In *Proceedings of the ACM SIGCOMM 2019 Workshop on Mobile AirGround Edge Computing, Systems, Networks, and Applications* (pp. 43-48).‏

[19]Aggarwal, J. K., & Cai, Q. (1999). Human motion analysis: A review. *Computer vision and image understanding*, *73*(3), 428-440.‏

[20]  Goodrich, M. A., & Schultz, A. C. (2008). *Human-robot interaction: a survey*. Now Publishers Inc.‏

[21]Alpaydin, E. (2020). *Introduction to machine learning*. MIT press.‏

 [22]Jensen, W., Hansen, S., & Knoche, H. (2018, April). Knowing you, seeing me: investigating user preferences in Drone-Human acknowledgement. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-12).‏

[23] Yu, F., Seff, A., Zhang, Y., Song, S., Funkhouser, T., & Xiao, J. (2015). Lsun: Construction of a large-scale image dataset using deep learning with humans in the loop. *arXiv preprint arXiv:1506.03365*.‏

[24]Grissom, C. K., Thomas, F., & James, B. (2006). Medical helicopters in wilderness search and rescue operations. *Air medical journal*, *25*(1), 18-25.‏

[25]Barekatain, Mohammadamin, et al. "Okutama-action: An aerial view video dataset for concurrent human action detection." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2017.‏

[26] Otto, A., Agatz, N., Campbell, J., Golden, B., & Pesch, E. (2018). Optimization approaches for civil applications of unmanned aerial vehicles (UAVs) or aerial drones: A survey. *Networks*, *72*(4), 411-458.‏

[27] Ingre, B., & Yadav, A. (2015, January). Performance analysis of NSL-KDD dataset using ANN. In *2015 international conference on signal processing and communication engineering systems* (pp. 92-96). IEEE.‏

[28]Bonetto, M., Korshunov, P., Ramponi, G., & Ebrahimi, T. (2015, May). Privacy in mini-drone based video surveillance. In *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)* (Vol. 4, pp. 1-6). IEEE.‏

[29]Chen, G. Y. (2016). Reforming the Current Regulatory Framework for Commercial Drones: Retaining American Businesses' Competitive Advantage in the Global Economy. *Nw. J. Int'l L. & Bus.*, *37*, 513.‏

[30]Karaca, Y., Cicek, M., Tatli, O., Sahin, A., Pasli, S., Beser, M. F., & Turedi, S. (2018). The potential use of unmanned aircraft systems (drones) in mountain search and rescue operations. *The American journal of emergency medicine*, *36*(4), 583-588.

[31]Rao, B., Gopi, A. G., & Maione, R. (2016). The societal impact of commercial drones. *Technology in Society*, *45*, 83-90.‏

[32] Yang, D., Li, L., Redmill, K., & Özgüner, Ü. (2019, June). Top-view trajectories: A pedestrian dataset of vehicle-crowd interaction from controlled experiments and crowded campus. In *2019 IEEE Intelligent Vehicles Symposium (IV)* (pp. 899-904). IEEE.‏

[33]  Beer, J. M., Fisk, A. D., & Rogers, W. A. (2014). Toward a framework for levels of robot autonomy in human-robot interaction. *Journal of human-robot interaction*, *3*(2), 74.‏

[34]Clarke, R. (2014). Understanding the drone epidemic. *Computer Law & Security Review*, *30*(3), 230-246.‏

[35] Perera, A. G., Wei Law, Y., & Chahl, J. (2018). UAV-GESTURE: A dataset for UAV control and gesture recognition. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 0-0).‏

 [36]Murphy, R. R., & Burke, J. L. (2005, September). Up from the rubble: Lessons learned about HRI from search and rescue. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 49, No. 3, pp. 437-441). Sage CA: Los Angeles, CA: SAGE Publications.‏

[37]Nievas, E. B., Suarez, O. D., García, G. B., & Sukthankar, R. (2011, August). Violence detection in video using computer vision techniques. In *International conference on Computer analysis of images and patterns* (pp. 332-339). Springer, Berlin, Heidelberg.‏

