4-27-2018

From Sensors to Knowledge: The Challenge of Training the Next Generation of Data Analysts

Sonya A. H. McMullen
Embry-Riddle Aeronautical University, hall76d@erau.edu

Troy Henderson
Embry-Riddle Aeronautical University, hendert5@erau.edu

Scott Burgess
Embry-Riddle Aeronautical University, burgesco@erau.edu

Patti J. Clark
Embry-Riddle Aeronautical University, clark092@erau.edu

David Ison
Embry-Riddle Aeronautical University, isond46@erau.edu

Scholarly Commons Citation

Sonya A. H. McMullen, Troy Henderson, Scott Burgess, Patti J. Clark, David Ison, "From sensors to knowledge: the challenge of training the next generation of data analysts," Proc. SPIE 10653, Next-Generation Analyst VI, 106530S (27 April 2018); doi: 10.1117/12.2304885

From sensors to knowledge: the challenge of training the next generation of data analysts

Sonya A. McMullen, Troy Henderson, Scott Burgess, Patti Clark, David Ison

Sonya A. H. McMullen, Troy Henderson, Scott Burgess, Patti J. Clark, David Ison, "From sensors to knowledge: the challenge of training the next generation of data analysts," Proc. SPIE 10653, Next-Generation Analyst VI, 106530S (27 April 2018); doi: 10.1117/12.2304885

SPIE.

Event: SPIE Defense + Security, 2018, Orlando, Florida, United States

From Sensors to Knowledge: The Challenge of Training the Next Generation of Data Analysts

Sonya A. H. McMullen*[a], Troy Henderson[b], Scott Burgess[a], Patti J. Clark[a], David Ison[a]
[a]Embry-Riddle Aeronautical University Worldwide, College of Aeronautics; [b]Embry-Riddle Aeronautical University, College of Engineering, 600 S. Clyde Morris Blvd., Daytona Beach, FL 32114-3900

ABSTRACT

With the advent of commercial-off-the-shelf sensors for use in a variety of applications, integration with analytical software tools, and expansion of available archived datasets, there is a critical need to address the problem of transforming resultant data into comprehensible, actionable information for decision-makers through rigorous analysis. In previous research the participating authors have emphasized that users are often faced with the situation in which they are "drowning in a sea of data" but still "thirsting for knowledge". The availability of analysis software, tools, and techniques provide opportunities for information collection of ever increasing complexity, but the need for the training of analysts to employ *appropriate* tools and processes to ensure accurate and applicable results has not been addressed. The purpose of this paper is to discuss the challenges and opportunities facing the training of effective analysts capable of handling a wide-range of data types in this era of dynamic tools and techniques.

1. INTRODUCTION

The Information Age has resulted in an explosion of data collection and processing capabilities through the miniaturization of computers and sensors, dramatic increases in computing power available to the public, and substantial cost reduction of sensors and computers. Further, small unmanned aerial systems (sUAS) with optical, infrared, and multispectral sensors, have put remote sensing capabilities into the hands of amateurs offering collection capabilities with little training on how to operate these devices. Capabilities that were once the purview of government agencies due to costs and the advanced nature of projects, land-, air-, and space-borne remote sensing are now within the realm of possibilities for the general public. As highlighted by Dr. David L. Hall, former Dean of the College of Information Sciences and Technology (IST) at The Pennsylvania State University (PSU), as well as Dr. Michael McNeese (PSU), these capabilities afford the potential for "every citizen to be a sensor" [1,2], especially with cellphone cameras, social media, and expanded internet access. Computer storage chips the size of finger nails now have vastly more storage capacities than the rockets that took men to the Moon in the late 1960s and early 1970s. In fact, IBM just announced the development of a x86 computer the size of a grain of salt, or 1 mm^2, containing 1 million transistors and costing under $0.10 to manufacture [3]. While data collection and storage are no longer challenges to remote sensing, the processing and analysis of these large data sets to glean relevant actionable knowledge, remains an obstacle to the exploitation of this information [4,2].

In this expanded environment of more data and sensor capabilities, there is a seemingly growing trend of U.S. intelligence failures, notably the events of September 11, 2001 (9/11) and the assessed weapons of mass destruction prompting the invasion of Iraq [5]. An obvious question is how could such intelligence failures occur with more advanced collection capabilities? Several researchers [6,7] have highlighted limited formalized intelligence analysis education and training and a need for the "professionalization" of intelligence analysis. It could also be argued that intelligence analysis is no longer the sole purview of the national security organizations, similar techniques are used in the private sector for competitive advantage assessments and the more obvious law enforcement and security [7]. Although the customers and resulting intelligence products might differ, there are similarities in processing given the availability of data and both sensor and analysis tools. This paper will explore the challenges in the evolution of the intelligence analysis task through a discussion of analysis processes, formal intelligence analysis education and training, evolution remote sensing and new technologies, and discuss current analysis training and research applications.

*sonya.mcmullen@erau.edu; phone 1 619 403-0611

Next-Generation Analyst VI, edited by Timothy P. Hanratty, James Llinas, Proc. of SPIE Vol. 10653, 106530S · © 2018 SPIE · CCC code: 0277-786X/18/$18 · doi: 10.1117/12.2304885

2. INTELLIGENCE ANALYSIS PROCESSES

The notion of intelligence analysis has expanded and is, "taking place in more contexts than ever before" [1]. The U.S. Intelligence Community (IC) has an established history of successes as well as what might be defined as failures in assessing complex source data to develop inferences [1] and the model of the processes are depicted in Figures 1 and 2. For the national security application, the role of the intelligence analyst is well defined and established and is a separate function from the decisionmaker [4]. In the intelligence cycle (Figure 1), the roles of each are fixed with the customer driving requests for information and the analysts working within the IC to develop collection and tasking requirements and conducting the analysis for reporting [8]. A more streamlined cycle is depicted in Figure 2 illustrating the continuous nature of the process as it is argued that the products provided by analysts to decisionmakers and the eventual action or inaction on that information results in the need for additional data collection and analysis as the intelligence process influences future events [5]. For instance, had the IC or law enforcement detected and prevented the events of 9/11, a new set of circumstances likely would have been placed in motion requiring additional collection and analysis of intelligence [5].

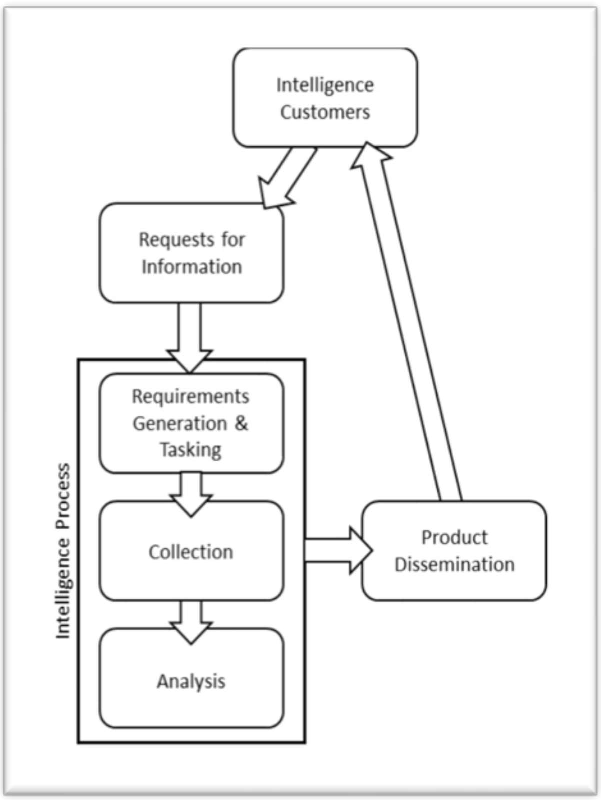


Figure 1. The intelligence cycle as defined by Berkowitz and Goodman [8].

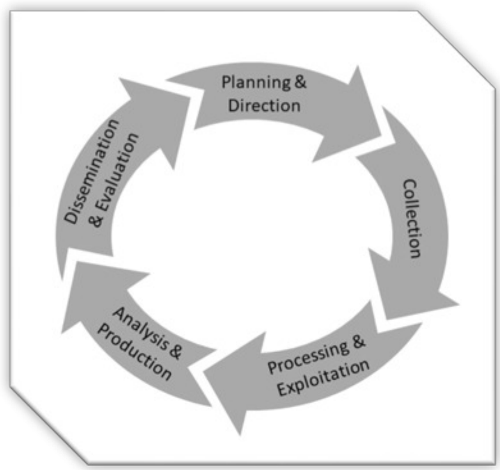


Figure 2. A modern depiction of the intelligence cycle derived from the Central Intelligence Agency [9] and PSU Department of Geology [10].

