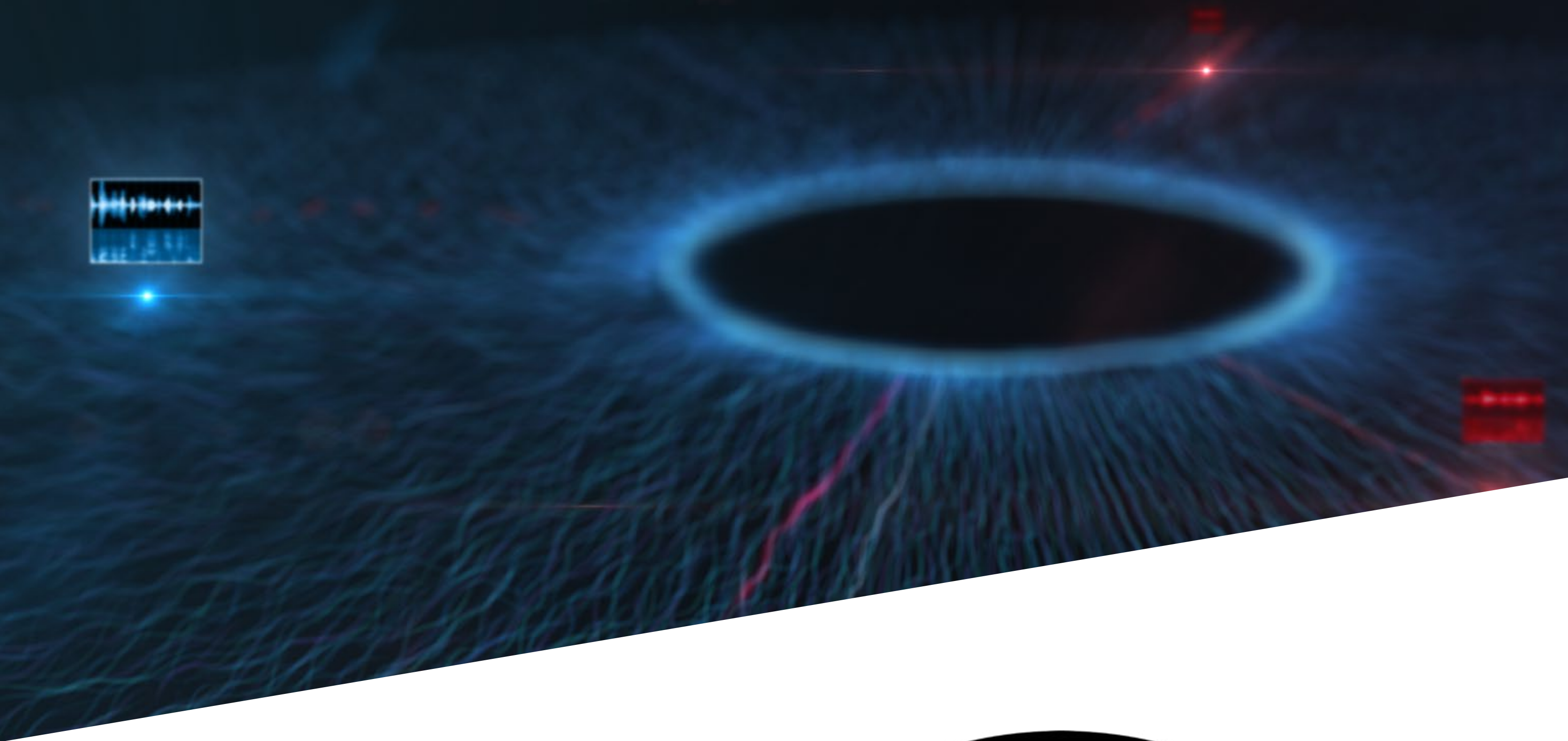

SIGNALEYE™

SignalEye™: Machine Learning Automation for SIGINT

May 2019

SignalEye™ Whitepaper Question

GENERAL DYNAMICS
Mission Systems

By David Ramirez

David Ramirez is a machine learning engineer and Marine Corps veteran at General Dynamics Mission Systems in Scottsdale, Arizona. He holds a bachelor's and master's degree in electrical engineering, and is pursuing a PhD with focus in artificial intelligence. David's academic research and professional experiences range across wireless communication systems, tactical waveforms, radar, media compression algorithms, and computer vision. He now focuses on the use of machine learning for these military applications. David sits on the General Dynamics Board of Machine Learning Subject Matter Experts advising use of machine learning throughout the company. He served in the Marine Corps from 2008-2012 advancing to the rank of Sergeant, and deployed with the 11th Marine Expeditionary Unit aboard USS New Orleans. As a trained MOS 2846 Ground Communication Electronics Repairer, David has supported a range of military engagements, from Amphibious Assault Vehicle landings to Maritime Raid Force pirate hunting.

