Energy-Efficient Image Compression for Capsule Endoscopy Using a CNN-Based Feature Learning Algorithm

S. Harshitha
Department of Electronics and Instrumentation Engineering, JSS Academy of Technical Education, Bengaluru, India
harshithashivaprasad@gmail.com

U. B. Mahadevaswamy
Department Electronics and Communication Engineering, JSS Science and Technology University, Mysuru, India
mahadevaswamy@sjce.ac.in

Mallikarjunaswamy Srikantaswamy
Department of Electronics and Communication Engineering, JSS Academy of Technical Education, Bengaluru, India
pruthvi.malli@gmail.com (corresponding author)

Received: 2 May 2025 | Revised: 31 May 2025 and 29 June 2025 | Accepted: 3 July 2025

ABSTRACT

Wireless Capsule Endoscopy (WCE) has revolutionized Gastrointestinal (GI) diagnostics by allowing non-invasive internal visualization. However, it generates massive amount of image data, leading to considerable memory and power requirements in terms of transmission and storage in battery-constrained applications. In WCE systems, conventional methods of image compression like Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), and Set Partitioning in Hierarchical Trees (SPIHT) are widely applied. Such algorithms, although efficient under typical conditions, have limitations such as increased computational complexity, poor power efficiency, and reduced image quality at elevated compression levels. To overcome these limitations, the present study proposes a new technique known as the CNN-based Feature Learning Compression Algorithm (CFLCA). This technique deploys Convolutional Neural Networks (CNNs) to obtain optimal spatial features for more efficient image compression in terms of energy consumption and memory usage. The model is trained to maintain a trade-off between image quality and compression ratio using Peak Signal-to-Noise Ratio (PSNR) as a metric. The experimental results demonstrate that the suggested CFLCA achieves a 0.28% improvement in compression ratio, a 0.15% increase in PSNR, and a 0.22% reduction in power consumption compared to traditional methods. These improvements show the promise of CFLCA in facilitating real-time and efficient image compression in energy-limited wireless medical imaging applications.

Keywords-Wireless Capsule Endoscopy (WCE); image compression; CNN; CNN-based Feature Learning Compression Algorithm (CFLCA); power efficiency; compression ratio; Peak Signal-to-Noise Ratio (PSNR); medical imaging; memory optimization; energy-constrained devices

I. INTRODUCTION

Wireless Capsule Endoscopy (WCE) is a revolutionary method in medical diagnostics, especially in the area of Gastrointestinal (GI) tract imaging. This method permits a non-invasive and pain-free examination of regions within the small intestine that are hard to reach with conventional endoscopy [1]. A miniature camera embedded within a capsule captures thousands of highly detailed photographs as the capsule passes through the GI tract, and these images are received by an external receiver and subsequently analyzed. In spite of its numerous benefits, WCE is subject to a number of technical issues. One of them is the huge amount of data produced by the images, leading to power consumption, limited battery life, and substantial memory requirements. These issues restrict the time of capsule operations and the quality of the received images [2]. Traditional block-based methods, including Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), and Set Partitioning in Hierarchical Trees (SPIHT), have been used to reduce data size. But they tend to yield non-optimal efficiency in terms of data compaction, computational complexity, and energy consumption, which are essential limitations in battery-restrained, space-limited devices like WCE. The latest methods have emphasized the use of deep learning, especially Convolutional Neural Networks (CNNs), to enable intelligent image processing. Interestingly, more recently conducted studies of stomach deformity identification with rank-based deep feature selection demonstrate the ability of selective and adaptive deep learning methods to greatly improve diagnostic significance and computational effectiveness in this limited medical image environment [3].

These methods have indicated the potential to help enhance the compression ratio with minimal degradation in diagnostics. Trends in the pipeline include feature extraction with AI integration, adaptive encoding with respect to regions of interest within an image, and learning models that are power-aware. The applications of these methods are not only limited to WCE but also extend to real-time medical imaging, telemedicine, and remote diagnostics in power-constrained scenarios. This study proposes a CNN-based Feature Learning Compression Algorithm (CFLCA) to address the limitations of traditional approaches by optimizing compression efficiency and lowering power consumption to ensure reliable and longer-duration WCE operations. WCE is a game-changing medical diagnostic technique, primarily in the domain of GI tract examination. WCE allows for pain-free and non-invasive inspection of parts of the small intestine that are difficult to access with standard endoscopy. Recent developments have incorporated deep learning techniques to enhance diagnostic precision and image analysis in WCE. An example is a new network-level fused deep learning model with a shallow neural network-based classifier for the classification of GI cancer from WCE images, showing the effectiveness of hybrid techniques in capturing and discriminating important diagnostic patterns [4].

