Proposed Solution - Non-Segmented Image Classification Using Radon Cumulative Distribution Transform

Abeed Pasha Sirajus Salekin Mohiuddin Sarker
2131353042 2132275642 1821404042

***Abstract*—In this study, we propose a novel and efficient image classification method that integrates the Radon Cumulative Distribution Transform with Nearest Subspace Classifier (RCDT-NS). Our approach extends RCDT-NS to non-segmented images by first converting them into segmented image data using Otsu's method for automatic image thresholding. We will evaluate the method's performance using the MNIST, Fashion MNIST, CIFAR-10, CIFAR-100, and Bangla Digits datasets. Our experiments will assess the effectiveness of R-CDT in representing image data for classification, aiming to achieve competitive accuracy while substantially reducing computational complexity. Additionally, we will provide a detailed comparison with convolutional neural networks (CNNs), highlighting the potential benefits of R-CDT in terms of interpretability, efficiency, and adaptability to non-segmented image classification.**

I. Introduction

Image classification of non-segmented images is challenging due to the complexity of identifying objects or patterns without predefined regions. Segmentation is a crucial preprocessing step that simplifies classification by dividing an image into meaningful regions. However, non-segmented images require more complex feature extraction, and imperfect segmentation—such as overlapping or under-segmented regions—can lower classification accuracy by introducing inconsistencies in the data.

Machine learning (ML) techniques help overcome these challenges by providing robust methods to handle segmented data variations. One promising approach is the Radon Cumulative Distribution Transform (RCDT) combined with a Nearest Subspace (NS) classifier. RCDT extracts essential geometric and spatial features from segmented images, while the Nearest Subspace classifier efficiently categorizes them, even in the presence of noise or segmentation imperfections.

In this study, we propose a novel classification method that integrates RCDT-NS for both segmented and non-segmented images. For non-segmented images, we apply Otsu's method, an automatic thresholding technique that converts them into segmented data before feature extraction with RCDT-NS. Our model incorporates a segmentation algorithm and the RCDT-NS classifier in a unified system, where non-segmented images pass through a segmentation layer before classification.

By utilizing the RCDT-NS approach, our method offers a more accurate and computationally efficient solution for classifying segmented images. It effectively addresses segmentation inaccuracies while enhancing interpretability and reducing computational complexity compared to deep learning methods like convolutional neural networks (CNNs). This makes it particularly well-suited for applications requiring high precision and efficiency.

II. Proposed Methodology

A. Pipeline Overview

The classification pipeline is composed of three main stages:

1) **Preprocessing**: This step involves normalizing and resizing the images to standardize their format.
2) **Segmentation**: Non-segmented images are run through the Otsu's thresholding algorithm and converted to segmented image data.
3) **R-CDT Transformation**: The Radon Cumulative Distribution Transform (RCDT) is applied to improve the separability of the image features.
4) **Classification**: The Nearest Subspace Classifier (NS) is employed to efficiently predict the image class.

B. Datasets and Preprocessing

1) **Datasets**: To comprehensively evaluate our methodology, we used a variety of datasets, including:
   - **MNIST**:
     – This dataset contains 28 × 28 pixel grayscale images of handwritten digits (0–9).
     – It is a widely used benchmark for image classification, comprising 60,000 training images and 10,000 testing images.
   - **Fashion MNIST**:
     – Fashion MNIST is a collection of 28 × 28 grayscale images representing various Zalando product categories like T-shirts, pants, and shoes.
     – This dataset presents a more complex challenge than MNIST, featuring more intricate patterns for testing advanced methods.
   - **CIFAR-10**:
     – The CIFAR-10 dataset consists of 60000 32x32 colour images in 10 classes, with 6000 images per class. There are 50000 training images and 10000 test images.
   - **CIFAR-100**:
     – This dataset is just like the CIFAR-10, except it has 100 classes containing 600 images each. There are 500 training images and 100 testing images per class. The 100 classes in the CIFAR-100 are grouped into

20 superclasses. Each image comes with a "fine" label (the class to which it belongs) and a "coarse" label (the superclass to which it belongs).

