

Problem statement 1:

Masked Face Recognition: Identifying People Behind the Mask

Difficulty: Intermediate

A) Background

Videonetics' Face Recognition System performs real-time identification across camera feeds for access control and watch-list matching, recognizing faces as small as 30×30 pixels from a gallery of up to a million faces. In practice, the enrollment gallery is built from **unmasked** images (ID photos, prior enrollments), but in the field — post-COVID norms, security concealment, weather, pollution — subjects frequently wear masks that occlude the lower half of the face (nose, mouth, chin), which is also the most discriminative region for conventional recognizers. A face model trained on full faces learns to attend to features like the eyes, nose, lips and face edges, so when a mask hides them the system can fail to identify the person. **The challenge is to close that gap:** enroll/train on unmasked faces, yet still recognize a person reliably when they are masked.

B) Precisely Scoped Task

Minimum viable deliverable: Train (or enroll embeddings for) a face recognition model on unmasked face images, then evaluate it on masked probe images matched against the unmasked gallery. Report rank-1 identification accuracy and/or 1:1 verification accuracy (ROC-AUC, TAR@FAR). The headline result is the masked-probe-vs-unmasked-gallery accuracy and the generalization gap relative to the unmasked-vs-unmasked baseline.

Ambitious version: Actively close the gap — synthetic mask augmentation during training, occlusion-aware or periocular (eye-region) modeling, attention to the visible upper face — and demonstrate robustness across mask types/colors and partial extra occlusion. Add 1:N watch-list search with confidence scoring.

C) Verified Public Datasets

- **LFW (Labeled Faces in the Wild):** 13,233 images, 5,749 identities; standard verification benchmark with matched/mismatched pairs. vis-www.cs.umass.edu/lfw
- **MaskTheFace (mask-synthesis tool, the key enabler):** dlib-landmark script that warps realistic masks onto any unmasked face to generate masked training/test pairs; reports ~38% TPR increase for FaceNet after retraining. github.com/aeqelanwar/MaskTheFace
- **MFR2 (real masked test set):** 53 identities, 269 images with both masked and unmasked faces per identity, pre-aligned 160×160, ready pairs file. Fetch via MaskTheFace.
- **MaskedFace-Net (large synthetic):** ~137,016 images (Correctly + Incorrectly Masked, ~70k each) from Flickr-Faces-HQ. github.com/cabani/MaskedFace-Net
- **RMFRD / RMFD:** real masked + normal faces with identity labels for recognition. github.com/X-zhangyang/Real-World-Masked-Face-Dataset

- **MAFA:** 30,811 images, 35,806 masked faces — detection only, no identity labels (face-detection stage only).

D) Suggested Baseline / Tech Stack

Use InsightFace/ArcFace (buffalo_l pre-trained pack) for detection + embeddings and cosine-similarity matching to establish the unmasked baseline. Then run MaskTheFace over the unmasked training set to create masked pairs, and either fine-tune the embedding model or train an occlusion-robust head; evaluate the closed gap on MFR2/RMFRD. DeepFace works as a quick wrapper. Sanity-check target: ArcFace reaches ~99.8% on unmasked LFW — the interesting number is how much of that you retain on masked probes.

E) Evaluation Rubric (100 pts)

- Masked face recognition accuracy — rank-1 identification and/or verification accuracy on masked probes vs unmasked gallery **(35)**
- Generalization gap & robustness — accuracy drop vs the unmasked baseline; consistency across mask types/colors **(15)**
- Technical execution & code quality **(15)**
- Innovation — occlusion-aware/periorcular architecture, augmentation strategy, smart training pipeline **(10)**
- Real-world applicability & deployment feasibility **(10)**
- Demo quality — live webcam: enroll a person unmasked, then recognize them masked in real time **(10)**
- Presentation & communication **(5)**