Home security system based on drone automation - IoT approach

Darko Kostadinov
School of Computer Science and Information Technology
University American College Skopje
Skopje, Macedonia
kostadinov.darko91@gmail.com

Veno Pachovski
School of Computer Science and Information Technology
University American College Skopje
Skopje, Macedonia
pachovski@uacs.edu.mk

Irena Stojmenovska
School of Computer Science and Information Technology
University American College Skopje
Skopje, Macedonia
irena.stojmenovska@uacs.edu.mk

***Abstract*—Drones have already established a significant role in many industries. As technology progresses, in the next few years, drones will be considered useful tools for a variety of consumers. The paper aims to show the benefits and advantages of using a drone as a dynamic security camera (strengthened by using image recognition software) for home security purposes. Two applications were created as proof of concept to present the idea. This approach can be beneficial not only for home security, but also for other industries such as: cinematography, agriculture, energy, civil engineering and constructions, safety, security in general, governmental use, oil refining and other environmental sciences.**

Keywords— Home security, IoT, drone, image recognition, web sockets, tensorflow.js, AWS, Node.js

I. Introduction

Internet is one of the most important things ever created which affected humanity and the modern lifestyle. Internet of things, or IoT, is a system of connected computing devices that have the ability to transfer data over a network without the need for human interaction [1]. They are smart devices that are using built-in sensors or other embedded systems to collect and send data that they have acquired from the environment around them [2]. In the last couple of years, the hype around the Internet of things has been constantly rising and will continue to do so. The number of connected devices installed worldwide is increasing and that trend won't stop any time soon. The total installed base of IoT connected devices is projected to amount to 75.44 billion worldwide by 2025, which is a fivefold increase in ten years [3]. The increasing number of connected devices helped the global market value for IoT to rise continuously. A report done by Fortune Business Insights, shows that the overall market for IoT, which was valued at $190 billion in 2018, is expected to reach $1,111.3 billion by 2026 at a yearly improvement rate of 24.7%. The banking and financial industries are expected to occupy the biggest part of that sum [4].

Drones have already been used for various purposes. However, the growth and variety of their usage will continue to rise in the following years; there will be increase in productivity, safety, wellbeing, and even new job openings for the drone industry. Some of other advantages are reduced manual labor and lower risk (since there will be no need for locating labor in dangerous areas), as well as getting a more extensive field of view that drones provide. Nowadays, drones are used not only for military purposes. They also made a big leap in the consumer market; forecast for the end of 2020 predicts $100 billion market value for drones which is generated by the increased demand from the commercial and civil government sectors. Although the defense sector will have the biggest share of the market, with a projected $70 billion by the end of 2020, consumer market is expected to be the runner up with $17 billion in market share, while the commercial/civil usage projection is evaluated to be $13 billion [5].

Home security plays an important role in our lives and in people's wellbeing in general. Modern era allows us to have UAV patrolling the house and letting us know of any inconsistencies and intruders. The idea of this paper is to give a contribution in that direction. The approach presented further on, relies on two custom applications that communicate with each other using web sockets. The drone app is communicating with the drone hardware and depending on the video stream and the data provided by the drone, makes decisions such as autopilot, auto-land, object detection, face detection, etc. The video frames are analyzed using tensorflow.js, which is an open-source library for building, training, or running machine learning models entirely in the browser [6]. The library is providing coordinates and sizes of the recognized objects in the frame and a given command is executed. Also, the drone application is sending notifications to the second application which serves as a mobile app, so that the user/landlord can preview flights or be notified if needed. The data provided by the drone is saved on Amazon Web Service (AWS), in particular the DynamoDb service and the Simple Storage Service (S3). DynamoDb is a managed NoSQL service that has strong consistency and stores the data as key-value pairs. Also, it doesn't require complex manual setup [7]. The drone application uses S3 to store the video data from the whole flight. Amazon S3 serves as a storage that is designed for large-capacity at a lower price and can be easily provisioned to other geographical regions [8].