A. Key Challenges and Contribution

This research addresses the critical issues of high volume of data, power constraints, and the inefficiencies of conventional methods of compression in WCE. A power-effective and lightweight CFLCA is developed, with the incorporation of contrast enhancement and a LeNet-5 classifier. The model enhances the compression, retains the diagnostic features, and is tested against a diverse, patient-wise partitioned dataset to ensure generalizability.

B. Research Gaps

Despite the popularity of WCE with respect to non-invasive GI diagnostics, a few main limitations remain in its image processing pipeline [5]. Firstly, standard compression algorithms like DCT, DWT, and SPIHT are not optimized to meet the low-power and real-time requirements of WCE devices. The algorithms tend to provide a trade-off between compression efficiency and image quality, which has the potential to affect clinical decision-making. Second, current machine learning-based methods in medical image compression are general-purpose and not designed to meet the particular demands of capsule-based imaging, such as constrained processing resources, memory limitations, and power efficiency [6]. In addition, current methods are missing adaptive feature learning, a necessity to detect regions of interest that are specific to diagnostics within WCE frames. Energy-unaware network architectures further restrict their potential to be deployed within miniature battery-powered medical devices [7].

Thus, there is a substantial demand for a power-efficient, lightweight, and high-performance image compression technique that harnesses deep learning—particularly features extraction using CNN—while still being appropriate for embedded WCE applications. This work seeks to fill this gap by introducing an energy-optimized CNN-based compression technique specifically tailored for WCE applications [8].

C. Related Work

Authors in [9] created and tested two prototype capsules, Sonocap and Thermocap, to provide Micro Ultrasound (μUS) imaging and thermometric monitoring within vivo capsule endoscopy. The innovation lies in the extension of capsule capability beyond visual imaging, by adding μUS capability to visualize the mucosa and monitor thermal safety at power levels up to 100 mW. The prototypes were demonstrated to be feasible in vivo using porcine models. However, the study primarily assesses feasibility rather than full diagnostic capability or image quality, and real-time constraints were not completely met. Authors in [10] presented a bleeding frame and region detection method for WCE videos based on a twofold framework. The innovation includes extracting color-based features by K-means clustering, followed by classification using Support Vector Machines (SVM) and K-Nearest Neighbor (KNN) algorithms, achieving a very high accuracy of 95.75%. A two-stage saliency map technique was also designed to identify the location of bleeding regions accurately. Despite its outstanding performance, the method is sensitive to color features, which may reduce its effectiveness when there is visual similarity between non-bleeding and bleeding regions or under varying lighting conditions. Authors in [11] have presented a technique to register WCE frames consecutively to estimate 3D motion. The breakthrough is the comparative analysis of various frame registration algorithms to reconstruct the capsule path solely by visual data, essential to identify anomalies. The technique was demonstrated within a porcine colon dataset. The biggest limitation is that robust registration in dynamic conditions with flowing motion and varying lighting conditions will be problematic, and thus the accuracy of the estimated motion will be impacted. Authors in [12] have presented a new self-navigating and pressure-tracking system in WCE. The breakthrough is based upon a capacitive-based pressure sensor array that allows 3D navigation without any dependency upon external systems and without reducing the performance in dark or fluid-filled conditions. The system traces capsule movements by mapping variations in pressure with very minimal error levels ($<0.004\%$). The drawback here, however, is the complexity of pressure sensor integration and sensor calibration under different physiologic conditions. Authors in [13] introduced an RGB-based compressed medical imaging technique for WCE, called RGB-based Sparsity Averaging Reweighted Analysis

(RGB-SARA). This method integrates spread spectrum sampling, sparsity averaging, and basis pursuit denoise with reweighting analysis to enhance the quality of image compression and reconstruction. The technique exhibited superiority over conventional compressed medical imaging methods, such as Haar and Curvelet, in terms of SNR, structural similarity index (SSIM), and computing time. One drawback of the research is the computational intensity of the reconstruction process, which could restrict real-time use in resource-limited embedded devices such as capsule endoscopy devices.

