A Report

on

PolypInsight: A Unique Unsupervised Deep Learning Framework for Cancer

carried out as part of the course Minor Project CS3270

Submitted by

Sainyam Acharya(229301524)

Akash Raj(229301533)

VI-CSE

in partial fulfilment for the award of the degree

of

BACHELOR OF TECHNOLOGY

In

Computer Science & Engineering

Department of Computer Science & Engineering,
School of Computer Science and Engineering,
Manipal University Jaipur,
April 2025

A Report

on

PolypInsight: A Unique Unsupervised Deep Learning Framework for Cancer

carried out as part of the course CSE CS3270 Submitted by

Sainyam Acharya(229301524)

Akash Raj(229301533)

VI-CSE

in partial fulfilment for the award of the degree

of

BACHELOR OF TECHNOLOGY

In

Computer Science & Engineering

Under the Guidance of :

Guide Name : …………………………………………….

Guide Signature(with date) : ……………………………………

Acknowledgement

This project would not have been completed without the help, support, comments, advice, cooperation and coordination of various people. However, it is impossible to thank everyone individually; I am hereby making a humble effort to thank some of them.

I acknowledge and express my deepest sense of gratitude to my internal supervisor ………………………………… for his/her constant support, guidance, and continuous engagement. I highly appreciate his technical comments, suggestions, and criticism during the progress of this project "…………………………………".

I owe my profound gratitude to ***Dr. Neha Chaudhary*** , Head, Department of CSE, for her valuable guidance and for facilitating me during my work. I am also very grateful to all the faculty members and staff for their precious support and cooperation during the development of this project.

Finally, I extend my heartfelt appreciation to my classmates for their help and encouragement.

Registration No. **Student Name**

MANIPAL UNIVERSITY
JAIPUR
(University under Section 2(f) of the UGC Act)

Department of Computer Science and Engineering

School of Computer Science and Engineering

Date: _____________

CERTIFICATE

This is to certify that the project entitled "***(Project title)***" is a bonafide work carried out as ***Minor Project Midterm Assessment (Course Code: CS3270)*** in partial fulfillment for the award of the degree of Bachelor of Technology in Computer Science and Engineering, by ***(name of the student)*** bearing registration number(_**Reg no. of student** _), during the academic semester *VI of year 2024-2025.*

Place: Manipal University Jaipur, Jaipur

Name of the project guide: _________________________

Signature of the project guide: ______________________

CONTENT

1. Introduction

Colon cancer remains one of the most prevalent and life-threatening cancers globally, underscoring the importance of early and accurate detection for improved patient outcomes. Traditional diagnostic procedures such as colonoscopy and histopathological examination, while effective, are time-intensive, susceptible to human error, and reliant on expert interpretation. In response, recent advancements in artificial intelligence have opened new avenues for enhancing diagnostic workflows. This research introduces a novel and unique hybrid unsupervised deep learning framework for automated colon polyp classification, which eliminates the need for human-annotated training labels and offers a scalable and efficient diagnostic support tool.

The proposed pipeline integrates multiple state-of-the-art components to achieve this goal. Initially, YOLOv8 is employed to detect and localize polyp regions from raw colonoscopy video frames. These polyp crops are then processed using the MobileNetV2 architecture to extract deep convolutional features that capture important visual characteristics. To reduce noise and emphasize informative features, Variance Thresholding is applied, followed by dimensionality reduction techniques such as PCA and UMAP for projecting the high-dimensional feature vectors into lower-dimensional space while preserving their structural integrity. Two clustering algorithms, K-Means and Agglomerative Hierarchical Clustering, are then independently applied to these reduced features to group polyps into potential categories such as cancerous and non-cancerous.

To further enhance clustering robustness and accuracy, an ensemble clustering strategy is implemented by fusing the outputs of K-Means and Agglomerative methods, with a filtering step to disregard uncertain or ambiguous samples. This ensemble approach significantly improves clustering performance across key internal metrics like Silhouette Score, Davies-Bouldin Index, and Calinski-Harabasz Index. The framework’s ability to categorize polyps without relying on human labels makes it particularly valuable in pathology-scarce environments. It also offers potential for real-time clinical deployment, reducing the workload on medical experts while maintaining diagnostic reliability.

1.1 Motivation

- **Early Detection Saves Lives:** Colon cancer ranks among the deadliest forms of cancer when detected in advanced stages. Timely identification of precancerous polyps significantly increases treatment success rates. Our automated unsupervised detection and clustering framework offers a powerful solution by rapidly identifying high-risk polyps, facilitating faster clinical decision-making and improving patient outcomes**.**

- **Limitations of Traditional Diagnostic Methods:** Manual diagnostic techniques like colonoscopy and histopathological analysis are time-consuming, labor-intensive, and susceptible to inter-observer variability. These limitations often delay treatment. Our proposed AI-based pipeline automates polyp localization and classification without requiring labeled data, effectively reducing diagnostic delays and minimizing human error.

