FICE: Fovea-Inspired Closed-Loop Edge-AI System for Enhanced Capsule Endoscopy Imaging

Tiancheng Cao[1], Hengming Zhang[1], Chen Shen[1], Hen-Wei Huang[1,2]
[1]School of Electrical and Electronic Engineering, Nanyang Technological University, Republic of Singapore
[2]Lee Kong Chian School of Medicine, Nanyang Technological University, Republic of Singapore

Email: tiancheng.cao@ntu.edu.sg

***Abstract*— The wireless feature of a capsule endoscope enables non-invasive gut imaging, but it results in limited power supply, poor imaging quality, and low frame rate. This paper proposes a Fovea-Inspired Closed-Loop Edge-AI (FICE) system to address these constraints without compromising the diagnostic quality. Specifically, we integrate a lightweight transformer-based model for onboard lesion detection and a Gated Corner Proposal Network for precise region-of-interest (ROI) localization. Once lesions are flagged, the system captures and transmits high-resolution ROI images while maintaining low-resolution frames elsewhere, effectively conserving energy and bandwidth. Key findings using the HyperKvasir and Kvasir-SEG datasets demonstrate a lesion detection recall and precision of 99.88% and 99.40%, respectively, along with a mean Intersection over Union (mIOU) of 0.804 for bounding box localization, while maintaining a processing speed of 9.58 fps at 320-pixel density under a maximum data transmission rate of 1248 KBps. The frame rate and pixel density are improved by five and three folds, respectively, compared to the state-of-the art capsule endoscope. This closed-loop approach mimics the human eye's fovea and saccadic movements, balancing the need for detailed imaging of critical areas with overall resource constraints.**

Keywords —Wireless Capsule Endoscopy, Fovea-Inspired Imaging, Closed-Loop Edge AI, Lesion Detection