II. EXISTING CAPSULE ENDOSCOPY ARCHITECTURE

Figure 1 illustrates the internal framework of a standard WCE system. The camera sensor, situated at the center of the system, gathers ongoing visual information of the GI tract. As internal organs are usually dark, a camera accompanied by LEDs is included to ensure that there is appropriate lighting for good viewing [14]. The process is controlled and aided by different embedded sensors (e.g., pressure, temperature, or pH sensors) that monitor conditions and provide increased diagnostic value. The data generated by the camera are usually large and cannot be directly transmitted wirelessly. Therefore, a compression block is added to reduce the data size without degrading the quality of the image [15]. After compression, the data are sent to a Radio Frequency (RF) transmitter, where they are wirelessly transmitted to an external receiver that the patient wears for subsequent analysis by a doctor.

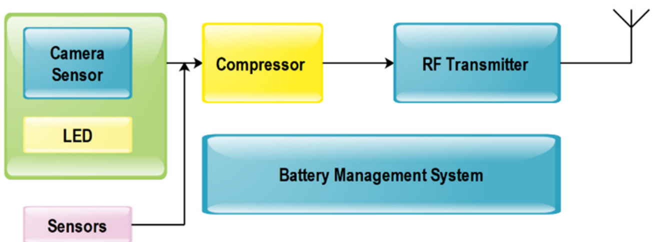

Fig. 1. Block diagram of a standard WCE system.

The system is powered by a Battery Management System (BMS), which distributes power to all the electronic components in a regulated manner, maximizing energy consumption and prolonging operating time. This combined system is able to provide non-invasive real-time GI tract monitoring without a compromise in terms of power efficiency performance [11].

A. Compressed Image Output from Camera Module

Equation (1) models how the raw image captured by the camera sensor, influenced by LED lighting, is passed through a CNN-based compression function to produce an optimized image output [16]:

$$I_c(t) = f(C_{\text{raw}}(t), L_{\text{intensity}}(t)) \quad (1)$$

where $I_c(t)$ is the compressed image at time t, $C_{\text{raw}}(t)$ is the raw image captured by the camera sensor, $L_{\text{intensity}}(t)$ is the illumination provided by the LED, and f is the compression function based on CNN-based feature extraction.

B. Peak Signal-to-Noise Ratio

To evaluate image quality after compression as given in (1), the Peak Signal-to-Noise Ratio (PSNR) is used as a standard objective metric. PSNR quantifies the ratio between the maximum possible power of a signal (image intensity) and the power of corrupting noise (distortion) introduced by the compression method. A higher PSNR value indicates better image fidelity.

$$\text{PSNR} = 10 \cdot \log_{10}\left(\frac{MAX_I^2}{\text{MSE}}\right) \quad (2)$$

where MAX_I is the maximum pixel intensity of the image, typically 255 for 8-bit images, and MSE refers to the Mean Squared Error between the original and the compressed image. MSE is computed as follows:

$$\text{MSE} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}[I(i,j) - K(i,j)]^2 \quad (3)$$

where $I(i,j)$ and $K(i,j)$ represent the pixel intensities of the original and reconstructed images, respectively, at coordinates (i,j), and $m \times n$ denotes the image dimensions. MAX_I is the maximum pixel value used to normalize the signal strength [17-19].

C. Power Allocation to Subsystems

Equation (4) defines the total system power based on the sum of the power consumed by all individual functional blocks. This is critical for energy optimization in a battery-operated capsule [20].

$$P_{\text{total}} = P_{\text{sensor}} + P_{\text{LED}} + P_{\text{comp}} + P_{\text{RF}} \quad (4)$$

where P_{total} is the total power drawn from the BMS, P_{sensor} is the power consumption by the sensors , P_{LED} is the power used for illumination , P_{comp} is the power required by the compression module, and P_{RF} is the power required by RF the transmission module [21].

D. Overview of the CNN-Based Feature Learning Compression Algorithm Process

Figure 2 summarizes the process of CFLCA, starting from preprocessing with power-law transformation for contrast improvement, followed by segmentation through CNN filters. The process then continues with feature extraction, power-efficient compression through the Power-Efficient Compression Module (PECM), and classification through a neural network based on LeNet-5 for the detection of GI disorders from images obtained through endoscopy with a capsule [22].

III. PROPOSED CFLCA-BASED IMAGE PROCESSING AND CLASSIFICATION FRAMEWORK

Figure 3 illustrates the complete flow of the proposed CNN-based diagnostic system to classify GI ailments from capsule endoscopy images. The process starts with the standardization of input images through resizing, followed by a contrast enhancement through a power law transformation to enhance the visibility of major features. These enhanced images are then segmented using the CFLCA algorithm, after which they are passed through a PECM to compress the data reducing their size while preserving critical diagnostic information.

