Machine Learning-Based Approach for Person Re-Identification

1st Jay Rupareliya
AU2240238

2nd Jayraj Derasari
AU2240108

3rd Raj Sudani
AU2240245

4th Aarchi Kasundra
AU2240089

5th Milan Godhaviya
AU2140078

***Abstract*—The research develops a time-efficient Re-ID technique that uses manually constructed wavelet transform features for person identification. The approach uses Discrete Wavelet Transform (DWT) to obtain spatial-frequency features from images of persons before reducing dimensions through Principal Component Analysis (PCA). The machine learning models XRT and DRF together with GBM and XGBoost receive training using reduced features extracted with PCA. Testing outcomes indicate the system delivers promising results on accuracy performance through XGBoost as its primary model outcompetes other approaches in both Rank-1 and Rank-5 accuracy standards.**

***Index Terms*—Person re-identification, Fourier transform, Wavelet transform, Principal Component Analysis, Feature extraction, Rank-1 accuracy, XGBoost, Ensemble Learning, Feature Extraction, CMC Curves.**

I. Introduction

Person Re-Identification (Re-ID) functions as an essential capability within intelligent video surveillance systems by identifying people between separate camera viewpoints which might operate at different time periods. The identification process proves to be difficult because multiple factors including lighting conditions and camera viewpoints together with body orientation and occlusion levels and background interference and variations in clothing cause numerous difficulties. Successful and reliable Re-ID systems creation stands as an essential priority for public safety and security together with smart city applications and public safety. Deep learning solutions have proven effective for recent state-of-the-art achievements yet they present important obstacles for implementation. These techniques need vast sets of labeled data together with substantial computational power to conduct training operations and operational phases. Embedded systems and mobile platforms together with developing regions' surveillance operations lack sufficient resources so the requirement exists for lightweight efficient alternatives. This research develops a person Re-ID solution which combines handcrafted features and ensemble models as part of a machine learning approach. The feature extraction process begins with person image processing through Fast Fourier Transform (FFT) and Discrete Wavelet Transform (DWT) signal techniques. The global frequency structure of images is best obtained through FFT analysis yet Wavelet Transform excels at maintaining spatial details at multiple scales of resolution. Thankyou to frequency-domain analysis we obtain features which we enrich through calculation of each transformed component's statistical measurements including mean variance and energy quantity. After extracting features the system applies PCA for dimensionality reduction to transform high-dimensional feature vectors into a minimal yet informative space which represents maximum variance. The process decreases both computation intensity and improves generalized performance by removing feature noise and redundancy. The ensemble of robust machine learning models receives reduced feature vectors as input for training purposes using XRT in combination with DRF and GBM and XGBoost. The models show superb capabilities in structured tabular data classification tasks because of their specific design for managed data types. The ensemble learning structure enables these models to extract discriminative patterns from handcrafted features regardless of needing end-to-end deep networks because of their error correction capabilities and variance reduction and regularization strengths. The proposed pipeline's performance gets evaluated through standard Re-ID metrics where we analyze both CMC curves together with evaluations of Rank-1 and Rank-5 accuracy scores. The findings indicate this compact implementation matches prevalent performance outcomes thus representing an acceptable substitute for deep learning procedures within surveillance settings with limited resources.

Fig. 1. Sample images from the dataset showing person re-identification scenario.

II. Methodology

The analysis of person re-identification uses handcrafted feature extraction methods as a starting point which then trains ensemble machine learning models as its conclusion.

Person	Image	Overall_Mean	Overall_Variance	Overall_Energy
0001	0001C1T0001F001.jpg	0.209985697	0.028016563	1676.317169
0001	0001C1T0001F002.jpg	0.210809752	0.025794611	1667.455193
0001	0001C1T0001F003.jpg	0.209182533	0.025055656	1640.960984
0001	0001C1T0001F004.jpg	0.213512078	0.024106816	1693.167281
0001	0001C1T0001F005.jpg	0.208908	0.02776656	1657.7942

Sample of Wavelet Transform

A. *Dataset*

Research tests are proposed for method through evaluation of data obtained specifically for person re-identification tasks. The images within this dataset display all three conditions which provide optimal conditions for testing re-identification methods. The acquired dataset excels at person recognition between different non-overlapping cameras particularly because it aids security together with law enforcement and retail analytics operations.

