

Classical Face Recognition Under Real-World Variations

Ayoub EL KBADI Fotios KAPOTOS Jean-Vincent MARTINI

CentraleSupélec

February 15, 2026

- Face recognition predates deep learning: explicit features + linear modeling.
- Classical methods remain relevant for:
 - interpretability and analyzable failure modes,
 - low compute / embedded settings,
 - limited training data regimes.
- Goal: compare global vs local representations under real-world degradations.

Problem Definition

Dataset $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^N$, $x_i \in \mathbb{R}^{H \times W}$, $y_i \in \{1, \dots, C\}$.

Learn a classifier $f : \mathbb{R}^{H \times W} \rightarrow \{1, \dots, C\}$.

Robustness via degradations $\delta_\alpha(\cdot)$ (illumination, noise, blur, rotation, crop, occlusion).

Accuracy:

$$\text{Acc} = \frac{1}{|\mathcal{D}_{\text{test}}|} \sum_{(x_j, y_j) \in \mathcal{D}_{\text{test}}} \mathbb{I}(f(\delta_\alpha(x_j)) = y_j).$$

Methods Overview

Global subspace methods

- **Eigenfaces (PCA)**: project faces into top-variance linear subspace; 1-NN in PCA space.
- **Fisherfaces (PCA+LDA)**: discriminative projection maximizing between-class / within-class scatter.

Local texture method

- **LBPH**: Local Binary Patterns + regional histograms; nearest-neighbor with χ^2 distance.

Baseline

- **TinyCNN** trained from scratch (no augmentation/pretraining).

Eigenfaces (PCA): Core Idea

Vectorize image $\mathbf{x} \in \mathbb{R}^D$. Mean $\boldsymbol{\mu} = \frac{1}{N} \sum_i \mathbf{x}_i$. Centered matrix $X = [\mathbf{x}_1 - \boldsymbol{\mu}, \dots, \mathbf{x}_N - \boldsymbol{\mu}] \in \mathbb{R}^{D \times N}$.

PCA subspace $U_K \in \mathbb{R}^{D \times K}$ from covariance $C = \frac{1}{N} X X^\top$.

Projection:

$$\mathbf{z} = U_K^\top (\mathbf{x} - \boldsymbol{\mu}).$$

Classification: 1-NN (Euclidean) in \mathbf{z} -space.

Fisherfaces (PCA + LDA)

Within/Between scatter:

$$S_B = \sum_{i=1}^c N_i (\mu_i - \mu)(\mu_i - \mu)^\top, \quad S_W = \sum_{i=1}^c \sum_{\mathbf{x} \in X_i} (\mathbf{x} - \mu_i)(\mathbf{x} - \mu_i)^\top.$$

Solve $S_B \mathbf{w} = \lambda S_W \mathbf{w}$ (max $c - 1$ directions).

Small sample size: S_W singular in pixel space \Rightarrow PCA to $k_{\text{pca}} \leq N - c$, then LDA to $k_{\text{lda}} \leq c - 1$.

Key practical point: stability strongly depends on controlling/capping k_{pca} .

LBPH: Local Binary Pattern Histograms

LBP code at (x, y) (basic 3×3):

$$\text{LBP}(x, y) = \sum_{p=0}^7 s(g_p - g_c) 2^p, \quad s(t) = \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases}$$

LBPH: split image into regions $\{R_j\}$, compute LBP histograms per region, concatenate.

Matching via χ^2 distance:

$$\chi^2(S, M) = \sum_{i,j} \frac{(S_{i,j} - M_{i,j})^2}{S_{i,j} + M_{i,j}}.$$

Hyperparameters: sampling points P , radius R , number of regions.

ORL (AT&T)

- 40 subjects, 10 images/subject, grayscale.
- Moderate variation: expression, slight pose, glasses.
- Good for low-data sensitivity analysis.

Yale

- Frontal faces with systematic illumination variations.
- Challenging for global appearance models; classic benchmark for illumination robustness.

- Accuracy vs # training images per subject (fixed test identities, repeated random draws).
- Robustness sweeps with controlled degradations:
 - blur, noise, photometric jitter,
 - rotation, flips,
 - cropping / partial occlusion.
- Report mean and uncertainty (std / CI depending on experiment).

