A Proposed Framework for UAS Positioning in GPS-Denied and GPS-Spoofed Environments

Jack L. Burbank
School of Electrical Engineering and Computer Science
University of North Dakota
Grand Forks, ND, USA
jack.burbank@und.edu

Landon Foust
School of Electrical Engineering and Computer Science
University of North Dakota
Grand Forks, ND, USA
landon.foust@und.edu

Trevor Greene
School of Electrical Engineering and Computer Science
University of North Dakota
Grand Forks, ND, USA
trevor.greene@und.edu

Dr. Naima Kaabouch
School of Electrical Engineering and Computer Science
University of North Dakota
Grand Forks, ND, USA
naima.kaabouch@und.edu

2024 Integrated Communications, Navigation and Surveillance Conference (ICNS) | 979-8-3503-9309-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICNS60906.2024.10550601

Abstract*—Unmanned Aerial Systems have become ubiquitous and are now widely used in commercial, consumer, and military applications. Their widespread use is due to a combination of their low cost, high capability, and ability to perform tasks and go places that are not easy or safe for humans. Most UAS platforms are dependent on Global Navigation Satellite Systems (GNSS), such as the Global Positioning System (GPS), to provide positioning information for navigation and flight control. Without reliable GPS signals, the flight path cannot be trusted, and flight safety cannot be assured. However, GPS is vulnerable to several types of malicious attacks, including jamming, spoofing, or physical attacks on the GPS constellation itself. Additionally, there are environments in which GPS reception is not always possible, a key example being urban canyon areas where line-of-site to the GPS satellite constellation may be blocked or obscured by large obstacles such as buildings. Numerous methods have been proposed for position estimation in GPS denied environments. However, these methods have significant limitations and typically exhibit poor performance in certain environments and scenarios. This paper analyzes the strengths and weaknesses of existing alternate positioning methods and describes a framework where multiple positioning solutions are jointly employed to construct an optimal position estimate. The proposed framework aims to reduce computation complexity of and yield good positioning performance across a wide variety of environments.