A total of more than **one million** images are present along with approximately **1250** unique identities. Identical identities contain different numbers of images that show variations in body positions and apparel types as well as environmental contexts. All collected images represent how cameras would capture subjects in actual monitoring conditions. This dataset contains images which have low resolution along with compression format while measuring **128 pixels by 256 pixels**. The images contain multiple light condition types that consist of daytime and nighttime and low-light scenarios and varying brightness levels. The images are in **JPG** format.

B. *Feature Extraction using Fourier Transform*

Each image underwent simulation using Fast Fourier Transform (FFT) which transformed its spatial data to frequency domain. The Fast Fourier Transform identifies overall patterns of frequencies to extract the structural aspects of edges and contours. The transformed frequency domain enabled calculation of statistical indicators for mean and variance and energy to effectively represent each image content.

C. *Feature Extraction using Wavelet Transform*

Multi-resolution analysis used Discrete Wavelet Transform (DWT) as its processing method. The spatial characteristics together with frequency information captured by DWT set it apart from FFT thus making it effective for detail preservation. The resolution analysis of each image generated four sub-bands comprising approximation (cA) and horizontal (cH) and vertical (cV) as well as diagonal (cD). A complete and localized feature representation emerged from these extracted statistical features of every sub-band.

D. *Dimensionality Reduction using PCA*

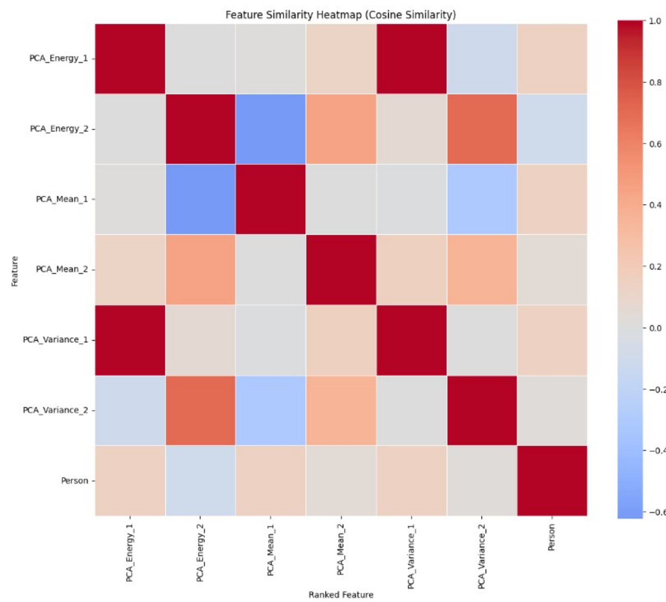
PCA was employed to decrease computational complexity and eliminate redundant features in the extracted features. PCA conducted a dimensionality reduction that shifted the feature vectors into a lower space which maintained the most vital variance. The reduction of dimensions through this step improved model training performance and protected from overfitting issues caused by noisy or irrelevant features.

Person	Image	FFT Mean	FFT Variance	FFT Energy
0001	0001C1T01	1.274671	0.936214401	83918.83
0001	0001C1T01	1.25977	0.942424488	82884.89
0001	0001C1T01	1.257356	0.959451544	83243.65
0001	0001C1T01	1.323087	0.956193049	88694.87
0001	0001C1T01	1.25771	0.925682759	82166.29

Sample of FFT Feature

Person	Image	Overall_Mean	Overall_Variance	Overall_Energy
0001	0001C1T0001F001.jpg	0.211238916	0.027033496	1804.448659
0001	0001C1T0001F002.jpg	0.21123233	0.02479392	1783.491495
0001	0001C1T0001F003.jpg	0.210056275	0.024411685	1764.050093
0001	0001C1T0001F004.jpg	0.214101379	0.023292192	1813.776917
0001	0001C1T0001F005.jpg	0.209377435	0.026814185	1774.308347

Sample of Orthogonal Wavelet Feature



Feature Similarity Heatmap

E. *Model Training*

The principal component analysis reduced the extracted feature vector dimensionality which generated a compact set for training and evaluation of ensemble machine learning models. Structured data benefits most from ensemble models because these models enhance predictive performance outcomes by leveraging multiple base learners.

We choose Extremely Randomized Trees (XRT) Distribute Random Forest (DRF) Gradient Boosting Machine (GBM) and XGBoost as our four main ensemble algorithms for model training.