For more information, visit our website or email us and ask for:
Josh Smith, *Product Owner and Founder*
Skip Foote, *Solutions Architect*
Tom Triebwasser, *Solutions Architect*

SignalEye.io
SignalEye@gd-ms.com

General Dynamics Mission Systems, Inc. • 12450 Fair Lakes Circle, Fairfax, Virginia 220033

Table of Contents

Existing methods for acquiring and processing SIGINT as well as conducting Electronic Warfare will not scale to current operating environment. We need to use 21st century tools and technologies to meet this 21st century threat.

1. Executive Summary

Electromagnetic warfare and wireless communications continue to evolve. An opportunity exists to leverage modern technology to improve military signals intelligence to enable our own forces to be more lethal, agile, and survivable. Massive amounts of wireless data can overwhelm traditional SIGINT methods. Automation through machine learning algorithms speeds detection and classification of signal data. We designed SignalEye™ to automate tedious SIGINT tasks, freeing time for complex human decision making. It also accelerates engagements to address vexing short dwell targets. SignalEye integrates into 3rd party software defined receiver platforms, turning any radio into a SIGINT platform. It also processes backlogs of signal data, such as months of data stored on a commodity server. SignalEye ingests radio data, automatically detects and isolates observed signals, and automatically determines the modulation type. SignalEye allows for faster outcomes than manual methods through automation. Confidence measures, reported to the users, back up machine learning results. Our patent pending machine learning framework can specialize to specific Area of Responsibility (AOR) and terrain. This tool revolutionizes the processing of signal data into actionable intelligence.

2. Revolutionizing SIGINT

2.1 Problem Statement

Operating in the electromagnetic spectrum keeps getting more demanding. In foreign countries, tactical communications might once have been the only signal on the airwaves. Today, the use of wireless technologies has exploded worldwide. The number of mobile devices has surpassed the number of humans on this planet [1]. Civilian cell service, Wi-Fi, satellites, and radio waves penetrate every corner of the globe. The majority of these devices operate on expected frequency bands, but many do not. Especially in developing markets with minimal regulation, the prevalence of interfering transmitters is growing rapidly.

These new signals complicate military operations for friendly communications planning and enemy signals intelligence (SIGINT) collection. An allied communication plan must account for existing civilian communications patterns. Finding these patterns can consume vast amounts of time for the few available highly trained military personnel. Identifying and separating civilian from enemy communications is another huge challenge. In many cases a huge backlog of data is available for exploitation, but opportunities are missed due to a lack of resources. A missed opportunity could result in degraded or dropped communications at a critical time. In hostile environments, it could mean a nearby enemy transmission goes unnoticed; missing the precursor to an attack.

The United States is lagging in electronic warfare technology. The 20-year counter insurgency (COIN) in the Middle East has shifted focus away from near-peer adversaries. The United States has focused on the present here and now conflict. Great powers competitors, such as China and Russia, are major adversaries who have focused on the future of warfare. We must improve our SIGINT technology to maintain superiority on the battlefield of tomorrow.

Current SIGINT methods require a highly trained operator to turn data into actionable intelligence, thus extending the "kill chain" and complicating engagements of short dwell targets. The analysis of existing data backlogs requires large numbers of people. Responding to mission critical events is a constant race against time. For mission support, dedicated personnel need to always be available in real-time. These resources are better used with modern technologies to save time where possible. Furthermore, a SIGINT collection system miles away from combat does have the same precision as a frontline collection element. The rapidly changing electromagnetic landscape is best navigated by an expert in the right place with the right tools. For those forward deployed units without the availability of dedicated expert signals analysts, the mission needs other means to give front line warfighters a tactical SIGINT edge.