[38]Zhu, P., Wen, L., Du, D., Bian, X., Ling, H., Hu, Q., ... & Ma, W. (2018). Visdrone-det 2018: The vision meets drone object detection in image challenge results. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 0-0).‏

[39] What is data, and why is it important? June 28, 2018 [Import.io](https://www.import.io/post/author/import-io/)

[40]- Vikram Puri, Anand Nayyar & Linesh Raja (2017) Agriculture drones: A modern breakthrough in precision agriculture, Journal of Statistics and Management Systems, 20:4, 507-518, DOI: 10.1080/09720510.2017.1395171

[41]- George Bekey (Panel chair) Robert Ambrose Vijay Kumar Art Sanderson Brian Wilcox Yuan Zheng INTERNATIONAL ASSESSMENT OF RESEARCH AND DEVELOPMENT IN ROBOTICS, World Technology Evaluation Center, Inc.

[42] - Anke Brock, Jessica Cauchard, Markus Funk, Jérémie Garcia, Mohamed Khamis, et al.. iHDI: First International Workshop on Human-Drone Interaction. CHI 2019, Conference on Human Factors in Computing Systems Extended Abstracts, May 2019, Glasgow, United Kingdom. ACM, pp.art. no. 3299001., 2019, CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI’19 Extended Abstracts), ff10.1145/3290607.3299001ff. ffhal-01990121f

[43] - Silverstein, D. Amnon, Shun Wai Go, Suk Hwan Lim, Timothy J. Millet, Ting Chen, and Bin Ni. "Modeless video and still frame capture using interleaved frames of video and still resolutions." U.S. Patent 9,344,626, issued May 17, 2016.

[44] [Martin Heller](https://www.infoworld.com/author/Martin-Heller/), What is deep learning? Algorithms that mimic the human brain, InfoWorld.com,MAY 24, 2019

[45] What is Machine Learning? A Definition, Expert System Team, MAY 6,2020

[46] Karen Hao, What is machine learning?, MIT Technology Review, November 17, 2018

[47] [Bernard Marr](https://www.forbes.com/sites/bernardmarr/), 7 Amazing Examples Of Computer And Machine Vision In Practice, Forbes, April 8, 2019

**References - Protocol:**

[1] Rothenbücher, Dirk, Jamy Li, David Sirkin, Brian Mok, and Wendy Ju. "Ghost driver: A field study investigating the interaction between pedestrians and driverless vehicles." In 2016 25th IEEE international symposium on robot and human interactive communication (RO-MAN), pp. 795-802. IEEE, 2016.

[2] Israel Drone Organization - www.icd.org.il/

**Appendix**

**טופס הסכמה להשתתפות במחקר**

מטרת המחקר - מטרת המחקר היא לבחון את דפוסי ההתנהגות של עוברי אורח בעת המפגש עם רחפן, בכדי להבין כיצד עוברי אורח מגיבים למפגש זה ולמצוא פתרונות להטסת רחפנים באזורים מאוכלסים. טופס זה נכתב בלשון זכר אך מיועד לשני המינים.

במסגרת מחקר שנערך על ידי ד"ר ג'סיקה קושארד מאוניברסיטת בן גוריון בנגב, אנו מנסים להבין את תגובת עוברי האורח למפגש עם רחפן במרחב הציבורי. בניסוי תיעדנו ע"י מצלמת הרחפן את תגובתך למפגש עם הרחפן, השימוש בתיעוד יהיה לצורך מחקר על אינטראקציה בין אדם לרחפנים.  אנו מבקשים ממך לענות על מספר שאלות אשר יסייעו לנו להבין טוב יותר את דפוסי ההתנהגות שלך בעת המפגש עם הרחפן. במידה ותסכים לקחת חלק נשמח אם תחתום על טופס ההסכמה הנ"ל.

כל המידע והשאלות שיאספו לא יכללו אף פרט אישי שיכול לזהות אותך. זהותך תישאר אנונימית, מלבד הוידאו עצמו.

במידה ומכל סיבה שהיא תחוש אי נוחות, זכותך  לבחור לא לקחת חלק בפרויקט הניסוי.   
אנו מודים לך על השתתפותך בסקר, ומעריכים את עזרתך.