- **Breakthroughs in Unsupervised Deep Learning:** With advancements in object detection (YOLOv8), deep feature extraction (MobileNetV2), and unsupervised learning, we can now uncover latent patterns in colonoscopy images without ground truth annotations. Our ensemble clustering approach (K-Means + Agglomerative) coupled with UMAP-based dimensionality reduction provides a novel, robust methodology to identify cancerous versus non-cancerous polyps. This represents a unique and label-free direction for AI-assisted colon cancer screening, reducing the annotation burden while maintaining high diagnostic relevance.

1.2 Objectives

- **Early Detection Saves Lives:** Colon cancer ranks among the deadliest forms of cancer when detected in advanced stages. Timely identification of precancerous polyps significantly increases treatment success rates. Our automated unsupervised detection and clustering framework offers a powerful solution by rapidly identifying high-risk polyps, facilitating faster clinical decision-making and improving patient outcomes.

- **Limitations of Traditional Diagnostic Methods:** Manual diagnostic techniques like colonoscopy and histopathological analysis are time-consuming, labor-intensive, and susceptible to inter-observer variability. These limitations often delay treatment. Our proposed AI-based pipeline automates polyp localization and classification without requiring labeled data, effectively reducing diagnostic delays and minimizing human error.

- Breakthroughs in Unsupervised Deep Learning: With advancements in object detection (YOLOv8), deep feature extraction (MobileNetV2), and unsupervised learning, we can now uncover latent patterns in colonoscopy images without ground truth annotations. Our ensemble clustering approach (K-Means + Agglomerative) coupled with UMAP-based dimensionality reduction provides a novel, robust methodology to identify cancerous versus non-cancerous polyps. This represents a unique and label-free direction for AI-assisted colon cancer screening, reducing the annotation burden while maintaining high diagnostic relevance.

2. Literature Review

2.1 Review

- The detection and diagnosis of colon cancer have seen significant advancements in recent years, primarily due to the application of machine learning (ML) and deep learning (DL) techniques. Several studies have explored the potential of these technologies in improving diagnostic accuracy and streamlining the clinical workflow. Histopathological image analysis, endoscopy, and immunohistochemical techniques have become crucial components in early detection.

- Stahl et al. (2024) explored the use of convolutional neural networks (CNNs) for detecting cancerous tissues in colon cancer cases, achieving high accuracy levels. The study further emphasized the need for explainable AI (XAI) methods to make model predictions interpretable to healthcare professionals. Despite the promising results, they highlighted the limitation of insufficient annotated datasets, suggesting the need for larger, diverse datasets to ensure better generalization of models across populations.

- Kumar and Lee (2023) focused on enhancing the sensitivity and specificity of colorectal cancer (CRC) detection using various diagnostic modalities, including endoscopy and histopathological images. Their study highlighted how deep learning techniques significantly outperformed traditional diagnostic methods. They called for the creation of a standardized dataset to allow for a more accurate benchmark of ML models and to facilitate their deployment in clinical settings.

- Zhang et al. (2024) took a different approach by utilizing immunohistochemical (IHC) images combined with machine learning to improve CRC detection workflows. The study showed how key features extracted from IHC images could boost diagnostic accuracy compared to conventional methods. However, the authors pointed out the need for validation of their approach on external datasets and a deeper integration of this model into clinical practice.
- Patel and Singh (2023) compared several ML models for predicting CRC risk, finding that machine learning models outperformed traditional statistical methods. Their work stressed the importance of addressing the interpretability of these models, as clinicians must understand how predictions are made in order to trust their outputs. They also recommended validating these models on diverse demographic datasets to ensure their robustness across populations.

- Ahmed et al. (2023) explored hybrid models that combine deep learning with traditional machine learning classifiers for CRC detection. Their research found that such hybrid models significantly improved classification accuracy, particularly when using transfer learning techniques. The study called for future work to optimize computational efficiency and test the models on multi-center datasets to ensure their real-world applicability.

- Chen and Roy (2024) provided a comprehensive review of the challenges facing AI-based CRC detection systems, such as data scarcity, regulatory issues, and model interpretability. They emphasized the need for more expansive datasets, improved explainability, and better integration of these AI tools into existing clinical workflows.

2.2 Tabular Review

Paper Title	Author(s) & Year	Key Findings	Future Work
Machine Learning-based Lung and Colon Cancer Detection using Deep Feature Extraction and Ensemble Learning Detection[1]	Md. Alamin Talukder et al., 2022	Proposed a hybrid ensemble model integrating deep feature extraction and ensemble learning. Achieved high accuracy rates: 99.05% (lung cancer), 100% (colon cancer), and 99.30% (both). Evaluated on LC25000 dataset with various TL models (VGG16, VGG19, MobileNet, DenseNet169, DenseNet201) and ML classifiers (RF, SVM, LR, MLP, XGB, LGB)..	Future work includes exploring new lung and colon cancer datasets with improved preprocessing techniques to enhance model performance further.
Deep learning for colon cancer histopathological images analysis[2]	A. Ben Hamida et al., [2021]	Evaluated CNN models (ALEXNET, VGG, RESNET, DENSENET, INCEPTION) for patch-level colon cancer classification. Used transfer learning with ImageNet to enhance feature extraction. RESNET achieved up to 99.98% accuracy on public datasets. SEGNET outperformed UNET in pixel-wise segmentation, achieving 99.12% accuracy on NCT- CRC-HE-100K.	Future work includes optimizing SEGNET for a better balance between accuracy and computational cost, improving performance on sparsely annotated datasets, and reducing misclassification of tumor tissues with immune/necrotic regions.
Feature Extraction using Convolution Neural Networks (CNN) and Deep Learning[3]	Manjunath Jogin et al., [2018]	Explores CNN-based feature extraction for image classification. Achieved 85.97% accuracy using deep learning techniques. Applications include home automation, industrial automation, facial recognition, surveillance, UAV navigation, medical diagnostics, and 5G network management.	Future work includes extending CNN applications to video surveillance, security-related tasks, and feature extraction for classification in autonomous systems.

