Iowa Initiative for Artificial Intelligence
Final Report

Project title:	Detecting Video-Patterns
Principal Investigator:	Sang Jung Kim, Assistant Professor, School of Journalism and Mass Communication
Prepared by (IIAI):	Avinash Reddy, Mudireddy
Other investigators:	
Date:	July 05, 2025

Were specific aims fulfilled:	Yes
If yes … Planned submission date	
Funding agency	
Grant mechanism	
If no … Why not? What went wrong?	

Brief summary of accomplished results:

This project successfully developed a machine learning model to detect specific video patterns in images. The primary model, a YOLOv11n, was trained to identify these patterns in video imagery. An initial goal was to augment the training data using the HandsOff method to generate synthetic labeled images; however, this was not fully achieved due to the data requirements of the GANs involved. Despite this, the trained YOLOv11n model demonstrates promising results in detecting the targeted video patterns.

Research report:

Aims (provided by PI):

The original aim of this research was to investigate how are specific video patterns published. This was to be achieved by developing methodologies to identify video-pattern imagery and then analyzing the frequency of this published imagery.

The IIAI-specific goals were refined to focus on the technical implementation of the detection model. The primary aims became:

- To build machine learning classifiers capable of detecting specific video patterns.
- To use the HandsOff method to generate a synthetic, labeled training dataset of video patterns for model training.

Data: The dataset used for this project consisted of 641 images containing 111 instances of specific video patterns.

AI/ML Approach:

The AI/ML approach involved two main stages:

1. **Labeling and Training:** We utilized Azure Machine Learning's Assisted Machine Learning tool to label the bounding boxes for the video pattern in the image dataset. A **YOLOv11n model** was then trained on this labeled data to perform video pattern detection.
2. **Synthetic Data Generation (Attempted):** To increase the number of training samples, we explored the **HandsOff method**. The HandsOff framework is designed to generate a large, labeled synthetic dataset from a small number of existing labeled images by leveraging Generative Adversarial Networks (GANs). The implementation involves using GAN inversion to map real, labeled images into the latent space of a pre-trained GAN and then training a "label generator" to associate these latent representations with their corresponding labels. This allows for the generation of new, diverse, and accurately labeled synthetic images. However, we were unable to successfully implement this method as the GANs required more initial data than was available to produce high-quality synthetic images.

Experimental methods, validation approach:

The YOLOv11n model was trained and validated on the described dataset. The performance of the model was evaluated using standard object detection metrics, including precision (P), recall (R), and mean Average Precision (mAP) at different IoU (Intersection over Union) thresholds.

Results:

The trained YOLOv8n model achieved the following performance:

Model Summary:

- Layers: 100
- Parameters: 2,582,542
- GFLOPs: 6.3

Performance Metrics:

- Precision (P): 0.665
- Recall (R): 0.594
- mAP@.5: 0.641
- mAP@.5-.95: 0.388

Detailed mAP Scores:

- mAP50-95: 0.361
- mAP50: 0.607
- mAP75: 0.335

Inference Speed:

- Preprocessing: 0.1ms per image
- Inference: 0.5ms per image
- Postprocessing: 1.0ms per image

These results indicate a moderately effective model for detecting specific video patterns, with room for improvement, particularly in recall and mAP at higher IoU thresholds.