SignalEye™

SignalEye.io

Use Artificial Intelligence and Machine Learning (AI/ML) to automatically provide end users with the right data at the right time - at the speed the mission needs.

2.2 Solution

Automation can greatly improve communications planning and SIGINT capabilities. Artificial intelligence augments human capabilities to reduce latencies and improve outcomes. An automated algorithm detects and identifies signals in sensor data much faster than a highly trained operator. Signal detection from massive amounts of stored data are like searching for a needle in a haystack. An operator controlled autonomous agent finds incoming signals, automatically determine signal type, and provides an analyst with reasons why a determination was made. The system automates the low-level detection and classification tasks. This frees up military personnel to focus on higher level tactical decision making. This way, the system becomes another team member, with a supervising human in the loop to authorize the appropriate military response. In addition, a commander can gain an "EM signature picture" of his forces as they are arrayed in the battlespace. This way, he can glean valuable information on his own EM signature and use that information to improve or implement additional passive and active actions to increase survivability.

3. SignalEye™

SignalEye™ automates detecting, isolating, and classifying RF signals using advanced machine learning technology. It receives intercepted RF data, detects signals that are present, isolates each individual signal, classifies each signal by modulation type, and reports the detected signals and metadata for further exploitation. Automating this simplifies the most labor intensive and time consuming parts of the SIGINT process.

SignalEye provides automated spectrum situational awareness for RF signals using AI/ML. Users don't need to be RF experts.

General Dynamics Mission Systems developed the SignalEye ™ software for ease of integration with existing systems. Our software enables integration into existing 3rd party military hardware and software platforms, including embedded radio receiver/transceivers, existing software suites, and server class data analysis. We work with system manufacturers and developers to integrate, test, and deploy our software. Our software-only plugin framework enables increased capabilities unrestricted to a particular hardware platform. Machine learning computations are done on both commodity servers as well as embedded processors. This allows machine learning enabled performance on the edge, detached from traditional compute resources. The runtime varies based on the amount of processing horsepower available. In just a few seconds a battlefield radio extracts tactically relevant signal information, immediately empowering forces at the tip of the spear. Figure 1 shows how SignalEye integrates into your existing system.

SIGNALEYE BENEFITS

- *Automates signal detection and isolation, minimizing probability of false alarm (PFA)*
- *Automates modulation classification using machine learning*
- *Automatically reports relevant signal metadata*
- *Provides classification confidence and certainty*
- *User informed of anomalies and trends*
- *Easily integrates with your existing systems as a hardware agnostic, software-only product*
- *Support operations in any mission environment:*
  - *Embedded forward deployed devices that quickly process mission critical data*
  - *Servers back home that process large quantities of data for national missions*

SignalEye™

SignalEye.io

SignalEye can run on almost any RF receiver or with existing saved data through preprocessing IQ streams in the open VITA-49 format.