II. The Architecture of the Model

A. *System Overview*

The location of the research is a backyard with different objects added to create continuous changing scenery, and thus enriching the viewport of the image recognition software during the automated flights of the drone. The hardware used in the research is: Dell XPS 15 laptop serving the drone application hosted on localhost with a node/express.js application and also a mobile application which is presenting the data fetched from AWS DynamoDB and AWS S3 bucket. The actual drone hardware is a DJI Tello which is providing a wireless network that can be accessed using three ports: video stream port, control commands port and read commands port.

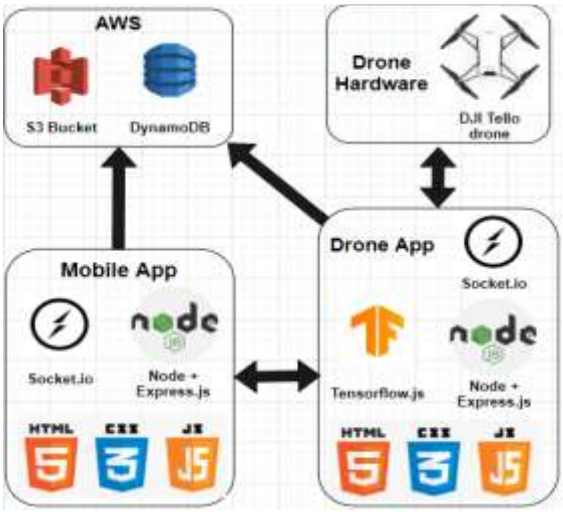

Fig. 1. Overview of software architecture

The drone SDK is providing various text commands that can be used to control the drone [9].

Drone Application

When the application is bootstrapped for the first time, the autopilot path is initialized with all the necessary variables. Then, the models from tensorflow.js are fetched including faceapi.js which is built as a wrapper from tensorflow.js to help with the face recognition feature in order to differentiate if the person is known or unknown. The recognized faces are first fetched from DynamoDB and later from an S3 bucket.

Next, a web socket connection is created with the drone; one for the video feed and another for read commands and control communication. The drone is flying on a predefined path around the house on autopilot, but when it detects an unknown person, it sends a notification to the user.

The landing is also dependent on image recognition. If the landing pad sign is in sight and the flight time is near the end, the drone starts adjusting for better accuracy before landing, so it can land on the same spot where it took off. The flight, with all the executed commands and detected faces during flight are saved on DynamoDB while the video of the flight is saved on the S3 bucket.

B. Mobile application

This application is used by the user/landlord and it has the features for previewing history of flights and familiar faces. Moreover, it has a feature for adding a new face which is directly updated and used instantaneously by the drone application.

III. RESULTS AND ANALYSES

A. Object recognition

Object recognition is one of the most important features of this application. Therefore, it requires a systematic approach. So far, there are three classes of 'familiar' objects, in which the stop sign is a class by itself. At present, there is only one image with that classification, but there could be much more, for reasons of security or obscurity.