REAL-Colon: A dataset for developing real- world AI applications in colonoscopy[4]	Carlo Bif et al., [2024]	"Presents REAL-Colon: 2.7M frames from 60 full-resolution colonoscopies with 350k expert- annotated boxes, enabling AI- driven CADe/CADx research and early polyp detection."	Future work includes exploring augmentation techniques, improving object detection models, and refining training strategies to enhance early detection while reducing false positives in complex real-world conditions.
Y-Net: A deep Convolutional Neural Network for Polyp Detection[5]	Ahmed et al., 2023	Introduces Y-Net, a dual-encoder deep learning model with sum-skip- concatenation and encoder-based learning rate, achieving 7.3% F1- score and 13% recall improvement.	Future work includes exploring alternative pre-trained models (e.g., VGG16) and optimizing encoder-decoder architectures for better polyp detection with limited data.

2.3 Outcome of Literature Review

The reviewed studies showcase recent progress in AI-powered colorectal cancer detection, particularly leveraging unsupervised and ensemble learning techniques for polyp classification. Talukder et al. (2022) introduced a hybrid system integrating deep feature extraction with traditional machine learning classifiers, underscoring the effectiveness of ensemble models for cancer detection tasks. Ben Hamida et al. (2021) explored RESNET for feature extraction and reported high accuracy in colorectal image analysis, though the heavy computational footprint remains a challenge. Jogin et al. (2018) highlighted the versatility of CNNs in cross-domain image classification, validating their robustness for high-level feature extraction in medical imaging. Bif et al. (2024) presented the REAL-Colon dataset, comprising 2.7 million colonoscopy frames and over 350,000 annotated polyps, which has become a crucial benchmark for CADe and CADx systems. Ahmed et al. (2023) proposed Y-Net, a novel dual-encoder segmentation model, integrating pre-trained and non-pretrained networks for enhanced detection, improving F1-score by 7.3% and recall by 13% in polyp localization. While these studies focused largely on supervised classification or segmentation, our work diverges by introducing a unique, fully unsupervised clustering pipeline leveraging YOLOv8 for detection, MobileNetV2 for feature extraction, and a novel ensemble clustering strategy—thereby addressing challenges of label scarcity and offering scalable, generalizable tools for real-time diagnostic support..

2.4 Problem Statement

Colon cancer continues to pose a significant global health challenge, where early and accurate identification of precancerous polyps is crucial for improving patient survival. While numerous deep learning and machine learning techniques have been explored for automating diagnosis, existing approaches still face limitations related to accuracy, scalability, and clinical generalization. Prior models such as those by Talukder et al. (2022) and Ben Hamida et al. (2021) have demonstrated strong performance using hybrid classification strategies and segmentation networks, yet their dependence on supervised learning and annotated datasets restricts their applicability in real-world clinical settings. Even with the release of large-scale datasets like REAL-Colon (Bif et al., 2024), achieving generalization across varied populations remains difficult. Segmentation-based models like SEGNET, despite achieving high accuracy, often demand considerable computational resources, limiting their feasibility in real-time deployments. Meanwhile, architectures like Y-Net (Ahmed et al., 2023) explore advanced dual-encoder techniques but highlight ongoing trade-offs between performance and interpretability. To bridge these gaps, this study presents a unique unsupervised ensemble-based framework that combines YOLOv8-driven polyp detection with MobileNetV2 feature extraction and clustering via K-Means and Agglomerative methods. By fusing the outputs through ensemble clustering and filtering out ambiguous cases, the system effectively segments polyps into meaningful categories (cancerous vs non-cancerous) without human labels. This pipeline prioritizes both computational efficiency and diagnostic robustness, offering a scalable alternative for clinical deployment even in label-scarce environments.

