Reproduction of Application of Linear and Nonlinear PCA to SAR ATR

Amit Kumar Mishra and Tshiamo Motaung

May Yuan-Klitzner

January 15, 2026

1 Executive Summary

Successfully reproduced paper's key results with similar agreement:

- **Linear PCA (20 Principal Components):** 86.78% vs paper's 85.84% (+0.94%)
- **Polynomial KPCA (16 Principal Components):** 93.64% vs. paper's 92.67% (+0.97%)
- **Gaussian KPCA (16 Principal Components):** 83.42% vs. 85.79% (-2.37%)

Key finding validated: The polynomial KPCA (non-linear PCA) outperforms both Linear PCA and Gaussian KPCA for SAR target classification.

2 Dataset

Table 1: Parameters used in simulation with its corresponding value

Parameter	**Value**
Source	MSTAR-PublicMixedTargets-CD1 and MSTAR-PublicMixedTargets-CD2
Target Classes	5 (BRDM_2, D7, T62, ZIL131, ZSU_23_4)
Training Dataset	17 degree depression, 1495 total target clips (299 clips per class)
Testing Dataset	15 degree depression, 1369 total target clips (274 clips per class)
Image Size	96×96 pixels (center-cropped from 128×129)
Features	9216 per image (96×96 flattened)

Table 2: Number of Target Clips per class at different elevations

Target Class	**15 degree elevation**	**17 degree elevation**
t001 - BRDM_2	274	299
t005 - D7	274	299
t016 - T62	273	299
t025 - ZIL131	274	299
t026 - ZSU_23_4	274	299

3 Main Results: Table II Reproduction

Table 3: Comparison of results to paper's % Pcc and % Error rate

PCA System	**No. of PCs**	**Paper % Pcc**	**Paper % Error rate**	**My % Pcc**	**My % Error rate**	**Difference %**
Linear PCA	20	85.841	14.159	86.78	13.209	+0.94
Polynomial KPCA	8	89.2308	10.7692	89.34	10.6692	+0.10
Polynomial KPCA	12	92.1612	7.8388	92.48	7.5188	+0.32
Polynomial KPCA	16	92.6740	7.326	93.64	6.356	+0.97
Gaussian KPCA	8	79.2674	20.7326	78.45	21.5526	-0.82
Gaussian KPCA	12	85.0549	14.9451	81.88	18.1151	-3.17
Gaussian KPCA	16	85.7875	14.2125	83.42	16.5825	-2.37

Key Findings Validated:

- Polynomial PCA consistently outperforms Gaussian KPCA and Linear PCA
- There was no significant performance improvement shown by the Gaussian KPCA system compared to the Linear PCA system
- Best system: Poly KPCA with 16 PCs (93.64%)
- Paper's claim validated: Nonlinear ¿ Linear PCA

Table 4: Comparison of results to paper's Computation Time

PCA System	**No. of PCs**	**Paper's Computation Time (s)**	**My Computation Time (s)**
Linear PCA	20	3.4529	6.9043
Polynomial KPCA	8	4.7414	1.8074
Polynomial KPCA	12	4.6606	1.5948
Polynomial KPCA	16	4.8182	1.4400
Gaussian KPCA	8	702.1897	1.6838
Gaussian KPCA	12	702.8210	1.6668
Gaussian KPCA	16	702.7004	1.7557

Key Findings:

- Average computation time for Polynomial KPCA: 1.6140 s
- Average computation time for Gaussian KPCA: 1.7021 s
- The runtime of Linear PCA was the slowest and the runtime of Poly KPCA was close to the runtime of Gaussian KPCA

Possible reasons for discrepancies:

- Difference in hardware: Intel i5 CPU@3.40GHz system vs. AMD Ryzen 7 5700U @1.80 GHz system
- Use of built-in Python libraries could have sped up computation process (scikit-learn)
- Images were initially 128×129 pixels, by centre-cropping the images to be 96×96 pixels could have differed the data that was used in the paper
- Original dataset found could have been similar to dataset used in paper but possibly not the exact same database