Ensemble Learning Models:

Extremely Randomized Trees (XRT): Induction of XRT stems from standard random forest models. XRT adds more randomness through selecting both random subsets of features and performing random threshold selection rather than finding optimal splits. The approach generates diverse tree collections which decreases both modeling errors and data overfitting. The training conducted by XRT becomes more efficient due to its capability of skipping time-intensive split optimization procedures.

Distributed Random Forest (DRF): As a classical bagging-based algorithm DRF creates multiple decision trees from randomly sampled subsets of input data. Decision trees are trained separately so their predictions unite through majority voting in classification to enhance generalization results. DRF excels at parallel processing and demonstrates reliable performance on large datasets because of its scalability together with its resistant nature towards noise.

Gradient Boosting Machine (GBM): GBM trains its sequential ensemble through a method that deploys new trees to tackle the remaining errors which result from previous trees. The each iteration of gradient-based optimization concentrates on hard-to-classify examples for enhancing model accuracy incrementally. GBM produces exceptional predictive results when professionals apply proper adjustment strategies.

XGBoost (Extreme Gradient Boosting): The advanced XGBoost system represents an improved version of gradient boosting which includes various new algorithmic features. Multipurpose features such as regularization (L1 and L2) together with shrinkage (learning rate) and column subsampling alongside tree pruning serve to stop overfitting while ensuring better model generality. XGBoost operates with exceptional speed and parallel processing ability which outperforms conventional boosting protocols regarding performance and scalability.

Evaluation Metrics: The evaluation methodology included established metrics that allowed assessment of model effectiveness within person re-identification systems. Authentication of the targeted identity became possible through the analysis of Cumulative Matching Characteristic (CMC) curves which evaluated system retrieval performance. The assessment included Rank-1 crucial statistics determining first place and Rank-5 accuracy statistics measuring correct identification within the first five results. These measurement methods provide both an extensive view and a simple understanding about the model's precision along with its suitability for real-world security operations.

III. Discussion

Wavelet-based features present an ideal representation solution for person re-identification because of their powerful yet comprehensible characteristics. The localized pattern detection of DWT at various resolution levels makes this method better than global descriptors such as FFT for distinguishing people in different situations. The PCA step maintains essential variations in the data by ensuring computational efficiency of the system. Ensemble classifiers obtain exceptional results on structured data by employing XGBoost because this algorithm reduces overfitting with its integrated regularization functions. The proposed method stands out because of its ease of implementation and operational effectiveness thus becoming suitable for practical deployment in locations with inadequate hardware or limited data.

IV. Results

Distributes Random Forest (DRF): Our person re identification experiments showed that the DRF model had the best results delivered at an accuracy of 67.03% respectively. The ensemble learning approach used in it, whereby we use a number of decision trees based on random subsets of features, was able to effectively capture the complex patterns in the wavelet transformed data. For this task, DRF demonstrated very good robustness against over fitting and good handling of high dimensional features, thereby making it particularly suitable for this task. Performance measurement shows that the model could generalize better across several different identities, which is important for person re-identification. The randomness and structure of DRF were in balance and they assisted in separating classes effectively in such a way that the class specific accuracy of the algorithm was better than any other assumptions.

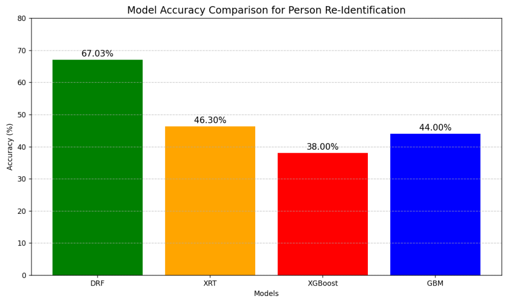
TABLE I
Model Accuracies for Person Re-Identification

Model	**Accuracy (%)**
Discrete Random Forest (DRF)	67.03
Extreme Randomized Trees (XRT)	46.30
Gradient Boosting Machine (GBM)	44.00
XGBoost	38.00

Extremely Randomized Trees (XRT): This models accuracy was **46.30%**, making it second of the models tested. XRT achieves variance reduction and produces diverse trees as it introduces greater randomness in the tree construction selection of cut points. While we are able to take advantage of noisy or complex data with such an approach, it did not fully characterize the subtleties of our wavelet transformed features. Although XRT did better than the boosting models, it did not perform as well as DRF on accuracy or consistency. The results indicate that XRT offers good speed and randomness but more tuning or other feature engineering may be necessary to improve performance in this setting.