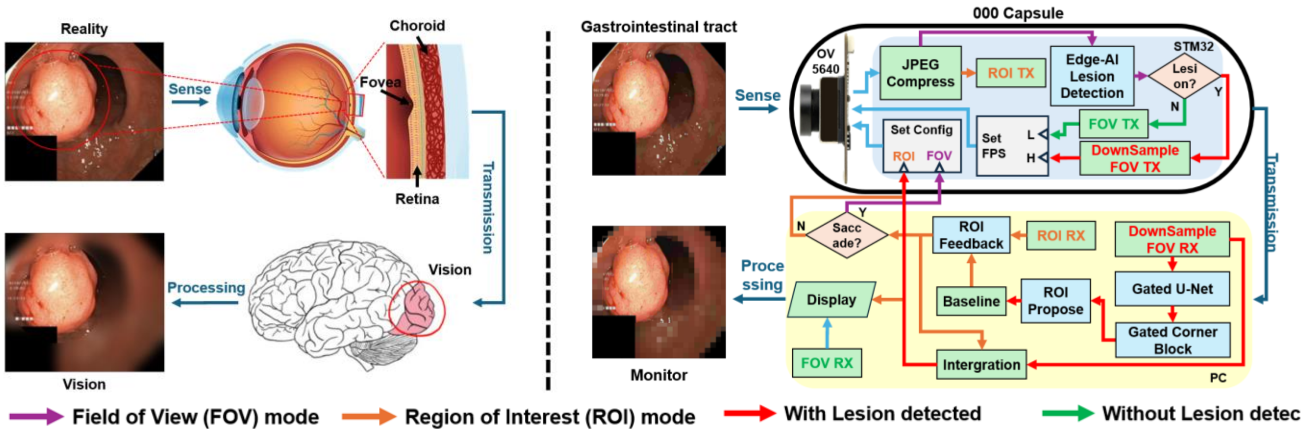
I. INTRODUCTION

Wireless capsule endoscopy (WCE) has revolutionized gastrointestinal diagnostics by providing a non-invasive method to visualize the small intestine [1, 2]. Patients swallow a pill-sized device equipped with a miniature camera that captures images as it naturally moves through the gastrointestinal tract, greatly improving the detection and management of conditions such as obscure gastrointestinal bleeding, Crohn disease, ulcerative colitis, and small intestinal tumors. Recent advancements have focused on enhancing image quality and diagnostic accuracy through improvements in camera technology, data transmission, and image processing algorithms [3-5]. However, inherent trade-offs persist between image quality, frame rate, and battery life due to the limited transmission bandwidth and power resources available within the capsule. Addressing these challenges necessitates innovative approaches that optimize these parameters without compromising device performance.

This work is supported by the Ministry of Education, Singapore, under the Award No. RG71/24.

Recent research has explored various solutions to address these challenges, including energy-efficient hardware designs [6, 7], advanced image compression algorithms [8], and frame rate control mechanisms [9]. Energy-efficient hardware, such as low-power CMOS image sensors [7], aims to reduce power consumption within the capsule; however, the trade-off in processing capability makes it difficult to maintain the high frame rate required for accurate diagnostics. Advanced image compression algorithms [8] help decrease the amount of data transmitted, conserving energy, but they can lead to a loss of image quality. Adaptive frame rate control [9] adjusts the frame rate based on movement but lacks specificity in targeting lesions. These unresolved issues highlight the need for a solution that simultaneously optimizes diagnostic accuracy and power consumption.

The human eye is one of the most efficient imaging systems, featuring a unique characteristic: only a small area called the fovea is perceived in high resolution and vivid color. This mechanism cooperated with visual integration and saccade optimizes visual perception while reducing overall energy consumption. Inspired by this and advanced research on biomedical edge intelligence [10, 11], we have significantly revised the WCE system by integrating a high-resolution (5 MP pixels) CMOS sensor that is 40 folds higher than a PillCam (Medtronic). In this paper, we propose a fovea-inspired closed-loop edge-AI (FICE) system that performs lesion detection within the capsule using edge computing. When potential abnormalities are detected, compressed images are transmitted to an external processor, where a specially designed AI algorithm combined with conventional image processing, Gated Corner Proposal Network, determines the precise location of the target area. The capsule then combines high-resolution images of the lesion with low-resolution backgrounds to create interpolated frames transmitted at relatively high speeds. This approach enables high-quality, high-frame-rate transmission for critical areas to achieve accurate diagnosis while reducing the frame rate and power consumption for healthy regions. Validation using the HyperKvasir Dataset [12] demonstrates that the proposed system achieves an onboard diagnostic recall and precision of 99.88% and 99.40%, respectively, along with a Region of Interest (ROI) detection mean Intersection over Union (mIOU)



(a) (b)

Fig. 1. The proposed fovea-inspired closed-loop edge-AI (FICE) system. The "000 Capsule" integrates an OV5640 camera sensor, JPEG compression, and edge-AI lesion detection capabilities within an STM32 microcontroller. The capsule detects lesions in real-time and transmits region-of-interest (ROI) data or full field-of-view (FOV) data based on lesion presence. A closed-loop mechanism enables the PC-based processing module, which utilizes image reconstruction and lesion analysis, to guide the sensor configuration.

of 0.804, while achieving 9.58 fps with 320 pixel density under a maximum data transmission rate of 1248 KBps. This innovative approach simultaneously optimizes image quality, frame rate, and power consumption, addressing longstanding trade-offs in capsule endoscopy and enhancing diagnostic accuracy without compromising device performance.

The rest of the paper is organized as follows. Section II provides a comprehensive overview of the proposed fovea-inspired closed-loop edge-AI (FICE) system. Section III presents the structures of lesion detection algorithm and the proposed Gated Corner Proposal Network. Section IV presents the results on HyperKvasir Dataset and Kvasir SEG Dataset while Section V concludes the paper.