Keywords*—Unmanned Aerial System, Positioning, Navigation, GPS-denied Operations

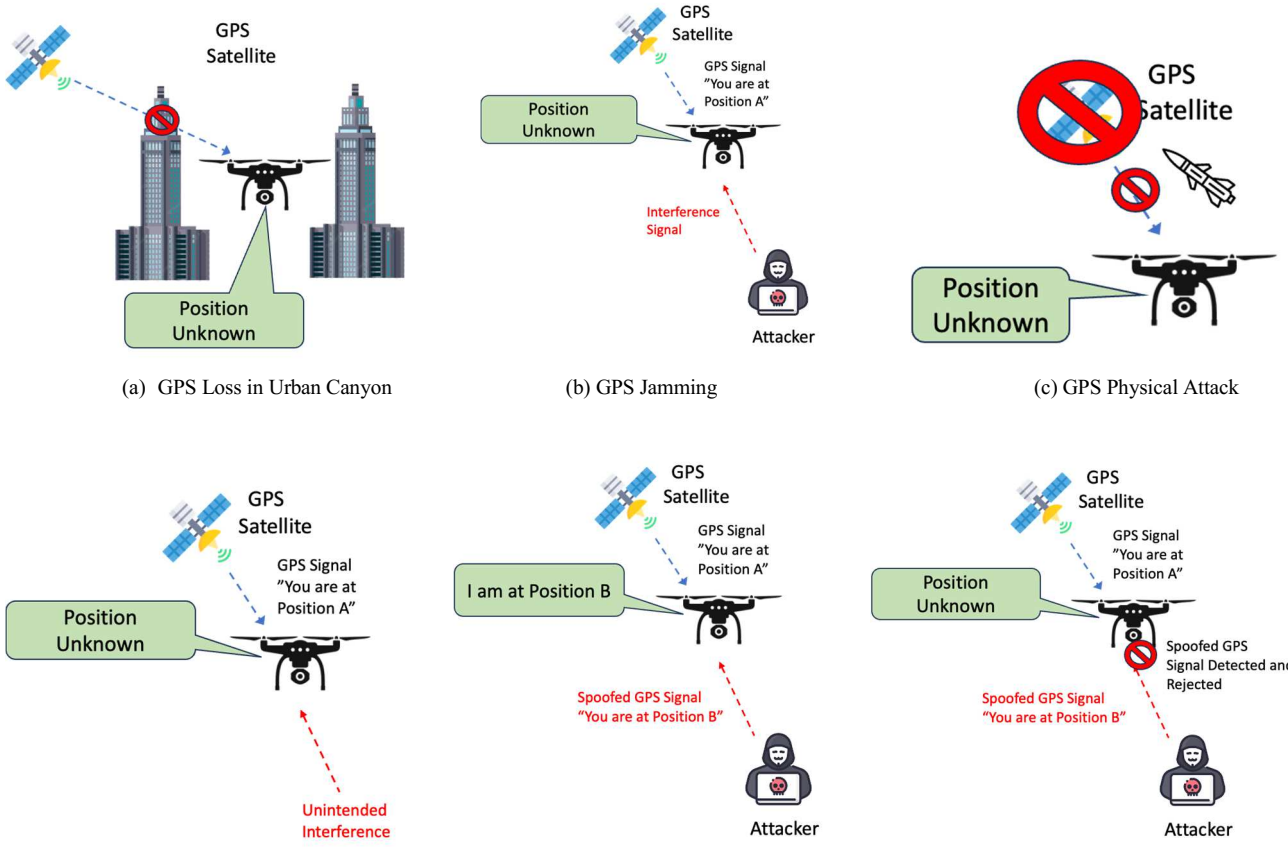
I. Introduction

Unmanned Aerial Systems (UAS) have become ubiquitous and are now widely used in commercial, consumer, and military applications. Their widespread use is due to a combination of their low cost, high capability, and ability to perform tasks that are tedious or not safe for humans. Most UAS, particularly those used for commercial or Government applications, are dependent on Global Navigation Satellite Systems (GNSS), such as the Global Positioning System (GPS), to provide positioning information for flight control. Without reliable GPS, the flight path cannot be trusted and flight safety cannot be assured. However, there are several environments in which GPS reception is not always possible. Fig. 1 shows six examples of GPS-denied scenarios. An example is urban canyon areas where line-of-site (LOS) to the GPS satellite constellation may be obscured by large obstacles such as buildings. GPS signals are also sensitive to radio frequency (RF) intentional or unintentional interference. GPS operates on spectrum allocated and reserved for its own use. Spectrum utilized by GPS includes the L1 carrier operating at 1575.42 MHz while the L2 carrier operates at 1227.60 MHz. Newer GPS satellites operate at other frequencies, such as the L5 carrier at 1176 MHz. While these frequencies are reserved for GPS use, there is always the potential that other RF devices could inadvertently emit signals in these reserved bands (either due to poor quality or malfunctioning hardware components). GPS signals are extremely weak on Earth, often on the order of -130 dBm or lower. A relatively low power emitter can interfere, intentionally or unintentionally, with proper GPS signal reception. This is exacerbated by the fact that commercial GPS receivers use widely varying quality of hardware components, which can lead to poor performance for some receivers in non-pristine electromagnetic environments.

Commercial GPS signals are not authenticated or protected in any way from malicious attacks. The most common attack is GPS spoofing, where an attacker broadcasts its own falsified GPS signal, leading a receiver to believe it is at a different position. This type of attack is made more practical due to the low signal strength of the downlink GPS signal, making it easier for the malicious attacker's signal to overwhelm the actual GPS signal at the receiver. It is also important to note the ease with which a GPS spoofing attack can be delivered [1, 2]. While there are many methods for GPS spoofing detection that have been proposed in the literature [3], these approaches do not go on to provide another alternate forms of positioning. GPS spoofing detection is important since the UAS will not be following fraudulent position information. However, even if an effective GPS spoofing detection algorithm is employed, the UAS can find itself in the equivalent of a GPS denied environment and the UAS is left without positioning information.

II. Previous work

Positioning, navigation, and timing (PNT) in GPS-denied environments has received significant research interest over the past two decades [4]. Consequently, there have been many proposed methods for positioning in GPS-denied environments. These various proposed methods typically fall into one of six categories of approach, as summarized in Fig. 2. Inertial Measurement Unit (IMU)-based approaches utilize the onboard suite of devices, including accelerometers, gyroscopes, and magnetometers, to measure roll, pitch, yaw, velocity, and altitude. Additionally, these devices measure changes in magnetic and gravitational forces. They can be used to detect and track motion such that UAS position relative to starting point can be tracked [4]. These approaches are well understood



(a) GPS Loss in Urban Canyon (b) GPS Jamming (c) GPS Physical Attack

(d) Unintended Interference (e) GPS Spoofing (f) GPS Spoofing with Detection

Fig. 1: Different Types of GPS-Denied Scenarios

and can provide very accurate initial results. However, their accuracy decreases over time due to error accumulation during the integration of angular and linear velocity. These errors fall into one of three categories: biases, scale-factor errors, and misalignment errors. They can be quite large for commercial IMU equipment, with significant inaccuracies resulting after as little as one minute of use. IMU calibration procedures can significantly lower these errors. However, error accumulation limits the timeframe during which accurate results can be maintained. There is the emergence of the concept of quantum accelerometers that would produce negligible error such that error accumulations would be insignificant. In this case, IMU-based positioning and navigation would be sufficient. However, quantum accelerometers do not yet exist in a form practical for UAS applications.

Many approaches for positioning and navigation have been proposed in literature that rely on the use of landmarks. These approaches generally fall into one of two categories: 1) RF landmarks and 2) visual landmarks. RF landmark-based approaches use existing terrestrial cellular infrastructure for position calculation by the UAS [5]. Previous studies have shown good accuracy that can be obtained; however, the accuracy is dependent on the environment (e.g., rural versus urban) and position accuracy is sensitive to interference. Other approaches have been proposed where known signal emitters are pre-placed along the flight path. This requires a priori knowledge of the flight path and requires installation of ground-based infrastructure. These approaches have been shown to be effective and may prove valuable in applications where a UAS always performs a fixed flight path. However, they may not be practical for generalized UAS operations.

One potentially viable approach for positioning based on RF-based landmarks is the use of signals from the terrestrial cellular communications infrastructure [5]. There are two fundamental approaches that can be taken: 1) cooperative cellular-based positioning, or 2) uncooperative cellular-based positioning. In the case of cooperative cellular-based positioning, the UAS is equipped with a cellular transceiver and a member of the cellular network. In this scenario the UAS could receive position via the cellular network using existing methods (e.g., in 4G cellular using the LTE Positioning Protocol (LPP)). The cooperative approach has been shown to yield good results but is sensitive to jamming and interference [6]. If this approach became common, the result could be a significant number of cellular users at high elevations with visibility to many cell towers. This will stress the cellular network's interference model and could negatively impact the overall network [7]. In the uncooperative cellular-based positioning approach, the UAS is not a user of the cellular network (i.e., no onboard cellular transceiver). Rather, the UAS