4 Linear PCA Saturation Analysis

Table 5: Percentage of Correct Classification vs. Number of Principal Components

No. of PCs	% Pcc	Avg. Gain (%)	Cum. Gain (%)
1	25.20	0.000	0.00
5	52.01	6.703	26.81
10	76.63	5.714	51.43
15	85.03	4.274	59.83
20	86.78	3.241	61.58
25	87.87	2.611	62.67
30	90.36	2.247	65.16
35	91.16	1.970	+65.96 (peak)
40	91.02	1.688	+65.82 (overfitting)
60	90.94	1.114	65.74
80	90.58	0.828	65.38
100	89.63	0.651	64.43

Key Findings Validated:

- **Peak found at 35 Principal Components:** 91.16%
- After 20 Principal Components, incremental gain per PC is around 2% and decreases per PC, therefore saturation around 20 PCs is plausible as it balances accuracy and efficiency.
- After 35 Principal components, accuracy decreases due to overfitting.

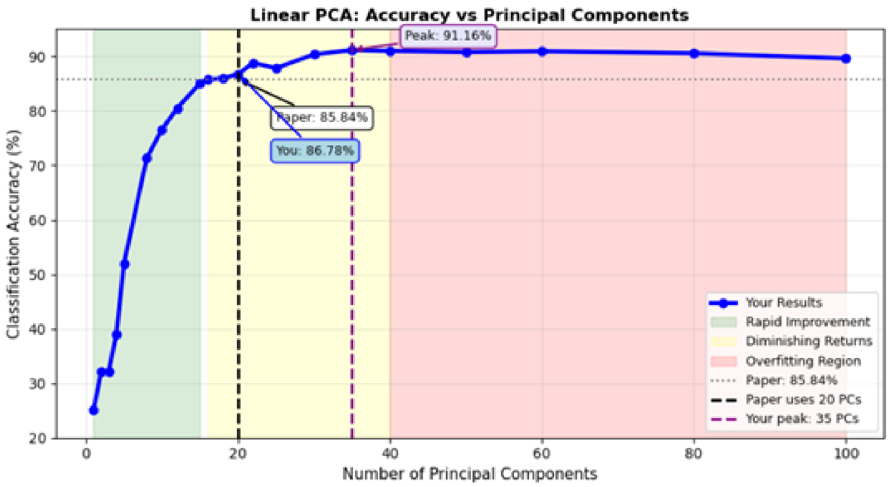

Figure 1: Accuracy vs Number of Principal Components

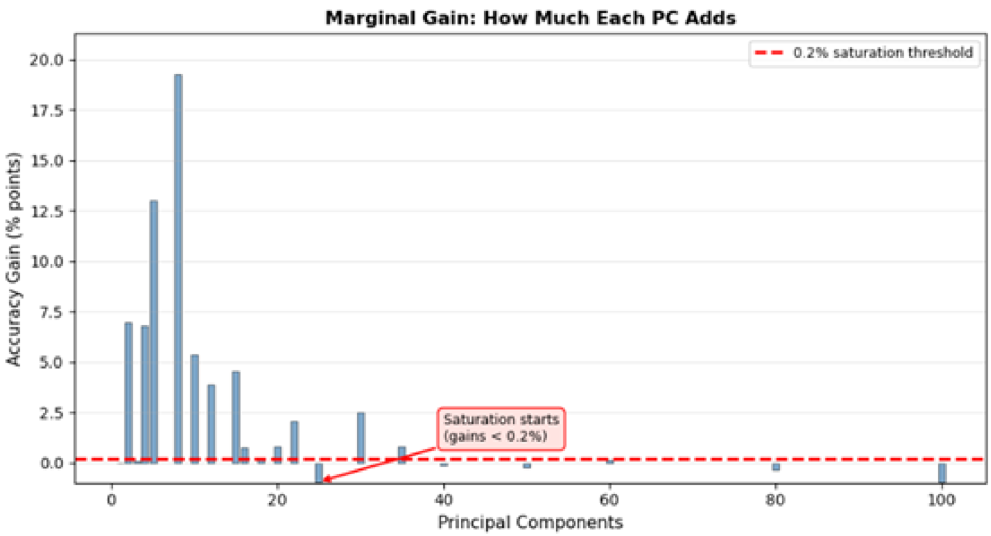

Figure 2: Marginal Gain – How much accuracy each PC adds

5 Reduced Training Dataset Results: Robustness to incomplete training data

Table 6: Linear PCA 20 PCs – Reduced Training Dataset

Reduction Factor	No. of Training Samples	% Pcc
r = 1	1495	86.7787
r = 2	747	84.5873