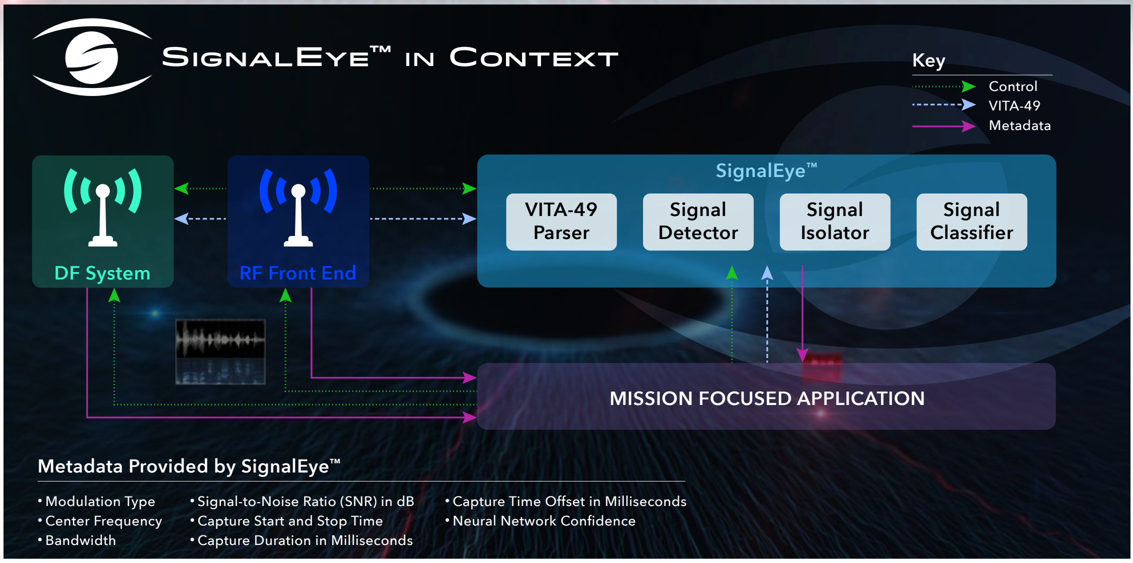

Figure 1. SignalEye and Your System. Our software-only tool works with your existing DF system and RF front-end, and displays data on your existing mission focused application using our open API framework.

3.1 Machine Learning

Our machine learning framework includes the following major elements: Preprocessing, Feature Extraction, Machine Learning Model, Training, Validation, and Testing.

3.1.1 Preprocessing

Preprocessing received signals enables SignalEye to accurately and quickly classify the data. SignalEye supports open standards such as VITA-49 as an input to enable preprocessing of data. At the most fundamental level, all digital electromagnetic receivers use similar data. Whether the system is a radio transceiver or a specialized SIGINT sensor platform, all systems measure electromagnetic waves through IQ data. In-phase and quadrature-phase (IQ) data comes from two synchronized analog-to-digital converters (ADC) circuits. The sampling rate of these ADCs sets how much observable electromagnetic frequency can be viewed at a time. IQ digital data are then processed in different ways to detect RF energy. In more advanced platforms, multiple ADC channels perform direction finding or digital beam forming towards certain signals of interest. IQ digital data is streaming live through many electronic devices, and in some cases it is also stored for future processing.

The SignalEye™ platform takes in IQ data in the VITA-49 format to identify present waveforms and modulations. This open standard allows our software to run with almost any software defined radio (SDR) receiver or with existing saved IQ data. Using digital signal processing (DSP), our system then detects and isolates any signals present within the IQ data. Our framework only requires the IQ data and information on the receiver's tuned frequency and bandwidth. The detected signals each have unique frequency and bandwidth signatures. This information allows isolation of each signal into a new IQ data record.

3.1.2 Feature Extraction

Our approach to feature extraction focuses on efficient computation of the right features for the specific neural network-determined modulation class in order to provide high accuracy for modulation orders. IQ and frequency information on their own do not tell the whole story. Identification of unique waveforms often requires additional processing and transformations to produce distinguishing features. Many significant features exist to separate one signal type from another. This computation can be costly and time consuming; a brute force comparison across all possible features is not feasible in most cases.

SignalEye™

SignalEye.io

We train SignalEye to recognize signals by giving it thousands of examples of signals, both pristine signals from a lab and ones with noise encountered in real operating environments. SignalEye learns from what we provide it to recognize similar signals in the field. This process is called supervised machine learning.

General Dynamics Mission Systems has a long history of producing SIGINT systems for the U.S. Military. Our many decades of experience allow for insights on best features to use. GDMS knows how to integrate into embedded platforms with disadvantaged processing resources. This is part of what allows our system the capability to run on platforms both big and small. This computationally efficient feature extraction process gives our system a clear advantage.

3.1.3 Machine Learning Model

Many different machine learning methods exist, from the simple to the complex. Our application focuses on a set of techniques called "supervised machine learning". This involves defining both input and output for the system, and training the machine learning algorithm to learn the relationship. Given a new input, the trained system is able to perform this learned transform to predict the output. The training process requires many data examples with both signal observations (input) and the ground truth classification (output). These training examples are iteratively processed, and the system parameters are incrementally adjusted to better fit this input/output relationship. During this iterative process the machine learning model is monitored to ensure performance is consistent. Once this relationship is trained, a new input is tested, and the correct output calculated. These training, validation, and inference steps are typical for machine learning systems.

Modern machine learning approaches commonly use neural networks. Neural nets are so called due to the interconnected neuron elements used for calculations. This is where the similarities with the human brain end. Layers of neurons are computed as realization of matrix math. Matrix math is done very efficiently on modern processors and hardware accelerators.