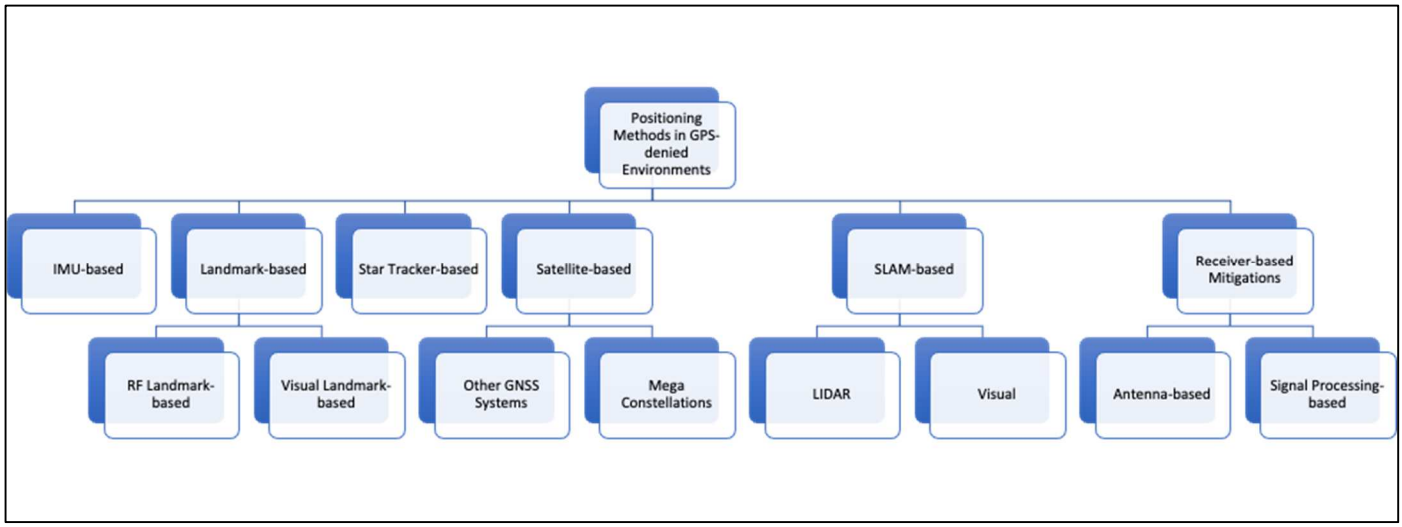


Fig. 2: High-Level Taxonomy of Existing Methods for Positioning in GPS-Denied Environment

is passively observing and using the cellular signals.

Visual landmark-based approaches use a camera onboard the UAS to visually detect natural or manmade landmarks for position determination [8, 9]. These approaches are typically limited to well-known existing landmarks, or reliant on pre-placed objects along the flight path, limiting the generalization of these approaches. Approaches that rely on the pre-placement of landmarks along a pre-planned flight path could prove valuable in applications where a UAS flies a fixed, known flight path. However, these types of approaches are not practical for many UAS applications. Approaches that utilize existing landmarks have been shown capable of good results [8]. Furthermore, the combination of visual landmark-based approaches with IMU-based approaches has also shown promising results [10]. In these types of approaches, visual recognition provides an accurate dead reckoning that can be then used to reset IMU integration errors back to their initial state. Within regions dense with visually identifiable landmarks, this approach works well [8, 10]. However, in regions void of unique landmarks, these approaches can perform poorly.

Star tracker-based approaches utilize onboard optical devices that measure the position of stars in the sky to determine their position. These approaches are well known and established in nautical and space-based positioning and navigation and have a long history of working well in low light, clear sky scenarios. However, these approaches have limited utility in daytime or obstructed views (e.g., cloud cover).

Several approaches are based on the utilization of other satellite systems for the purpose of positioning and navigation. These methods fall into one of two categories: 1) the use of other GNSS systems, or 2) the use of non-GNSS mega constellations. There are other GNSS systems, such as GLONASS, Galileo, and BeiDou. Many commercial chipsets include receivers for multiple of these GNSS systems. These commercial chipsets somewhat mitigate the threat of physical GPS attack since it is less likely that all GNSS systems will be simultaneously attacked. However, these other GNSS systems, like GPS, are susceptible to jamming and spoofing attacks. A growing area of research is in the use of non-GNSS satellite systems for position determination, such as Starlink, OneWeb, OrbComm, and Iridium [11], with particular interest in the use of mega-constellation systems such as Starlink [12]. The extremely large size of the Starlink constellation and the Low Earth Orbit (LEO) nature of the constellation would make a physical attack scenario extremely unlikely. However, Starlink signals will likely also be susceptible to jamming and/or spoofing attacks.

Simultaneous Localization and Mapping (SLAM) is another category for navigation and positioning. SLAM-based methods aim to characterize the environment surrounding the autonomous vehicle via its onboard sensors and determine the vehicle's position within that environment. Most SLAM methods fall within two categories: 1) Light Detection and Ranging (LIDAR)-based SLAM [13] and 2) Visual SLAM (vSLAM) [14]. These SLAM approaches attempt to detect features in the environment, building up a feature map such that it can navigate the vehicle's surroundings. These approaches have shown very promising results, but typically are used for localized environments and serve the purpose of obstacle avoidance or navigation along specific pre-planned paths. They typically do not provide absolute positioning (i.e., latitude and longitude) for generalized scenarios.

Generalized vision-based approaches, that can operate on general terrain rather than simple landmarks, and that can provide absolute positioning (i.e., latitude and longitude) have received limited research [15]. In this category of approaches, the UAS-based camera observes the ground below. Algorithms onboard the UAS then attempt to recognize the terrain to determine the UAS position. These approaches have been demonstrated to be achievable [15]. However, a major limitation of these approaches is their complexity due to the size

of image dataset for training and validating the AI algorithms. For a single global training dataset that would encompass all locations on Earth, an 80-meter resolution would require a minimum dataset of 419.9 billion images (assuming a single image representing each (latitude, longitude) coordinate separated by 80 meters). However, a more realistic dataset for 80-meter resolution would be approximately 4.2 trillion images. This is because the training dataset requires diversity to accommodate different scales (i.e., altitudes), different rotations and orientations, different environmental conditions (e.g., partial occlusions due to cloud cover, shadowing efforts due to time of day), and different degradations that may appear in the UAS-based imagery (e.g., Additive White Gaussian Noise (AWGN), Salt and Pepper Noise, blurring). This estimation assumes at least 10 images per coordinate would be required to provide sufficient diversity within the training dataset. For a globally applicable dataset (i.e., a dataset representing all positions on Earth), the relationship between estimated training dataset size and achievable spatial resolution is depicted in Fig. 3. Even with a very large Graphical Processing Unit (GPU) infrastructure to support algorithm training, this dataset size is not practical.