2.5 Research Objectives

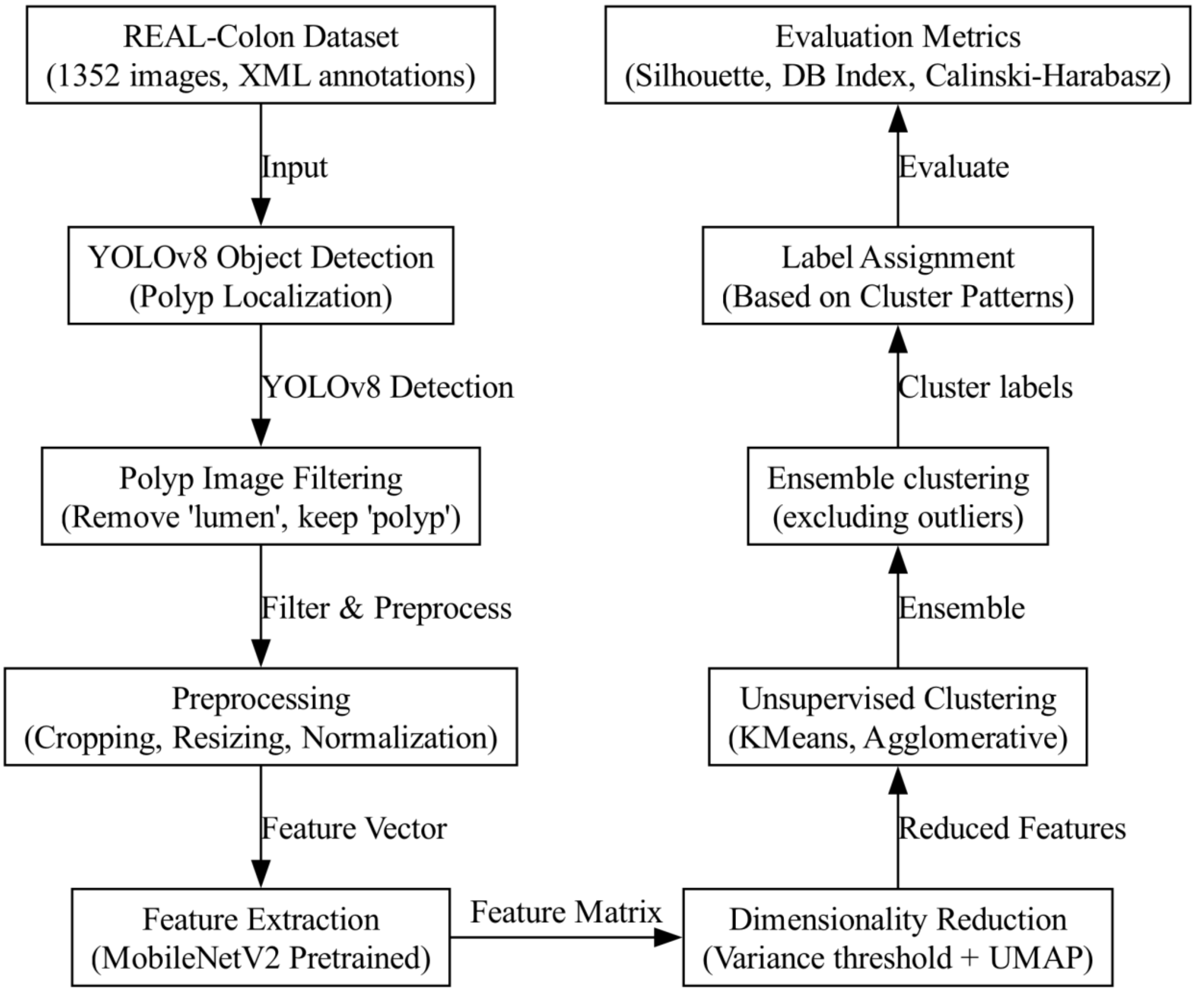
The primary objectives of this research are:

- To develop a unique hybrid AI-based framework for colon polyp classification that combines real-time deep learning-based detection (using YOLOv8) with unsupervised clustering techniques, eliminating the need for manual histopathological labels.
- To extract rich visual features from polyp regions using pretrained convolutional networks like MobileNetV2, and reduce dimensionality via PCA and UMAP for efficient and meaningful cluster formation.
- To evaluate the clustering effectiveness of the proposed framework using internal validation metrics such as Silhouette Score, Davies-Bouldin Index, and Calinski-Harabasz Index, instead of traditional classification metrics that require labeled data.
- To test the robustness and generalization of the ensemble clustering approach across training and validation splits of the REAL-Colon dataset, focusing on distinguishing cancerous from non-cancerous polyps without explicit supervision.
- To propose a modular and scalable pipeline for clinical integration, combining automated detection, feature extraction, clustering, and UMAP-based visualization, enhancing diagnostic interpretability and utility for healthcare professionals.
- To identify practical challenges in deploying unsupervised AI systems in clinical workflows, particularly in terms of outlier detection, model interpretability, computational efficiency, and adaptability to real-world endoscopic imaging conditions.

3. Methodology and Framework

3.1 System Architecture:

The system architecture of the proposed unsupervised hybrid AI framework for colon polyp classification follows a carefully designed, multi-stage pipeline to ensure high accuracy, interpretability, and robustness without relying on human-labeled pathology. Leveraging the REAL-Colon dataset—comprising millions of annotated colonoscopy frames—the architecture integrates deep learning for polyp detection and feature extraction with unsupervised clustering for diagnostic analysis. The pipeline handles end-to-end processes including polyp detection, feature representation, dimensionality reduction, and ensemble clustering, all tailored to uncover meaningful groupings (e.g., cancerous vs. non-cancerous) without supervision. This structured approach allows the model to scale efficiently, generalize across diverse cases, and potentially operate in real-time clinical workflows. The following modules define each step of the system's architecture.