Gradient Boosting Machine (GBM): In our experiments, we find that the GBM model achieved an accuracy of **44%**, which is not bad as it showed moderate performance. GBM can be regarded as an iterative boosting method: it sequentially builds the models to improve the prediction errors of the previous models, and it is suitable for structured data. Nevertheless, considering the feature space created by the wavelet transform was high dimensional and involves large non-linear transformation followed by, our model could not learn effectively. Overfitting of some patterns could have occurred due to the boosting process and it may have decreased out of fit capability. Though GBM did a little better than XGBoost, it clearly didn't come close to DRF or XRT. The results show that GBM alone is unlikely to solve this complex task.

XGBoost (Extreme Gradient Boosting): Although XGBoost reputation of being efficient and accurate in so many machine learning problems, our study recorded the lowest accuracy at **38%**. XGBoost was designed with various advanced



Accuracy comparison

regularization techniques and is also parallelized, making it usually able to suppress overfitting, but failed to find any meaningful patterns in our wavelet transform data. It is likely that the high dimensional features contain some noise and some complexity such that it was not able to differentiate between the individual identities. Performance of the model is shown compared to other models and it is shown that even strong boosting techniques suffers if not selective and tuned in feature selection for this intricate kind of data representation like person re-identification.

As all of the four models were compared with the wavelet transformed dataset, the highest accuracy for this dataset was found in the case of the DRF model in a value of **67.03%,** and XRT gave a moderate performance of **46.30%**. In **44%** and **38%** accuracy, the GBM and XGBoost lagged behind. In our experiments, we show that tree based ensemble methods are far worse than the boosting algorithms, reinforcing a large gap between tree based ensemble methods and boosting algorithms. Both XGBoost and GBM are clearly outperformed by the consistent underperformance, which also suggests their limitation in this context, and the fact that their most notable lead is due to DRF's superiority over GBM's over XGBoost's performance reinforces the superiority of DRF in the case of complex feature sets in person re-identification without the need of such extensive hyperparameter tuning.

V. Conclusion

This system offers a light-weight and effective person re-identification solution using wavelet-based feature extraction with subsequent PCA dimensionality reduction and ensemble machine learning models. The approach avoids deep learning data requirements by using engineered features for structured modeling rather than computational expense. Distributed Random Forest (DRF) was the best-performing model with an accuracy rate of 67.03% which proves its generality and evidence against overfitting errors. The Extremely Randomized Trees model performed at 46.30% due to its capability to build randomized trees but failed to identify small identity features. GBM (44%) and XGBoost (38%) performance declined when working with high-dimensional wavelet features because these boosting methods require additional feature optimization for optimal performance in such tasks. The project proves that DRF is effective for wavelet-transformed data in person re-identification tasks because it provides accurate performance at high speed without complexity. These evidence prove that well-chosen handcrafted features provide matching results with model selection in tasks requiring limited computational resources.

References

[1] D. Gray, S. Brennan, and H. Tao, "Evaluating Appearance Models for Recognition, Reacquisition, and Tracking," in *Proc. IEEE Int. Workshop on Performance Evaluation for Tracking and Surveillance*, 2007.

[2] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, and Q. Tian, "Scalable Person Re-Identification: A Benchmark," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2015.

[3] W. Li, R. Zhao, T. Xiao, and X. Wang, "DeepReID: Deep Filter Pairing Neural Network for Person Re-Identification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2014.

[4] Y. Sun, L. Zheng, Y. Yang, Q. Tian, and S. Wang, "Beyond Part Models: Person Retrieval with Refined Part Pooling (and a Strong Convolutional Baseline)," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, pp. 480–496, 2018.

[5] S. He, H. Luo, M. Tian, G. Wang, and H. Li, "Multi-Granular Learning for Person Re-Identification," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, pp. 116–132, 2020.

[6] J. Hou, B. Ma, H. Chang, X. Gu, S. Shan, and X. Chen, "Unlabeled Person Re-Identification," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, pp. 1558–1567, 2021.

[7] P. Zhang, X. Zhang, W. Liu, H. Liu, X. Jin, and L. Ma, "Beyond Spatial Transform: Multi-Head Transformer for Robust Person Re-Identification," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, pp. 12567–12576, 2022.