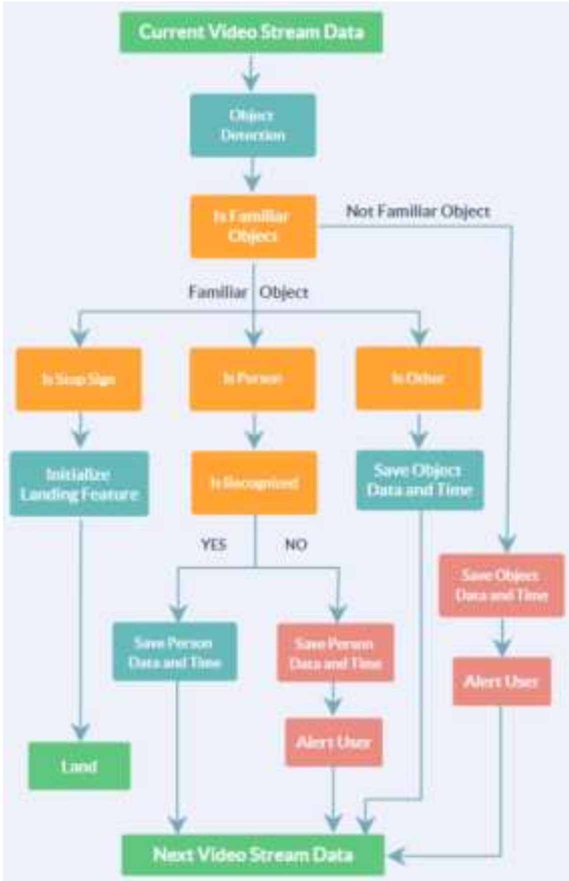

Fig. 3. Workflow of the image recognition

For the purposes of this research, the drone application is tested using dogs as objects. The results are based on the analysis of 100 screenshots taken from the drone camera - both moving and non-moving objects and 4 different dogs, all of them also from different breed. The research was done using the coco-ssd model which is capable of detecting 90 types of objects. It managed to recognize 73% of the objects provided as type dog. Each of the predictions has an appropriate confidence score, presented in the table below.

TABLE I. OBJECT RECOGNITION

Confidence Score	Dog
<0.6	4%
0.6-0.7	8%
0.7-0.8	15%
0.8-0.9	34%
0.9>	39%

Fig. 2. DOG images – an example of recognition

The efficiency of the system depends on the object current rotation or the drone perspective which is constantly changing, so in the next frame the recognition would get the correct object type. When tested in a video of 1 minute, while the object was in the viewport of the camera and constantly moving, it was correctly recognized in more than 90% of the time.

B. Face Recognition

In order to determine if a person is known or unknown, the faceapi.js library is used, built also with tensorflow.js which is based on the familiar faces fetched from AWS DynamoDB and S3 bucket. It waits until an object of type person is in the viewport and then activates the face recognition feature which tries to compare familiar faces with the one in the viewport. This was tested by having a person in front of the drone. It was successful in recognizing a person at distances of 1 and 3 meters, but it failed at 5 meters.

All of the data provided above is generated with the built-in camera of the drone which has a resolution of 720p. If equipped with a better camera, the results can be significantly improved; with a clearer image, it will be easier for tensorflow.js to predict objects.

C. Landing analyses

Since the drone does not have any GPS hardware, the landing feature is based on image recognition in order to land on the same point it took off. The data provided by tensorflow.js for each of the prediction boxes and its dimensions, are used to determine if any adjustment is needed (up, down, left, or right). Then the drone moves forward until it reaches the distance that is needed to position in front of the sign for landing, and, if everything is ok, it executes the landing.

Fig. 4. Drone in front of the landing image

The landing feature of the drone application has a result of 98% successful landings and in only 2%, where the drone missed the landing pad. The following heat map shows the points of (each of the) 50 landings done on landing pad with dimensions 70x70 centimeters:

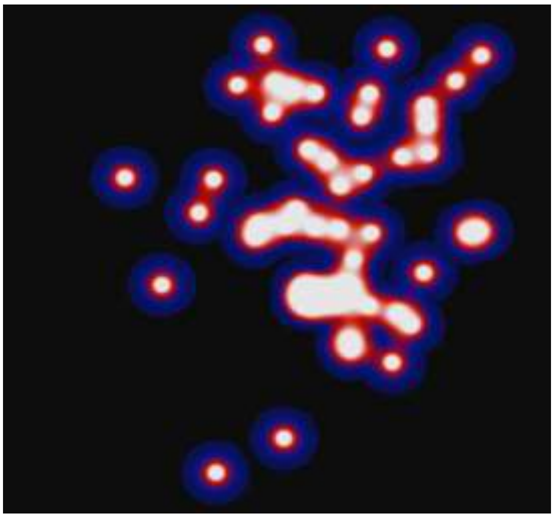

Fig. 5. Heatmap containing each of the landing points

With the current implementation of the drone application software and taking into consideration that the DJI Tello drone is a small one, when even light weather conditions can easily affect it, it still managed to execute more than 90% of the landing on a dimension of less than 50x50 centimeters

D. Aws Rekognition

AWS Rekognition is a visual analysis service from Amazon that uses deep neural network models, so it can detect objects, faces, and text in your images [10]. Since all of the results presented in Table I were done by using tensorflow.js, and in order to make a comparison between open source and proprietary software, the images were put to the test but this time using AWS Rekognition. It produced much better results and recognized 94% of dog type. The recognition results are presented in the table below.