• **Bangla Digits**:
– This dataset consists of handwritten Bangla digits with relatively lower resolution images.
– The distinct script and unique characteristics of Bangla digits make it an ideal choice for evaluating the method's ability to generalize across different languages and writing systems.

2) **Preprocessing**: Several preprocessing steps were undertaken to ensure consistency and compatibility across the datasets:
   a) **Conversion to .mat Files**:
   – The original image files (from MNIST and Fashion MNIST) and CSV files (from Bangla Digits) were converted into MATLAB-compatible .mat files.
   – This format change allowed for smoother integration into our processing pipeline.
   b) **Key Extraction**:
   – Keys corresponding to labels and metadata were identified and extracted for structured access during both training and testing.
   c) **Size and Dimension Adjustment**:
   – Image sizes were standardized (28 × 28 pixels) as needed.
   – For datasets with multiple channels or varying resolutions, adjustments were made to ensure all images were in grayscale and consistent with our methodology.
   d) **Thresholding**: – Images that are not already segmented are passed through the thresholding algorithm and converted to segmented data using the OTSU method.

C. R-CDT Transformation

The R-CDT transforms the image data into a new space where geometric separability is significantly improved, enhancing the ability to distinguish between different classes.

III. Challenges and Solutions

A. Dataset Compatibility Issues

A key challenge faced was ensuring the datasets were compatible with the RCDT-based model. This involved thoroughly checking the data formats and preprocessing scripts. The issue was addressed by standardizing the input dimensions and validating the outputs at each preprocessing stage.

B. Preprocessing Errors

Initial attempts at preprocessing the Bangla dataset revealed inconsistencies in image scaling and noise levels. These errors were addressed by implementing automated validation checks and refining the preprocessing pipeline to handle diverse image qualities.

C. Training Limitations

The initial training runs revealed that the model's performance could be further improved through hyperparameter optimization and data augmentation. These findings will inform future work to enhance the accuracy and robustness of the model.

IV. Future Work

Looking ahead, the focus will be on:

- **Enhancing Feature Selection for Complex Data Structures:** Refining the R-CDT transformation to emphasize the most relevant features while filtering out irrelevant or redundant information in non-segmented images.
- **Exploring Applications in Medical Imaging and Beyond:** Applying the method to fields such as medical diagnostics, where images are often non-segmented (e.g., full-body X-rays or unsegmented MRI scans). Testing the method on real-world datasets with minimal preprocessing, thereby reducing reliance on manual or semi-automated segmentation techniques.
- **Expanding Dataset Scope:** Additional datasets will be incorporated to evaluate the generalizability of the model in various tasks.
- **Performance Benchmarking:** Comparing the performance of the R-CDT + NSC pipeline with segmentation-free approaches in deep learning, such as fully convolutional networks (FCNs) or vision transformers, to highlight its advantages and pinpoint areas for further improvement.

V. Conclusion

RCDT-NS shows significant promise as a segmented image classifier with both performance and effectiveness as well as speed of computation. An all-encompassing system that can prepare non-segmented image for the model as well can prove to be superior to other image classification models allowing for versatility backed up by effectiveness. Further development and optimization may prove to make this approach a more feasible approach compared to traditional machine learning approaches.

References

[1] M. Shifat-E-Rabbi et al., "Radon Cumulative Distribution Transform Subspace Modeling for Image Classification," Journal of mathematical imaging and vision, vol. 63, no. 9, pp. 1185–1203, Aug. 2021, doi: https://doi.org/10.1007/s10851-021-01052-0.

[2] M. Park, E. K. Ryu, S. Kolouri, and G. K. Rohde, "The Radon Cumulative Distribution Transform and Its Application to Image Classification," arXiv preprint, arXiv:2004.03669, 2020. Available: https://arxiv.org/abs/2004.03669.

[3] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[4] RohdeLab, "RohdeLab GitHub Repository," Available: https://github.com/rohdelab. [Accessed: 26-Nov-2024].

[5] [1]A. Krizhevsky, "Learning Multiple Layers of Features from Tiny Images," May 08, 2012. https://www.researchgate.net/publication/265748773$_{Learning_Multiple_Layers_of_Fea}$