Lastly, several methods in literature propose to mitigate the effects of a jammer or spoofer through advanced antenna or signal processing techniques to reject the energy of the attacker [16-18]. Many of these approaches have shown promising results. However, these methods introduce hardware and/or processing complexity into the UAS, which makes many of these methods impractical for low-cost commercial UAS platforms.

The advantages and disadvantages of these various positioning approaches are summarized in Table 1.

III. Proposed Approach

The proposed framework, depicted in Fig. 4, not only incorporates positioning estimates from multiple sources, but also uses certain individual position estimates as aids for other positioning methods (i.e., to calculate conditional position estimates based on information learned from other sensors/approaches). These position estimates are combined in a confidence-based weighting approach to produce an overall position estimate.

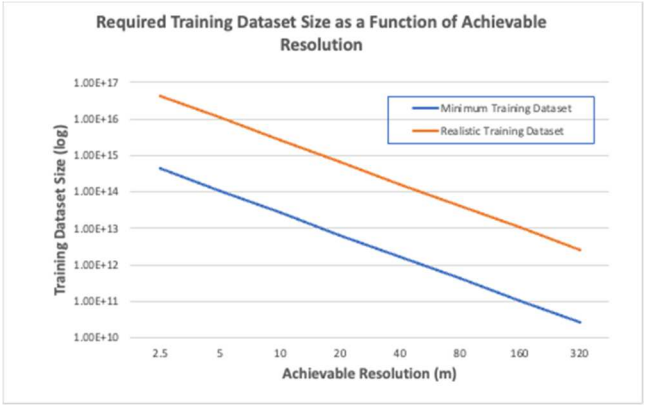


Fig. 3: Relationship between Training Dataset Size and Achievable Spatial Resolution

The proposed confidence-based weighting approach is intended to address the weaknesses and shortcomings of the previously mentioned approaches While a UAS may have multiple sensors and positioning methods available to it, not all methods are ensured to work well within a given circumstance. IMU-based approaches may work well for short flights but have lower confidence as time-of-flight increases. Star tracker-based approaches may work well during nighttime environments with clear skies but will have low confidence during daytime operations. RF-based landmark produces good results in some environments, but not other environments. The proposed framework aims to utilize confidence-based weighting to accommodate these performance trends.

A key aspect of this framework is a cellular-aided visual recognition position determination approach. The goal here is to utilize cameras that are likely already onboard the UAS that can observe the ground beneath the UAS and then associate that observation with a location, all while mitigating the complexities of previous approaches. This cellular-aided visual recognition approach does not require any pre-placed anchor points or a priori information. While previous research has demonstrated it is possible to determine position by using ground images (i.e., image of the ground directly beneath the UAS) [15], these methods suffer from high computational complexity and/or require an impractically large onboard image library.

The proposed cellular-aided visual recognition approach mitigates this issue. It is an uncooperative cellular-based positioning approach that serves as a coarse position estimation. That coarse position is then used to perform region-specific vision-based position estimation. Instead of a single vision-based detection algorithm trained against a global dataset, we employ multiple detectors each trained against smaller, more tenable regional datasets. Both position estimates, along with confidence scores, and along with estimates from other sources (e.g., IMU) into a weighted combination function.

Weighting functions are calculated based on location estimates and confidence scores.

$$w_i = f(L_c, C_{Li}) \tag{1}$$

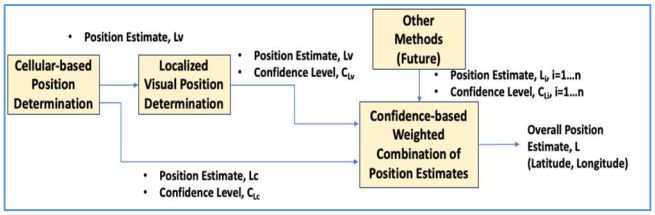


Fig. 4: Proposed Framework for Hybrid RF Landmark / Computer Vision Positioning

TABLE I. STRENGTHS AND WEAKNESSES OF ALTERNATE POSITIONING APPROACHES

Positioning Method	Hardware Requirement	Advantages	Disadvantages
GPS	GPS receiver	• Low Size, Weight, and Power (SWaP) hardware • Low-cost hardware • Good performance in most benign environments	• Vulnerable to spoofing and Denial-of-Service (DoS) attacks • Poor performance in some environments
IMU	IMU	• Good performance for short flight times • Utilizes hardware already onboard UAS	• Error accumulations degrades performance over time
RF Landmark	RF receiver	• Good performance possible, particularly in dense RF environments	• Some approaches require pre-placed signals along flight path • Potentially requires additional radio receiver • Performance dependent on environment and geometry
Visual Landmark	Camera	• Good performance possible • Minimal hardware requirements	• Requires pre-placed physical objects along flight path
Star Tracker	Star tracker	• Good performance in certain conditions	• Limited utility in daytime or when night sky obscured • Requires additional star tracker hardware
Alternate GNSS	GNSS receiver	• Low-cost, low-SWaP hardware, often same hardware as GPS receiver	• Vulnerable to spoofing and DoS attacks • Poor performance in some environments
Mega-Constellation	RF receiver	• Good performance possible • Mega-constellations less vulnerable to physical attacks	• Vulnerable to spoofing and DoS attacks • Additional hardware required
LIDAR SLAM	LIDAR transceiver	• Good performance possible	• Expensive hardware • Computationally expensive • Typical application is regional mapping, not absolute position for navigation
Visual SLAM	Camera	• Good performance possible • Minimal hardware requirements	• Computationally expensive • Typical application is regional mapping, not absolute position for navigation
Receiver-based	RF hardware (e.g., antenna)	• Good performance possible	• Additional hardware and/or computational complexity
Generalized Vision	Camera	• Good performance possible • Minimal hardware requirements	• Computational expense for algorithm training