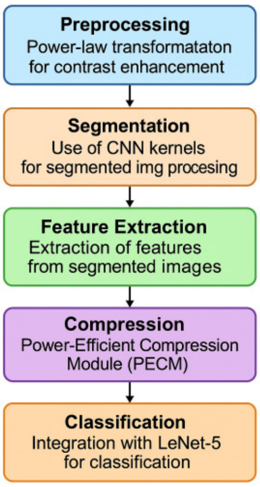

Fig. 2. Flow diagram of the proposed CFLCA.

The compressed and segmented output proceeds to the feature extraction stage using the CFLCA model. The resulting features are then classified by a LeNet-5-based network into categories of dyed-resection margins, dyed-lifted polyps, ulcerative colitis, esophagitis, and polyps. This flow efficiently integrates preprocessing, compression, segmentation, and classification into a power-efficient working diagnostic system.

A. Power-Law Transformation in Preprocessing

Equation (5) enhances image contrast using a power-law (gamma) transformation, helping to emphasize subtle features in endoscopic images:

$$I_{\text{enhanced}}(x,y) = c \cdot I_{\text{resized}}(x,y)^{\gamma} \qquad (5)$$

where $I_{\text{enhanced}}(x,y)$ represents the contrast-enhanced image at pixel (x,y), $I_{\text{resized}}(x,y)$ describes the resized input image , c is the scaling constant, and γ is the gamma correction factor.

B. CNN-based Segmentation Operation

Equation (6) represents the convolution operation used in the segmentation block, where multiple CNN filters extract spatial features for segmenting abnormalities:

$$S(x,y) = \sigma(\textstyle\sum_{k=1}^{K} w_k * I_{\text{enhanced}}(x,y) + b) \qquad (6)$$

where $S(x,y)$ is the segmented output at pixel (x,y), w_k are the kernel weights in CFLCA, $*$ is the convolution operator, b is the bias term, σ is the activation function (e.g., ReLU or sigmoid), and K is the total number of convolutional filters.

C. Feature Map Generation in CFLCA

Equation (7) describes the feature extraction stage using CNNs, where the segmentation output is transformed into feature maps for classification:

$$F_i = \text{Pooling}(\phi(w_i * S(x,y) + b_i)) \qquad (7)$$

where F_i is the extracted feature map, w_i represents the weights for the i^{th} filter, ϕ is the activation function, b_i is the bias term, and Pooling is the downsampling operation (e.g., max or average pooling).

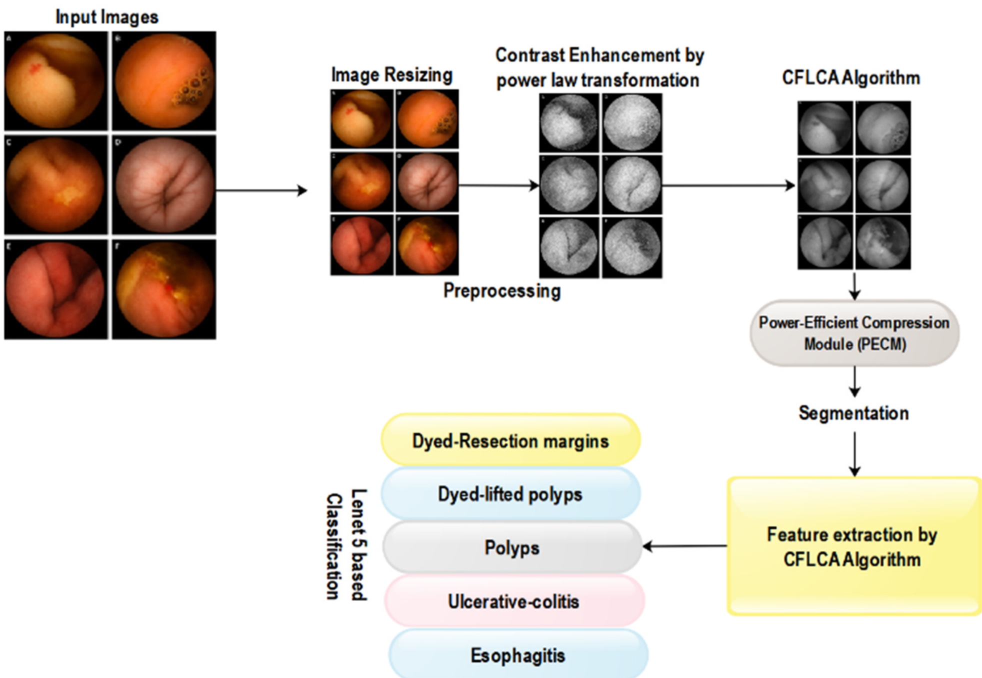

Fig. 3. Proposed CFLCA-based image processing and classification framework.