Global Comparison (Low-Data)

ORL

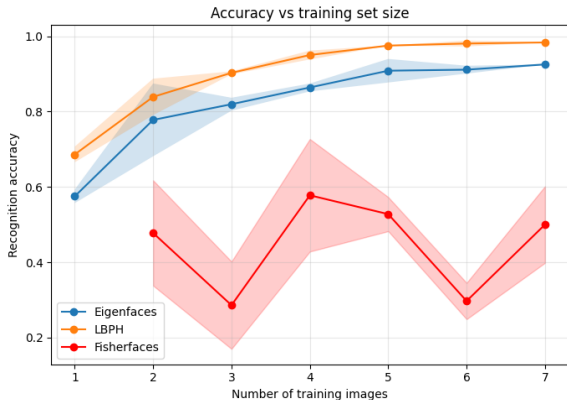
- Eigenfaces: strong dependence on train size; saturates $\sim 92\text{--}93\%$.
- LBPH: consistently higher; saturates early $\geq 97\%$.
- Fisherfaces: config-sensitive; capping k_{pca} stabilizes performance.

(Insert your figures on the next slide.)

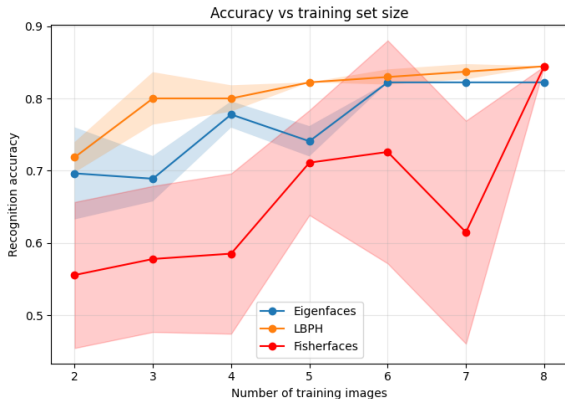
Yale

- Overall lower due to illumination.
- Eigenfaces saturates $\sim 82\%$.
- LBPH $\sim 84\text{--}85\%$ with better robustness.
- Fisherfaces: benefits from more data; still sensitive to geometry shifts.

Global Comparison Figures



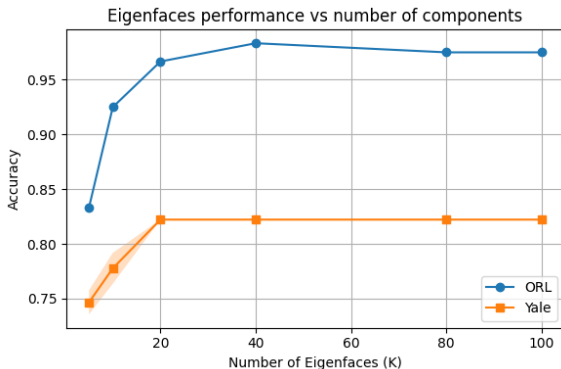
ORL: Eigenfaces vs LBPH



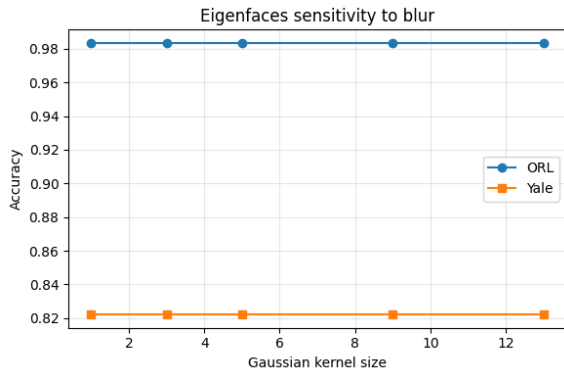
Yale: Eigenfaces vs LBPH

Eigenfaces: Robustness Findings

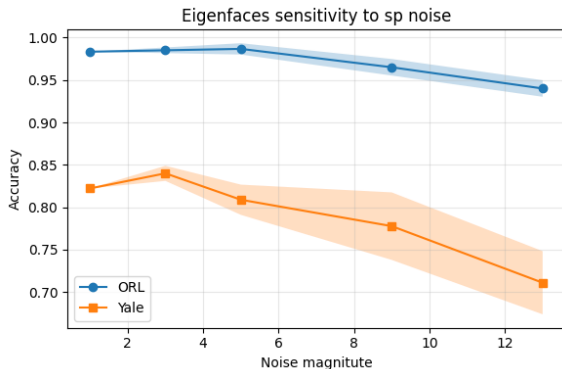
- **Hyperparameter K :** accuracy improves quickly then saturates beyond $K \approx 50$.
- **Most destructive:** geometric misalignment (rotation, crop).
- **Moderate effect:** photometric jitter + salt-and-pepper noise (global corruption impacts projection).
- **Less effect:** blur (in your experiments: relatively mild degradation).