In order to evaluate quality of analysis processes and resulting products, measures of quality and accuracy are required, and this is the source of some debate among intelligence experts. First, the end goal of the analysis must be identified. Marrin [5] argues three overarching goals of intelligence analysis: accurate assessments with which to make decisions, preventing surprise, and influence on policy. Garst and Gross identified key pitfalls for seemingly failed intelligence, "biases, stereotypes, mirror-imaging, simplistic thinking, confusion between cause and effect, bureaucratic politics, group-think, and host of human failings" [5]. Intelligence experts have argued that the historical methods of analysis with emphasis in subject matter expertise with little foundation in advanced methodologies such as statistical analysis result in inferior products to those developed using regression or even random chance [6].

Prior to the advent of the Information Age and the expanded available information, both from IC collection methods as well as open sources, intelligence experts promoted the need for formalized education and training in analysis techniques. The recommendations include hypotheses development and testing, theoretical and methodological preparation, social science analysis techniques, information validation and synthesis, and statistical methods [6,5,7,1]. These methods are not dissimilar from formal research methods taught in many undergraduate and graduate programs in the United States. Figure 3 is a simplified illustration of business research methods described by Cooper and Schindler [11]. While Figure 3 includes additional steps in the process, the model is very similar in nature to the intelligence cycles depicted in both Figures 1 and 2, specifically the formulation of the problem and research questions, data collection, analysis, reporting, and dissemination of results. Revisions of the graphic for specific quantitative and qualitative methods as compared to current intelligence processes would further exemplify the similarities, specifically with exploratory research tools.

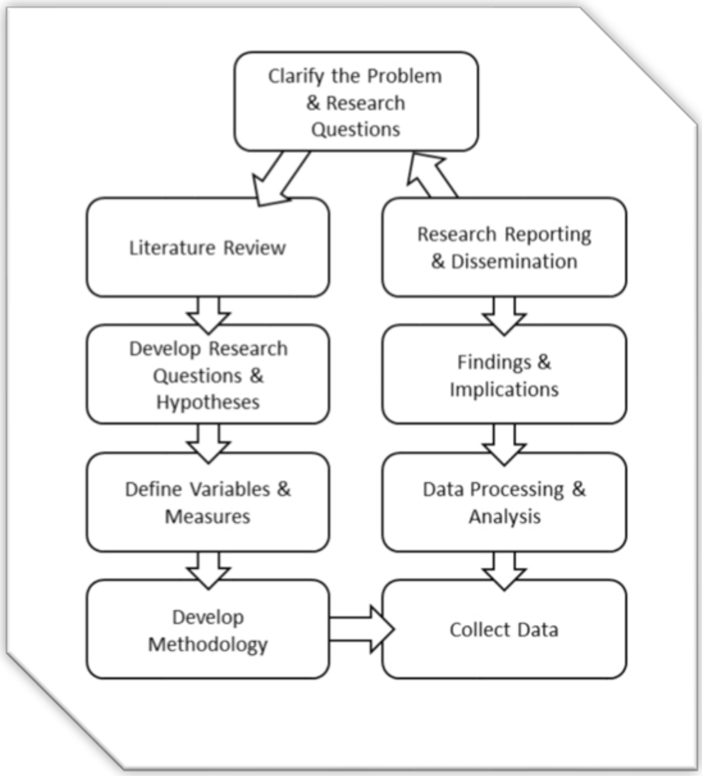


Figure 3. Modified illustration of Cooper and Schindler's [11] research process model for business researchers.

3. FORMAL INTELLIGENCE ANALYSIS EDUCATION AND TRAINING

Prior to the events of 9/11, much of the intelligence analysis education and training consisted of in-house programs conducted by the IC agencies themselves and were heavily reliant upon on the job training. Landon-Murray's [6] survey of intelligence-focused education programs included a detailed Appendix highlighting existing security and intelligence programs available at the time. Since that publication, additional programs are available to include undergraduate degrees such The Pennsylvania State University's Security and Risk Analysis Bachelor's Degree which features a concentration in Intelligence Analysis and Modeling [12] and Embry-Riddle Aeronautical University (ERAU) Worldwide [13] Bachelor of Science in Aviation Security featuring 18 credit hours in Security and Intelligence. In addition, the National Intelligence University offers opportunities for IC, law enforcement, and military personnel to conduct research at the Top Secret-Special Access levels as part of a degree program [14].

Geographic information systems (GIS) are ever valuable tools for planning and analysis and graduates of GIS degree programs are employed in web application design, environmental monitoring, transportation, and a host of other jobs [15]. Beyond traditional applications of GIS techniques for land use and environmental monitoring to include forest fire modelling [16], there is growing interest in spatial-temporal storytelling for intelligence analysis[17]. Police departments around the globe are increasingly using GIS-based analysis tools to track events to become less reactive and increasing proactive in identifying events to "reduce, disrupt, and prevent crime [18]. Allen, Tsou, Asiam, Nagel, and Gawron [19] are exploring the use of GIS methods to exploit Twitter data to track the spread of influenza. As a result of these expanded analysis capabilities of these techniques, there has been a trend of increasing GIS degrees offered and conferred, most notably at the Masters' degree level [20]. Table 1 is a summary of undergraduate GIS programs that specifically highlight remote sensing as part of the degree program and Lukinbeal and Monk [21] provide a detailed graduate-level program outline. The data in Table 1 also highlights programs with extensive sensor and remote sensing analysis curriculum. The proliferation of remote sensing data from a variety of sources, to include personal devices (discussed in more detail in the COTS Sensors and Tools) lends itself to additional training and education to effective exploit such sources of data. Further, applications such as active fire monitoring and hotspot detection [16] and agriculture, to name a few that are rapidly being aided with sUAS sensors, could be enhanced be GIS analysis techniques.

Table 1. Summary of U.S. undergraduate GIS degree programs with emphasis in remote sensing.

Bachelors-Level GIS Degrees	University
Geographic Information Systems (GIS)	Mississippi State University
Geospatial Technologies	University of Vermont
Geographic Information Science	SE Missouri State University
Location Intelligence	North Park University
Social and Behavioral Sciences – GIS Concentration	California State University Monterey Bay
Graphic Information Science	North West Missouri State University
Geography, GIS Concentration	University of Central Arkansas
Geography – GIS and Analysis	California State University Sacramento
Geography with Emphasis in GIS	University of Missouri Columbia
Geography – GIS	Central Connecticut State University[1]
Environmental Studies – Planning and GIS Concentration	Richard Stockton College[2]
Geographic Information Science	University of Northern Alabama
Geographic Science – Applied Geographic Information Science	James Madison University
Geography – Applied Geographic Technology	Slippery Rock University
Geography – Geographic Information Science Emphasis and Remote Sensing	University of Cincinnati[3]
Geographic Information Science	Ball State University
Geography - Information Science	University of West Georgia
Geospatial Sciences	Metropolitan State University of Denver
Geology, Geographic Information Analysis Concentration	Austin Peay State University
Geography, Geographic Information Systems (GIS)	California State University Los Angeles
Geographic Information Science	Michigan State University[1]
Geography, Geographic Information Science	California State University Northridge
University Studies, Concentration in Geographic Information Science and Technology	Texas A&M[2]
Geography, Geographic Information Technology	University of Southern Mississippi[1]
Cartography and GIS	Salem State University
Geography – GIS	Bemidji State University[2]
Geographic Information Systems (GIS)	Illinois State University
Geographic Information Science and Technology	Mansfield University
Environmental Studies – GIS and Spatial Analysis	Gettysburg College[2]
Geographic Information Science	Texas State University
Geography, Geographic Information Concentration	Kent State University
Geographic Information Systems (GIS)	Auburn University

