Applying Deep Learning to Improve Maritime Situational Awareness

August 14, 2016
KDD 2016 Workshop On Large-scale Deep Learning for Data Mining

Kathy Tang
Stottler Henke Associates, Inc.

David Crandall
School of Informatics and Computing

Overview

- Describe a computer vision system, called ExPATSS, which uses deep learning techniques to classify ships in videos
- What does ExPATSS do and why is it important
- How ExPATSS uses deep learning
- Data and training methodology
- Results
- Future work

What is ExPATSS?

- ExPATSS = Extensible Platform for Automated Tactical Sensor Screening
- Computer vision system being developed for the Navy to aid maritime surveillance of threats
- ExPATSS automatically detects and classifies ships in video streams in real time
- Processes infrared and visible light data

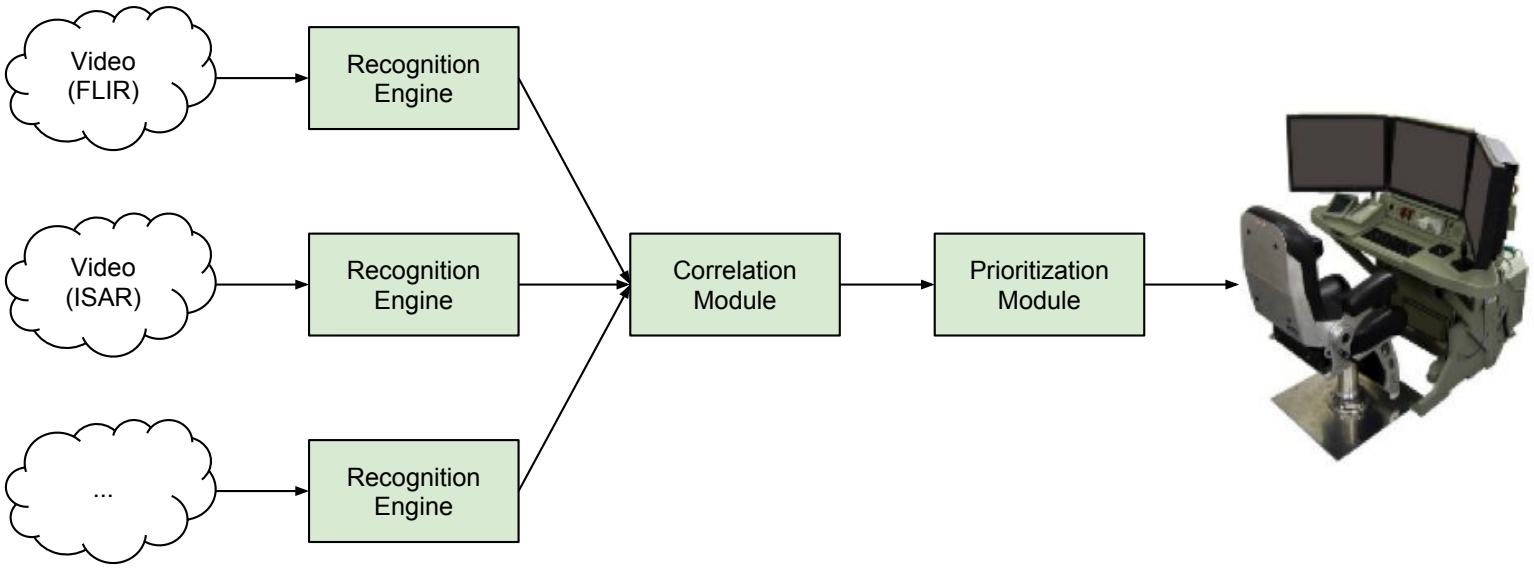
Why is ExPATSS Important?

- ExPATSS is targeted to run on the central tactical support center onboard an aircraft carrier
- Multiple sensor gathering data simultaneously
- Currently – one sensor operator/analyst at each sensor location
- Future – all sensor data streamed to centralized location on carrier, single operator/analyst

Why is ExPATSS Important?