Eigenfaces: Photometric Robustness

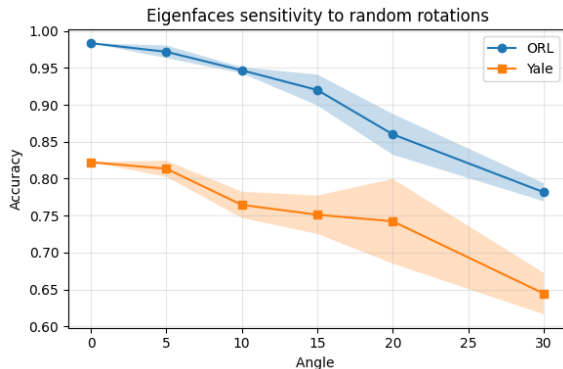


Gaussian blur

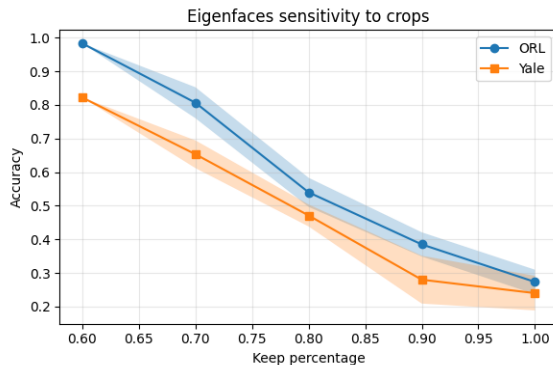


Salt-and-pepper noise

Eigenfaces: Geometric Robustness



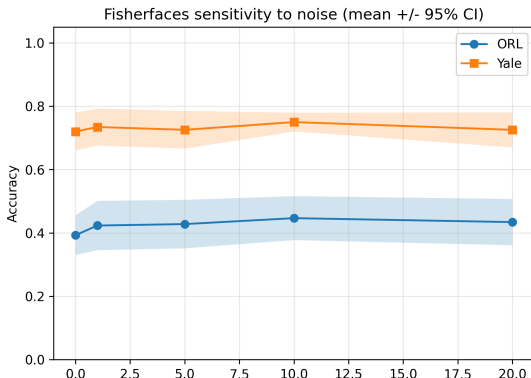
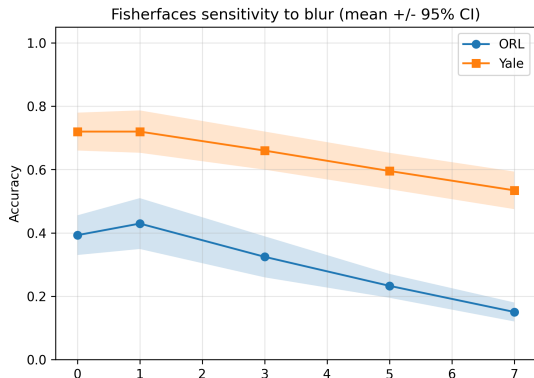
Rotation



Cropping

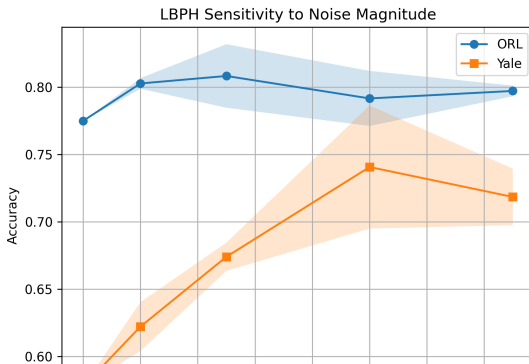
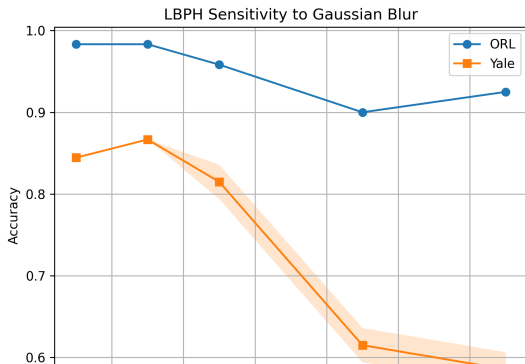
Fisherfaces: Key Results

- Dominant design choice: k_{pca} .
 - Paper-aligned $k_{pca} = N - c$ can induce peaking/non-monotonicity (esp. ORL).
 - Fixed/capped k_{pca} improves stability and accuracy.
- Test-only degradations: blur, Gaussian noise, rotation, brightness scaling, flips.
- Major bottleneck: **geometry** (rotation/flip) \gg small additive noise.