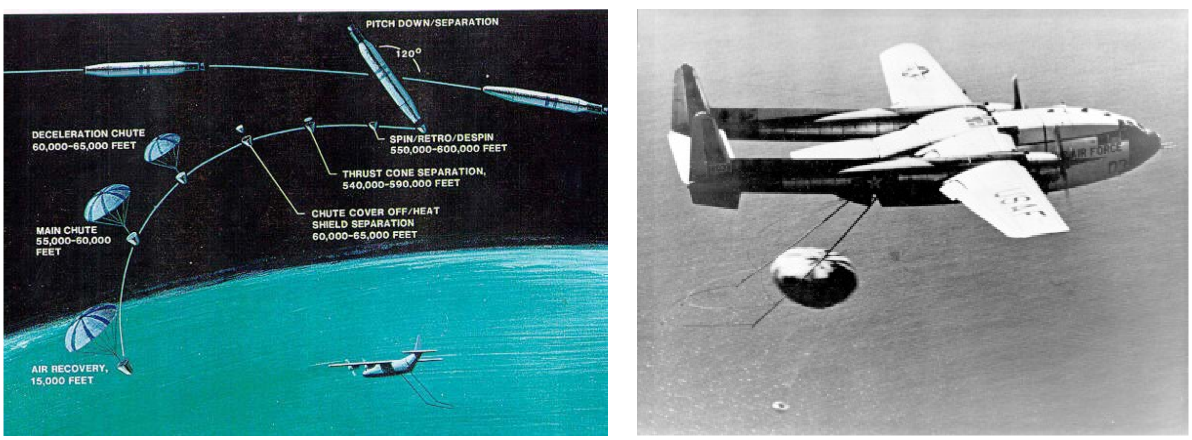
Notes: This table is not inclusive of all available degree programs.
[1] Indicates Advanced Remote Sensing.
[2] Indicates Interdisciplinary degree
[3] Remote Sensing is a separate major

The formation of new degrees and transformation of other traditional disciplines to employ complex analysis tools, such as in the geography, as well as the reduction in "stove-piped" research and curriculum is indicative of the needed transformation of the intelligence analysis. Historically, intelligence analysis been a cognitive task assigned to humans because it could not be automated due to computing limitations [4]. The model of analysis placed the human in the center between data and the development of products for the customer as well as archiving of the information [4]. The advent of enhanced and low-cost computing has effectively changed that equation not only providing the potential to automate some aspects of the cognitive task but in a sense requires a level of automated processing due to the expansion of the datasets [4]. As a result, the methods of analysis and functions of the analysts must evolve with the technology [22,4]. Hence, the role of the analyst has been removed in part from much of the initial data processing such as searching, data of interest will more often be "pushed" to the analyst, and products to the analyst will include maps, timelines, and other visualizations [4]. This process will be reliant on existing and future data fusion tools to exploit both hard and soft sensor data.

As a result of these processing capabilities, the role of the analysts will evolve beyond situational awareness-type problems to more complex intelligence tasks such as the development of forecasts [4]. It is also likely that analysts will work more collaboratively with customers to determine the decision-problem and other opportunities for exploitation [4]. These tasks will require additional skills, not only with the use of current and future tools, but also an understanding and command of methodologies to generate theories [4]. These skills are more indicative of higher-level researchers.

4. HISTORICAL REMOTE SENSING DEVELOPMENTS

The transformation of data availability, particularly that of sensors begs for a brief review of remote sensing capabilities. Historically, remote sensing has been the purview of government agencies, with military mission and national security needs driving many requirements. From the airborne and space-based sensing perspective, it could be argued that the first remote sensing operations were the use of balloons for observation during the French Revolution [23]. The development of fixed-wing aircraft permitted aerial mapping and continued reconnaissance by air [23]. The vulnerability of manned aerial reconnaissance missions in part drove the development of early space-based surveillance satellites as well as the eventual employment of UAS [24,25]. Early imaging and remote sensing included significant delays for the return of film from these platforms and as well as human processing times. The CORONA satellite program, the United States' early space-based reconnaissance platform was reliant on film cannisters dropped from orbit and retrieved in flight by airplane [24] as shown in Figure 4.



(a) Film ejection trajectory (b) Airborne film cannister capture

Figure 4. CORONA film capsule retrieval.

In addition to the multitude of military and national security applications, remote sensing has been used to explore and sense the Earth and other planetary bodies in the Solar System. For the purposes of this discussion, Earth remote sensing is the focus. Well prior to the first the successful orbital rocket launch, a patent was issued in 1891 to the Ludwig Rahrmann of Germany for a rocket-propelled imaging system recovered by parachute [26], although the first orbital test flights of US early reconnaissance missions (Discoverer Program) did not begin until 1959[24].

The Landsat program is perhaps the longest and most storied Earth sensing program with eight satellites launched to date and over 3200 peer-reviewed publications resulting from the data since 1972 [27]. One of the most notable techniques developed for the National Aeronautics and Space Administration's remote sensing programs is hierarchical segmentation (HSEG) which has been employed in diagnostic imaging to aid in the detection of breast cancer as well as for data mining tasks [28]. The value of sensor data in a variety of applications cannot be understated nor the potential to transfer techniques from one detection and analysis task to another. Currently, multiple companies from Google Earth to TerraServer share new images of the Earth on a daily basis.

Much of the collected remote sensing data is more readily available through global communications networks, and to the general public, allowing for more timely application of data gathered. In the case of sUAS, often imagery and sensor data can be viewed in real-time while the aircraft is in flight with the use of WiFi and tablets or augmented reality googles. Figure 5 illustrates multiple real-time sensor images of an aircraft propeller and aileron. The sUAS operator is viewing the sensor data as it is being collected. This is an experimental application for aircraft crash search and rescue operations and National Transportation Safety Board (NTSB) investigators to locate key components of the wreckage for analysis. The authors are collaborating with the NTSB on sUAS data collection and analysis of crash sites.

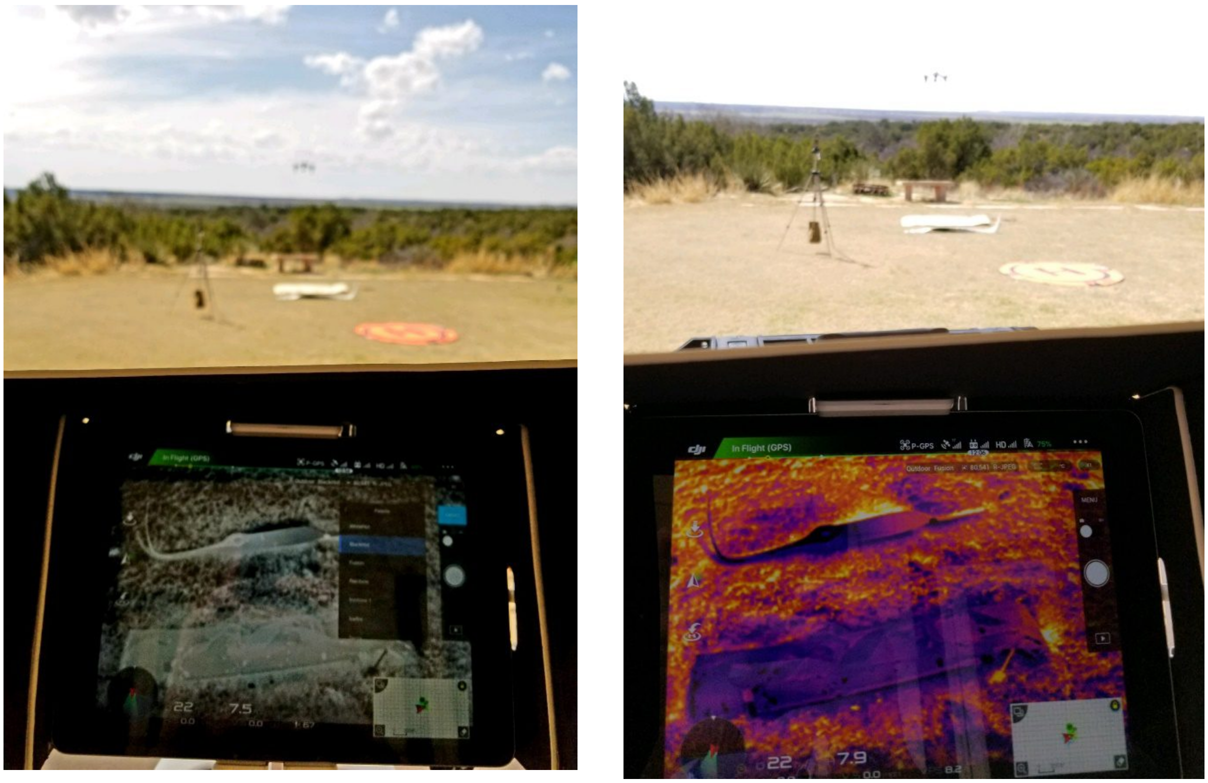


Figure 5. sUAS tablet display of real-time sensor data while the vehicle is in flight.