TABLE II. AWS REKOGNITION

Confidence Score	**Dog**
<0.6	2%
0.6-0.7	3%
0.7-0.8	5%
0.8-0.9	24%
0.9>	66%

Compared with the confidence score (Table I), this test shows much higher confidence.

Again, as previously mentioned, with a better camera which has better image stabilization, the results would have been significantly improved (with a clearer image, either tensorflow.js or AWS Rekognition will be more precise at predicting objects).

Due to the drone camera capabilities, all of the image data tested was done during daylight. For night testing , there are more advanced drone cameras with night sight mode. where these methods can be applied and consequently, see what kind of results will they provide.

IV. ADDITIONAL FEATURES AND IMPROVEMENTS

The drone application also has other features such as voice recognition and speak functionalities, by using the built-in objects from the browser: SpeechRecognition and SpeechSynthesisUtterance. Currently, the drone has the feature of understanding English sentences and executing drone commands by a given user command, and also replying back in English for each of the commands. This could be helpful for industries where a user needs to exercise a voice control on the drone (by a set of predefined voice commands). Finally, text generation feature could be included; to generate a full report about the outcome of the patrol - flight path registered and/or images discovered.

In order to make the recognition more accurate, an implementation of AWS Rekognition can be added as a feature to the drone application, so that the user can have a more detailed analysis.

Next step in improving this home security system would be to add a wireless charging functionality. A charging coil could be added to the drone and a wireless charging pad to the landing location, so that the whole system would become autonomous.

V. Conclusion

Research presented in this paper was based on two custom web applications connected to AWS and a DJI Tello drone. The data that was provided by the drone video stream was analyzed by image recognition software on different video frames during the drone flights. Moreover, there have been numerous tests for the landing feature of the drone application, which used only image recognition to decide/detect the point of landing. Recorded videos and images were also run through AWS rekognition in order to compare open source and a proprietary image recognition software. Finally, possible improvements and ideas for further investigations were suggested.

References

[1] Margaret Rouse , "Internet of things", 2016, [Online] Available: https://internetofthingsagenda.techtarget.com/definition/Internet-of-Things-IoT , [Accessed: 06- Mar- 2020]

[2] Jen Clark , "What is the Internet of Things", 2016, [Online] Available: https://www.ibm.com/blogs/internet-of-things/what-is-the-iot/ , [Accessed: 06- Mar- 2020]

[3] Statista Research Department, "Internet of Things (IoT) connected devices installed base worldwide from 2015 to 2025", 2016, [Online] Available: https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/ , [Accessed: 06- Mar- 2020]

[4] fortunebusinessinsights.com , "Internet of Things (IoT) Market", 2019 [Online] Available: https://www.fortunebusinessinsights.com/industry-reports/internet-of-things-iot-market-100307 , [Accessed: 06- Mar- 2020]

[5] goldmansachs.com , "Drones reporting for work", [Online] Available: https://www.goldmansachs.com/insights/technology-driving-innovation/drones/ , [Accessed: 06- Mar- 2020]

[6] Josh Gordon and Sara Robinson , "Introducing TensorFlow.js: Machine Learning in Javascript", 2018 [Online] Available: https://medium.com/tensorflow/introducing-tensorflow-js-machine-learning-in-javascript-bf3eab376db , [Accessed: 06- Mar- 2020]

[7] Chandan Patra , "Amazon DynamoDB: 10 Things You Should Know", 2019, [Online] Available: https://cloudacademy.com/blog/amazon-dynamodb-ten-things/ , [Accessed: 06- Mar- 2020]

[8] Hemant Sharma , "Deep Dive into Amazon Simple Storage Service", 2019, [Online] Available: https://www.edureka.co/blog/s3-aws-amazon-simple-storage-service/ , [Accessed: 06- Mar- 2020]

[9] ryzerobotics.com , "Tello SDK", [Online] Available: https://dl-cdn.ryzerobotics.com/downloads/tello/20180910/Tello%20SDK%20Documentation%20EN_1.3.pdf , [Accessed: 06- Mar- 2020]

[10] amazon.com , "Amazon Rekognition FAQs", [Online] Available: https://aws.amazon.com/rekognition/faqs/ , [Accessed: 06- Mar- 2020]