II. Fovea-Inspired Closed-Loop Edge-AI (FICE) System

A. FICE Architecture Overview

The imaging process of the human eye is illustrated in Fig. 1(a), where the fovea is a small central region of the retina densely packed with photoreceptor cells, allowing it to capture images in high resolution and vivid color. While the fovea processes detailed information in the center of the visual field, the surrounding regions provide lower-resolution input, which the brain integrates into a coherent image with a sharp central focus and less-detailed periphery.

Inspired by this, we propose the FICE system, with the data flow illustrated in Fig. 1(b). After the image sensor captures the initial image, it undergoes JPEG compression and is processed by onboard edge AI to determine the presence of potential lesions. If no lesion is detected, the system transmits images at a lower frame rate to conserve energy. If a lesion is identified, the compressed image is further downsampled and transmitted to an external computer for Region of Interest (ROI) detection.

The external system utilizes the proposed Gated U-Net to extract key point features, followed by the Gated Corner Block to predict the top-left and bottom-right corners of the lesion, thereby defining a bounding box. This bounding box serves as feedback to the image sensor, directing it to capture

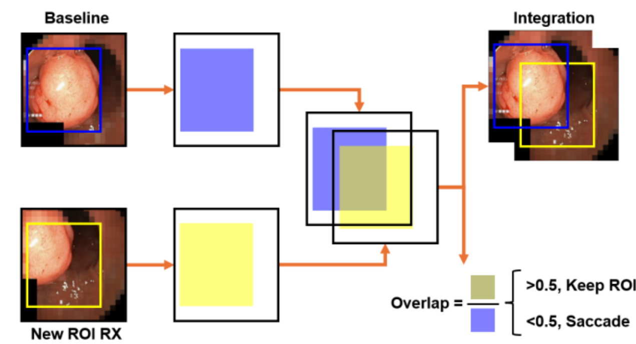


Fig. 2. Scan mechanism based on ROI image feedback.

a high-resolution image of the identified Region of Interest (ROI). The high-resolution ROI image is then transmitted and seamlessly integrated with the low-resolution background on the external computer, where fovea imaging interpolation is produced as the final output. This synthesized output prioritizes high-resolution details in critical regions while maintaining lower resolution in less important areas, achieving an optimal balance between diagnostic accuracy, energy efficiency, and transmission frame rate.

B. Lightweight Edge AI for Lesion Detection

Onboard edge AI is essential in this system for minimizing data transmission, optimizing power consumption, and enabling real-time adaptation. However, conventional AI algorithms, with their extensive parameters and computational demands, are not well-suited for this application. To address this, we adapted a lightweight transformer-based algorithm, PoolFormer [13], which has shown significant success in prior edge AI implementations, tailoring it to meet the specific requirements of this system. 6-layer ParaPoolFormer is adapted here with only 45KB parameters and achieves 99.88% lesion detection recall and 99.40% precision on HyperKvasir dataset. During the training of the lesion detection model, 1,000 lesion images and 1,000 non-lesion images are randomly sampled from HyperKvasir. The resulting dataset is then split into training,

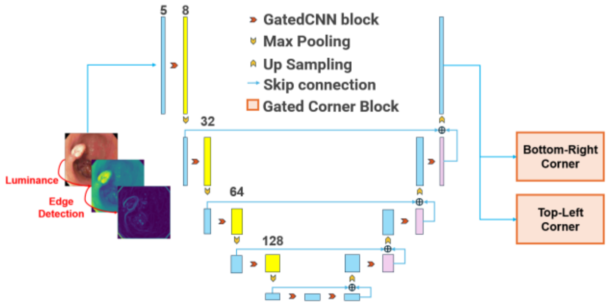


Fig. 3. The proposed Gated U-Net feature extraction with preprocessing.

validation, and test sets with a ratio of 8:1:1. By processing images locally, it reduces the need to transmit large volumes of data and conserving energy.