In 2016, a team at Embry-Riddle Aeronautical University Worldwide College of Aeronautics conducted an analysis of commercially available sUAS and produced a consumer guide. This guide rates sUAS on criteria such as operational ease, speed, endurance, payload capacity, camera quality, communications range, performance metric data availability and accuracy, price, and user support [29]. All aircraft were under $3500 total cost for equipment and case, with the mean of the evaluated systems calculated to be $839.84 [29]. Potential sensor payloads included GoPro cameras; integrated cameras, some of which are high definition (1080p) or ultra-high definition still and video cameras; and GPS live mapping [29].

Mount kits as well as 3D printed mounts can be employed to add sensor packages to sUAS. Sentera [30-36] offers drone sensors with a wide-range of capabilities to include 3D imaging, red edge, multispectral, normalized difference vegetation index (NDVI), and normalized difference red edge (NDRE), as well as an incident light sensor to compensate for sunlight variations. These sensors can be mounted on common sUAS platforms such as the DJI Phantom, DJI Inspire, DJI Solo, or Sentera platforms, among others [30-36]. Common uses of the sensors include agricultural monitoring, field scouting, geographic volume measures, and surface data analysis [30].

Recently, numerous high school and university teams have successfully launched and recovered high-altitude balloons. These balloons typically cost under $1500 total and have been used to capture images of Earth's curvature with reasonably high resolution. Dr. Henderson lead a team of undergraduates who successfully launched a balloon to over 104,000 ft. Figure 6 shows a typical image of Earth from a balloon.

Figure 6. Earth as seen from a high-altitude balloon.

5. COMMERCIAL OFF THE SHELF SENSORS AND SOFTWARE

As previously highlighted with the Sentera [30] and fire hotspot detection [16] examples, commercial sensors are becoming more widely available and employed. Additional enabling technologies and reduced cost components such as data storage [4], computing capabilities, smart and programmable sensors, and commercial analysis software serve to produce extraordinary data generation and analysis potential. Very low-cost but capable computers such as Raspberry Pi and Arduino offer opportunities for students, hobbyists, and professional to use open-source software tools to integrate sensors, input devices (mouse, keyboard), and displays to construct devices to control cellphone cameras; control smart home systems; detect earthquakes, radiation, and flooding; just to name a few [37]. Initial Raspberry Pi and Arduino computers can be purchased for under $50 and sensor kits are similarly priced.

Arguably these tools have reignited interest in programming, as both the Raspberry Pi and Arduino, were designed to do. Production of Raspberry Pi devices was 15,000 per day in 2017 alone with over 14 million sold [38]. Arduino has been used for fast prototyping and has gained additional attention as an education tool [39]. The Creative Technologies in the Classroom 101 (CTC 101) is a course developed for secondary education Science, Technology, Engineering, Arts, and Mathematics (STEAM) programs [40]. The program has been used in 735 schools with 90% of students surveyed indicating interest in additional programming education [40]. Another example is the development of a cyber-intrusion detection system using an Arduino developed by students in the Master of Professional Studies in Information Sciences at The Pennsylvania State University World Campus, demonstrating the potential of these capabilities, particularly for distributed teams.

Software such as C, Python, and Linux can be used to program a number of devices to include the Raspberry Pi and Arduino. Python and Linux are both open-source and available for no charge to users. Linux is promoted by the developers of Raspberry Pi as the features include cybersecurity as well as safeguards to ensure new programmers do not override critical device programming, making for an enhanced and low-threat learning experience [41]. Several books [42] and online tutorials are available for newer programmers for each of these tools. In addition, full project [37] examples can be found on the internet and in published texts to aid in skill development.

MATLAB is a tool long been used in engineering applications for simulation, modeling, algorithm development, visualization, and more. An additional feature of MATLAB is the variety of add-on tools, some of which are demonstrated in the Image Processing in MATLAB section, to expand the functionality to include applications for multisensor data fusion [43]. Recognizing the value of the tools and associated skill-sets, Embry-Riddle Aeronautical University has recently expanded access to all faculty and students in order to integrate the tool into more areas of the curriculum, beyond the traditional engineering applications. For example, MATLAB can be used for inventory analysis, employee turnover tracking, financial planning for investment and purchase versus lease decisions, beyond obvious engineering and scientific applications [44].

Additionally, there are several software options available for sensor imaging analysis and developed to function with sUAS sensor data collection platforms. Two tools will briefly be discussed here: Pix4D and Agisoft Photoscan. Pix4D is imaging processing software aids in the development of 2D and 3D images using multiple frames captured via from sUAS operations. Additional capabilities include Pix4D mapper to produce georeferenced images from a variety of sensors to include LIDAR and multispectral and features image integration from multisource cameras; aerial crop analysis; and site schedule comparisons [45]. Pix4D is a tool used by ERAU's sUAS Flight Department to teach students to process and analyze captured data. Agisoft Photoscan is a commercial GIS tool used to build 3D models from digital images to produce spatial data for analysis [46]. While Agisoft could be used in conjunction with sUAS imagery, applications are not limited to that source as a variety of source data could be processed with this software.

These are just a few examples of the available hardware and software to collect and process complex data. The next two sections will explore such applications in detail, particularly as they relate to existing education and research initiatives at Embry-Riddle Aeronautical University.

6. sUAS OPERATIONS AND SENSOR DATA ANALYSIS

Small UAS operations are increasingly used for a variety of private sensing activities. These activities include real estate sales, urban planning, infrastructure inspections, fire fighting operations, security, search and rescue, and post-disaster surveys, to name a few. Commercial UAS systems currently available offer inexpensive platforms from small sensors with some sensors embedded into the airframe. In addition, these aircraft require little training to operate, although arguably more training is preferred to ensure safety. Although several of the sUAS manufacturers provide software for sensor data processing, there is often little formal training for the tools to maximize analysis capabilities. The following subsections describe the sUAS training for data collection offered through Embry-Riddle Aeronautical University Worldwide College of Aeronautics UAS Flight Program.

6.1 Data Collection

In the analysis of an intended target for collection, there may be numerous software programs available to process collected data, depending on the purpose of the flight. In this light, pilots and analysts must be versed in a myriad of commercial off-the-shelf (COTS) systems, sensors, and software. Selection of a best-case use platform and sensor is essential to fly a mission and collect data that can be efficiently processed and effectively used in a timely manner. In general, most pilots are not well trained on analysis tools, and those who are typically do not have training in the mathematical rigor of analysis tools.

Pre-flight planning processes are a time-honored practice from the earliest days of manned aviation and essential to building efficient collection and effective results that present actionable information. These processes, if operationalized, enable the data assessment team and flight crews to rapidly assess the situation and get to work. Operationalizing the process brings best practices and tailoring to the needs of the organization and their missions. Once this occurs, pre-flight planning times are significantly reduced sometimes to minutes and collection efforts are organized.

Mission operations occur in a timeline that is limited by platform capabilities. Lithium-Polymer battery technology in COTS systems is improving. Flight times are platform dependent, however a typical multi-rotor system can operate in flight times less than 60 minutes. Many variables exist. Missions can be rapidly loaded and flown, usually within a few minutes provided data capabilities and GPS coverage are present (research is being conducted by many institutions, including ERAU, on GPS degraded or denied navigation,). Important to note is that some COTS systems are becoming more advanced in that the manufacturer is embedding restrictions to flight based on location and airspace. As an example, today if an Asiana type accident occurred, the use of any DJI product would be restricted unless prior waivers were completed. This is something emergency services entities can pre-waive however. Once a mission is complete, the best practice of data validation while still on site for accuracy and acceptability is vital. Should the data not provide needed results, the mission may very well need to be re-flown. Depending on the systems used, this task could be completed in minutes.

6.2 Data Processing

At the point where collected data is deemed accurate and functional, the analysts assume control and begin processing. A caveat may be that a pilot and analyst may be the same person depending on the size of the organization and availability of properly trained individuals. The data must be more finitely scrubbed to remove derelict imagery and errant data. Much of this task may require human, rather than a software algorithm review and these processes will likely continue to evolve.