Finally overall position estimates (latitude and longitude) are calculated using the weighting function given by:

$$L = w_c \cdot L_c + w_v \cdot L_v + \sum_{i=1}^{n} w_i \cdot L_i \qquad (2)$$

This paper does not define the weighting function, nor does it define the method to calculate confidence scores. Rather, the paper focuses on investigating the viability of cellular RF-based positioning and CV-based positioning.

A. Cellular RF-based Course Position Estimation

The location of all commercial 4G/5G infrastructural emitters (i.e., base stations) are publicly known. Thus, it would be possible to build an onboard database for positioning purposes. Each base station in the cellular network uniquely identifies itself over the Radio Access Network (RAN), so a mobile receiver will be able to definitively determine whose signals it is receiving. Combining that knowledge with a priori knowledge of the location of those emitters, the problem to solve is given a received signal from a known source, estimating the receiver's location based on characteristics of that received signal. In other words, the UAS will know the locations of the surrounding base stations (i.e., those within range of the UAS), and the UAS will have received signals that are attributable to those base stations. The UAS will use that information to estimate its own position; the UAS will attempt to geolocate itself based on the known signals around it. It is interesting to note that this is the opposite of the typical geolocation problem. In most geolocation problems, you have a network of sensors attempting to determine the location of a single emitter. In this case, we have a network of known emitters and attempting to calculate the position of the single sensor. However, standard geolocation approaches can still be applied to this problem.

There are numerous geolocation methods that can be used to determine location based on received signal attributes, including Time Difference of Arrival (TDOA), Frequency Difference of Arrival, and Power Difference of Arrival (PDOA). These approaches are predicated on the scenario of multiple distributed spectrum sensors attempting to locate a single emitter by comparing the differences in the versions of the same signal independently received at these distributed sensor locations.

The proposed framework aims to employ a Received Signal Strength (RSS)-based approach due to its simplicity. In this approach, these known identifiable signals will be received with some RSS. Based on that RSS along with the known position of the emitters, the UAS would then estimate the distance between itself and the known transmitter location. If at least three emitters (i.e. base stations) can be observed, then the UAS can then use reverse triangulation to determine its own position. Previous research has demonstrated that RSS-based approaches can provide accurate results [19]. This scenario is depicted in Fig. 5.

This RSS-based approach is a non-cooperative method in which the UAS is not actually a user on the terrestrial cellular network. It is a receive-only capability receiving well-defined 4G/5G signals, and then estimating the UAS position based on the RSS of these signals. For 4G LTE, we propose the use of the Primary Synchronization Signal (PSS) and Secondary Synchronization Signal (SSS), as these are well known and will convey the identification of the transmitter.

For the case of 5G cellular, we propose the use of the Downlink Position Reference Signal (DL PRS), which is used by the 5G network itself in its own positioning protocol.

B. Vision-based Position Estimation

This study employs You Only Look Once (YOLO) due to its computational efficiency and ability to run in embedded hardware. YOLO is an object detection algorithm widely used in the Computer Vision community and is also widely used in commercial applications, including security and surveillance applications as well as autonomous vehicles. While there are many differences across the various versions of YOLO, the main characteristic that makes the YOLO model different than many other approaches is that it is a single-stage detector versus a two-stage detector. The basic object detector architecture is depicted in Fig. 6 [20].

The input stage in Fig. 6 is the actual input image fed into the detector. The backbone stage is a multi-layer deep learning (DL) network that acts to extract features from the input image. There are numerous backbone architectures that are available to use, including VGG16 [21], DARKNET-19 [22],

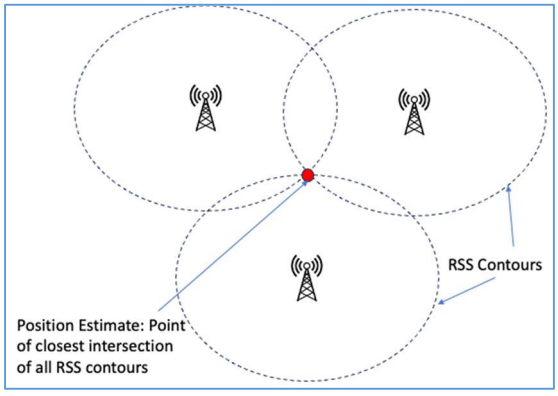


Fig. 5: Position Determination Process using RSS-based Triangulation

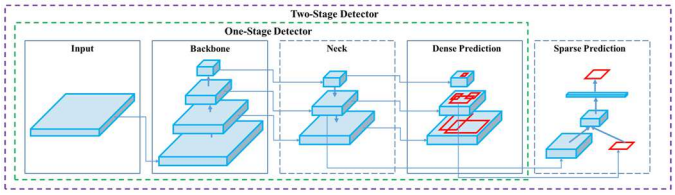


Fig. 6: High-Level Architecture of One-stage versus Two-State Detectors [20]

DARKNET-53 [23], among many others. Here, the number in each generally refers to the number of layers in a convolutional neural network (CNN). The neck stage connects the backbone to the head (Dense Prediction stage in Fig. 6) and performs the function of feature map concatenation from the different layers of the backbone stage and presents that information as inputs to the head, which performs the object detection based on those feature maps. The neck can consist of many different approaches, including Feature Pyramid Networks (FPN), Path Aggregation Network (PAN), and Receptive Field Block (RFB). The head can take the form of YOLO or Single-Shot Detector (SSD) [24]. In a two-stage detector approach, such as Faster R-CNN [25], there is an additional sparse detection stage. Note this last stage is not present in a single-stage detector.