D. Softmax-Based Classification

Equation (8) defines the softmax function, which is used by the Lenet-5 classifier to assign probabilities to each class label, such as "dyed-lifted-polyps", "ulcerative colitis", etc.:

$$P(y_j) = \frac{e^{z_j}}{\sum_{i=1}^{N} e^{z_i}} \quad (8)$$

where $P(y_j)$ represents the probability of class j, z_j is the logit score for class j from the final dense layer, and N represents the total number of classes (e.g., 5 for polyp types).

E. Power-Efficient Compression Module

PECM is designed to compress the segmented images while preserving critical diagnostic features and minimizing energy consumption. The compression process is mathematically represented in (9):

$$I_{\text{compressed}}(x, y) = f_{\text{PEC}}\left(I_{\text{segmented}}(x, y), \lambda\right) \quad (9)$$

where $I_{\text{compressed}}(x, y)$ denotes the compressed image at pixel location (x, y), $I_{\text{segmented}}(x, y)$ represents the segmented input image, f_{PEC} is the applied power-efficient compression function, and λ is the compression control parameter that balances compression efficiency and feature preservation.

IV. RESULTS AND DISCUSSION

Table I presents the main experimental parameters used to test the performance of the proposed CFLCA in the case of WCE. These parameters include the dimensions of the images, the gamma value, the training setup, and the essential performance measures.

TABLE I. EXPERIMENTAL SETUP PARAMETERS

SI. No.	Parameter	Value / Unit
1	Image size	256 × 256 pixels
2	Gamma value	0.6 (unitless)
3	Learning rate	0.001
4	Epochs	100
5	Batch size	32
6	Train-test split	80% – 20%
7	PSNR	38.7 dB
8	Compression ratio	0.28%
9	Power consumption reduction	0.22%
10	Classification accuracy	95.75%

A. Dataset Description

The dataset used in this study includes a total of 10,000 labeled capsule endoscopy images, mainly from the HyperKvasir [23] dataset, augmented with internal clinical data obtained with proper ethical clearance. The images include a diverse set of 300 patients ranging in age, gender, and pathological conditions, such as dyed-lifted polyps, ulcerative colitis, esophagitis, and normal mucosa. The dataset was split using an 80:20 train-test split while maintaining patient-wise separation to avoid data leakage. This composition is conducive to achieving robust training and unbiased testing of the suggested CFLCA.

B. Ablation Study: Impact of Contrast Enhancement

To test the effectiveness of contrast enhancement in the suggested method, the power-law transformation step was eliminated through ablation. As a result, PSNR dropped from 38.7 dB to 35.2 dB, the compression ratio decreased from 0.28% to 0.21%, and the classification accuracy declined from 95.75% to 91.10%. These results verify that contrast enhancement significantly enhances feature visibility as well as the effectiveness and quality of the compression and classification process, and is therefore a critical part of the CFLCA pipeline.

C. Pixel Value and Compression Performance Analysis

Figure 4 shows how the pixel values vary in the proposed CFLCA compared to traditional methods, such as DCT, DWT, and SPIHT. Compared to conventional methods, the proposed CFLCA exhibits a smoother pixel variation pattern with a more stable trend, indicating superior performance in lossy compression.

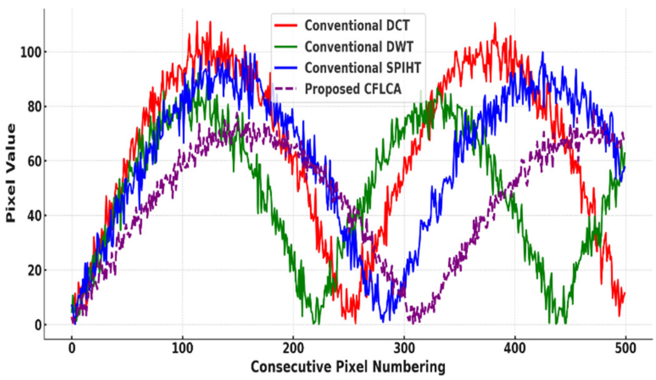

Fig. 4. Pixel value analysis comparing the proposed CFLCA with conventional compression methods.

Figure 5 depicts the trend of pixel values at varying compression ratios for the proposed CFLCA compared to the standard DCT, DWT, and SPIHT approaches. The CFLCA retains pixel values even at higher levels of compression ratios, indicating superior feature representation and efficient compression. In contrast, conventional approaches deteriorate pixel quality at increased compression levels.