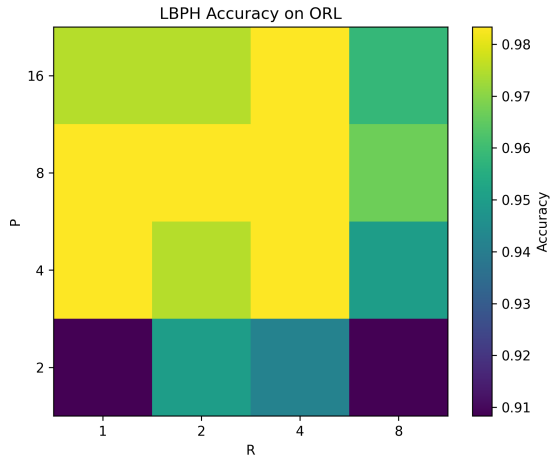


LBPH: Key Results

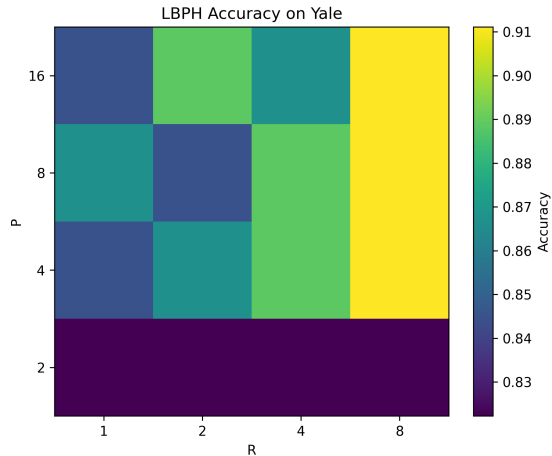
- Hyperparameters P, R : relatively insensitive; good region around $P \in [4, 8]$, $R \in [2, 4]$.
- Robust to photometric jitter and salt-and-pepper noise (local encoding).
- Rotation tolerance up to moderate angles; degrades beyond $\sim 15^\circ$.
- Main failure mode: **cropping** (removes critical facial regions / breaks spatial layout).
- Also sensitive to heavy blur (esp. Yale).



LBPH: Hyperparameter Sensitivity



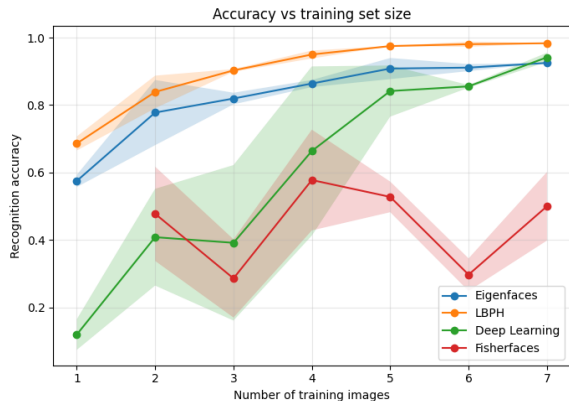
ORL: P, R sweep



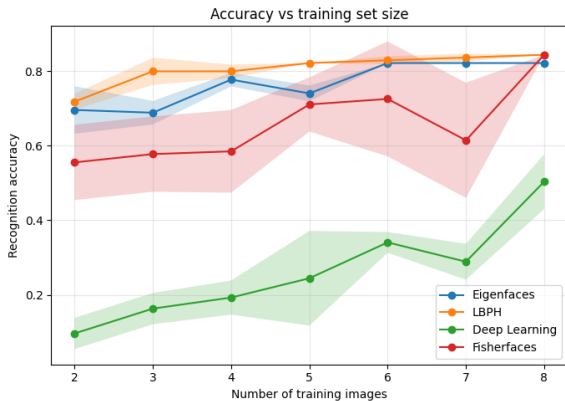
Yale: P, R sweep

Deep Learning Baseline (TinyCNN)

- Trained from scratch (grayscale, CE loss, Adam), no augmentation/pretraining.
- ORL: improves with train size, but weak in low-data; likely overfits.
- Yale: consistently low + high variance \Rightarrow data scarcity hurts generalization.



ORL



Yale

- Clear hierarchy in your regime:
 - **Eigenfaces**: fragile under misalignment and illumination shifts.
 - **Fisherfaces**: better discrimination/illumination handling, but stability hinges on k_{pca} .
 - **LBPH**: most reliable overall in low-data and illumination variability; fails under heavy crop/blur.
- Deep learning (from scratch) underperforms without enough data or stronger regularization.
- Practical message: classical methods remain useful baselines with interpretable failure modes.

Code: <https://github.com/fotisk07/VIC-Project>

References (selected) I

 M. Turk and A. Pentland. Eigenfaces for recognition. 1991.

 P. Belhumeur et al. Eigenfaces vs. Fisherfaces. 1997.

 T. Ahonen et al. Face recognition with LBP. 2004.

 F. Samaria and A. Harter. Parameterisation of a stochastic model for human face identification. 1994.

 A. Georgiades et al. From few to many: illumination cone models. 2001.

 Y. Taigman et al. DeepFace. 2014.

 F. Schroff et al. FaceNet. 2015.

 J. Deng et al. ArcFace. 2019.