This study uses the YOLOv4 implementation included as part of the MATLAB CV Toolbox, and its architecture is shown in Fig. 7. In this implementation, the backbone uses a DarkNet-53 architecture that is a 53-layer CNN. The neck employs Spatial Pyramidal Pooling (SPP) and PAN. The head consists of three separate YOLOv3 [26] head modules that takes the feature maps provided by the neck and produces predicted bounding boxes and confidence scores.

Typical YOLO usage involves the training of YOLO on an image dataset containing many different objects of different types (e.g., dogs, bicycles, cars) with that training data labeled with bounding boxes and descriptive labels for each object type. The trained algorithm is then used to detect arbitrary objects of those same types contained within input images (images not part of the original training dataset). In this study, rather than detecting specific objects, YOLO is employed to detect a grid of ground area from an aerial image (e.g., multiple city blocks vs. a single building). Objects, in this case, are defined by bounding boxes to reflect an entire geographic grid area and labeled with the coordinates of the center of the bounding box, as depicted in Fig. 8.

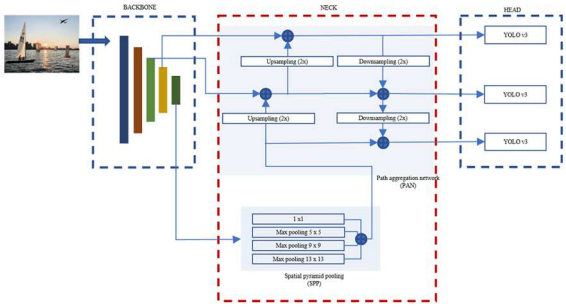


Fig. 7: YOLOv4 Architecture as Implemented in MATLAB [27]

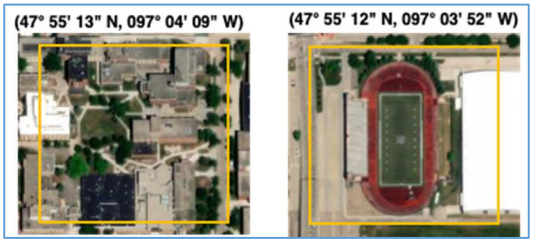


Fig. 8: Example of Using Object Definition to Represent Geographic Grid Coordinates

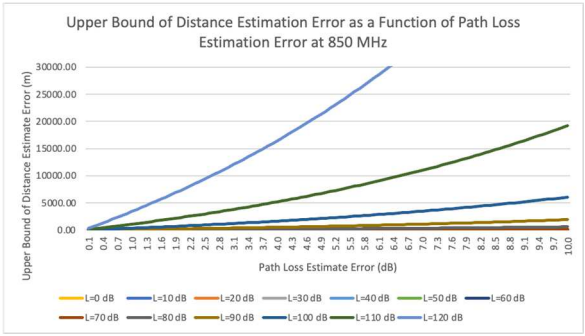
In this study, the YOLO algorithm is trained with a dataset that contiguously covers a geographic region, with each image representing an x by y grid of space within that geographic region. Ground images taken from the UAS are then fed into this trained YOLO detector, and a detection represents a match with one of these grid image objects, thus yielding the GPS coordinate of the UAS platform.

IV. RESULTS AND DISCUSSION

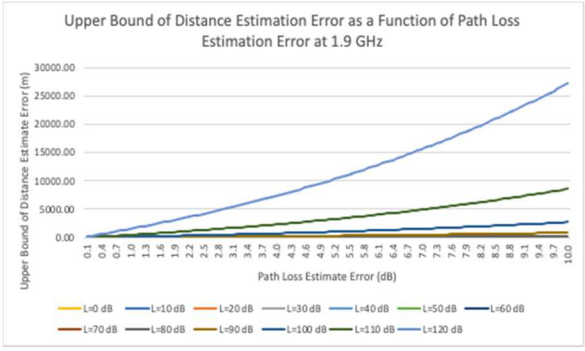
Fig. 9 shows the relationship between the path loss estimation error (i.e., the difference between estimated and the actual path loss between the UAS and the cellular transmitter) and distance estimation error (i.e., the difference between the estimated and actual path distance between the UAS and the cellular transmitter). The subplots in Fig. 9 shows this relationship for three common cellular frequency bands: 850 MHz (Fig. 9a), 1900 MHz (Fig. 9b), and 27 GHz (Fig. 9c). The curves in each subplot of Fig. 9 represent a set of true propagation losses between the UAS and the cellular transmitter, which is directly related to the distance between them.

Making a worst-case assumption that distance estimation errors align in the worst possible way (i.e., errors are additive), these cumulative distance estimation errors will directly equate to position estimate errors. The results in Fig. 9 shows that position estimate error increases with increased RF path loss estimation error, which places a premium on accurate RF propagation path loss estimation. Note that the results in Fig 9. assume free space path loss to relate path loss estimation error to distance error. Overall, across all frequencies it was found that relatively good positioning can be achieved if the path loss estimation error is less than or equal to approximately 5 dB. For path loss estimation errors greater than 5 dB, the distance error increases, leading to a larger position estimate error. Because the UAS will generally have LOS conditions to the terrestrial towers, previous research has shown that the propagation channel may follow a LOS free-space model [7], which means that achieving a 5 dB agreement may be possible.

