

Faculty of Engineering and Technology  
Department of Electrical and Computer Engineering

## **Computer Vision - ENCS5343**

---

# **Arabic Handwritten Text Identification Using Local Feature Extraction Techniques**

---

Prepared by:

Mohammed Abed Alkareem - 1210708

Supervisor:

Dr.Aziz Qaroush

Birzeit University

December 4, 2024

# 1. Objective

The objective of this report is to evaluate and compare the performance of two local feature extraction techniques, SIFT (Scale-Invariant Feature Transform) and ORB (Oriented FAST and Rotated BRIEF), for identifying Arabic handwritten text. The focus is on assessing their accuracy, efficiency, and robustness across a variety of transformations and conditions.

## 2. Introduction

Feature extraction is a crucial step in computer vision tasks. It enables systems to identify and differentiate patterns effectively. SIFT and ORB are popular local feature extraction techniques, each offering unique advantages. This report evaluates their performance on the AHAWP dataset, a benchmark dataset for handwritten Arabic text. The study aims to assess:

- The quality of features extracted by each algorithm.
- Their robustness to transformations like noise, rotation, scale, and illumination changes.
- Their computational efficiency and generalization capabilities.

## 3. Dataset

The AHAWP dataset consists of **8144 grayscale images** of **10 unique Arabic words**, handwritten by **82 individuals**. The dataset was divided into **60% training (4886 images)**, **20% validation (1629 images)**, and **20% testing (1629 images)**.

Preprocessing steps included the conversion of all images to grayscale for uniformity, augmentation of the test data with transformations such as Gaussian noise, rotation, scaling, and illumination changes to evaluate robustness, and the assignment of labels corresponding to the handwritten Arabic word represented in each image.

## 4. Methodology

### 4.1 Feature Extraction

- **SIFT:** Extracts 128-dimensional descriptors for each key point, ensuring robustness to scale and rotation changes.
- **ORB:** Uses binary 32-dimensional descriptors for computational efficiency.

### 4.2 Feature Vector Formation

- Visual Bag of Words (vBoW) representation was constructed using KMeans clustering.
- TF-IDF was applied to generate fixed-length feature vectors, corresponding to the number of clusters (set as a hyperparameter).
- FAISS (Facebook AI Similarity Search) was employed for GPU-accelerated clustering.

### 4.3 Classification

Logistic Regression was used to classify the feature vectors. Classifiers were trained on the training set and evaluated on the validation set.

### 4.4 Robustness Testing

The test set was augmented with:

- Gaussian Noise,

- Random Rotation (up to  $\pm 30^\circ$ ),
- Scaling (0.5x to 1.5x),
- Illumination Changes (random contrast and brightness adjustments).



Figure 1: Augmented Datasets

## 5. Metrics for Comparison

The following metrics were used:

- **Accuracy:** Percentage of correctly classified test cases.
- **Efficiency:** Execution time for feature extraction, clustering, and feature vector formation.
- **Robustness:** Performance on augmented test sets.
- **Number of Key Points:** Average number of key points detected per image.

## 6. Results and Analysis

### 6.1 Efficiency Comparison

Metric	SIFT	ORB
Number of Processed Images	8144	8144
Average Key Points per Image	90.99	175.94
Average Feature Extraction Time (s/image)	0.01072	0.00135
Clustering Time (1000 Clusters) (s)	2.54	0.95
Feature Vector Formation Time (s)	2.79	5.64

Table 1: Efficiency comparison of SIFT and ORB.

The efficiency results highlight a trade-off between speed and descriptor quality for SIFT and ORB. ORB processes features faster, with an average extraction time of **0.00135 seconds per**

**image** compared to SIFT’s **0.01072 seconds**, and achieves faster clustering (**0.95 seconds versus 2.54 seconds**). This efficiency stems from ORB’s binary descriptors and higher average key point detection (**175.94 vs. 90.99**). However, the feature vector formation time for ORB (**5.64 seconds**) is longer than SIFT’s (**2.79 seconds**) due to its higher key point count.

## 6.2 Clustering and Accuracy Analysis

Algorithm	Number of Clusters	Training Accuracy (%)	Validation Accuracy (%)
SIFT	100	19.00	11.00
SIFT	1300	60.00	21.00
SIFT	4300	84.00	24.00
ORB	100	10.00	6.00
ORB	1300	30.00	9.00
ORB	4300	47.00	9.00

Table 2: Accuracy comparison for different cluster sizes.

