Computer Vision - 2025

Lecture #07. Face Recognition

Lectures by Alexei Kornaev [1,2,3]
Practical sessions by Kirill Yakovlev [2]

[1]AI Institute, Innopolis University (IU), Innopolis
[2]Robotics & CV Master's Program, IU, Innopolis
[3]RC for AI, National RC for Oncology, Moscow

March 2, 2025

Agenda

Section 1. Outcomes

Outcomes

This week on Face Recognition aims to provide students with a comprehensive understanding of face detection and recognition techniques. By the end of this lecture, students will be able to:

1. Estimate the challenges in the domain, including pose variation, illumination changes, occlusion, ethics, and adversarial attacks.
2. Understand and compare different face recognition architectures.
3. Analyze loss functions like contrastive loss, triplet loss, and ArcFace loss for face recognition models.
4. Implement face recognition systems using deep learning frameworks and estimate the results.

Key Takeaway: Face recognition is a crucial technology in security, authentication, and social applications, but it also raises ethical concerns that must be addressed.

Section 2. Challenges in Face Recognition

Challenges in Face Recognition

Overview of Challenges

Face recognition systems face several challenges that can significantly impact their performance. These include pose variation, illumination changes, occlusion, ethical concerns, and adversarial attacks.

- **Pose Variation:** Faces can appear in different orientations, making it difficult to match them.
- **Illumination Changes:** Variations in lighting conditions can alter the appearance of faces.
- **Occlusion:** Parts of the face may be obscured by objects or other faces.
- **Ethics:** Privacy concerns and biases in face recognition systems.
- **Adversarial Attacks:** Deliberate attempts to fool face recognition systems.

Pose Variation

Pose Variation

Pose variation refers to the different orientations of a face relative to the camera. This can include frontal, profile, and oblique views.

- **Impact:** Reduces the accuracy of face recognition systems.
- **Solutions:** Use of 3D models, multi-view training data, and pose-invariant features.

Illumination Changes

Illumination Changes

Illumination changes refer to variations in lighting conditions that can alter the appearance of a face.

- **Impact:** Can cause significant changes in the appearance of facial features.
- **Solutions:** Normalization techniques, use of infrared imaging, and deep learning models trained under various lighting conditions.

Occlusion

Occlusion occurs when parts of a face are obscured by objects (e.g., masks, sunglasses), other people, or environmental factors. It degrades recognition accuracy by hiding critical facial features.

Understanding Occlusion

Figure: Examples of occluded face images from the MAFA dataset [Zeng et al., 2021].

- **Impact:**
  - Loss of key features (e.g., eyes, mouth).
  - State-of-the-art models like DeepFace or ArcFace suffer >30% accuracy drop under heavy occlusion.
- **Common Occlusions:** *Masks* (e.g., medical masks). *Accessories* (e.g., sunglasses, scarves). *Environmental* (e.g., hands, hair).
- **Challenges:**
  - Partial face data limits recognition methods.
  - Dynamic occlusions require real-time adaptation.

Solutions for Occlusion Handling

Technical Approaches

Addressing occlusion requires robust feature extraction and reconstruction techniques.

- **Partial Face Recognition:**
  - Train models on cropped facial regions (e.g., eyes-only, mouth-only). Use *metric learning* to match partial and holistic embeddings.
- **Attention Mechanisms:**
  - Focus on visible regions, e.g. *Transformer Networks*) may prioritize non-occluded areas with attention [Liu et al., 2023].
- **Generative Models:**
  - Reconstruct occluded regions using GANs, Diffusion models.

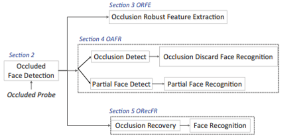

Figure: FR under occlusion [Zeng et al., 2021]

Ethics: Privacy, Bias, and Misuse

Face recognition systems raise critical ethical concerns, including privacy violations, algorithmic bias, and misuse in surveillance. These challenges undermine trust and fairness.

- **Privacy Violations:**
  - *Issue:* Unauthorized data collection and mass surveillance.
  - *Approach:* Implement strict data governance, anonymization, and user consent mechanisms.
- **Algorithmic Bias:**
  - *Issue:* Higher error rates for minorities, women, and darker-skinned individuals due to imbalanced training data.
  - *Example:* Buolamwini & Gebru (2018) showed 34% error for dark-skinned women vs. 0.8% for light-skinned men.
  - *Approach:* Curate diverse datasets, audit models for fairness, and adopt bias-mitigation techniques (e.g., reweighting losses).

Ethical Challenges: Solutions and Regulations

Regulatory Frameworks

Addressing ethical issues requires technical, legal, and societal interventions.