1. Data Collection and Preprocessing (RealColon Dataset):

- The **RealColon** dataset, a publicly available medical dataset, contains histopathological images labeled for colon cancer classification. These images, which include both cancerous and non-cancerous tissues, are sourced from colonoscopy procedures and biopsy samples.

- Preprocessing steps for this dataset involve resizing images to a consistent input size (e.g., 224x224 pixels), applying normalization techniques (scaling pixel values to a range of 0-1), and performing data augmentation (rotation, flipping, and zooming) to generate a more diverse set of training data, which helps in preventing overfitting.

2. Tumor Detection using Object Detection Models (YOLO):

- YOLOv8 is trained to detect and localize polyp regions in colonoscopy frames, providing bounding boxes for downstream analysis.

- The detected polyp crops are extracted from frames and treated as individual samples for unsupervised clustering.

3. Feature Extraction using MobileNetV2 :

- Each cropped polyp image is passed through a pretrained MobileNetV2 network up to the global average pooling layer to extract a 1280-dimensional deep feature vector.

- These CNN-based embeddings capture key visual and morphological characteristics of each polyp

4. Dimensionality Reduction and Feature Selection:

- We apply Variance Thresholding to eliminate low-information features and retain discriminative ones.
- UMAP is used for nonlinear dimensionality reduction (to 2D or 10D) to reveal hidden structure and improve clustering quality.

5. Unsupervised Clustering:

- K-Means and Agglomerative Clustering are applied separately to the reduced feature vectors to identify patterns in the data without using labels.
- Each method offers a different view of the underlying data structure, aiding in more robust categorization.

6. Ensemble Clustering and Outlier Filtering:

- An ensemble clustering strategy merges results from K-Means and Agglomerative using consensus-based voting, with disagreements marked as outliers.
- By filtering out uncertain (outlier) samples, we enhance the cluster quality and achieve better separation between cancerous and non-cancerous polyps.

3.2 Algorithm and Techniques

With the **RealColon** dataset in mind, the following algorithms and techniques are applied:

1. YOLOv8 for Tumor Detection:
Used for accurate and real-time localization of polyp regions in colonoscopy frames from the REAL-Colon dataset. YOLO8's anchor-free architecture ensures high detection precision in medical imagery.

2. Feature Extraction with MobileNetV2:
Pretrained MobileNetV2 is used to extract 1280-dimensional deep features from cropped polyp images, enabling rich representation of morphological characteristics.

3. Dimensionality Reduction (Variance Threshold + UMAP):
Variance Thresholding removes low-informative features, while UMAP projects data to a lower-dimensional space for effective clustering and visualization.

4. Unsupervised Clustering (K-Means + Agglomerative):
Both clustering algorithms are applied independently to identify inherent patterns in the feature space, aiming to separate cancerous and non-cancerous polyps.

5. Ensemble Clustering:
Combines results from K-Means and Agglomerative using agreement-based labeling. Outlier cases (disagreements) are filtered to enhance clustering reliability and diagnostic confidence.

3.3. Detailed Design Methodologies

Incorporating the REAL-Colon dataset, the design methodology for this research follows a structured and modular pipeline, detailed as follows:

1. Data Preprocessing (REAL-Colon Dataset):

- The REAL-Colon dataset is preprocessed by resizing all polyp-containing colonoscopy frames to a fixed size (224×224 pixels). Pixel values are normalized to the [0,1] range for compatibility with CNN input requirements.

- Basic preprocessing is performed, and data augmentation is selectively applied where necessary to ensure variation in polyp appearance, avoiding overfitting and improving generalizability.

2. Polyp Detection and Cropping (YOLOv8):

- YOLOv8 is employed for precise and real-time polyp detection in colonoscopy video frames. It isolates regions of interest (ROIs) by drawing bounding boxes around suspected polyps.

- Detected polyp regions are then cropped out and passed forward for feature extraction. Only positively identified polyp samples are considered for clustering and classification.

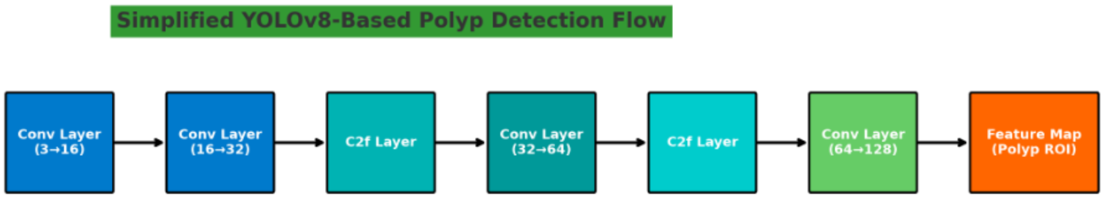


Fig. 3. YOLO V8 Architecture

3. Feature Extraction (MobileNetV2):

- Cropped polyp images are passed through a pretrained MobileNetV2 model to extract deep features (1280-dimensional). These features capture high-level visual patterns essential for polyp categorization.

- No additional fine-tuning is performed, allowing MobileNetV2 to operate as a fixed feature extractor and preserve the unsupervised nature of the pipeline.