r = 3	498	80.6428
r = 4	374	78.816
r = 5	299	72.2425

Table 7: Polynomial KPCA 8 PCs – Reduced Training Dataset

Reduction Factor	No. of Training Samples	% Pcc
r = 1	1495	89.3353
r = 2	747	87.6552
r = 3	498	85.8291
r = 4	374	84.8064
r = 5	299	85.0256

Table 8: Polynomial KPCA 12 PCs – Reduced Training Dataset

Reduction Factor	No. of Training Samples	% Pcc
r = 1	1495	92.4763
r = 2	747	90.2118
r = 3	498	89.5544
r = 4	374	87.7283
r = 5	299	87.6552

Table 9: Polynomial KPCA 16 PCs – Reduced Training Dataset

Reduction Factor	**No. of Training Samples**	**% Pcc**
r = 1	1495	93.6450
r = 2	747	91.2345
r = 3	498	89.9927
r = 4	374	88.1665
r = 5	299	87.5822

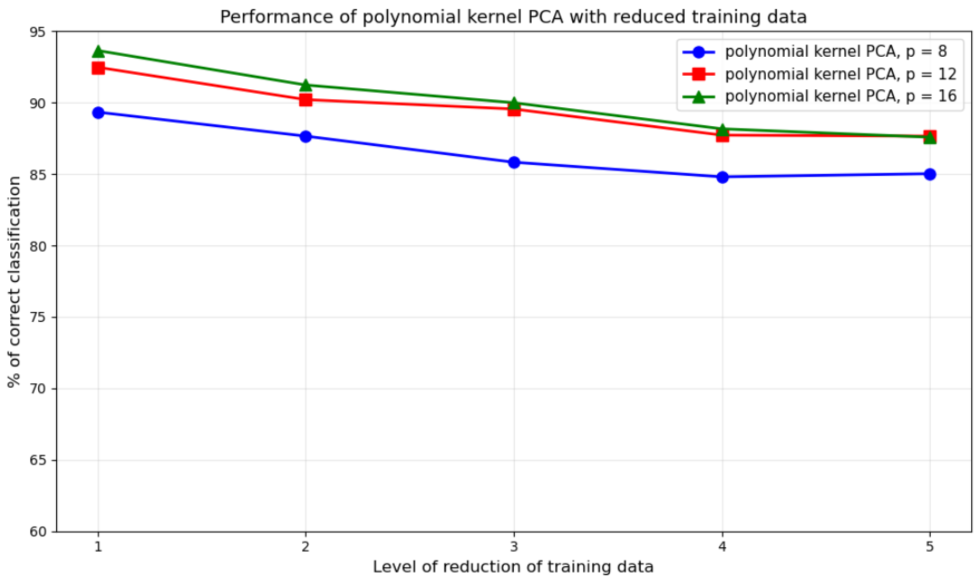

Figure 3: Performance of polynomial kernel PCA with reduced training data

Table 10: Gaussian KPCA 8 PCs – Reduced Training Dataset

Reduction Factor	**No. of Training Samples**	**% Pcc**
r = 1	1495	78.4514
r = 2	747	76.0409
r = 3	498	70.7085
r = 4	374	70.3433
r = 5	299	68.0789

Table 11: Gaussian KPCA 12 PCs – Reduced Training Dataset

Reduction Factor	**No. of Training Samples**	**% Pcc**
r = 1	1495	81.8846
r = 2	747	78.9627
r = 3	498	74.7261
r = 4	374	72.0964
r = 5	299	70.4894

