Liveliness Detector for Face Recognition System: Fake VS Real

Computer Vision and Pattern Recognition

Salma El Sheshtawi, Jana Abdelqader

Contact Information:
01005880154
01014812057
Email:
salma.elsheshtawi@ejust.edu.eg
jana.abdelqader@ejust.edu.eg

Abstract

The proliferation of deepfake technologies and AI-generated facial imagery presents increasing challenges in digital media authentication. This project proposes a real-time fake face detection system that combines deep learning with classical image analysis techniques. By integrating the YOLOv8 object detection model with a suite of handcrafted heuristics—focused on texture, blur, and color variation—we aim to enhance the reliability of face authenticity classification. Inspired by Li et al. (2018), our approach can detect visual artifacts introduced by face synthesis methods. The system achieves real-time performance and demonstrates promising accuracy in distinguishing between genuine and artificially generated faces.

Introduction

The advent of generative models, particularly GANs (Generative Adversarial Networks), has led to a significant rise in the production of highly realistic synthetic facial content. While these technologies have valuable applications, they also pose serious risks in the domains of misinformation, identity theft, and digital manipulation. Traditional detection systems rely heavily on deep learning models, which can sometimes fail under adversarial or ambiguous conditions. Motivated by the work of Li et al. (2018), which identifies face warping artifacts as telltale signs of manipulation, we designed a hybrid system. Our method leverages both YOLOv8 for robust object detection and classical image-based heuristics to evaluate the authenticity of detected facial regions in real time.

Main Objectives

1. **Develop a Real-Time Detection System:** Implement a system capable of identifying fake faces from live webcam input with minimal latency.
2. **Integrate Deep Learning with Heuristics:** Combine a state-of-the-art object detector (YOLOv8) with statistical image analysis to enhance decision reliability.
3. **Improve Explainability:** Provide visual cues such as bounding boxes, labels, and confidence scores for interpretability of detection results.
4. **Evaluate System Robustness:** Assess the system's accuracy and reliability in different lighting, motion, and resolution conditions.
5. **Apply and Translate Research:** Operationalize the theoretical insights from Li et al. (2018) into a practical software prototype.

Mathematical Section

The proposed system evaluates facial authenticity using several image-processing metrics after face detection:

1. Laplacian Variance (Texture and Blur Assessment)

To assess image sharpness and texture presence:

$$\text{Blur} = \text{Var}(\nabla^2 I)$$

Where $\nabla^2 I$ denotes the Laplacian of the image I. A low variance implies a blurred or texture-deficient image, which is often characteristic of synthetic content.

2. Color Saturation Standard Deviation (Color Variation)

To measure the degree of natural color variation within the face:

$$\sigma_S = \text{Standard Deviation}(S)$$

Where S is the saturation channel from the HSV color space. A low value indicates unnatural uniformity, typical in AI-generated imagery.

3. Face Area Threshold

To avoid processing small, unreliable detections:

$$A = w \times h \quad \text{such that} \quad A < A_{\text{min}} \Rightarrow \text{Fake}$$

Where A_{min} is a manually set threshold (15,000 pixels in our case).

4. Combined Heuristic Rule

A face region is classified as "FAKE" if any of the following conditions are met:

- Low Laplacian variance (blurred)
- Low texture in grayscale image
- Low color variation in HSV space
- Insufficient face area

Results

The proposed fake face detection system was evaluated using a variety of input sources, including live webcam feeds, AI-generated images, and deepfake video sequences. The system demonstrated strong performance in detecting synthetic faces that exhibit low texture, limited color variation, or noticeable blur—common artifacts in generated or manipulated imagery. In real-time testing with genuine human faces under good lighting conditions, the system consistently classified them as "REAL," achieving high accuracy. Deepfake videos with facial warping or blending inconsistencies were also effectively detected. However, the system exhibited some false positives in scenarios involving low-resolution input or motion blur, where genuine faces occasionally failed the heuristic checks. Despite these limitations, the system maintained an average frame rate of approximately 10 frames per second (FPS), providing smooth and responsive real-time detection. The fusion of YOLOv8 with handcrafted verification techniques yielded a practical and explainable solution for identifying facial forgery.