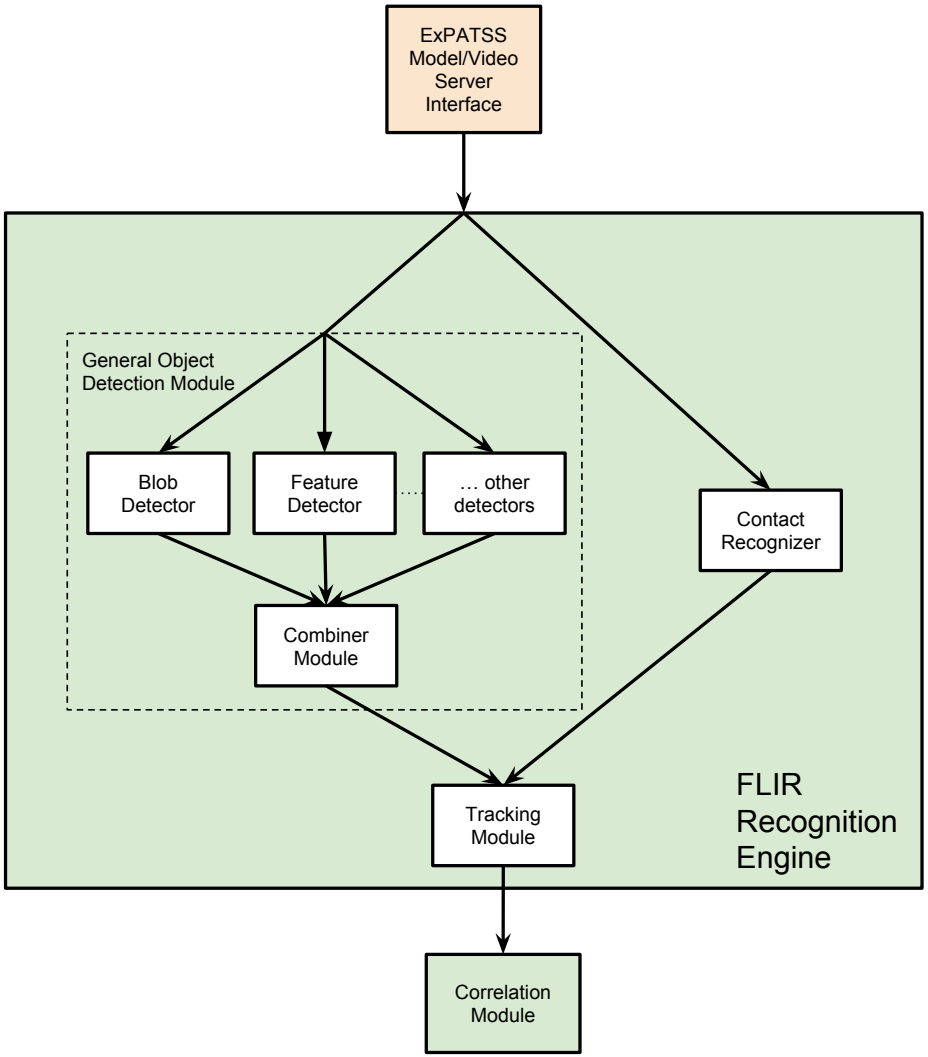
- Monitoring multiple sensor streams can quickly lead to attention overload
- ExPATSS automatically processes each individual streams for important events
- Assigns each stream a priority number to help direct the analyst's attention
- More quickly identify threats
- Generate a more complete and accurate picture of the tactical situation

ExPATSS System Overview



- Processes multiple streams
- Aggregates activity from each stream
- Prioritizes the importance of each stream for the operator
- Computer vision and deep learning processing happens in Recognition Engine

ExPATSS Computer Vision Module



- Main computer vision module that monitors for important events
- Important events are defined as streams containing ships
- Detects and tracks objects and classifies frames
- Several object detectors working in parallel to detect activity
- Tracking Module keeps track of object locations and filters out sporadic false positives
- Contact Recognizer classifies frames using CNN

Contact Recognizer – Applying Deep Learning

- Convolutional Neural Networks are used in the Contact Recognizer
- Performs classification at the frame level
- Classifies the frame into the following six classes
  - Open water (no visible ships)
  - Warship
  - Speedboat
  - Sailboat
  - Merchant ship
  - Cruise ship

CNN – Training Data

- Real data is difficult to acquire, and expensive to collect
- Gathered publicly available images from Flickr
- Searched Flickr for images of each of six classes
- Used a training set of 2,128 images, and 456 images for validation, and 456 images for testing