The results indicate evidence of *overfitting* as the number of clusters increases, particularly for **SIFT**. While SIFT’s training accuracy improves significantly from **19.00%** at 100 clusters to **84.00%** at 4300 clusters, the corresponding validation accuracy increases more modestly, from **11.00%** to **24.00%**. This widening gap suggests that the model may be memorizing training data rather than generalizing effectively to unseen validation data. For **ORB**, overfitting is less pronounced, but its validation accuracy stagnates at **9.00%**, even as training accuracy rises to **47.00%** at 4300 clusters. This indicates that ORB’s descriptors are not robust enough to capitalize on the additional clusters, resulting in limited generalization capacity. These results emphasize the need to balance cluster complexity to avoid overfitting while maintaining generalization.

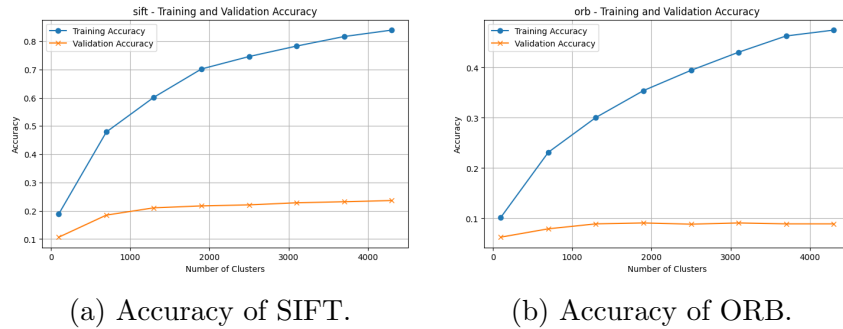


Figure 2: Comparison of SIFT and ORB for training and validation accuracy.

## 6.3 Robustness Evaluation

Algorithm	Test Set	Original	Gaussian Noise	Rotation	Scale	Illumination
SIFT	Accuracy (%)	21.00	10.00	18.00	16.00	17.00
ORB	Accuracy (%)	8.00	8.00	9.00	6.00	5.00

Table 3: Robustness comparison of SIFT and ORB.

The robustness results highlight that SIFT outperforms ORB across all test conditions. On the original test set, SIFT achieves significantly higher accuracy (**21.00%**) compared to ORB (**8.00%**). Under transformations such as Gaussian noise, rotation, scaling, and illumination changes, SIFT demonstrates superior resilience, achieving accuracies of **10.00%**, **18.00%**, **16.00%**, and **17.00%**,

respectively. In contrast, ORB’s performance remains consistently low, with accuracies peaking at **9.00%** for rotation and dropping as low as **5.00%** for illumination changes. These results underline SIFT’s robustness and adaptability to real-world variability, whereas ORB is limited to controlled environments due to its less discriminative binary descriptors.

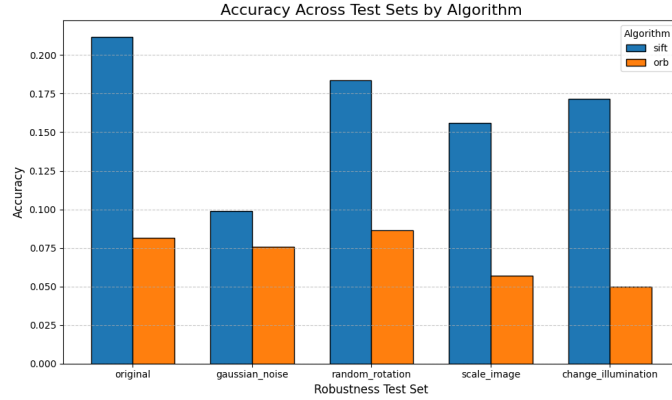


Figure 3: Robustness comparison of SIFT and ORB across transformations.

## 7. Conclusion

This study highlights the strengths and limitations of SIFT and ORB for handwritten Arabic text identification:

- **SIFT:** Superior robustness, accuracy, and generalization make it suitable for high-precision tasks. Its ability to handle transformations such as rotation and illumination changes demonstrates its effectiveness for challenging scenarios.
- **ORB:** Faster processing and computational efficiency make it ideal for applications prioritizing speed. However, its limited descriptor quality reduces its effectiveness for complex or noisy data.

Future work could explore hybrid approaches that combine SIFT’s robustness with ORB’s efficiency. Additionally, integrating preprocessing techniques to handle noisy or distorted images could further improve overall performance.

## 8. References

1. Lowe, D. G. (2004). "Distinctive Image Features from Scale-Invariant Keypoints." *International Journal of Computer Vision*.
2. Rublee, E., et al. (2011). "ORB: An efficient alternative to SIFT or SURF." *Proceedings of ICCV*.
3. AHAWP Dataset. <https://data.mendeley.com/datasets/2h76672znt/1/files/9031138a-b812-43>
4. Johnson, J., et al. (2019). "Billion-scale similarity search with FAISS." *IEEE Transactions on Big Data*.