As shown in Fig. 1(b), the onboard lesion detection model not only identifies the presence of lesions in the current frame, but also dynamically adjusts the camera's frame rate and activates the off-board localization model. When no lesion is detected in the current frame, the OV5640 captures a full FOV image at a low resolution of 500×500, which is then wirelessly transmitted. Simultaneously, the frame rate is reduced to below 1 fps. According to [14], the peristaltic frequency of the gastrointestinal tract ranges from 0.03 to 0.2 contractions per second, making a frame rate of 1 fps sufficient to capture relevant motion. If a lesion is detected in the low-resolution FOV image, the onboard lesion detection model triggers the off-board localization model to generate the ROI coordinates. Simultaneously, the OV5640 camera is configured to its native 5MP resolution and captures an image of the ROI. The MCU calculates the number of pixels in the ROI, estimates the image size, and adjusts the OV5640 frame rate according to the maximum bandwidth available from the wireless communication module.

C. Visual Integration and Saccades by ROI Feedback

Building on the fovea mechanism, the brain visual integration capability continuously overlays and combines high-resolution images captured by the fovea with surrounding background information, stitching together successive high-resolution focal points into a complete visual scene. In the absence of targets in the ROI, the brain initiates saccades, rapid eye movements, to scan the environment, bringing new ROI into the visual field to acquire more information.

Inspired by this efficient visual mechanism, a novel ROI-based feedback design is proposed, as illustrated in Fig. 2. Building on the process outlined in Section II.A, the first interpolated frame generated by the system serves as the baseline for feedback. Due to the natural movement of the capsule endoscope within the gastrointestinal tract, driven by peristalsis, the same ROI area set by the image sensor captures different image details over time. As the image sensor continues to collect new ROI data, this information is used to generate subsequent interpolated frames, which are then compared to the baseline. By comparing the newly acquired ROI data with the baseline, the system performs visual integration, stitching together the ROI regions to form a more comprehensive image over time.

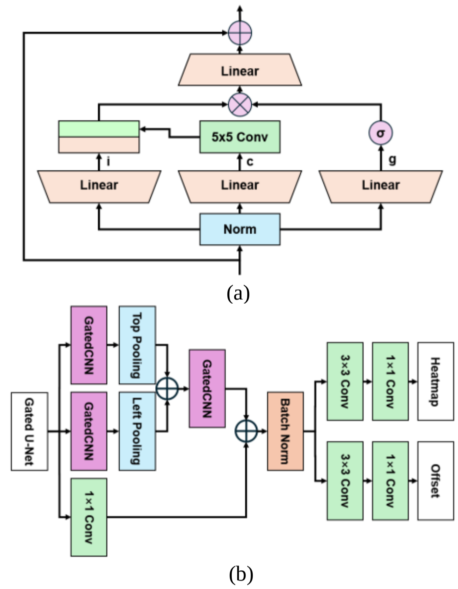


Fig. 4. (a) The structure of Gated CNN block. (b) The structure of Gated Corner Block.

This process continues until the overlap between the newly captured ROI and the baseline ROI falls below a predefined threshold, set at 0.5 in this study. At this point, the system provides feedback to the image sensor, prompting it to switch back to full-view image capture to re-detect lesion areas, mimicking the saccadic movement of the human eye to rapidly scan new regions of the visual field and refocus on areas of interest, thereby closing the loop.

This dynamic feedback mechanism leverages ROI-based feedback and visual integration to continuously update and incorporate high-resolution information while adapting to changes in the capsule position and environment. By dynamically stitching high-resolution ROI data and making adjustments only when necessary, the system eliminates the need for frequent reconfiguration of the image sensor. This approach ensures efficient adaptation to capsule movement while maintaining high-quality imaging for accurate and reliable diagnosis.

III. Gated Corner Proposal Network

In this work, only bounding box detection is needed, and most existing object detection algorithms, such as YOLO, are overly computationally intensive. Therefore, we adopt a strategy inspired by the CornerNet architecture [15], dividing the process into feature extraction and corner proposal as illustrated in Fig. 3.