With the data through initial analysis, it must then be compartmentalized for processing. Approved data can include hundreds of images and associated flight log information and segregating it for rapid processing is essential. Analysts then establish the software processing parameters to manipulate the data as needed to obtain accurate and applicable results. Within approximately 20 minutes of pulling data from an aircraft, the data processing can be started. Running processes can take between several minutes with low resolution imagery, to several hours for more complex data depending on the software used. As an example, regarding still images using a Zenmuse XT FLIR Camera (any resolution), can take just minutes to open and use instantly. The Parrot Sequoia multispectral sensor can take the same amount of time reviewing still shots or if flown autonomously and depending on the size of the flight pattern and images collected, can be processed rapidly and analyzed in less than 30 minutes. Finally, if a high amount of detail is needed, high resolution RGB imagery over a larger (several acre) site, can take several hours to process. Only then can a detailed analysis be engaged.

6.3 Lessons Learned

Regardless of COTS system and sensor package, one can identify that processes are essential for problem design and data collection for a timely, efficient, and applicable mission. Further, it follows that an evaluation of benefits and challenges for each mission profile should occur. This is not an evaluation in a vacuum, but with the team involved. Teams should collectively identify the successes and challenges and follow a specific process previously developed. Workflows must be validated for efficiency and accuracy. For issues needing resolution, it is important to briefly identify what the issue was, a discussion of factors surrounding the issue, and finally a recommended solution. From this process, changes to procedures must then be implemented in a continual process improvement. When recommendations for change include application of resourcing, leadership must be present to fully comprehend limitations that could negatively affect future operations. This process should be a detailed assessment with an estimation of cost to the operation both internally and externally and always include the human factor, and safety.

7. IMAGE PROCESSING USING MATLAB

Photo enhancements (such as filters or Photoshop effects) have thrived, in many cases due to the rise of social media and smartphones. However, many users have little to no understanding of the mathematical underpinning of the algorithms, and simply look to visually enhance photos. The application of rigorous analysis tools to image sets has not been widely exploited, whereas the proliferation of inexpensive sensors has. The authors at ERAU have developed, and applied, image processing techniques in MATLAB to the images shown in the examples below. Recently, a lab has been set up at ERAU specifically for virtual and physical analysis of image sets. The Multi-Spectral Sensing and Data Fusion (MSSDF) Lab is a joint effort between multiple campuses and departments at ERAU. The following examples are results from ongoing research within the MSSDF, using MATLAB and student projects.

Figure 7 shows a Falcon 9 upper stage as imaged from a 10-inch telescope. The general shape is obvious, and the object is lightly resolved (i.e., not a point source). The center of mass of the object is estimated along with the major axis, as shown, using MATLAB's Image Processing Toolbox. Many spacecraft and rockets travel along the line of the major axis (rocket bodies are long and thin), therefore, there is a high likelihood that the velocity direction is also identified. The method applied was center of mass centroiding based on the image intensity per pixel. The major axis length and direction was identified by computing the second central moment.

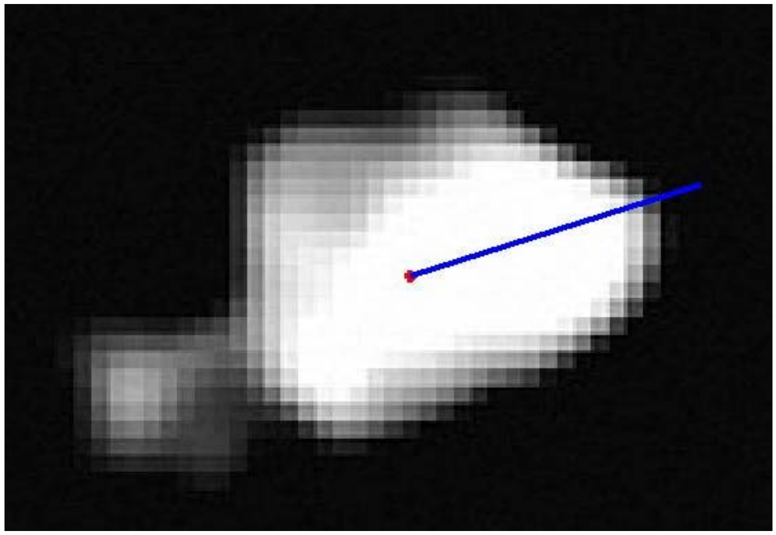


Figure 7. Falcon 9 upper stage

Figure 8 shows power lines autonomously identified in an image taken from a UAS at an altitude of 100 feet. The original, color image is first converted to gray scale. The Canny edge detection [1] method is applied to find edges via local maxima of the image gradient. Line segments are then extracted via a Hough transform [2]. In this case, *a priori* knowledge that the power lines spanned the image was applied to only highlight the lines longer than the height of the image, as shown. This method could be applied for disaster recovery, utility inspection, or autonomous navigation.

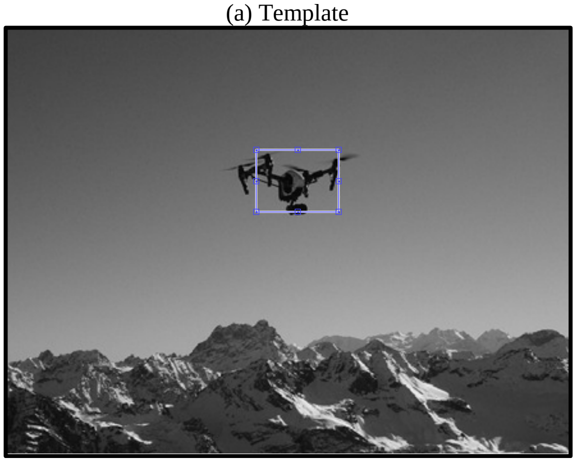
(a) Original image (b) Identified power lines

Figure 8. Identification of power lines

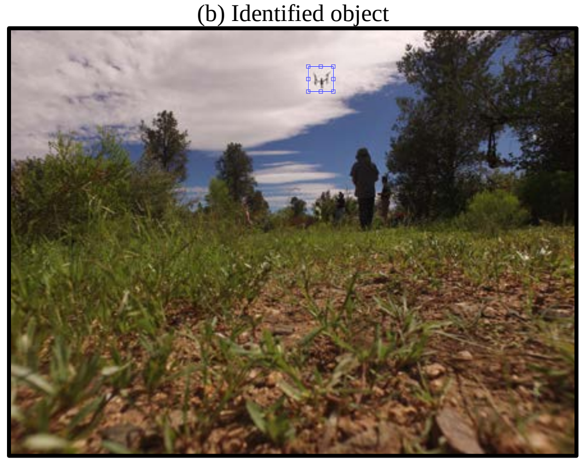
In the image sequence below, a single image was used of a DJI *Inspire*, collected from product literature, which served as a template for finding similar objects in a set of images. Figure 9 shows the template, and three test images, all with successful identification of the UAS. In order to accomplish this, the normalized cross-correlation [3] of the template and the current image were computed. This process takes into account the potential scale difference and did not matter if it was performed on the color or grayscale image. The maximum correlation was identified, which then defined the bounding boxes in the original image where the object should be located. It should be noted that the purpose of this task was to identify if the *Inspire* was in an image or not, and to locate the area of interest within the image.

(a) Template

(b) Identified object



(c) Identified object



(d) Identified object

Figure 9. Template matching to find UAS in imagery

Perhaps a more challenging problem is to identify, and then track, an unknown or uncooperative UAS in flight. Figure 10 shows the application of feature points, using the features from accelerated segment test (FAST) [4]. Using a cluster of feature points, and windowing around the clusters, unknown objects of interest are identified. Figure 10 shows an air-to-air encounter where the targets were a DJI *Inspire* (left) and a DJI *Phantom 4* (right), and the background clutter is minimal (i.e., mostly sky). The application to scenes with more complex scenes is still ongoing.

Figure 10. FAST feature points on two unknown objects

By windowing around the feature clusters, we can now assume a number of objects, and identify the centroid of them by computing the center of mass. Figure 11 shows color image, which was the same shown in Figure 10, and computes the object centroid in two ways. On the left, the *Inspire* window is converted to grayscale and the centroid identified. On the right, the *Phantom 4* window is converted to black and white (binary) and the centroid is computed, simply to show multiple versions of analysis capability. It should be noted that the importance of identifying the centroid of the object is in tracking the object. Many algorithms for tracking and intercept/avoidance rely on the dynamic motion of the centroid according to Newton's Laws of motion.