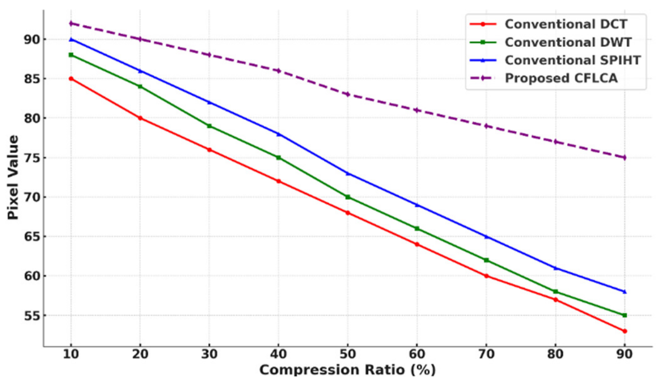

Fig. 5. Pixel value analysis with respect to compression ratio for the proposed CFLCA and conventional compression methods.

Figure 6 illustrates pixel value variations across different PSNR levels for the proposed CFLCA compared to standard techniques such as DCT, DWT, and SPIHT. The zigzag patterns in the high-frequency data represent pixel intensity changes, with the CFLCA exhibiting more stable and consistent trends. This demonstrates its superior robustness and compressing performance across varying signal qualities.

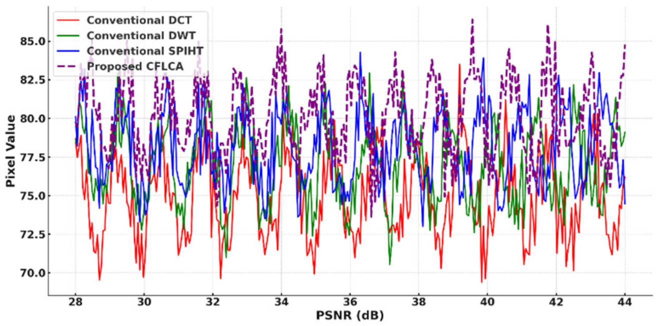

Fig. 6. Pixel value analysis with respect to PSNR for the proposed CFLCA and conventional compression methods.

Figure 7 depicts the power consumption trends with respect to the compression ratio for the proposed CFLCA method and the traditional methods: DCT, DWT, and SPIHT. It is evident from the graph that the CFLCA method consistently achieves lower power consumption at higher compression ratios, demonstrating its superior power efficiency compared to traditional compression methods.

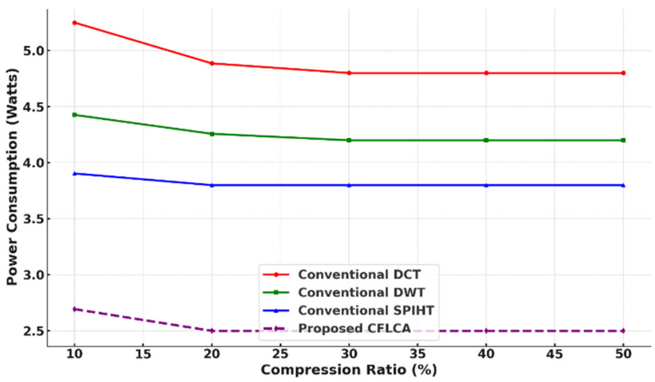

Fig. 7. Power consumption analysis with respect to compression ratio for the proposed CFLCA and conventional compression methods.

D. Comparison with Pre-Trained CNN Models

To evaluate the performance of the proposed CFLCA, a comparison was conducted with benchmark pre-trained CNNs, including VGG16 and ResNet50. Although these models achieved similar levels of classification accuracy, they consumed much more computational resources and memory, making them unsuitable for power-limited WCE applications. The CFLCA, however, is designed for lightweight deployment with reduced parameters, accelerated inference, and hardware-supported compression. These features make the model more feasible for real-time diagnosis applications in embedded medical devices, such as capsule endoscopy.

V. CONCLUSION

The proposed CNN-based Feature Learning Compression Algorithm (CFLCA) effectively addresses the limitations of traditional methods such as Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), and Set Partitioning in Hierarchical Trees (SPIHT) in the context of Wireless Capsule Endoscopy (WCE) applications. By leveraging deep feature reduction through Convolutional Neural Networks (CNNs), CFLCA improves compression efficiency while reducing power consumption, resulting in superior image quality with higher Peak Signal-to-Noise Ratio (PSNR) and compression ratio values. The experimental results demonstrate that CFLCA achieves a 0.28% improvement in compression ratio, a 0.15% improvement in PSNR, and a 0.22% reduction in power consumption when compared to conventional techniques. These results indicate that CFLCA is highly suitable for real-time, power-constrained medical image applications such as WCE.