במידה ויש לך שאלות לגבי המחקר תוכל לפנות אלינו בכל זמן לכתובות המייל הבאות:

רועי יוגב:                           יהונתן להט:                     אלון קרמונה:

[royyoge@post.bgu.ac.il](mailto:royyoge@post.bgu.ac.il),  [lahatjo@post.bgu.ac.il](mailto:lahatjo@post.bgu.ac.il), alonkar@post.bgu.ac.il.

הצהרת המשתתף:

אני החתום/ה מטה מצהיר/ה כי קראתי את טופס ההסכמה במלואו, וכי בחתימתי אני מביע/ה את הסכמתי להשתתף בסקר. אני מסכים לכך שיעשה שימוש אקדמי בתיעוד שבו אני מופיע (כגון מחקר אקדמי, מצגות בכנסי מחקר ובעבודות מחקר).

**חתימת הנבדק: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**שעה: \_\_\_\_\_\_\_\_\_\_\_\_\_**

**מגדר: \_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**גיל: \_\_\_\_\_\_\_\_**

הצהרת המראיין:

אני החתום/ה מטה מצהיר/ה כי אני מחייב לשמור על סודיות מלאה. לא אעביר בשום צורה ואופן כל מידע שעולה בסקר.

**חתימת הנסיינ/ית: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**שאלון לנבדק**

1. גיל**:\_\_\_\_\_\_\_\_**
2. האם ראית את הרחפן מתקרב לכיוונך.

     כן                     לא

           במידה וכן, באיזה שלב ראית אותו?  **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

1. האם ראית רחפן בעבר.

     כן                     לא

במידה וכן, איפה?  **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

1. סמן (ניתן יותר מאחד) את התחושות אותן חשת בעת המפגש:

          פחד                  כעס              שמחה           עצב             גועל             הפתעה

5. תאר את תחושותיך בעת המפגש עם הרחפן בכמה מילים:

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Consent form for participation in the study**

Study goal - The goal of the study is to examine the behavioral patterns of passersby during the encounter with a drone, in order to understand how passersby react to this encounter and find solutions for flying drones in populated areas.

In a study, conducted by Dr. Jessica Cauchard, at Ben Gurion University of the Negev, we try to understand the reaction of passersby to an encounter with a drone in public spaces. In the experiment, we documented in a video camera located on the drone, your reaction to encountering with it. The use of the video feed will be for the purpose of researching the interaction between Humans and Drones. We would be like to sign the consent form if you agree to us using this video for research and educational purposes.

All information and questions collected will not include personal details that may identify you beyond the video itself.

If for any reason, you feel uncomfortable, you have the right to not participate in the experiment. We thank you for your participation in the survey and appreciate your help.

If you have any questions about the research, you can contact us at any time at the following email addresses:

Roy Yogev:                       Jonathan Lahat:                  Alon Karmona:

[royyoge@post.bgu.ac.il](mailto:royyoge@post.bgu.ac.il) , [lahatjo@post.bgu.ac.il](mailto:lahatjo@post.bgu.ac.il)  , [alonkar@post.bgu.ac.il](mailto:alonkar@post.bgu.ac.il) .

**Participant statement:**

I, the undersigned, declare that I have read the full consent form, and that in my signature I express my consent to participate in the survey. I agree that my image will be used for scientific dissemination, such as presentations in research conferences and research papers.

**Participant signature: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Time: \_\_\_\_\_\_\_\_\_\_**

**Gender: \_\_\_\_\_\_\_\_**

**Age: \_\_\_\_\_\_\_**

**Interviewer Statement:**

I, the undersigned, declare that I must maintain complete confidentiality. I will not in any way pass on any information that comes up in the survey.

**Interviewer signature: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Questionnaire**

1. Age:  \_\_\_\_\_\_\_\_
2. Did you notice the drone approaching you?

Yes                 No

If yes, at what point?  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Have you ever encountered a drone before?

Yes              No

If yes, where?  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Select the emotions that you felt whilst the encounter (one or more):

          Fear         Anger         Joy          Sadness         Disgust          Surprise

1. Describe the encounter with the drone in your own words:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Video number | Date | Time | Location | Description | Consent (X/V) |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

**Interactions Tracking**

**Ethic committee letter**