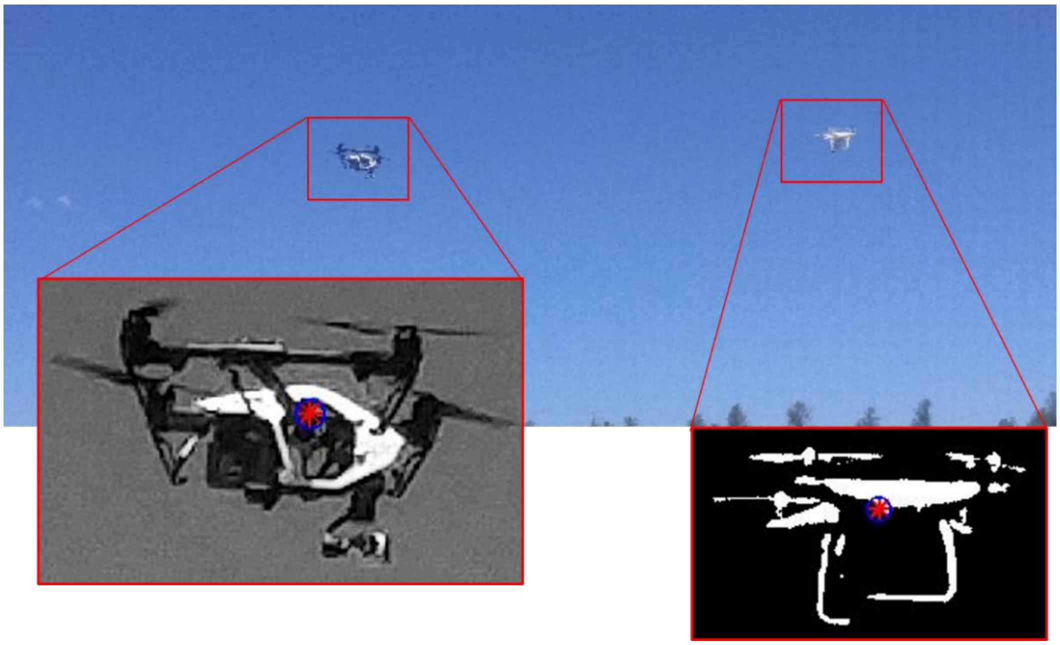


Figure 11. Identified centroids of unknown objects

Figure 12 shows that features can be identified, using FAST, in infrared (IR) images. The proliferation of inexpensive IR cameras, particularly on UASs, necessitates this scenario. Notice the ground and utility poles are in the image, but the strongest features are on the airborne sUAS itself.

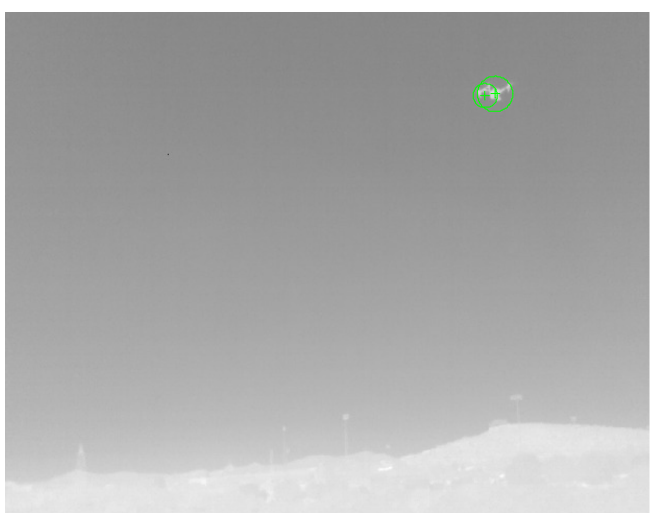


Figure 12. FAST features of UAS identified in IR imagery

Another method of identifying areas of interest, particularly in grayscale or IR imagery, is the use of a quadtree. The quadtree algorithm [5] first computes the gradient of the image and then divides the image into four equivalent squares (which may be generalized to rectangles). The gradient within each of these squares is compared within the region. If the standard deviation is within a user-specified tolerance, then the region is left alone. Otherwise, the square is sub-divided into four more squares. This process continues until the square sizes reach a minimum level. Figure 13 shows the same IR image from Figure 12, and the sUAS is easily identifiable. In this image, there is a non-working pixel (registers as black) near the left edge that is also identified.

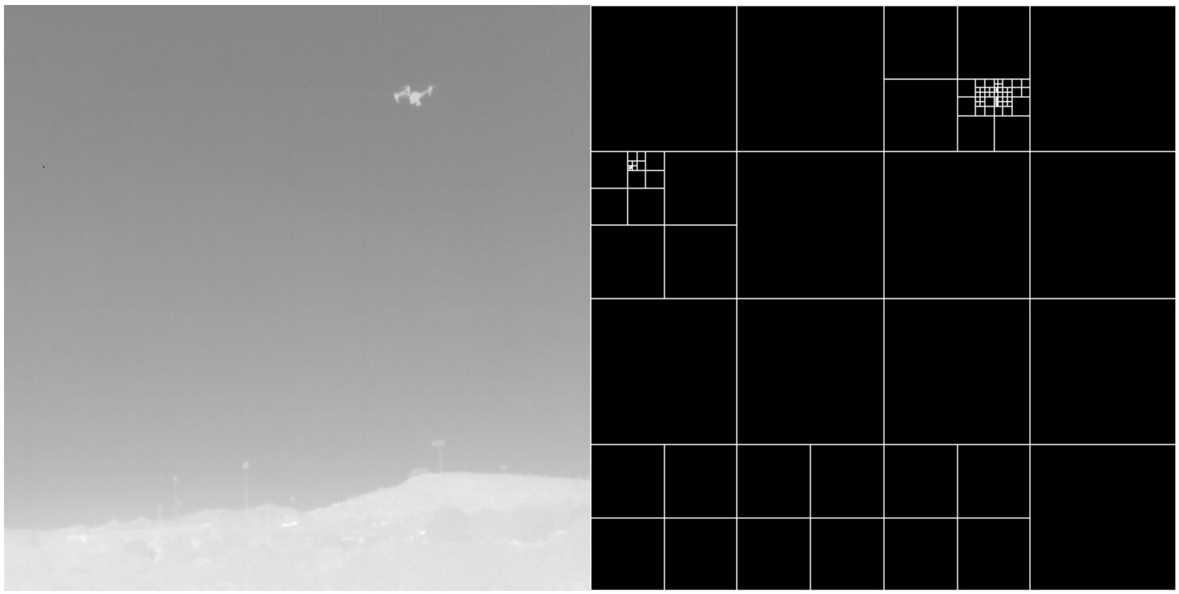


Figure 13. Quadtree deconvolution of IR image, identifying a UAS and zero pixel

Recent advances in machine learning and deep learning techniques have allowed for neural networks to be trained to identify and classify objects in a set of images. Neural networks have become increasingly popular for their reduction in deliberate and intentionally outlined pattern-recognition algorithms and domain expertise that was required for typical machine learning techniques. This allows for complex learning algorithms to be developed in an effective, timely, and inexpensive manner, while improving the accuracy for a variety of operations [6]. Where a typical machine learning algorithm has a few layers of internal operations, a deep-learning algorithm has many hidden layers of internal operations as shown in Figure 14.

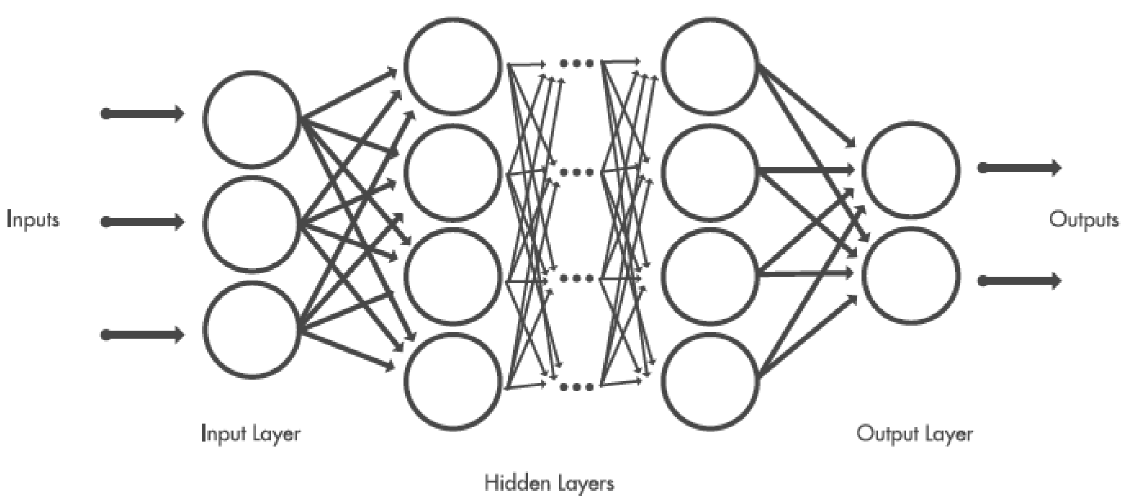


Figure 14. Internal structure of neural network [7]

