

An Architectural Model for Web-Based Technologies to Enhance Text-Image Capabilities in Detecting Sex Trafficking Cases

Summary

While humans are naturally able to make connections across modalities (e.g., text and imagery), this remains a hard problem for machines. This project is developing algorithms, computational methods, and prototype code to enable such cross-modality reasoning for AI/ML tools, specifically applied to the problem of detecting sex trafficking cases but applicable to a wide range of domains and problems.

Problem addressed

There are multiple state-of-the-art technologies, such as “my Domain-Specific Insight Graph (myDIG)”, available to perform detection of sex trafficking (ST) victims through analyzing texts from advertisements on the web. These technologies lack image analysis, which could increase the detection of latent cases visible as valid escorts and sex workers, but these technologies usually do not support cross-modal retrieval tasks. This study will contribute to this gap through research on Artificial Intelligence and Machine Learning (AI/ML) techniques by building on the advancements in multi-modal AI/ML. The purpose of this project is to solve the problem of computational understanding of cross-modal data through the development of prototypical computer code and AI/ML models for narrowly defined use cases (ST) with simulated data.

Approach

The main approach of this work extends recent advances in Deep Learning to build a pipeline that integrates bottom-up and top-down information flow. Solutions to the research problem of learning from cross-modal data can be utilized to build applications to enhance rapid and effective decision-making. This work will accelerate “data to decisions” timelines by leveraging machine reasoning and an end-to-end approach to enhance the performance of analysts to detect ST cases. The proposed prototypical software (PS), AI/ML model development, and algorithms will support the extension of the cross-modal technologies (i.e., image surveillance and text analysis) in one of the state-of-the-art meta-engines like myDIG to reduce the current gap in technologies.

Results

The research is expected to produce the following software and AI/ML models which could be used in ST detection:

- 1) Ontology of Contexts and Environments. This will be an extension of the representations and processes incorporated into the Image Surveillance Architecture for future software development.

- 2) Models of Scenes and Spatial Relations. This model will be capable of extracting data from scenes and spatial relations utilizing symbols and attributes.
- 3) Models of Scenes and Temporal Relations. This supplements Outcome#2.
- 4) Object Classification. This model will be based on contexts utilizing a Convolutional Neural Network (CNN) or Regional-CNN.
- 5) Natural Language Description (Caption) Matching & Generation. This model will be capable of producing natural language descriptions of visual information.

Anticipated Impact for DHS

- 1) This research helps identify security threats by developing prototypical software and AI/ML model detection technology.
- 2) This leads to dismantling transnational criminal organizations on the premise of ST.
- 3) The prototypical software (PS) and AI/ML model can be the basis of future software development. The developed software can be incorporated into detection technology to support the enforcement of immigration and other laws.
- 4) The PS and AI/ML models in the cross-modal retrieval capability operating on simulated data can be the basis of models with real data. The software and models can influence border security, federal investigative and law enforcement agencies, and future research.

Research Products:

Presentations:

[Final Results Briefing: An Architectural Model for Web-Based Technologies to Enhance Text-Image Capabilities in Detecting Sex Trafficking Cases](#)