Table 12: Gaussian KPCA 16 PCs – Reduced Training Dataset

Reduction Factor	**No. of Training Samples**	**% Pcc**
r = 1	1495	83.4186
r = 2	747	80.7889
r = 3	498	74.8722
r = 4	374	73.6304
r = 5	299	71.7312

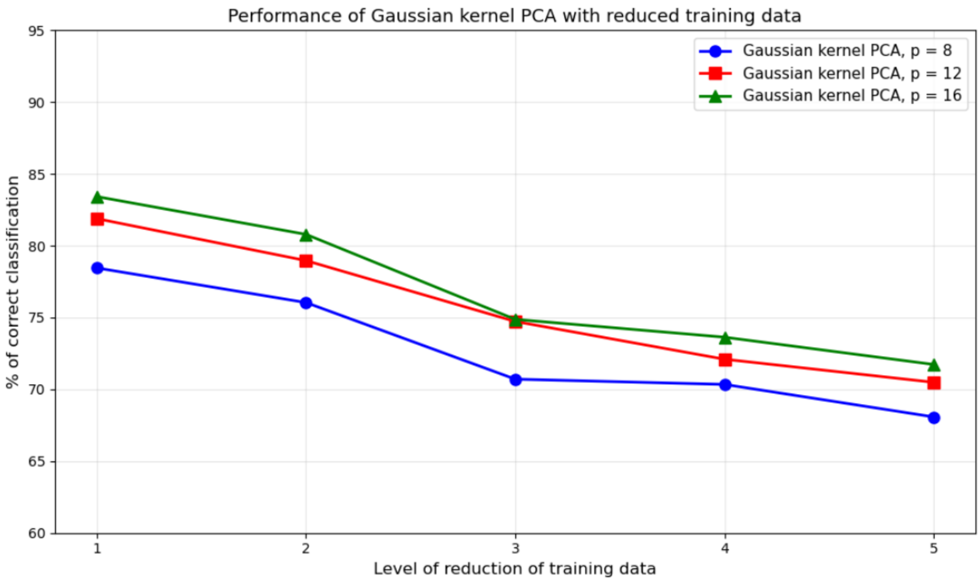

Figure 4: Performance of gaussian kernel PCA with reduced training data

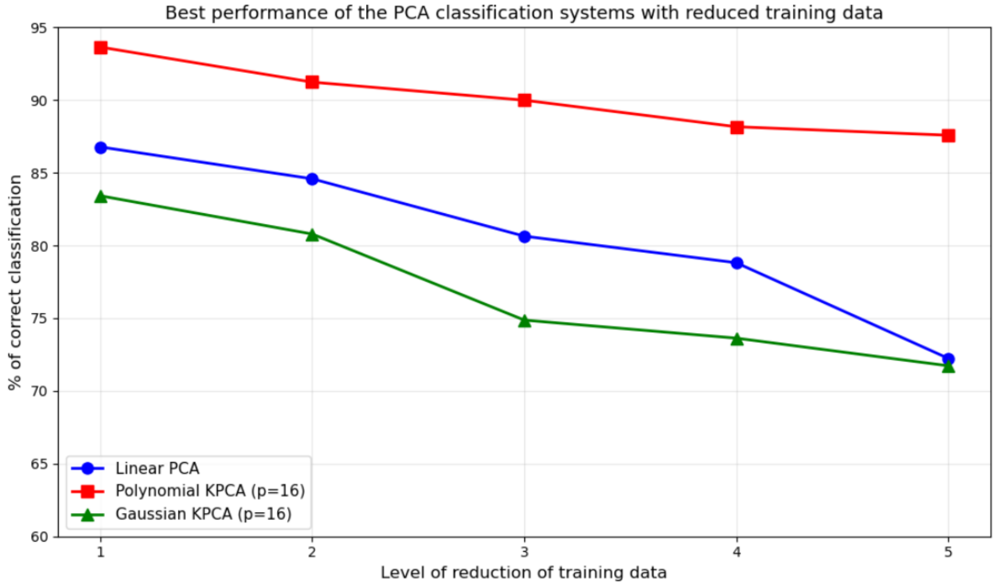

Figure 5: Best performance of the PCA classification systems with reduced training data

