Detecting and Blurring Picture Objects Using YOLO and Streamlit

An AI-Powered Solution for Privacy Protection

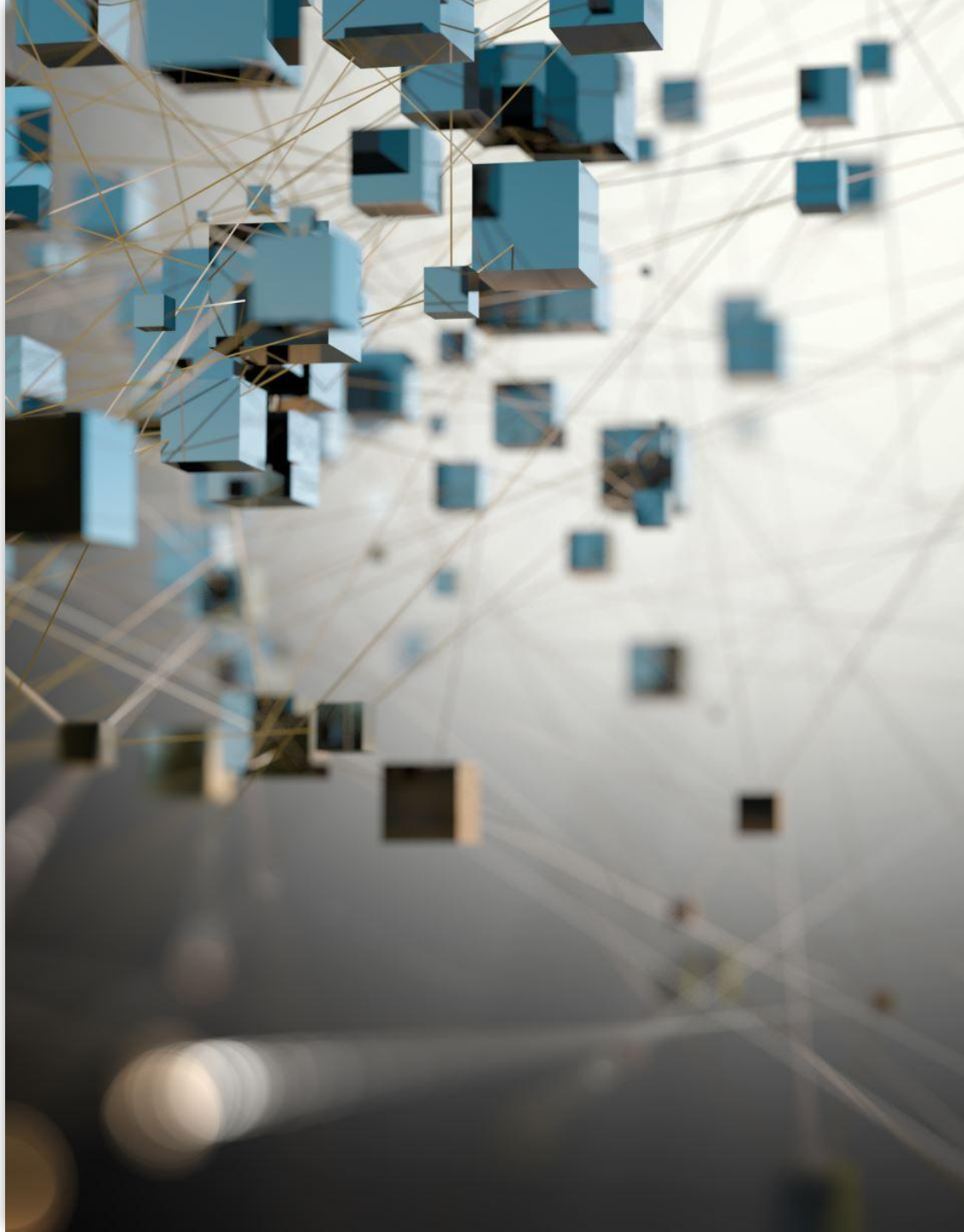
Authors: Silas Phillips, Keegan Nohavec, and Christoph Guenther

Introduction

Problem: Personal photos, art, and other picture objects in images and videos raise privacy concerns.

Solution: Use an Artificial Intelligence (AI) model to automatically detect and obfuscate these objects.

Technology Stack: Roboflow, YOLO, Streamlit, Python.



Goals

Main Goal: Develop an AI model to detect and blur picture objects in images and videos.

Sub-goals:

- High- accuracy detection of picture objects.
- Efficient blurring of detected objects maintaining image quality.
- Fast processing of images/videos without sacrificing accuracy.
- User interface for uploading and processing images/videos.

Our Process

- Data Sourcing and Preparation
- Model Selection
- Model Training
- Video Processing
- Web Application Development

Data Preparation

Data Sources: JPEG images of indoor spaces (Kaggle).

Image Preparation: Tool – Roboflow.

- 1 class (Pictures-on-Wall).
- Standard size: 224 x 224 pixels.
- Annotation – Identify, draw bounding boxes, and label picture objects.
- Create Train, Validation, and Test datasets.
- Augmentation – Crop, rotate, vary brightness, and blur images to increase training dataset.

Dataset Size: 3415 images split into

- Train: 2390 images (7170 after augmentation).
- Validation: 602 images.
- Test: 423 images

Model Selection

YOLO (You Only Look Once)

Why Yolo?

- Speed – Performs object detection and framing in one pass.
- Accuracy – Comparable or better compared to slower models (such as R-CNN based models).
- Flexibility – Variety of different models with different complexity.
- Convenience – Saves best parameters from each training run.