- **Technical Solutions:**
  - *Fairness-aware training:* Use fairness constraints during model optimization.
  - *Explainability:* Tools like LIME or SHAP to interpret model decisions.
- **Legal Solutions:**
  - *Bans on mass surveillance:* EU's GDPR, proposed U.S. FR Act.
  - *Transparency laws:* Require disclosure of face recognition use cases.
- **Societal Solutions:**
  - Public awareness campaigns and ethical AI certification programs.

Adversarial Attacks: Threat Models

CV-2025

A.Kornaev,
K.Yakovlev

Outcomes

Challenges in Face Recognition

Conclusion & Discussion

Adversarial Attacks in Face Recognition

Adversarial attacks manipulate inputs to deceive face recognition systems. Common types include:

- **Evasion Attacks:**
  - *Goal:* Fool the system during inference (e.g., adversarial glasses).
- **Poisoning Attacks:**
  - *Goal:* Corrupt training data (e.g. injecting malicious samples).
- **Model Inversion Attacks:**
  - *Goal:* Reconstruct training data from model outputs (privacy breach).

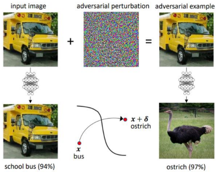

Figure: Adversarial noise attack.

INNOPOLIS UNIVERSITY

Defending Against Adversarial Attacks

Robustness Strategies

Mitigating adversarial attacks requires proactive defenses and robust model design.

- **Adversarial Training:**
  - Train models on adversarial examples to improve robustness.
  - *Limitation:* Computationally expensive; does not generalize to all attacks.
- **Input Preprocessing:**
  - Apply noise reduction, quantization, or JPEG compression to inputs.
- **Detection Mechanisms:**
  - Deploy detectors to flag adversarial inputs (e.g., using inconsistency in feature space).
- **Certified Defenses:**
  - Use mathematically proven robust models (e.g., randomized smoothing).

Paper reading

LUVLi Face Alignment: Estimating Landmarks' Location, Uncertainty, and Visibility Likelihood

Modern face alignment methods have become quite accurate at predicting the locations of facial landmarks, but they do not typically estimate the uncertainty of their predicted locations nor predict whether landmarks are visible. In this paper, we present a novel framework for jointly predicting landmark locations, associated uncertainties of these predicted locations, and landmark visibilities. We model these as mixed random variables and estimate them using a deep network trained with our proposed Location, Uncertainty, and Visibility Likelihood (LUVLi) loss. In addition, we release an entirely new labeling of a large face alignment dataset with over 19,000 face images in a full range of head poses. Each face is manually labeled with the ground-truth locations of 68 landmarks, with the additional information of whether each landmark is unoccluded, self-occluded (due to extreme head poses), or externally occluded. Not only does our joint estimation yield accurate estimates of the uncertainty of predicted landmark locations, but it also yields state-of-the-art estimates for the landmark locations themselves on multiple standard face alignment datasets. Our method's estimates of the uncertainty of predicted landmark locations could be used to automatically identify input images on which face alignment fails, which can be critical for downstream tasks Kumar et al. [2020].

Section 3. Conclusion & Discussion

Conclusion & Discussion

Strengths and Limitations

- **Stereo Vision:** Good accuracy but needs two cameras.
- **Monocular Depth Estimation:** Works with single images, but lacks absolute depth accuracy.
- **LiDAR:** High precision but expensive and sparse.

Future Trends

- Neural Radiance Fields (NeRF) for photo-realistic depth estimation Mildenhall et al. [2020].
- Transformer-based Keypoint Detection for more robust landmark estimation.

Open Challenges

- Depth estimation under low-light or occlusions.
- Real-time landmark detection with high precision in occluded faces.

Bibliography

Abhinav Kumar, Tim K Marks, Wenxuan Mou, Ye Wang, Michael Jones, Anoop Cherian, Toshiaki Koike-Akino, Xiaoming Liu, and Chen Feng. Luvli face alignment: Estimating landmarks' location, uncertainty, and visibility likelihood. In **Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition**, pages 8236–8246, 2020.

Chang Liu, Kaoru Hirota, and Yaping Dai. Patch attention convolutional vision transformer for facial expression recognition with occlusion. **Information Sciences**, 619:781–794, 2023.

Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis, 2020. URL https://arxiv.org/abs/2003.08934.

Dan Zeng, Raymond Veldhuis, and Luuk Spreeuwers. A survey of face recognition techniques under occlusion. **IET biometrics**, 10 (6):581–606, 2021.