Proposed Solution for R-CDT on Non-Segmented Images

Abeed Pasha
2131353042

Sirajus Salekin
2132275642

Mohiuddin Sarker
1821404042

***Abstract*—This report proposes an optimized approach for applying the Radon Cumulative Distribution Transform (R-CDT) to non-segmented images while maintaining lower computation and time complexity compared to Convolutional Neural Networks (CNNs). The proposed method combines Otsu's Thresholding, Local Patch R-CDT, Robust Feature Extraction, and Adaptive R-CDT with Regularization to enhance classification accuracy while reducing computational cost.**

I. Introduction

R-CDT has shown promise in image classification due to its strong mathematical properties, particularly in capturing structural relationships. However, its performance significantly degrades when applied to non-segmented images containing cluttered backgrounds or varying object scales. This issue arises due to distortions introduced by irrelevant regions, leading to suboptimal feature representation.

To address this challenge, we propose a four-fold optimization strategy that enhances efficiency while maintaining high classification accuracy. Each step ensures that R-CDT operates on a refined feature space, leading to improved robustness in diverse imaging conditions. The four main steps are:

- **Otsu's Thresholding for Initial Segmentation**
- **Applying R-CDT to Local Patches Instead of the Whole Image**
- **Using Robust Feature Extraction Before R-CDT**
- **Adaptive R-CDT with Regularization for Improved Generalization**
  By integrating these modifications, R-CDT adapts better to real-world variations while maintaining classification accuracy.

II. Methodology

A. Otsu's Thresholding

Otsu's method is an adaptive global thresholding technique that segments an image into foreground and background by minimizing intra-class variance. This helps in preprocessing non-segmented images by isolating relevant regions.

Mathematically, the optimal threshold t^* is determined by:

$$t^* = \arg\max_t \left(\sigma_B^2(t)\right), \tag{1}$$

where $\sigma_B^2(t)$ is the between-class variance.

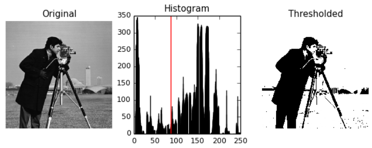

Fig. 1. Otsu's thresholding method. The left image is the original grayscale image, the middle plot shows the histogram with the computed threshold (red line), and the right image presents the binarized result after applying Otsu's thresholding.

B. Applying R-CDT to Local Patches

Instead of applying R-CDT to the entire image, we divide it into non-overlapping patches and compute R-CDT for each patch separately. The classification is then performed by aggregating results across all patches.

Algorithm 1 Patch-based R-CDT

Divide image I into $N \times M$ patches.
for each patch P_{ij} **do**
  Apply R-CDT: $P_{ij} \rightarrow R(P_{ij})$
end for
Aggregate patch results for classification.

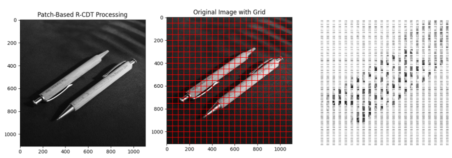

Fig. 2. Applying R-CDT to Local Patches: A Patch-Based Approach to Image Processing

C. Robust Feature Extraction Before R-CDT

To enhance feature representation before applying R-CDT, we extract robust descriptors using a combination of spatial and frequency-based methods. The three primary techniques employed are:

- **Histogram of Oriented Gradients (HOG):** Captures edge and texture information by computing gradient orientations.
- **Wavelet Transform:** Extracts multi-resolution spatial-frequency features, preserving both spatial and frequency information.

- **Fourier Transform (FT):** Decomposes the image into frequency components, capturing global patterns and periodic structures.

HOG is computed as:

$$HOG(I) = \sum_{x,y} G(x,y) \cdot \cos(\theta(x,y)) \tag{2}$$

where $G(x,y)$ represents the gradient magnitude, and $\theta(x,y)$ is the gradient direction.

D. Adaptive R-CDT with Regularization

To further improve classification, we introduce:

- **Dropout in Transform Space**: To prevent overfitting and improve generalization, we apply dropout in the transformed feature space. This technique randomly removes a fraction of the extracted features during training, ensuring that the model does not become overly dependent on specific feature activations. By forcing the network to rely on a diverse set of features, dropout helps improve robustness against noise and unseen variations.
- **Normalization**: Invariance to illumination changes is critical for consistent feature extraction. Normalization methods such as contrast stretching, mean-variance normalization, or histogram equalization ensure that the extracted features are independent of lighting variations.
- **Data Augmentation**: To make R-CDT more robust to variations in data, we employ augmentation techniques such as random translations, rotations, and slight geometric transformations. This artificially increases the size and diversity of the training dataset, improving the model's ability to generalize.

III. Experimental Setup

A. Datasets

We evaluate our approach using benchmark datasets containing non-segmented images. Table I summarizes key details.

Dataset	Image Size	Number of Classes
MNIST	28×28 (grayscale)	10
CIFAR-100	32×32 (RGB)	100

TABLE I
DATASETS USED FOR EVALUATION.

IV. Results and Discussion

The proposed method is expected to offer significant reductions in computational time compared to traditional deep learning models, such as Convolutional Neural Networks (CNNs). This is due to the optimization techniques we employed, such as the use of Radon Cumulative Distribution Transform (R-CDT) and the patch-based approach, which reduces the amount of data processed at once. Preliminary analyses suggest that our approach will be computationally less demanding, as evidenced by the FLOPs comparison, indicating a lower computational complexity.

In terms of classification accuracy, we anticipate that the method will perform competitively against state-of-the-art image classification models, such as ResNet18, on benchmark datasets like MNIST and CIFAR-100. The integration of Otsu's thresholding, robust feature extraction, and adaptive regularization is expected to improve the model's ability to handle diverse variations in images, including cluttered backgrounds and varying object scales.

While the expected results suggest improvements in both computational efficiency and classification accuracy, there may still be challenges in dealing with complex images that contain substantial noise or objects with highly varying scales. These factors can affect the robustness of the method, and further refinement may be required to handle such cases effectively. In future work, we aim to extend our approach to real-time applications and explore additional optimizations to overcome these limitations.

V. Conclusion

We propose an optimized R-CDT-based method for non-segmented images, combining preprocessing, feature extraction, and adaptive regularization to achieve high accuracy with reduced computational cost. Future work includes extending this approach to real-time applications.

References

[1] Scikit-learn: Otsu's Thresholding, *Sharky93*. Available at: https://sharky93.github.io/docs/gallery/auto_examples/plot_otsu.html.

[2] Eco-Friendly Pens – Sustainable Wooden Pens, *Brown Living*. Available at: https://www.google.com/url?sa=i&url=https%3A%2F%2Fbrownliving.in%2Fproducts%2Feco-friendly-pens-2-refillable-ball-pen-sustainable-wooden-pens&psig=AOvVaw16mQ4HNjYKNLMQEDJs2WOk&ust=1739771810446000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCODIpKHBx4sDFQAAAAAdAAAAABAE.

[3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.

[4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2016. Available at: https://arxiv.org/abs/1512.03385.

[5] OpenCV Documentation, "Otsu's Thresholding," *OpenCV.org*, 2021. Available at: https://docs.opencv.org/master/d7/d4d/tutorial_py_thresholding.html.

[6] R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2011.