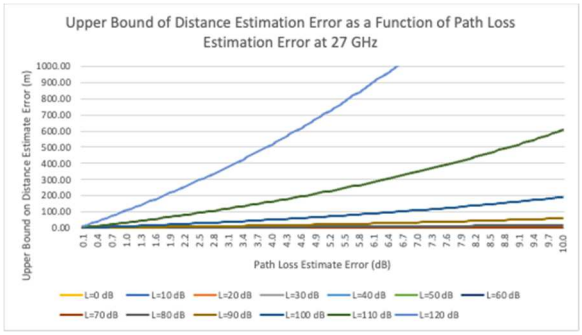
Fig. 9 also shows that the positioning accuracy improves with higher frequencies. The highest frequency band considered (27 GHz in Fig. 9c) provides the best positioning performance and lower cellular bands provide decreased positioning performance. In the 27 GHz band, most path loss curves (i.e., distances) between the UAS and the cellular tower resulted in position estimation errors under 100 m even for a



(a) Position Estimation Error Function for 850 MHz



(b) Position Estimation Error Function for 1900 MHz



(c) Distance Estimation Error Function for 27 GHz

Fig. 9: Distance Estimation Error as a Function of Path Loss Estimation Error for non-Cooperative Cellular RF-based Positioning

relatively large path loss estimation error. For the lowest frequency band considered (850 MHz in Fig. 9a), large path loss estimation errors resulted in position estimation errors of 1 km or greater, particularly for large distances between the UAS and the cellular transmitter.

It was found that positioning accuracy worsens with increased distance to emitters. At larger distances, the position error is more sensitive to the path loss estimation error. This can be seen in Fig. 9 by the curves representing lower path loss that are grouped closely together at low distance error values while the curves representing higher path loss (e.g., 120 dB) quickly grows as the path loss estimate error increases. Relatively good performance can be achieved under the assumed conditions for distances up to 1 km in most cases, assuming free space path loss, with significant degradations at distances greater than 1 km. These results suggest that within relatively close distances to cell towers ($\leq$ 1 km) and path loss estimation error of $<$ 5 dB,

that positioning accuracy of 100-meter resolution is achievable. While 100-meter resolution is not sufficient for safe flight, it is likely sufficient for coarse positioning to direct vision-based positioning. Even at greater distances or cases of a larger path loss estimation error, where positioning error will be greater, the position is likely sufficient to support a coarse positioning function.

We next investigate the viability of using YOLO for vision-based positioning. A small training dataset was constructed consisting of 90 images [28]. These images were taken from three different types of environments: 1) urban (downtown Chicago, IL), 2) suburban (University of North Dakota (UND) campus), and 3) rural (unpopulated desert region of Arizona). These three image classes attempt to represent regions with different feature densities (i.e., manmade or natural landmarks). A total of 30 images were taken for each of these three regions. Each image was appended with a bounding box and labeling data consisting of latitude and longitude corresponding to the center point of the image. All 30 images from each of these three environments were included in the training dataset. Twelve different images corresponding to these areas (three from each area) were then drawn from a different source (Google Earth) to serve as surrogate UAS camera images and then used as the test dataset. Examples of these test images for each region is shown in Fig. 10.

These test images were fed into the YOLO detector that was trained with the 90-image training dataset, with detection results shown in Table 2. This table shows that a good performance was observed for urban and suburban cases, with successful detections with a high confidence (>0.9) for all urban and suburban test cases. The results for the desert test cases were mixed. A portion of the desert test cases provided detection with moderate confidence (>0.7), while several other desert cases resulted in no detection. False positive detections were also observed in the desert test images (i.e., images mistakenly classified as the wrong location). However, even in the case of mistaken classification, the model still predicted that the location was in the correct region (i.e., predicted Arizona desert).

A key driver in the detection with high confidence appears to directly relate to the number of features present within the image. Images without distinct feature sets, such as Arizona Desert 1, exhibited poor performance, more features lead to

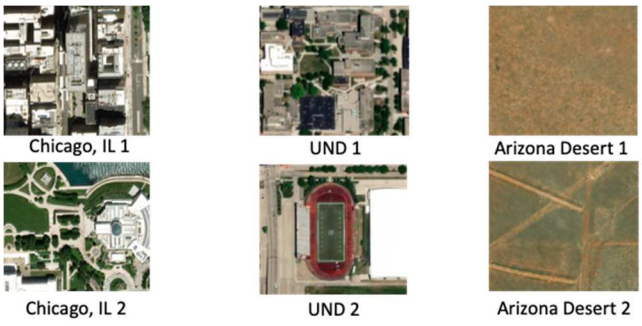


Fig. 10: Test Images from Urban, Suburban, and Desert Environments

TABLE II. ACHIEVED RESULTS FROM YOLO-BASED VISUAL POSITIONING

Position (Image)	Accurate Detection	Confidence Level
Chicago 1	Yes	0.97
Chicago 2	Yes	0.99
Chicago 3	Yes	0.91
Chicago 4	Yes	0.89
UND 1	Yes	0.96
UND 2	Yes	0.99
UND 3	Yes	0.92
UND 4	Yes	0.90
Arizona 1	No	N/A
Arizona 2	Yes	0.81
Arizona 3	Yes	0.68
Arizona 4	No	N/A

better detection performance. These results are promising for urban and suburban areas as well as rural areas with geographic or manmade features (e.g., rivers, roads). The results also suggest that this approach is not well suited for areas with monotonic feature space (e.g., desert). In these desert environments (i.e., low feature density environments), it is expected that cellular RF-based positioning will yield more accurate positioning performance than visual methods.