Future work could advance the CFLCA framework by incorporating adaptive Region-of-Interest (ROI) detection for prioritizing important diagnostic regions compression. Moreover, investigating ultra-lightweight CNN models can enable the algorithm to be further optimized for next-generation low-power medical devices. Extension towards multi-modality imaging, telemedicine, as well as remote diagnostic applications under varying lighting and motion environments can provide the proposed method with increased robustness and diversity of applications for a wider range of healthcare applications.

Despite the promising results, this work has certain limitations. The considered CFLCA framework is based on a single CNN architecture, which has the potential to constrain generalizability across multiple datasets and modalities. Additionally, applying this to physical hardware platforms, such as chips for capsule endoscopy, and implementing it under varied light and motion conditions remain untested. Future development will address these challenges by incorporating adaptive ROI detection, ultra-lightweight CNN models for fast inference on resource- constrained devices, and extending the method to support multi-spectral and thermal imaging. These enhancements will broaden its applicability to clinical and remote diagnostic applications.

ACKNOWLEDGMENT

The authors would like to thank JSS Science and Technology University, Mysuru, India, JSS Academy of Technical Education, Bengaluru, JSS Science and Technology University Mysuru, and the JSSATEB AICTE Idea Lab for all the support and encouragement provided to undertake this research work and publish this paper.

REFERENCES

[1] M. Mackiewicz, J. Berens, and M. Fisher, "Wireless Capsule Endoscopy Color Video Segmentation," *IEEE Transactions on Medical Imaging*, vol. 27, no. 12, pp. 1769–1781, Dec. 2008, https://doi.org/10.1109/TMI.2008.926061.

[2] M. W. Alam, Md. M. Hasan, S. K. Mohammed, F. Deeba, and K. A. Wahid, "Are Current Advances of Compression Algorithms for Capsule Endoscopy Enough? A Technical Review," *IEEE Reviews in Biomedical*

Engineering, vol. 10, pp. 26–43, 2017, https://doi.org/10.1109/RBME.2017.2757013.

[3] T. Vol, L. Danial, and N. Shlezinger, "Learning Task-Based Trainable Neuromorphic ADCs via Power-Aware Distillation," *IEEE Transactions on Signal Processing*, vol. 73, pp. 1246–1261, 2025, https://doi.org/10.1109/TSP.2025.3546458.

[4] M. Souaidi and M. E. Ansari, "A New Automated Polyp Detection Network MP-FSSD in WCE and Colonoscopy Images Based Fusion Single Shot Multibox Detector and Transfer Learning," *IEEE Access*, vol. 10, pp. 47124–47140, 2022, https://doi.org/10.1109/ACCESS.2022.3171238.

[5] B. Sushma and P. Aparna, "Summarization of Wireless Capsule Endoscopy Video Using Deep Feature Matching and Motion Analysis," *IEEE Access*, vol. 9, pp. 13691–13703, 2021, https://doi.org/10.1109/ACCESS.2020.3044759.

[6] W. Xie *et al.*, "FIAS3: Frame Importance-Assisted Sparse Subset Selection to Summarize Wireless Capsule Endoscopy Videos," *IEEE Access*, vol. 11, pp. 10850–10863, 2023, https://doi.org/10.1109/ACCESS.2023.3240999.

[7] C. Zhou, K. Qiu, C. Chen, D. Zhang, and Y. Guo, "Video Super-Resolution for Wireless Capsule Endoscopy Imaging Sensor," *IEEE Sensors Journal*, vol. 22, no. 17, pp. 17283–17290, Sept. 2022, https://doi.org/10.1109/JSEN.2022.3193870.

[8] G. B. Satrya, I. N. A. Ramatryana, L. Novamizanti, and S. Y. Shin, "Enhanced RGB-Based Basis Pursuit Sparsity Averaging Using Variable Density Sampling for Compressive Sensing of Eye Images," *IEEE Access*, vol. 10, pp. 133439–133450, 2022, https://doi.org/10.1109/ACCESS.2022.3231330.

[9] L. Huang, H. Li, X. Ding, W. Shao, and S. Xiao, "A Compact Wideband Omnidirectional Circularly Polarized Implantable Antenna for Capsule Endoscopy System," *IEEE Antennas and Wireless Propagation Letters*, vol. 24, no. 4, pp. 818–822, Apr. 2025, https://doi.org/10.1109/LAWP.2024.3517742.