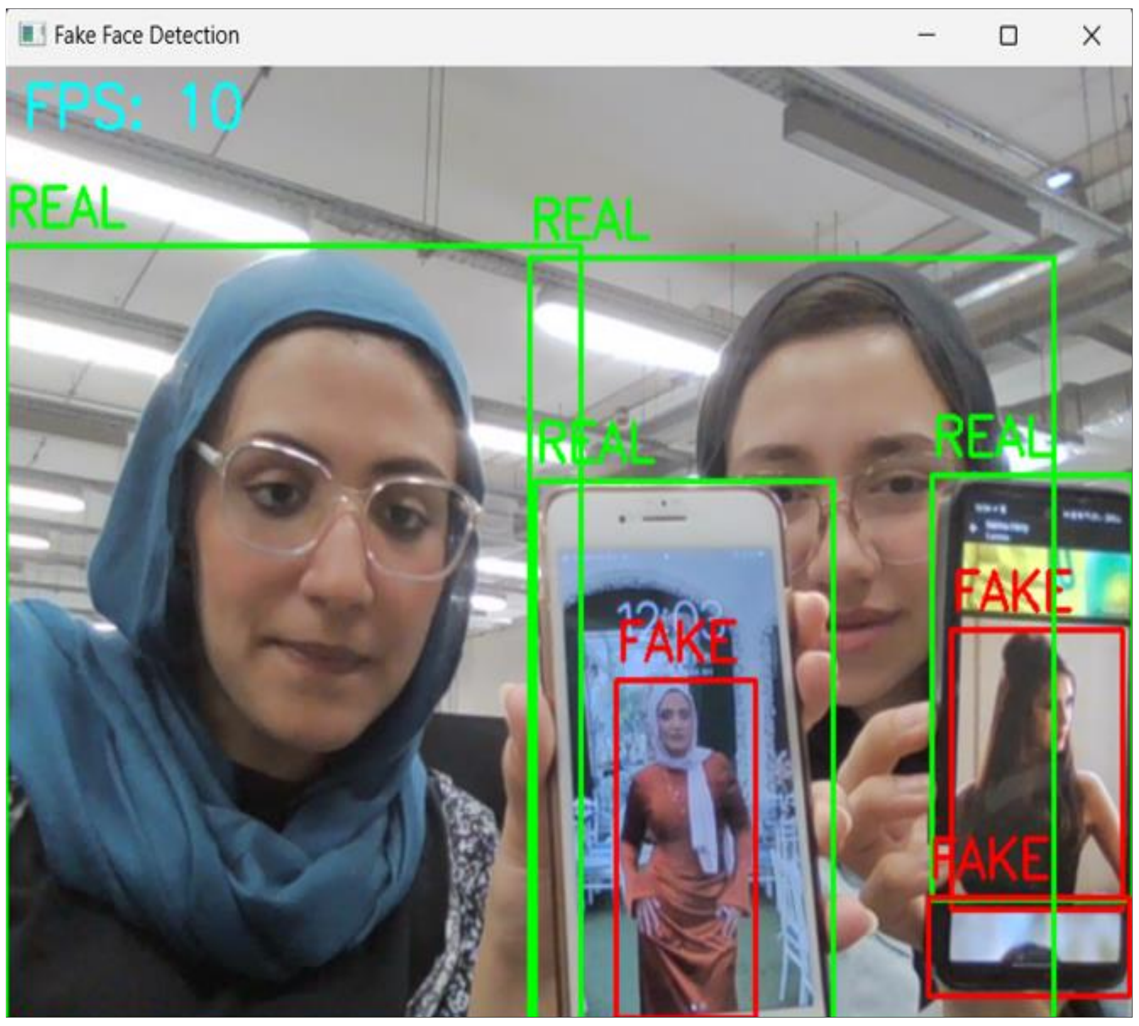

Figure 1: Depth analysis of detection parameters based on blur and color metrics.