WHAT ARE CONVOLUTIONAL NEURAL NETWORKS (CNN)?

CNNs are very similar to traditional neural networks with one major distinction: CNNs employ sliding filters instead of static neuron connections. These spatially sliding filter windows equate to the convolution DSP operation. These sliding elements, which allow for spatial ambiguity within the input, are the pronounced advantaged of CNNs. For instance, in the image data case, the object of interest resides anywhere spatially within the image frame. In the case of traditional neural networks, the target must reside within a specific trained region. This spatial ambiguity also extends to variation in signal timing and frequency information.

SignalEye™ uses proven Convolutional Neural Network (CNN) technology. Our machine learning model builds on academic and industry research to maximize performance. We instantiate this model into software using the TensorFlow™ and Keras open source, user-friendly, modular, and extensible machine learning frameworks.

3.1.4 Training

We train a supervised machine learning model with existing data before being deployed. Given a large set of input samples and corresponding outputs, the trained model predicts outputs for similar inputs. This is why machine learning emerged as a powerful approach over explicitly programmed rules: It is easier to get data than to engineer a specialized algorithm. For instance, instead of programming a complex decision algorithm, it is often easier to simply produce inputs and outputs and use machine learning to discover relationships in the data. Once this training is complete, the system generalizes to new data: given similar inputs, predict the output.

The performance of machine learning algorithms is directly tied to data. Deep machine learning models have shown world class performance when enough good data is available: the more data, the better these models do. For collected real-world data, labeling the data with the correct classification can be very time consuming. If there are not enough input/output

SignalEye™

SignalEye.io

MODULATION TYPES SUPPORTED

- *Analog Methods*
  - Amplitude Modulation (AM)
    - **- Single-Sideband Suppressed-Carrier (SSB-SC-AM)**
    - **- Double-Sideband Suppressed-Carrier (DSB-SC-AM)**
  - **Frequency Modulation (FM)**
- *Digital Methods*
  - Amplitude and Phase-Shift Keying (APSK)
    - **- APSK-16**
    - **- APSK-32**
  - Amplitude-Shift Keying (ASK)
    - **- ASK-2**
    - **- ASK-4**
    - **- ASK-8**
  - Continuous Phase Modulation (CPM)
    - Continuous-Phase Frequency-Shift Keying (CPFSK)
      - **- CPFSK-2**
      - **- CPFSK-4**
      - **- CPFSK-8**
    - **Gaussian Minimum Shift Keying (GMSK)**
  - Frequency-Shift Keying (FSK)
    - **- FSK-2**
    - **- FSK-4**
    - **- FSK-8**
  - **Orthogonal Frequency-Division Multiplexing (OFDM)**
  - Phase-Shift Keying (PSK)
    - **- 8PSK**
    - **- Binary Phase-Shift Keying (BPSK)**
    - - Quadrature Phase-Shift Keying (QPSK)
      - **- General QPSK**
      - **- π/4-QPSK**
  - Quadrature Amplitude Modulation (QAM)
    - **- QAM-8**
    - **- QAM-16**
    - **- QAM-128**
    - **- QAM-256**

SignalEye™

SignalEye.io

Figure 2. Current BluIris Signals.
Automatically classify twenty-four unique modulation types using machine learning.

SignalEye ensures accurate performance through being trained on real data gathered using trusted sources from extensive field testing.

examples or if the output labels are not accurate, the machine learning algorithm gets confused. Furthermore, trusting a new data source may be a security risk. Research has shown adversarial data can be slipped into training, allowing a backdoor to trick the system. This shows the importance of realistic data, available in large quantities, on demand, from only a trusted source.

Other systems rely only on simple synthetic data or limited collected data. Given the exact same neural network model, a machine learning system trained on data gathered from extensive field testing will have higher performance.

3.1.4.1 BluIris

BluIris, our patent pending data generation framework, allows for a limitless variety of IQ signal data for machine learning training. BluIris enables the generation of dozens of different signal types, with more being continually added. This IQ data generation framework replaces the need for numerous unique radio transmitters. Figure 2 details the modulation types that BluIris currently supports.