Key Findings Validated:

- **Polynomial KPCA (16 PCs)** is the best performer in being robust to variant limitations
- **Polynomial KPCA (8 PCs)** achieved a % correct classification greater than 80% for all reduction factors
- The polynomial kernel PCA system was able to produce higher results than linear PCA (20 PCs) and Gaussian KPCA, using fewer principal components (8 PCs)
- **Gaussian KPCA (8, 12, 16 PCs)** achieved a % of correct classification greater 70% for the first four reduction cases

6 Confusion Matrix of Best System

Table 13: Polynomial KPCA, 16 PCs - Per-Class Performance

	Predicted Class						
True Class	**t001**	**t005**	**t016**	**t025**	**t026**	**Total**	**% Pcc**
t001 (BRDM_2)	253	2	1	16	2	274	92.34
t005 (D7)	2	270	0	0	2	274	98.54
t016 (T62)	1	8	241	5	18	273	88.28
t025 (ZIL131)	5	1	3	261	4	274	95.26
t026 (ZSU_23_4)	1	5	8	3	257	274	93.80
						Overall Accuracy	**93.64**

Key Findings:

- **Best Classified:** D7 (t005) with 98.54% accuracy
- **Most Challenging:** T62 (t016) with 88.28% accuracy, primarily confused with ZSU_23_4 (18 misclassifications)

7 Implementation Details

Preprocessing:

- **Linear PCA:** Z-score standardization (StandardScaler library) - A. K. Mishra and B. Mulgrew, "Bistatic SAR ATR", IET Radar Sonar Navig., 2007, Vol. 1, pp. 459–469.
- **Kernel PCA:** Raw pixel values (no standardization)
- **Eigenvectors:** Not normalized

Table 14: Kernel Parameters

System	**Parameter Value**
Polynomial KPCA	d (degree) = 1/14
Gaussian KPCA	σ (bandwidth) = 5

Key Equations Implemented

1. **Gram Matrix Centering:** $\widetilde{K} = K - \mathbf{1}_n K - K\mathbf{1}_n + \mathbf{1}_n K \mathbf{1}_n$
2. **Polynomial kernel:** $k(\mathbf{x}, \mathbf{y}) = (\mathbf{x}^T \mathbf{y})^{1/14}$
3. **Gaussian kernel:** $k(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{2\times 5^2}\right)$

8 Conclusion

- Successfully reproduced paper's main findings with comparable accuracy scores
- Nonlinear PCA approach outperformed the linear PCA approach for SAR target classification
- The polynomial kernel PCA system achieved higher accuracy than linear PCA using fewer principal components (8 PCs)
- The optimal balance between accuracy and efficiency for linear PCA was achieved with 20 principal components
- The Gaussian kernel-based classification system did not offer significant performance improvement relative to the linear PCA approach
- Nonlinear Kernel PCA systems demonstrated greater robustness when classifying targets with reduced training datasets

Bibliography

References

[1] A. Kumar and T. Motaung, "Application of Linear and Nonlinear PCA to SAR ATR," 2015.

[2] A. K. Mishra and B. Mulgrew, "Bistatic SAR ATR," *IET Radar, Sonar & Navigation*, vol. 1, no. 6, pp. 459–469, Dec. 2007.

[3] Q. Wang, "Kernel Principal Component Analysis and its Applications in Face Recognition and Active Shape Models," Master's thesis, Rensselaer Polytechnic Institute, Troy, NY, 2012.

[4] I. T. Jolliffe, *Principal Component Analysis, Second Edition*, Springer-Verlag, New York, 2002.

[5] Air Force Research Laboratory, *MSTAR Public Targets*, Sensor Data Management System, `https://www.sdms.afrl.af.mil/index.php?collection=mstar&page=targets`,