Enhancing ROV/AUV Imagery Using Real-Time Underwater Optical Image Processing

Jack Wade
ZMicro, Inc., 9820 Summers Ridge Road, San Diego, CA 92121 USA
jack.wade@zmicro.com

Abstract —This paper introduces a new deep learning AI-enable processing platform that can be used to enhance optical imagery and to analyze and exploit video streams for military undersea applications including mine countermeasures (MCM) and anti-submarine warfare (ASW).

1 Purpose

Undersea visibility remains a challenge for defense applications including mine countermeasures, equipment inspections, surveillance, and search & rescue. With autonomous underwater vehicles (AUVs) and remotely operated underwater vehicles (ROV) playing a growing role in underwater missions, video enhancement and image classification using deep learning are now key enabling technologies.

This paper introduces a new deep learning AI-enable processing platform that can be used to enhance optical imagery and to analyze and exploit video streams for military undersea applications including mine countermeasures (MCM) and anti-submarine warfare (ASW).

2 Introduction

AUVs are outfitted with an array of sensors for navigation and other mission duties, but they still rely on visual sensing to perform vision-driven tasks. Human ROV pilots and mission personnel still rely primarily on undersea imagery as the basis for intelligent decision making. The problem is that the human eye and its analog, optical sensors see far better through air than through water. This is largely attributable to physics: water is 800 times denser than air. Consequently, when light enters water and interacts with the water molecules and particles, there is a loss of light intensity and contrast, changes in color, and diffusion which results in poor image quality that can hinder underwater missions.

3 Approach

Optical image sensors, such as those used in video cameras on underwater missions, capture far more information than can be perceived by the human eye. One of the most practical and effective methods to reveal hidden detail in underwater imagery is to apply image processing algorithms that, in essence, undo some of the undesirable effects of the physics of underwater vision.

We have developed algorithms that can improve visual acuity for users by enhancing contrast, redistributing brightness, reconstructing color, or eliminating noise in underwater imagery. Two of the most important algorithms are 1) the clarifier, which is a locally adaptive algorithm that can produce striking clarity by enhancing image contrast and 2) the global de-haze (a.k.a. True Color) which corrects color. These are described below.

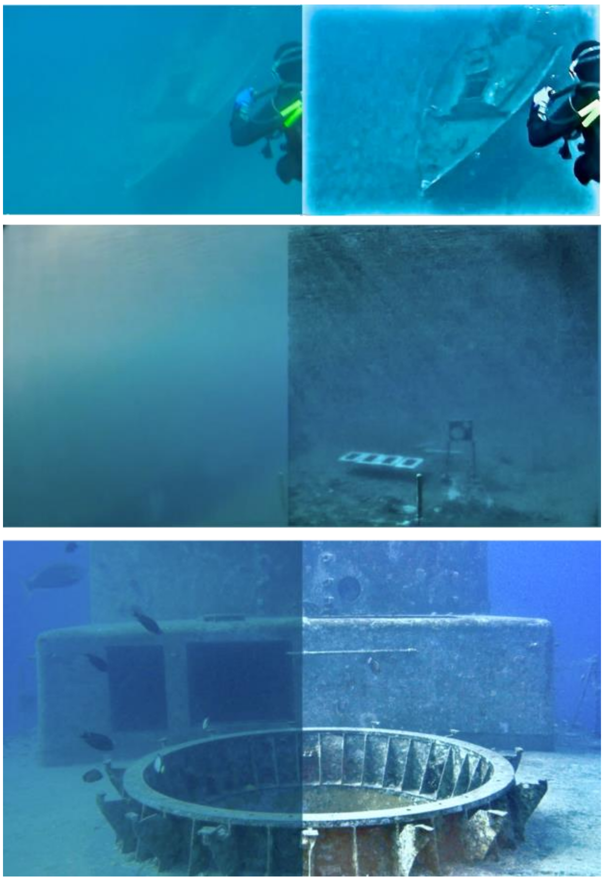

Fig. 1. Examples of images before and after enhancement algorithms have been applied.

A digital image is a two-dimensional array of values and is essentially just data that represents all of the pixels in a scene. That data can be transformed mathematically to create enhanced images that utilize all the information available in the original image, including some which may not have been visible to the human eye.