V. CONCLUSIONS AND FUTURE WORK

The use of UAS platforms has become ubiquitous and their use is expected to increase in the future. UAS platforms are heavily dependent on GPS for proper and safe operations. However, GPS is vulnerable to denial and spoofing attacks. Many alternative positioning solutions have been proposed in the literature, but all have significant limitations. This study described a framework where multiple positioning approaches would be jointly employed. This framework uses of a cellular RF-aided vision-based approach to positioning that utilizes the popular YOLO object detection algorithm for aerial landscape detection for positioning purposes. The results show that cellular-based RF signals may provide sufficient accuracy for coarse position estimates. Results also show that visual methods can provide accurate position estimates in many environments. Future work would include the evaluation of the proposed vision-based positioning approach with larger datasets across a wider set of environments and conditions. Future work would also mature the framework described in this paper by developing a confidence weighting factor scheme, as well as the position estimate combination function.

ACKNOWLEDGEMENT

This research was funded by the National Science Foundation (NSF), Award Number 2006674.

REFERENCES

[1] P. Betha et al., "Stealthy GPS Spoofing: Spoofer Systems, Spoofing Techniques and Strategies," 17th IEEE India Council International Conference (INDICON), February 2020.

[2] R. Ferreira et al., "Effective GPS Jamming Techniques for UAVs Using Low-Cost SDR Platforms," 2018 Global Wireless Summit (GWS), November 2018.

[3] R. Manesh et al, "Detection of GPS Spoofing Attacks on Unmanned Aerial Systems," 2019 Sixteenth IEEE Annual consumer communications and networking conference (CCNC), 2019.

[4] G. Balamurugan et al., "Survey on UAV Navigation in GPS-denied Environments," 2016 International Conference on Signal Processing,

Communication, Power and Embedded System (SCOPES), October 2016.

[5] E. Kim and Y. Shin, "Feasibility Analysis of LTE-Based UAS Navigation in Deep Urban Areas and DSRC Augmentation," MDPI Sensors 2019, 19, 4192, September 2019.

[6] S. Dwivedi et al., "Positioning in 5G Networks," IEEE Communications Magazine, Vol. 59, Issue 11, November 2021, pp. 38-44.

[7] S. Muruganathan et al., "An Overview of 3GPP Release-15 Study on Enhanced LTE Support for Connected Drones," IEEE Communications Standards Magazine, Vol. 5, Issue 4, pp. 140-146, December 2021.

[8] K. Sundar et al., "Landmark Placement for Localization in a GPS-denied Environment," 2018 Annual American Control Conference (ACC), June 27-29, 2018.X

[9] A. Badshah et al., "Vehicle Navigation in GPS Denied Environment for Smart Cities using Vision Sensors," Elsevier Journal of Computers, Environment and Urban Systems, Vol. 77, September 2019.

[10] U. Patel and I. Faruque, "Multi-IMU Based Alternate Navigation Frameworks: Performance and Comparison for UAS," IEEE Access, Vol. 10, January 2022, pp. 17565-17577.

[11] Z. Kassas et al., "Navigation with Multi-Constellation LEO Satellite Signals of Opportunity: Starlink, OneWeb, OrbComm, and Iridium," 2023 IEEE/ION Position, Location, and Navigation Symposium (PLANS), April 2023.

[12] M. Neinavaie and Z. Kassas, "Signal Mode Transition Detection in Starlink LEO Satellite Downlink Signals," 2023 IEEE/ION Position, Location, and Navigation Symposium (PLANS), April 2023.

[13] G. Ariante et al., "UAS for Positioning and Field Mapping using LIDAR and IMU Sensors Data: Kalman Filtering and Integration," 2019 IEEE 5th International Workshop on Metrology for AeroSpace, June 2019.

[14] D. Ruckert et al., "Snake-SLAM: Efficient Global Visual Intertial SLAM using Decoupled Nonlinear Optimization," 2021 International Conference on Unmanned Aircraft Systems (ICUAS), June 2021.

[15] H. Goforth and S. Lucey, "GPS-Denied UAV Localization using Pre-existing Satellite Imagery," 2019 International Conference on Robotics and Automation (ICRA), May 20-24, 2019.

[16] M. Abbasi et al., "GPS Continuous Wave Jamming Canceller using an ANF Combined with an Artificial Neural Network," 2020 8th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS), September 2020.

[17] V. Obi, G. Evans, and S. Lim, "Design of a Miniaturized, High-Gain, Anti-Jam Global Positioning System (GPS) Antenna," IEEE SouthEastCon 2022, April 2022.

[18] G. Kim et al., "A Reconfigurable RF Front-end System for GPS Anti-Jamming," 2019 Photonics & Electromagnetics Research Symposium – Spring (PIERS-Spring), June 2019.

[19] J. Ying and K. Pahlavan, "Precision of RSS-based Localization in the IoT," International Journal of Wireless Information Networks, 26, 10-23 (2019).

[20] A. Bochkovskiy, C. Wang, and H. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," arXiv: 2004.10934, April 2020.

[21] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv: 1409.1556, September 2014.

[22] J. Redmon and A. Farhardi, "YOLO9000: Better, Faster, Stronger," arXiV: 16612.08242, December 2016.

[23] J. Redmon et al., "You Only Look Once: Unified, Real-Time Object Detection," ArXiv: 1506.02640, June 2015.

[24] W. Liu, et al., "SSD: Single Shot Multibox Detector," arXiv: 1512.02325, December 2015.

[25] S. Ren, et al., "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," arXiv: 1506.01497, January 2016.

[26] J. Redmon and A. Farhardi, "YOLOv3: An Incremental Improvement," arXiV: 1804.02767, April 2018.

[27] "Getting Started with YOLO v4," Mathworks Help Center, https://www.mathworks.com/help/vision/ug/getting-started-with-yolo-v4.html, Accessed August 4, 2023.

[28] https://earthexplorer.usgs.gov