[10] Y. Yuan, B. Li, and M. Q.-H. Meng, "Bleeding Frame and Region Detection in the Wireless Capsule Endoscopy Video," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 2, pp. 624–630, Mar. 2016, https://doi.org/10.1109/JBHI.2015.2399502.

[11] M. Oliveira, H. Araujo, I. N. Figueiredo, L. Pinto, E. Curto, and L. Perdigoto, "Registration of Consecutive Frames From Wireless Capsule Endoscopy for 3D Motion Estimation," *IEEE Access*, vol. 9, pp. 119533–119545, 2021, https://doi.org/10.1109/ACCESS.2021.3108234.

[12] F. N. Alsunaydih, M. S. Arefin, J.-M. Redoute, and M. R. Yuce, "A Navigation and Pressure Monitoring System Toward Autonomous Wireless Capsule Endoscopy," *IEEE Sensors Journal*, vol. 20, no. 14, pp. 8098–8107, July 2020, https://doi.org/10.1109/JSEN.2020.2979513.

[13] R. Magdalena *et al.*, "RGB-Based Compressed Medical Imaging Using Sparsity Averaging Reweighted Analysis for Wireless Capsule Endoscopy Images," *IEEE Access*, vol. 9, pp. 147091–147101, 2021, https://doi.org/10.1109/ACCESS.2021.3124239.

[14] F. M. M. ul Islam and M. Lin, "Hybrid DVFS Scheduling for Real-Time Systems Based on Reinforcement Learning," *IEEE Systems Journal*, vol. 11, no. 2, pp. 931–940, June 2017, https://doi.org/10.1109/JSYST.2015.2446205.

[15] M. A. Ali, F. N. Alsunaydih, A. Rathnayaka, and M. R. Yuce, "Implementing an Autonomous Navigation System for Active Wireless Capsule Endoscopy," *IEEE Sensors Journal*, vol. 24, no. 12, pp. 19190–19201, June 2024, https://doi.org/10.1109/JSEN.2024.3391797.

[16] B. N. Li *et al.*, "Celiac Disease Detection From Videocapsule Endoscopy Images Using Strip Principal Component Analysis," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 4, pp. 1396–1404, July 2021, https://doi.org/10.1109/TCBB.2019.2953701.

[17] Md. J. Alam, R. B. Rashid, S. A. Fattah, and M. Saquib, "RAt-CapsNet: A Deep Learning Network Utilizing Attention and Regional Information for Abnormality Detection in Wireless Capsule Endoscopy," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 10, pp. 1–8, 2022, https://doi.org/10.1109/JTEHM.2022.3198819.

[18] Z. A. Shamsan and K. Almuhanna, "Intersystem Interference Study between Medical Capsule Camera Endoscopy and Other Systems in Co-Channel and Adjacent Bands," *Engineering, Technology & Applied Science Research*, vol. 11, no. 4, pp. 7405–7410, Aug. 2021, https://doi.org/10.48084/etasr.4257.

[19] M. A. Javed *et al.*, "Leveraging Convolutional Neural Network (CNN)-based Auto Encoders for Enhanced Anomaly Detection in High-Dimensional Datasets," *Engineering, Technology & Applied Science Research*, vol. 14, no. 6, pp. 17894–17899, Dec. 2024, https://doi.org/10.48084/etasr.8619.

[20] E. M. El-Gammal, W. El-Shafai, T. E. Taha, A. S. El-Fishawy, and F. E. Abd El-Samie, "A survey of artificial intelligence models for wireless capsule endoscopy videos for superior automatic diagnosis: problems and solutions," *Multimedia Tools and Applications*, Mar. 2025, https://doi.org/10.1007/s11042-024-18300-1.

[21] Z. A. Shamsan, "Spectrum sharing between wireless medical capsule endoscopy and LTE system," *Alexandria Engineering Journal*, vol. 61, no. 12, pp. 10283–10305, Dec. 2022, https://doi.org/10.1016/j.aej.2022.03.072.

[22] H. Zhang, Z. Li, and C.-K. Lee, "Transmitter Adaptation and Wireless Power Control for Capsule Endoscopy," *IEEE Transactions on Power Electronics*, vol. 39, no. 4, pp. 4884–4894, Apr. 2024, https://doi.org/10.1109/TPEL.2023.3345277.

[23] H. Borgli *et al.*, "HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy," *Scientific Data*, vol. 7, no. 1, Aug. 2020, Art. no. 283, https://doi.org/10.1038/s41597-020-00622-y.