BluIris also adds another unique and game changing advantage over a traditional transmitter. A usual radio transmitter, attempts to produce a pure, distortion free signal. This perfect signal does not include the typical distortions experienced in a real-world or military setting. Our BluIris framework replicates these random real-world distortions to produce realistic signal examples. Collecting a real-world dataset which is thoroughly expansive and complete would be nearly impossible. As you might guess, these distortion types can vary to include everything under the sun. General Dynamic Mission Systems has a long history of overcoming distortion effects for radio demodulators. We have expert knowledge on how best to emulate all significant real-world effects.

We validate SignalEye's performance using real world exercises and our neural network signal generation framework based on years of experience.

3.1.5 Validation

During training of a machine learning model, it is important to monitor incremental performance for a separate set of data. The validation set should be representative of real-world testable examples, but unique from the training data. This can be real-world data with known truth information for the signal type. When a machine learning model trains on a set, it reduced the error when classifying only that training set. A similar level of performance on the withheld validation data must be verified. This is a key challenge of machine learning: Generalization from operation on a small set of data to new input examples. In many cases a machine learning model over specializes to the training data, not able to generalize to new data. This overfitting problem is overcome through various machine learning architecture and training parameter methods.

SignalEye™ uses both our proprietary neural network signal generation framework, as well as data collected from real world exercises to validate system performance. We continually improve our training regime through an expanding set of testable data.

3.1.6 Testing

Once a neural network model has been trained with inline validation, it is important to continue to test performance against new examples. When a new input is presented to the system, a prediction is made. This computed inference should be reviewed by an expert to ensure correctness.

We continually test SignalEye™ to ensure detection performance and classification accuracy. We ensure all signals are being detected and isolated correctly, and the predicted machine learning classifications make sense. Should the system fall short, these failure examples are used to improve the system. We test signals at different signal-to-noise (SNR) levels and in the presence of interference or jammers.

General Dynamics Mission Systems is highly engaged with customers in testing SignalEye™. We participate in field operations to prove our system performs under extreme conditions and unique

SignalEye™

SignalEye.io

We continually test SignalEye, both internally and externally at field tests such as Vigilant Hammer and the Army RCO Challenge

electromagnetic environments. This includes Vigilant Hammer III (VH3), Army RCO Signal Classification Challenge, and NetModX. We have a proven track record of successful system operation, even under contested or congested electromagnetic environments. These exercises also serve as new data sources to further improve our system.

Our collaboration with 3rd party RF receiver manufacturers also provides a feedback loop for improving the system. During testing, any signal inputs which do not classify correctly are retrained into the machine learning model. This ensures a continually improving framework without re-engineering major components.

Our system has built-in checks to measure the machine learning model's prediction confidence. Should uncertainty exist within the machine learning model, this is reported to the user. This reduces the chance for a confusing input being misclassified and not inspected further.

4. Summary

SignalEye enables non-RF experts to identify signals rapidly and automatically, and immediately turn them into actionable intelligence in the field.

SignalEye enables warfighters and non-RF experts to rapidly and automatically identify signals of interest and immediately turn them into actionable intelligence in the field. It operates in all security domains, from Unclassified through TS/SCI. The SignalEye software integrates into systems big and small: almost any RF receiver can be enhanced to be a SIGINT platform. We work with partners to quickly integrate, test, and deploy SignalEye into existing systems. GDMS has supported successful, rapid integrations with RF receivers from iRF Solutions, Herrick, Systel, DRT and SilverPalm, as well as digital signal processors from Curtiss-Wright. SignalEye easily installs and runs on commodity servers. A SignalEye version is in development for ARM and FPGA within iRF Solutions' WIDERAIL receiver. We also plan to integrate SignalEye into SCEPTRE, the deployable, remoteable ISR collection software suite made by 3dB Labs.

Our machine learning technology outperforms traditional methods. Signal classification is typically a needle in the haystack problem: first finding a signal, then a brute-force pass with every likely demodulator imaginable. Our DSP algorithm automatically detects and isolates signals. Our machine learning model returns a modulation type with only one pass on the data; with fast, repeatable, and reliable results. The system provides insight into otherwise unknown signals by providing the modulation and additional signal metadata. SignalEye is a trusted tool that enhances the warfighter's effectiveness and survivability by providing spectrum situational awareness.

SignalEye™

SignalEye.io