The clarifier algorithm is an ultra-large convolution kernel that does locally-adaptive contrast enhancement. Each individual pixel is compared to its neighbor and the pixel values are adjusted to increase the subtle differences —all in real-time. Convolutions are computationally intensive and, therefore, most implementations use only small kernels (3 x 3, 9 x 9, 16 x 16). However, we have developed programming techniques that allow the clarifier algorithm to implement very large convolution kernels (400 x 400) that produce dramatically better results. Using a large kernel results in a much greater range of spatial frequencies and thus results in a clearer image with greater detail.

The clarifier enables viewers to discern subtle differences in brightness and contrast that were previously difficult to perceive. The result is a video clarification and enhancement that enables humans and machines to see more detail.

The global dehaze algorithm was designed to take into account the physical processes that go along with atmospheric haze and fog, and the different types of scattering of light underwater. It is based on knowledge and experience with how different media typically affect the light in terms of treating different wavelengths, or different colors differently, e.g., preferentially using more of the blue or the red, and so on. It then attempts to organize the reconstruction of the image in ways that make sense with the physics in the situation.

Like histogram-based algorithms, global dehaze looks at the statistics of the data from the entire image, determines what types of corrections to make, and then applies those corrections by doing the same operation to the entire image. The difference is that with global dehaze, the gains and offsets applied tend to be more linear because the processes in the physics that degrade image data are correspondingly linear, and the algorithm attempts to mirror those effects.

4 Results and Discussion

The success of this approach hinges on the availability of a specialized AI-enabled image processing platform. To address this need, ZMicro developed the Insight Vantage, a compact, rugged stand-alone image processing system that can be plugged into an existing ROV/AUV system architecture to enable image enhancement and deep learning capabilities. The Insight Vantage includes ZMicro's F-40 processor module that includes an FPGA to run the image enhancement algorithms in real time and a proprietary video controller to provide encoding/decoding of IP video streams, and the ability to fuse incoming video with external steams for chroma-key overlays such as GPS data and blending.

A key component to the system is the NVIDIA Jetson TX2 System-on-Module (SOM), which is a fast, power-efficient embedded AI computing device. The Jetson enables server-class AI compute performance for edge devices and is ideal for deploying advanced AI to remote field locations with limited internet connectivity. The Jetson comes with NVIDIA's Deep Learning GPU Training System (DIGITS), which enables engineers and data scientists to rapidly train the highly accurate deep neural networks (DNNs) for image classification, segmentation and object detection tasks.

DIGITS simplifies common deep learning tasks such as managing data, designing and training neural networks on multi-GPU systems, monitoring performance in real time with advanced visualizations, and selecting the best performing model from the results browser for deployment. DIGITS is completely interactive so that data scientists can focus on designing and training networks rather than programming and debugging.

Author/Speaker Biographies

Jack Wade is the founder and CEO of ZMicro, a trusted provider of MIL-SPEC computing solutions to U.S. military and intelligence system integrators for more than 30 years. He is credited as an inventor on more than a dozen patents. Jack attended Indiana State University.

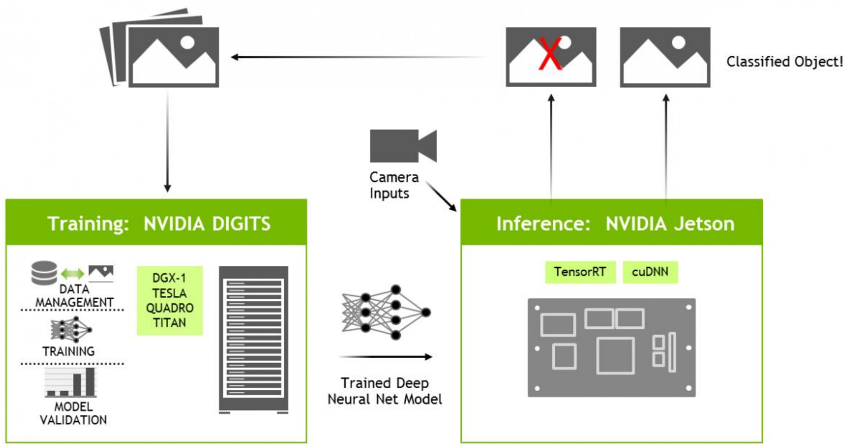

Fig. 2. NVIDIA provides computer vision primitives including image recognition, object detection +localization, and segmentation and neural network models trained with DIGITS. Developers can deploy these network models to Jetson for efficient deep learning inferencing with NVIDIA TensorRT.