A. Conventional Image Preprocessing

RGB images combine both color and brightness information, making it challenging for algorithms to isolate critical structural details like texture, edges, and shading necessary for lesion identification. This often reduces sensitivity to subtle variations in brightness and contrast, which are key indicators of abnormalities in medical imaging.

	[18]	[19]	Proposed
AI model	DSCNN	Mask R-CNN	PoolFormer + GCRN
Input Size	240*240*3	512*512*3	500*500*3
Onboard Memory Usage	3.2MB	NR	**0.54MB**
Quantization	INT8	NA	INT8
FOV (degree)	140	130	160
Pixel Density (PPD)	91.4	98.5	**320**
Frame Rate (fps)	2	2	**9.58**
RF module	BLE	BLE	BLE
Bandwidth Requirement	60Kpbs	273Kbps	1248Kbps

N.R.: Not reported

Fig. 5. The comparative analysis of the relative state-of-the-art capsule endoscope solutions.

To address this challenge, once the capsule endoscope transmits compressed images of suspected lesions, the system undergoes several preprocessing steps to enhance ROI detection. The image is first transformed into the Lab color space, where the luminance, L channel, is extracted to focus specifically on brightness and contrast variations, which are often more indicative of structural abnormalities. Sobel edge detection is then applied to the L channel to highlight significant edges and transitions, such as lesion boundaries or structural changes in the gastrointestinal tract, providing valuable spatial cues for accurate ROI localization. Finally, as shown in Fig. 3, the edge-detected L channel is concatenated with the original RGB and L images to form the input for the neural network, effectively integrating high-level contextual information with low-level structural details for improved accuracy and robustness in ROI detection.

B. Gated U-Net Feature Extraction

Compared to the Hourglass module used in CornerNet, the U-Net architecture, as illustrated in Fig. 3, is employed for feature extraction due to its skip connections between the encoder and decoder layers [16]. These skip connections facilitate the transfer of detailed spatial information to the upsampling path, preserving fine-grained details and enhancing localization accuracy. Moreover, U-Net effectively captures features at multiple scales, making it well-suited for both localization and segmentation tasks, which leads to more precise bounding box detection in medical imaging.

Further improvement is achieved by replacing the standard U-Net blocks with GatedCNN blocks shown in Fig. 4(a) [17]. GatedCNN blocks introduce learnable gating mechanisms that adaptively allow relevant features to pass through while filtering out unnecessary information. The mathematical formulation is described as

$$g, i, c = split(FC_1(BN(x)))$$

$$y = ReLU(g) \odot cat(i, conv(c))$$

$$z = x + FC_2(y)$$

where x is the input features, z is the output features and g, i, c are gate features, non-convolutional features, and convolutional features, respectively. $BN(\cdot)$ is the batch normalization layer, $FC(\cdot)$ is the fully-connected layer, $conv(\cdot)$ is the convolutional layer and $ReLU(\cdot)$ is the activation layer. $split(\cdot)$ and $cat(\cdot)$ are the split and concatenate operations.

This enhances the network ability to focus on lesion areas, improving the accuracy of bounding box detection. Moreover, GatedCNN blocks reduce noise and irrelevant background features, which is particularly important in medical imaging where images often contain artifacts or non-informative regions, resulting in a cleaner and more reliable feature representation.

C. Gated Corner Block

The corner block incorporates the concept of corner pooling [14], which captures contextual information by pooling features along horizontal and vertical directions to emphasize the edges of objects. As illustrated in Fig. 4(b), this structure is implemented as the gated corner block.

To efficiently determine corner positions, we modified the output heatmap to generate only the necessary components: a heatmap for the top-left corner, offsets in the x and y directions, and similarly for the bottom-right corner. This results in a total of six output channels for the heatmap, optimized for precise and rapid corner localization.