The convolutional neural network (CNN) is one of the most common deep-learning algorithms using a series of convolutional filters for feature extraction making it popular with image processing applications. The inner hidden layers of a CNN extract key features of the data set for classification by the final layers. These hidden layers are comprised of a series of functions to extrapolate all the relevant image data from the data set to accurately classify the image or image component. This series is comprised of a convolution to stimulate significant aspects of the image, pooling to reduce the parameters the network must learn, to facilitate faster and more effective training. This autonomous process available within most CNNs is what distinguishes this algorithm from its counterparts in machine learning. AlexNet, a pre-trained CNN available in the MATLAB Deep-Learning Toolbox [7], was used to extract and identify features from a spacecraft model as shown in Figure 15.

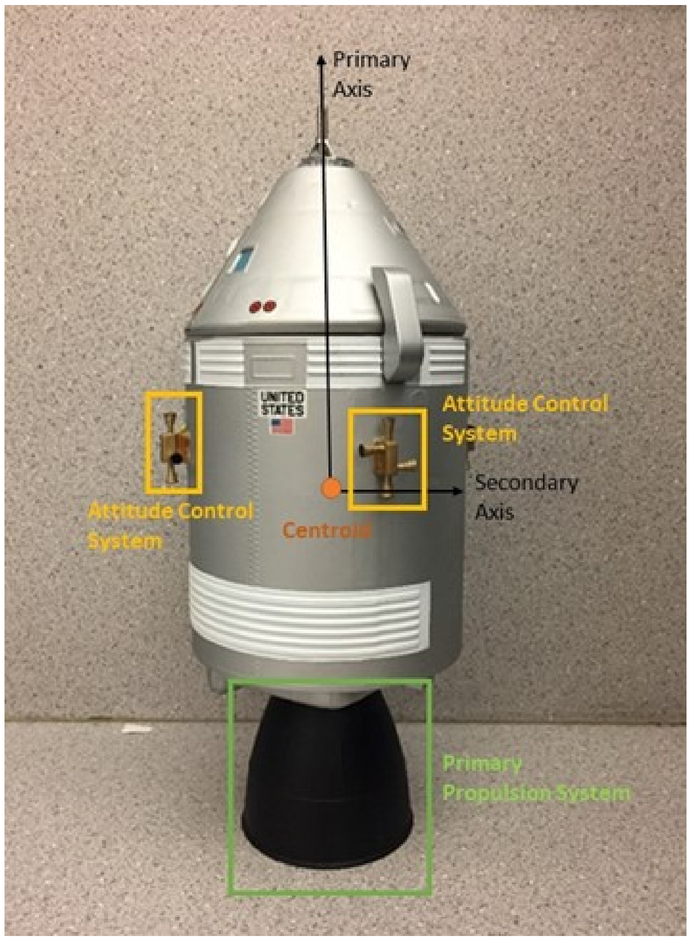


Figure 15. Detected and identified components of spacecraft model using CNN

Imaging through degraded atmosphere (e.g., fog, rain, low-light) has presented several issues for current, state-of-the-art imaging systems, especially when using only passive sensors (i.e., those not providing illumination). Relevant issues include surface and feature detection, the need for active illumination sources, and efficacy across a variety of obscuration and lighting conditions. Clouds and fog are composed of relatively large water droplets, larger than the wavelengths of visible light, and therefore, all wavelengths in the visible spectrum are scattered very efficiently by these large particles so that very little direct, visible wavelength sunlight reaches the observer. Infrared (IR) systems, particularly thermal IR, do well in detecting objects under such conditions but images are typically limited to a rough shape and contour identification. These images present little or no information contained within a surface/terrain. Preliminary research has been performed using a commercial sensor under laboratory simulated fog conditions. Using a proprietary differential polarization scheme and standard image processing algorithms, the lab results were able to reveal shape (IR) and surface marking details (polarization) in obscured/foggy conditions under ambient light as shown in Figure 16. The images were collected and processed in MATLAB. It should be noted that this was an oil-based fog machine, and as such the results will likely be different from atmospheric fog.

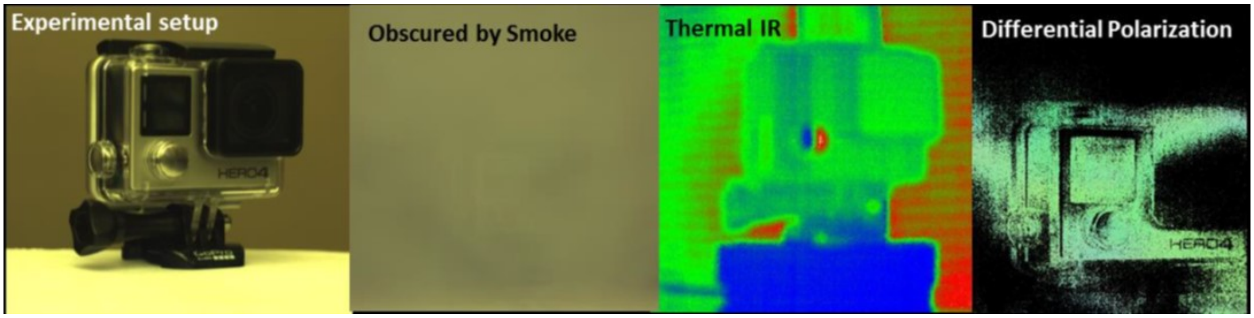


Figure 16. ERAU imaging through fog simulation

8. CONCLUSION

The affordability of sensors, sensor platforms, computing power, data storage, and advanced software has opened new doors for data collection, analysis, and exploitation. No longer are large government and private organizations the sole user or necessarily leader in data analysis. The applications for data and intelligence analysis using open-source data and these new capabilities is virtually unlimited. The challenge today is to modify the traditional notions of learning and training from the human being retaining large amounts of data, to the human becoming a modern analyst of data and maximize their employment of current and future tools.

9. ACKNOWLEDGEMENTS

This research has been funded in part by a FIRST grant from Embry-Riddle Aeronautical University. The authors thank LtCol (Ret.) Brian Bauries, U.S. Air Force for the loan of multiple aircraft components used in this and follow-on research. The team thanks Mr. Adolfo Ibarra, Dr. Erin Bowen, and Mr. Johnny Young at the ERAU Prescott, AZ Campus Robertson Aircraft Investigation Laboratory for the use of the lab and assistance in the initial sensor testing activities.

REFERENCES

[1] McNeese, N.J., Hoffman, R.R., McNeese, M.D., Patterson, E.S., Cooke, N.J. and Klein, G., "The human factors of intelligence analysis", Proceedings of the Human Factors and Ergonomics Society 59th Annual Meeting, 2015.

[2] McMullen, S.A.H., McMullen, M.J., Forster, P., Ison, D. and Clark, P.J., "Emergency management: Exploring hard and soft data fusion modeling with unmanned aerial systems and non-governmental human intelligence mediums", Proceedings of the SAI Intelligent Systems Conference, London, September 20-22, 2016.

[3] Thubron, R., "IBM's 'world's smallest computer' is the size of a grain of salt", Techspot [Online]. Retrieved from https://www.techspot.com/news/73770-ibm-world-smallest-computer-size-grain-salt.html. 2018, March 19.

[4] Hare, N. and Coghill, P., "The future of the intelligence analysis task", Intel and Nat'l Sec, 3*1*(6), 858-870 (2016).

[5] Marrin, S., "Evaluating the quality of intelligence analysis: By what (mis) measure?", Intel and Nat'l Sec, 27(6), 896-912 (2012).

[6] Landon-Murray, M., "Social science and intelligence analysis: The role of intelligence education", Journal of Applied Security Research, 6(4), 491-528 (2011).

[7] Bruce, J.B. and George, R., "Professionalizing intelligence analysis", Journal of Strategic Security, 8(3), 1-23 (2015).