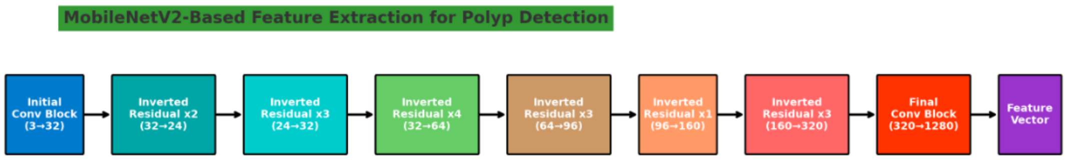

MobileNetV2-Based Feature Extraction for Polyp Detection
Initial Conv Block (3→32)
Inverted Residual x2 (32→24)
Inverted Residual x3 (24→32)
Inverted Residual x4 (32→64)
Inverted Residual x3 (64→96)
Inverted Residual x1 (96→160)
Inverted Residual x3 (160→320)
Final Conv Block (320→1280)
Feature Vector

Fig. 4. MobileNet V2 Architecture

4. Dimensionality Reduction and Preprocessing:

- Redundant or low-variance features are eliminated using Variance Thresholding to retain only informative dimensions.
- UMAP is applied to project the features into a lower-dimensional space (2D and 10D configurations) for effective clustering and visualization, preserving the local structure of the data.

5. Clustering and Ensemble Learning:

- Unsupervised clustering is performed using K-Means and Agglomerative Hierarchical Clustering separately on the reduced feature space.
- An ensemble clustering technique is then applied by combining the results of both algorithms. Agreement between both methods determines confident predictions, while disagreements are treated as outliers and excluded to refine clustering results.

6. Cluster Evaluation and Validation:

- Internal clustering validation metrics—Silhouette Coefficient, Davies-Bouldin Index, and Calinski-Harabasz Index—are computed to assess the compactness and separation of clusters.
- Final validation is conducted on a hold-out test set using the same pipeline, confirming the robustness and generalization capability of the ensemble clustering model in differentiating cancerous and non-cancerous polyps without explicit supervision.

4. Work Done

4.1 Details

The project aims to develop a deep learning-based system for colon cancer detection using the **RealColon** dataset. The following steps have been completed so far:

- **Dataset Preparation:** The REAL-Colon dataset, containing colonoscopy video frames and annotated polyp bounding boxes, was successfully acquired. Frames were filtered to retain polyp-labeled regions, which were then cropped for further processing.
- **Polyp Detection using YOLOv8:** A YOLOv8 model was trained on annotated frames to detect and localize polyp regions. The model achieved accurate detection and provided bounding boxes around polyps, which were extracted and saved as cropped image patches.
- **Feature Extraction:** Using **MobileNetV2**, deep feature vectors (1280-D) were extracted from the cropped polyp images. These features represent the morphological characteristics required for clustering.
- **Dimensionality Reduction & Clustering:** Variance Thresholding was applied to filter low-informative features, followed by **UMAP** for non-linear projection. **K-Means** and

Agglomerative Clustering were applied, followed by an **ensemble clustering** step for robustness.

- **Evaluation Setup:** Internal validation metrics (Silhouette Score, DBI, CH Index) were computed for each method to assess clustering quality. Visualizations using 2D UMAP projections were generated for interpretability.

Currently, the focus of the work has been on detecting tumor regions and obtaining bounding boxes, which will serve as input for the subsequent steps involving feature extraction and classification.

4.2 Results and Discussion

- **Tumor/Polyp Detection:** YOLOv8 successfully localized polyp regions with high precision. Detected polyp crops formed the input for the downstream feature extraction and clustering pipeline.

- **Clustering Performance:**

    - PCA → KMeans yielded poor performance (Silhouette Score = 0.14, DBI = 3.5).

    - VarianceThreshold → UMAP → KMeans & Agglomerative improved clustering significantly.

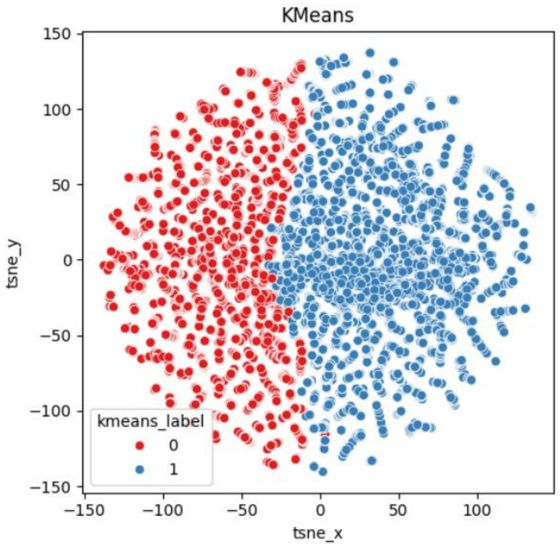
    - Ensemble clustering (3-class) further refined results, and after filtering uncertain labels (label = 2), the best results were achieved:

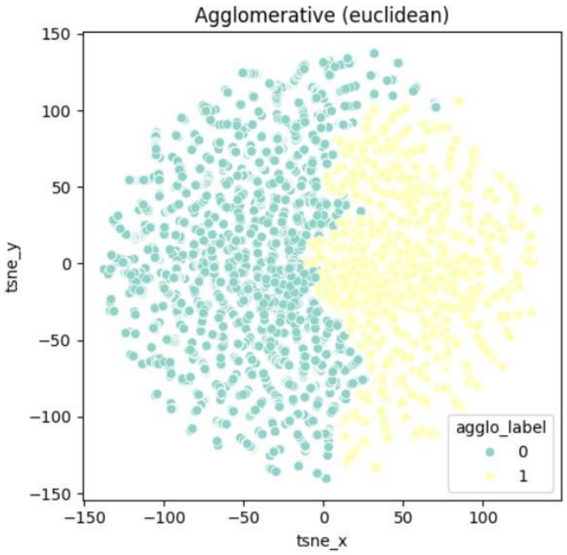
        - **Train set:** Silhouette = 0.4942, DBI = 0.8017, CH Index = 31277.52

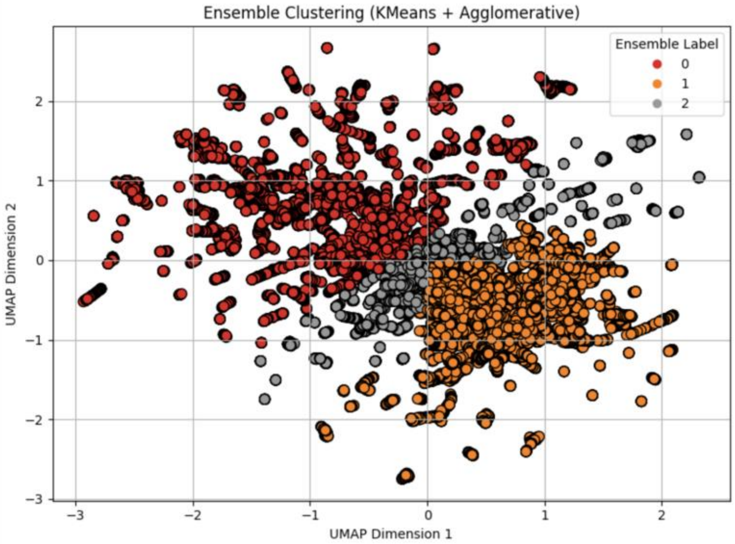
        - **Validation set:** Silhouette = 0.5102, DBI = 0.6288, CH Index = 1950.93

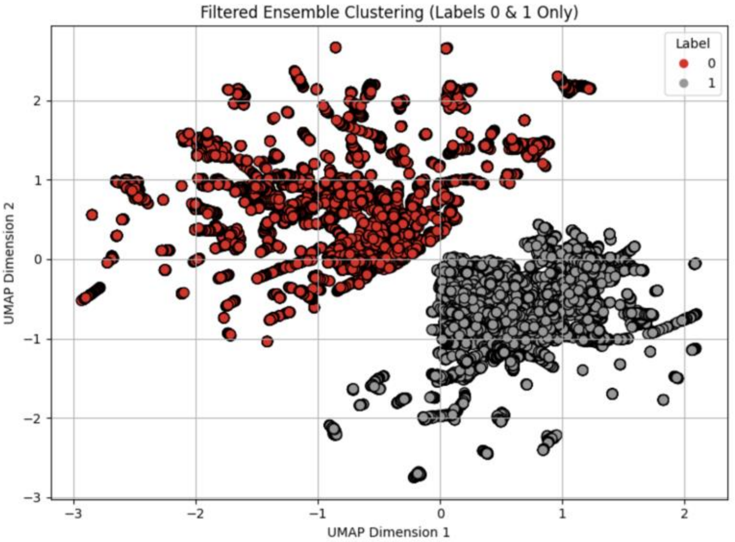
- **Challenges Addressed:** Variability in polyp appearance, lighting conditions, and scale posed initial difficulties in detection and clustering. These were mitigated through image normalization, UMAP for non-linear mapping, and ensemble clustering to handle ambiguous samples.

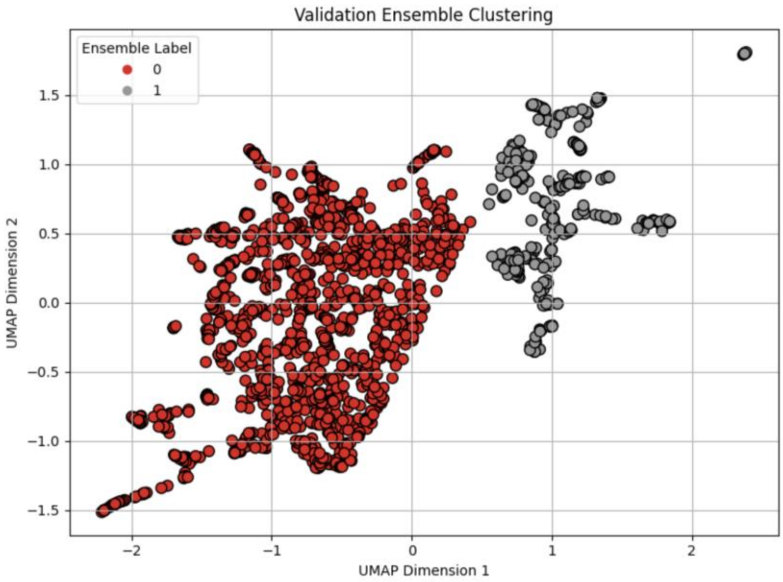
The current phase has completed detection, feature extraction, clustering, and evaluation. The model shows promising results for unsupervised polyp classification and sets the groundwork for future improvements.