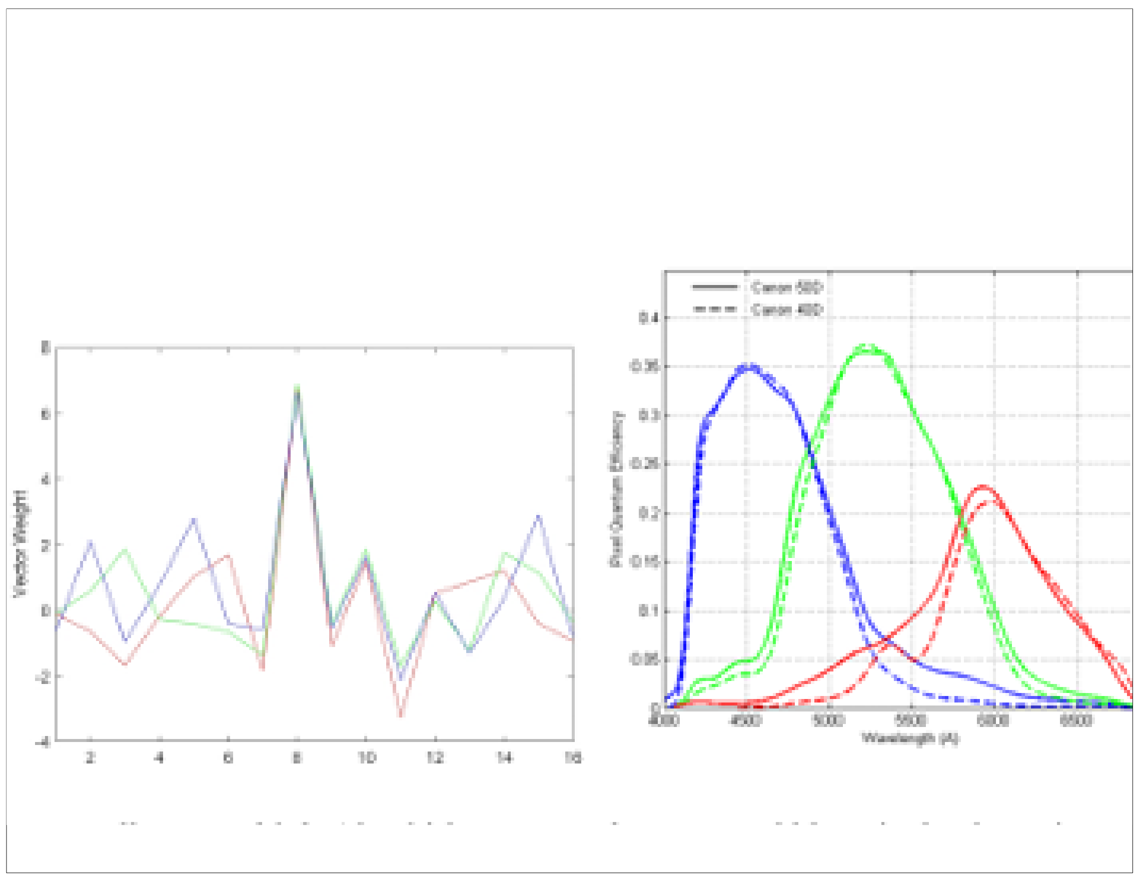

Figure 2: Frame rate performance (10 FPS) and classification outcomes showing "REAL" vs "FAKE" faces.

Conclusions

This project presents a hybrid approach to real-time fake face detection by integrating deep learning with traditional image analysis techniques. By employing the YOLOv8 model for face localization and supplementing it with handcrafted heuristics, such as texture analysis, blur detection, and color variation assessment, the system effectively distinguishes between real and artificially generated facial images. The combination of data-driven and rule-based methods enhances the robustness and interpretability of the detection process. While the system performs reliably under most conditions, certain limitations, such as false positives in low-light or blurred scenarios, highlight areas for future improvement.

References

- Li, Y., Chang, M. C., & Lyu, S. (2018). *Exposing DeepFake Videos by Detecting Face Warping Artifacts*. arXiv preprint arXiv:1812.08247.
- Ultralytics. *YOLOv8: Cutting-edge object detection model*.
- CVZone. *FaceDetectionModule for OpenCV*.