[8] Berkowitz, B.D. and Goodman, A.E., [Strategic intelligence for American national security], Princeton University Press, Princeton, NJ (1989).

[9] Central Intelligence Agency, "The Intelligence Cycle", Retrieved from https://www.cia.gov, (2013).

[10] Department of Geology, "Intelligence cycle and process", The Pennsylvania State University, University Park: PA, Retrieved from https://www.e-education.psu.edu/sgam/node/15, (n.d.).

[11] Cooper, D.R., and Schindler, P.S., [Business research methods, 10th ed.], McGraw-Hill Irwin, Boston, MA, (2008).

[12] College of Information Sciences and Technology, "Security and Risk Analysis (SRA)", The Pennsylvania State University, University Park, PA, Retrieved from https://ist.psu.edu/students/undergrad/majors/sra, (2018).
[13] Embry-Riddle Aeronautical University, "Bachelor of Science in Aviation Security", Embry-Riddle Aeronautical University, Daytona Beach, FL, Retrieved from erau.edu/degrees/bachelor/aviation-security/, (2018).
[14] Defense Intelligence Agency, "National Intelligence University (NIU)". Retrieved from http://www.dia.mil/Training/National-Intelligence-University-NIU/, (n.d.).
[15] Thomas, B., "Where are the jobs in the GIS field?" Tech Republic, Retrieved from http://www.dia.mil/Training/National-Intelligence-University-NIU/, (2013, June 25).
[16] Aden, S.T., Bialas, J.P., Champion, Z., Levin, E., and McCarty, J.L., "Low cost infrared and near infrared sensors for UAVs", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-1, ISPRS Technical Commission I Symposium, Denver, CO, (2014).
[17] Dos Santos, R. F., Shah, S., Boedihardjo, A., Chen, F., Lu, C.T., Butler, P., and Ramakrishnan, N., "A framework for intelligence analysis using spatio-temporal storytelling", Geoinformatica, 20, 285-326 (2016).
[18] Herchenrader, T., and Myhill-Jones, S., "GIS supporting intelligence-led policing", Police Practice and Research, 16(2), 136-147 (2015).
[19] Allen, C. Tsou, M.H., Aslam, A., Nagel, A., and Gawron, J.M., "Applying GIS and machine learning methods to Twitter data for multiscale surveillance of influenza", PLOS ONE, 11(7), 1-10 (2016).
[20] Revell, M., "Long-term trends in degrees conferred in geography", American Association of Geographers. Retrieved from http://www.aag.org/galleries/disciplinary-data/Longterm_trends_in_degrees_conferred_in_geography.pdf, (2017).
[21] Lukinbeal, C., and Monk, J. J., "Master's in Geographic Information Systems programs in the United States: Professional education in GIS and geography", The Professional Geographer, 67(3), 482-489 (2015).
[22] Odom, W.E., "Intelligence analysis", Intelligence and National Security, 23(3), 316-332 (2008).
[23] Lawrence, H., [Aviation and the role of legislation, 3rd ed.], Kendall-Hunt Publishing, Dubuque, IA, (2014).
[24] Day, D.A., Logsdon, J.M., and Latell, B. (Eds.)., [Eye in the sky: The story of the CORONA spy satellites], Smithsonian Institution Press, Washington, D.C., (1998).
[25] Taubman, P., [Secret empire: Eisenhower, the CIA, and the hidden story of America's space espionage], Simon & Schuster, New York, NY, (2003).
[26] Lillesand, T.M., and Kiefer, R.W., [Remote sensing and image interpretation, 4th ed.], John Wiley & Sons, New York, NY, (2000).
[27] National Aeronautics and Space Administration, "Landsat science: Practical Uses", Retrieved from https://landsat.gsfc.nasa.gov/about/applications/, (2018).
[28] National Aeronautics and Space Administration, "Hierarchical segmentation enhances diagnostic imaging", Retrieved from https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20080003922.pdf, (2007).
[29] Embry-Riddle Aeronautical University, "Small unmanned aircraft system consumer guide", Embry-Riddle Aeronautical University, Daytona Beach, FL, Retrieved from https://assets.erau.edu/cm1/Assets/worldwide/data/erau-suas-consumer-guide-june-2016-release.pdf, (2016).
[30] Sentera, "The world's best drone sensors", Sentera, Minneapolis, MN, Retrieved from https://sentera.com/sensors/, (2018).
[31] Sentera, "Sentera double 4K sensor: NDVI + NDRE", Sentera, Minneapolis, MN, Retrieved from https://sentera.com/wp-content/uploads/2018/01/NDVI-NDRE_Double4K_Ag_Lit4076A_WEB.pdf, (2018).
[32] Sentera, "Sentera multispectral double 4K sensor", Sentera, Minneapolis, MN, Retrieved from from https://sentera.com/wp-content/uploads/2018/01/SenteraMultispectralDouble4K_Ag_Lit4059B_WEB.pdf, (2018).
[33] Sentera, "Sentera multispectral double 4K inspection sensor", Sentera, Minneapolis, MN, Retrieved from Retrieved from https://sentera.com/wp-content/uploads/2018/02/Double4K_Inspection_Lit4022K_WEB.pdf, (2018).
[34] Sentera, "Sentera multispectral double 4K agriculture sensor", Sentera, Minneapolis, MN, Retrieved from https://sentera.com/wp-content/uploads/2018/03/SenteraDouble4K_Ag_Lit4023K_WEB.pdf, (2018).
[35] Sentera, "Sentera quad sensor", Sentera, Minneapolis, MN, Retrieved from https://sentera.com/wp-content/uploads/2018/01/Gen2QuadSensor_Lit4062E_WEB.pdf, (2018).
[36] Sentera, "The Sentera incident light sensor", Sentera, Minneapolis, MN, Retrieved from https://sentera.com/wp-content/uploads/2017/04/Incident_Light_Sensor_Lit4058A_WEB.pdf, (2018).
[37] Norris, D., [Raspberry Pi projects for the evil genius], McGraw-Hill, New York, NY, (2014).
[38] Tung, L., "Raspberry Pi: 14 million sold, 10 million made in the UK", ZDNet, Retrieved from http://www.zdnet.com/article/14-million-raspberry-pis-sold-10-million-made-in-the-uk/, (2017, July 27).

[39] Arduino, “What is Arduino?”, Retrieved from https://www.arduino.cc/en/Guide/Introduction, (2018).
[40] Arduino, “Teaching, Inspiring and Empowering!”, Retrieved from https://www.arduino.cc/en/Main/Education, (2018).
[41] Upton, E. and Halfacree, G., [Raspberry Pi user guide, 3rd ed.], Wiley Publishing, New York, (2014).
[42] Matthes, E., [Python crash course: A hands-on, project-based introduction to programming], No Starch Press, San Francisco, CA, (2016).
[43] McMullen, S.A.H., Sherry, R.E., and Miglani, S., “A survey of COTS software for multisensor data fusion”, In M.E. Liggins, D.L. Hall, & J. Llinas (Eds.), [Handbook of Multisensor Data Fusion, 2nd edition], CRC Press, Boca Raton, FL,677-699, (2008).
[44] Sizemore, J., and Mueller, J.P., “Top 10 uses of MATLAB”, Retrieved from http://www.dummies.com/programming/matlab/top-10-uses-of-matlab/, (n.d.).
[45] Pix4D, “Products”, Retrieved from http://pix4d.com, (n.d.).
[46] Agisoft, “Agisoft features”, Retrieved from http://www.agisoft.com/features/professional-edition/, (2018).
[47] Canny, J., "A Computational Approach to Edge Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-8(6), 679-698 (1986).
[48] Ballard, D.H., “Generalizing the Hough Transform to Detect Arbitrary Shapes,” Readings in Computer Vision, 23 September 1980.
[49] Haralick, R.M., and Shapiro, L.G., [Computer and Robot Vision, Volume II], Addison-Wesley, 316-317 (1992).
[50] Rosten, E., and Drummond, T., “Fusing points and lines for high performance tracking”, IEEE International Conference on Computer Vision (ICCV 2005), 2, 1508–1511 (2005).
[51] Khan, M., and Ohno, Y., "A hybrid image compression technique using quadtree decomposition and parametric line fitting for synthetic images", Advances in Computer Science and Engineering, 1(3), 263-283, (2007, November).
[52] LeCun, Y., Bengio, Y., and Hinton, G., “Deep Learning,” Nature, 521(7553), (2015).
[53] MathWorks, Inc., “Introducing Deep Learning,” (2017).