IV. Simulation Results and Discussion

We developed a prototype PCB for a wireless capsule endoscope to verify the functionality and dimensions of our system. The maximum diameter of the entire system is 10 mm, requiring all components on the PCB to fit within this constraint. The system is primarily composed of four parts: a high-resolution camera (OV5640), a Bluetooth BLE wireless communication module (NRF52840), a high-performance MCU (STM32H743 series), and a wireless power reception coil.

We chose the 5MP OV5640 to provide a higher pixel density under the same field of view (FOV) and viewing distance. This allows us to leverage our proposed retina-like ROI algorithm to enhance the resolution of the lesion areas without increasing the demand for wireless bandwidth. The MCU we selected is the STM32H743. This MCU is equipped with an 8-bit JPEG codec module, enabling real-time compression of images captured by the OV5640 to reduce bandwidth requirements. Additionally, it features 1MB of on-chip cache, providing sufficient resources for running the retina-like ROI algorithm and a lightweight lesion detection model. The NRF52840 module is a low-power Bluetooth communication module with a compact package size of just 4 × 4 mm. Based on our simulation validation, it can achieve a maximum data transmission rate of 1248 KBps for both image and control signals. All simulations were conducted based on the aforementioned hardware resources.

The neural network training was conducted using two specified models, leveraging an Intel 13600KF CPU and an RTX 3060Ti GDDR6X GPU with the PyTorch framework. For the lesion classification model, the primary focus was on recall, as correctly identifying positive cases is critical in medical imaging applications. The edge detection model achieved a recall of 99.88%, demonstrating its effectiveness in accurately identifying potential lesions. The precision is 99.40%. For the ROI detection model, the evaluation emphasized the overlap ratio between the predicted and ground truth regio ns. The overall mean Intersection over Union (mIOU) for the system reached 0.804, further validating its robustness and accuracy in both lesion detection and ROI localization tasks.

To benchmark the proposed system, we compared its simulation performance against state-of-the-art capsule endoscope solutions, as shown in Fig. 5. As depicted in the figure, our algorithm and onboard lesion detection model require only 0.54 MB of memory, which is 83.125% smaller than the system in [18]. This allows our entire system to operate based on near-memory computing, eliminating the need for DRAM access and, consequently, reducing both system latency and power consumption. Furthermore, with the integration of a high-resolution CMOS sensor, our system's pixel density is approximately three times greater than that of the systems in [18] and [19], providing significantly clearer visuals. The enhanced resolution of the lesion areas will greatly improve diagnostic efficiency and accuracy for healthcare professionals. Thanks to the closed-loop ROI algorithm, we can prioritize imaging and data transmission of the lesion regions with higher pixel density. Under the assumption that the lesion region is confined within a 500×500 pixel box, the entire system can achieve 9.58 fps, leveraging the maximum achievable data rate of the NRF52840.

The proposed Fovea-Inspired Closed-Loop Edge-AI (FICE) system addresses longstanding challenges in wireless capsule endoscopy by significantly improving image quality, frame rate, and power efficiency. Validated through the HyperKvasir dataset, the system achieved a diagnostic recall and precision of 99.88% and 99.40% respectively with edge-AI and a mean Intersection over Union (mIOU) of 0.804 in Kvasir SEG dataset, demonstrating its effectiveness in optimizing diagnostic precision. Compared to prior solutions, which often compromise image quality or frame rate to conserve energy, the FICE system leverages lightweight edge-AI and Gated Corner Proposal Networks to deliver superior lesion localization and energy efficiency. This advancement enhances the effectiveness of non-invasive gastrointestinal diagnostics and surgeries, while also paving the way for more advanced medical solutions. However, limitations include the need for further hardware validation, reliance on a single dataset, and predefined thresholds that may lack flexibility in diverse clinical contexts. Future research should expand testing across broader datasets, refine the feedback mechanism for greater adaptability, and explore real-time clinical implementation to ensure broader applicability and reliability.