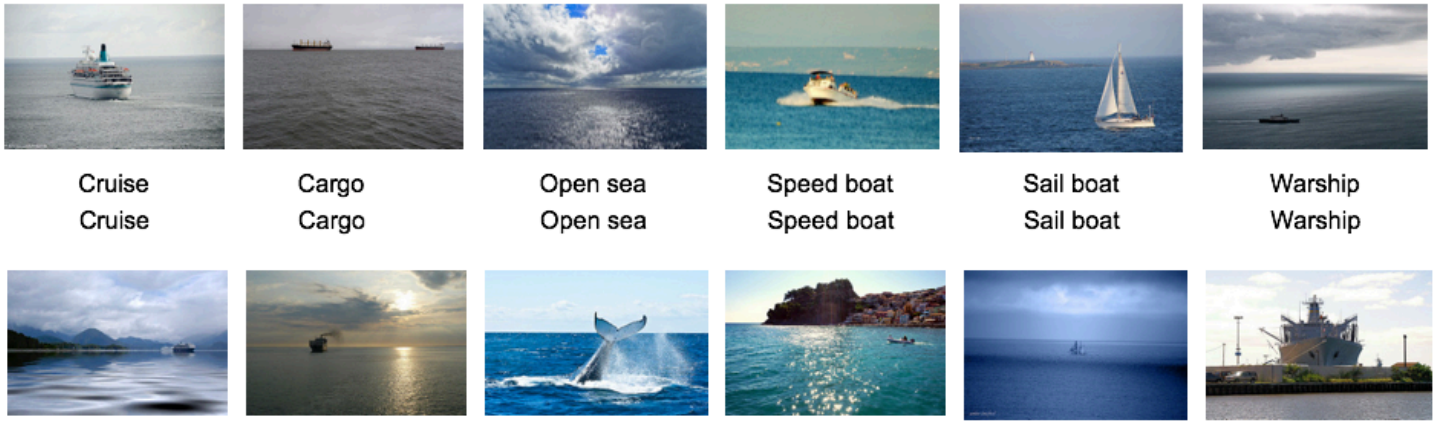
CNN – Training Process

- Using the open source deep learning framework called Caffe for training and classification
- Used the ImageNet dataset of millions of images to initialize the neural network
- Overcome limited training data
- Generic enough representation of the visual world to be useful across applications

CNN – Training Parameters

- Convolution Neural Network
  - AlexNet architecture
  - 5 convolutional layers with rectified linear activation functions, and max pooling
  - Followed by 3 fully connected layers
- 100,000 iterations
- Batch size of 50
- Base learning rate of 0.001
- Step size of 20,000
- Momentum of 0.9
- Weight decay of 0.005

CNN – Results on Flickr Data



Predicted:	Cruise	Cargo	Open sea	Speed boat	Sail boat	Warship
Correct:	Cruise	Cargo	Open sea	Speed boat	Sail boat	Warship

Predicted:	Open sea	Cruise	Sail boat	Open sea	Warship	Cargo
Incorrect:	Cruise	Cargo	Open sea	Speed boat	Sail boat	Warship

- Sample classification results
- Performs well on a variety of images
- Struggles when the ship is low resolution
- Struggles when object doesn’t belong to any of the classes it was trained on

CNN – Results on Flickr Data

		Detected class					
		Cruise	**Cargo**	**Open**	**Sail**	**Speed**	**War**
Actual class	Cruise	86.4%	1.2%	7.4%	0.0%	4.9%	0.0%
	Cargo	1.5%	88.1%	4.5%	0.0%	3.0%	3.0%
	Open	0.0%	0.0%	97.8%	1.1%	1.1%	0.0%
	Sail	0.0%	1.2%	4.8%	91.4%	1.2%	1.2%
	Speed	0.0%	1.5%	3.0%	1.5%	94.0%	0.0%
	War	1.4%	2.8%	4.2%	0.0%	1.4%	90.1%