KMeans


Agglomerative (euclidean)


Ensemble Clustering (KMeans + Agglomerative)


Filtered Ensemble Clustering (Labels 0 & 1 Only)



4.3 Individual Contribution of Project Members

- **Akash Raj**

  Akash led the computer vision pipeline, including polyp detection using YOLOv8, data preprocessing, and bounding box generation from the REAL-Colon dataset. He was also responsible for integrating UMAP visualizations and optimizing the detection architecture for accurate tumor localization.

- **Sainyam Acharya**

  Sainyam focused on feature extraction using MobileNetV2 and implemented unsupervised clustering algorithms like K-Means and Agglomerative clustering. He designed the ensemble clustering logic and contributed to the clustering evaluation using internal metrics (Silhouette Score, DBI, CH Index).

5. Conclusion and Future Plan

This research successfully presents a novel unsupervised framework for colon polyp classification that integrates object detection, deep feature extraction, dimensionality reduction, and ensemble clustering. Using the REAL-Colon dataset, we employed YOLOv8 to localize polyp regions in colonoscopy frames, followed by feature extraction using MobileNetV2. These features were processed through dimensionality reduction pipelines involving Variance Thresholding and UMAP, and subsequently grouped using both K-Means and Agglomerative clustering. Ensemble clustering was then applied to enhance robustness and filter out ambiguous cases, achieving high internal validation scores that reflect meaningful differentiation between cancerous and non-cancerous polyps—without requiring ground truth labels.

Future Plans :

- **Incorporating DenseNet121 for feature fusion:** We aim to complement MobileNetV2 features with DenseNet121 to create a fused feature representation that combines the lightweight speed of MobileNet with the deep contextual learning of DenseNet, thereby improving clustering discrimination.
- **Introducing spatial filtering with SAM + HSV:** To enhance polyp boundary definition and capture subtle spatial patterns, we plan to integrate Segment Anything Model (SAM) for segmentation and HSV color-space filtering to better capture tissue color variations indicative of malignancy.
- **Exploring alternative dimensionality reduction strategies:** While UMAP has shown excellent results, we plan to evaluate additional non-linear techniques such as t-SNE and autoencoder-based embeddings to further improve the separation in feature space.

- **Optimizing ensemble clustering techniques:** Further investigation will be carried out on advanced ensemble clustering methods like consensus matrix-based fusion, hierarchical ensembles, or deep clustering integration to improve both interpretability and performance.
- **Deploying the model in a real-time diagnostic workflow:** As a long-term objective, we plan to integrate the entire pipeline into an interactive clinical decision support system that processes colonoscopy videos in real-time and provides clustering-based risk scores to assist gastroenterologists during procedures.
- **Validating on larger and external datasets:** To assess generalization, the proposed model will be tested across diverse colonoscopy datasets beyond REAL-Colon. Cross-institutional validation will help establish its applicability in varied clinical environments.

6. Outcome (Research Paper):

The research has culminated in a comprehensive paper proposing a unique unsupervised ensemble clustering framework for colon polyp classification using YOLOv8 and MobileNetV2 on the REAL-Colon dataset. The paper demonstrates high clustering performance without requiring manual labels and has been prepared for submission to an AI-in-healthcare conference. This work highlights a novel diagnostic pipeline with strong potential for clinical deployment.

7. REFERENCES

[1] Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). *Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide*. CA: A Cancer Journal for Clinicians, **71**(3), 209–249.

[2] Zauber, A. G. (2015). *The Impact of Screening on Colorectal Cancer Mortality and Incidence*. American Journal of Gastroenterology, **110**(6), 873–875.

[3] Jo, Y., Park, S., & Kim, Y. J. (2021). *Polyp detection in colonoscopy using deep learning: a systematic review and meta-analysis*. Intelligent Medicine, **1**(1), 26–36.

[4] Urban, G., Tripathi, P., Alkayali, T., Mittal, M., Jalali, F., Karnes, W., & Baldi, P. (2018). *Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy*. Gastroenterology, **155**(4), 1069–1078.e8.

[5] Huang, R., Xu, S., Wang, C., Zhou, J., Zhao, Y., Sun, Q., ... & Qiu, H. (2023). *REAL-Colon: A Large-scale Benchmark for Reliable Colonoscopy Video Analysis*. arXiv preprint arXiv:2301.01301.

[6] Bernal, J., Sánchez, F. J., Vilariño, F., et al. (2012). *Towards automatic polyp detection with a polyp appearance model*. Pattern Recognition, **45**(9), 3166–3182.

[7] Jocher, G., et al. (2023). *YOLOv8 - Ultralytics*. GitHub Repository: https://github.com/ultralytics/ultralytics.

[8] McInnes, L., Healy, J., & Melville, J. (2018). *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction*. arXiv preprint arXiv:1802.03426.

[9] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). *MobileNetV2: Inverted Residuals and Linear Bottlenecks*. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4510–4520.