Training

Setup:

- Jupyter notebook on Google Colab.
- GPU.
- Multiple processors.

Process:

- Load datasets from Roboflow.
- Train model.
- Evaluate performance on
  - Precision.
  - Recall.
  - Mean Average Precision.
  - Confusion Matrix.
- Repeat, varying these parameters
  - YOLO version.
  - YOLO model.
  - Cache.
  - Batch size.
  - Number of epochs.
  - Patience.
  - Intersection over Union (IoU).

Training Outcome

Model Chosen:

- YOLOv10m: 16,451,542 parameters.

Hyperparameters:

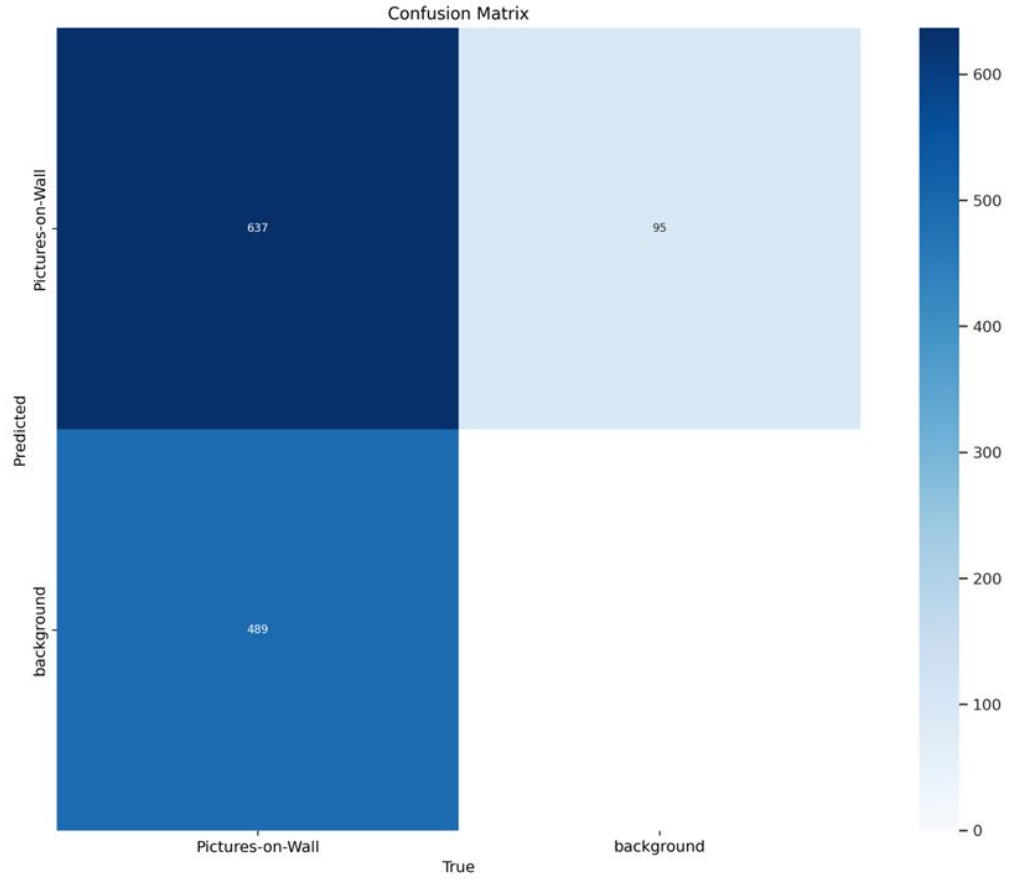
- Epochs = 600,
- Cache = True,
- Batch Size = -1,
- Patience = 100,
- IoU = 0.5.

Performance:

- Precision = 0.792,
- Recall = 0.552,
- mAP50 = 0.637,
- mAP50-95 = 0.425.

Performance - cont'd:

- Confusion Matrix



Video Processing

Improve processing speed by

- Applying object detection to a subset of video frames only.
- Use tracker to track objects in other frames.

Blurring:

- Apply Gaussian blur to bounding box.

Algorithm:

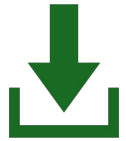
1. Instantiate chosen YOLO model.
2. Convert video to list of frames.
3. Every 5 frames (starting at frame 0)
    a. Detect picture objects with bounding boxes.
    b. Add new bounding boxes to tracker.
    c. Remove bounding boxes for objects no longer present from tracker.
    d. Apply Gaussian blur to each bounding box in tracker.
4. Reassemble list of processed frames into output video.

Streamlit User Interface

Why Streamlit?

Rapid development of web interfaces.

Easy integration with Python.



Features:

Upload images and videos.

Display original and processed images/videos.

Option to download processed videos.

UI Design:

Intuitive layout for user interaction.

Progress bar for processing status.


Streamlit Video Upload
Demo


reate your
ream room
or under
$680

Conclusion

- Successfully trained a YOLO model to detect picture objects with bounding boxes in images and videos.
- Applied model to video frames.
- Applied Gaussian blur to bounding boxes corresponding to detected objects.
- Built web application with Streamlit allowing users to submit videos for processing and download processed videos.

Future

Improve model accuracy

- More data, images and video, in a variety of formats and resolutions.
- Systematic hyperparameter tuning.
- Smarter video processing.

Improve processing speed

- Smarter video processing.
- Use multiple processors.