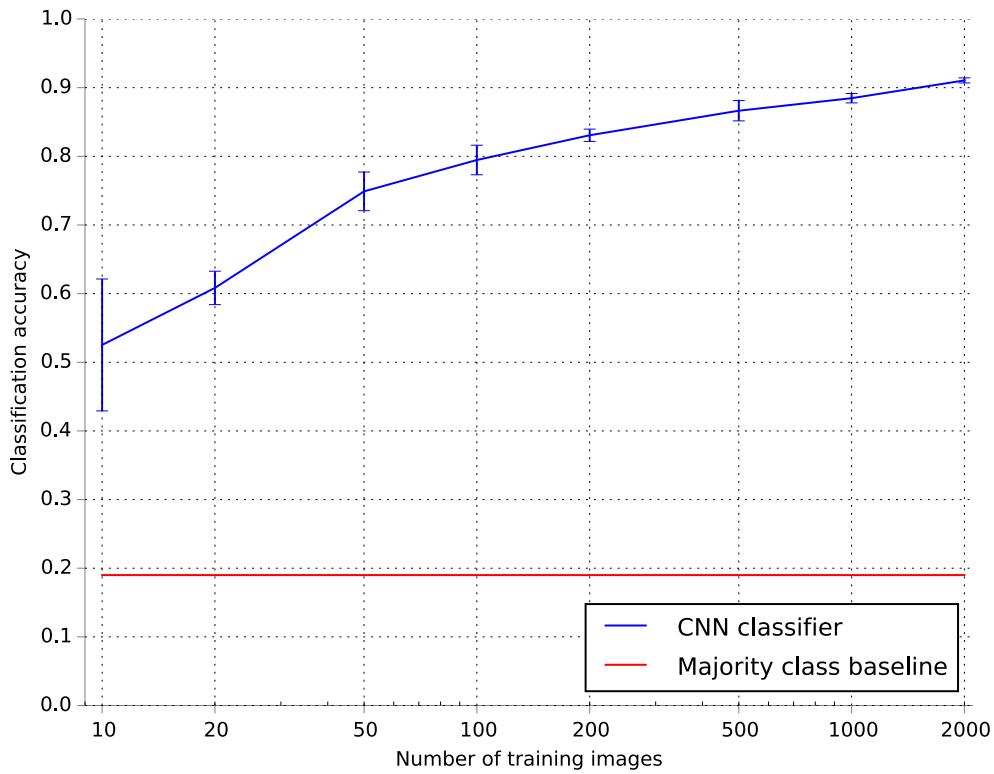
- Quantitative classification results
- Confusion matrix presented above
- Overall classification rate of 91.4% over 456 images

CNN – Training Set Size

- Typically, CNN requires large amounts of training data
- Investigated how the training set size would affect the classification accuracy
- Tested model performance with training set varying from only 10 training images to 2,000 images
- Performed 5 trials for each training set size

CNN – Training Set Size

- Six-way classification accuracy on the Flickr dataset, as a function of total training set size, averaged over five trials.



ExPATSS System Results

Correct Classification

ExPATSS System Results Cont'd

Incorrect Classification

ExPATSS System Results Cont'd

Classification results on ExPATSS data

Video	Number of frames	Correct Classification Rate
speed_boat_1	3,625	0.79
speed_boat_2	4,320	0.68
warship_1	750	0.91
sail_1	570	0.00
warship_2	600	0.97
speed_boat_3	480	0.67
warship_3	1,890	0.14
warship_4	1,980	0.91
open_water_1	114,497	1.00
stormy _waters_1	7,223	1.00
stormy_waters_2	441	1.00

ExPATSS System Results Cont'd

- The system uses CNNs for image classification and achieves higher than 79% correct classification rate for 7 out of 11 videos of real maritime data
- The system classifies long segments of open water (both stormy and calm) with 100% accuracy over 114,197 fames

Conclusions

- Designed and implemented a system to automatically detect and classify ships in maritime scenes
- The detection and classification results are used to cue operators of important events
- Can be used during video playback to skip uninteresting segments of footage, saving large amounts of human-hours

Future Work

- Improve the current CNN model by using a larger and more diverse training set
- Train new CNN models with more specific ship classes
- Apply CNN-based approaches directly to the object detection problem, to not only classify images but also identify where in the image the ships of interest are

ACKNOWLEDGMENTS

This material is based upon work supported by the Naval Sea Systems Command under Contract No. N00024-14-C-4082.

DISCLAIMER

Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Sea Systems Command.