V. CONLUSION

In this work, we introduced a Fovea-Inspired Closed-Loop Edge-AI (FICE) system that draws inspiration from the human eye's fovea mechanism and leverages lightweight transformer-based onboard lesion detection along with a Gated Corner Proposal Network for precise bounding-box localization. Our core contribution lies in capturing and transmitting high-resolution data selectively when a lesion is detected, significantly reducing energy consumption while preserving diagnostic accuracy. Experimental results on the HyperKvasir dataset demonstrate a 99.88% recall and a precision of 99.40% in lesion detection and an mIOU of 0.804 for ROI localization, while achieving 9.58 fps with 320 pixel density under a maximum data transmission rate of 1248 KBps, underscoring the system's reliability and efficiency in balancing image resolution, power usage, and transmission rate. Beyond improving clinical decision-making by ensuring targeted imaging, FICE also sets the stage for more advanced, closed-loop WCE solutions that incorporate adaptable AI-driven feedback. Future research will involve integrating more advanced wireless modules, expanding testing to diverse datasets, and exploring real-world clinical deployment to further validate and refine this innovative approach.

REFERENCES

[1] Q. Cao, R. Deng, Y. Pan, *et al.*, "Robotic wireless capsule endoscopy: recent advances and upcoming technologies," *Nat. Commun.*, vol. 15, no. 4597, 2024, doi: 10.1038/s41467-024-49019-0.

[2] A. Abdigazy, M. Arfan, G. Lazzi, *et al.*, "End-to-end design of ingestible electronics," *Nat. Electron.*, vol. 7, pp. 102-118, Feb. 2024, doi: 10.1038/s41928-024-01122-2.

[3] Y. Almalioglu *et al.*, "EndoL2H: deep super-resolution for capsule endoscopy," *IEEE Trans. Med. Imaging*, vol. 39, no. 12, pp. 4297-4309, Dec. 2020, doi: 10.1109/TMI.2020.3016744.

[4] H. Zhang, Z. Li and C. -K. Lee, "Transmitter adaptation and wireless power control for capsule endoscopy," *IEEE Trans. Power Electron.*, vol. 39, no. 4, pp. 4884-4894, April 2024, doi: 10.1109/TPEL.2023.3345277.

[5] B. Sushma and P. Aparna, "Recent developments in wireless capsule endoscopy imaging: compression and summarization techniques," *Comput. Biol. Med.*, vol. 149, Art. no. 106087, Oct. 2022, doi: 10.1016/j.compbiomed.2022.106087.

[6] T. Cao *et al.*, "A non-idealities aware software–hardware co-design framework for edge-AI deep neural network implemented on memristive crossbar," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 12, no. 4, pp. 934-943, Dec. 2022, doi: 10.1109/JETCAS.2022.3214334.

[7] F. Tang, D. G. Chen, B. Wang, and A. Bermak, "Low-power CMOS image sensor based on column-parallel single-slope/SAR quantization scheme," *IEEE Trans. Electron Devices*, vol. 60, no. 8, pp. 2561-2566, Aug. 2013, doi: 10.1109/TED.2013.2268207.

[8] Y. Xing, Z. Zhang, Y. Qian, Q. Li and Y. He, "An energy-efficient approximate DCT for wireless capsule endoscopy application," *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, Florence, Italy, 2018, pp. 1-4, doi: 10.1109/ISCAS.2018.8351769.

[9] J. Melson *et al.*, "Video capsule endoscopy," *Technol. Status Eval. Rep.*, vol. 93, no. 4, pp. 784-796, Apr. 2021, doi: 10.1016/j.gie.2020.12.001.

[10] T. Cao, Z. Zhang, W. L. Goh, C. Liu, Y. Zhu, and Y. Gao, "ECG classification using binary CNN on RRAM crossbar with nonidealities-aware training, readout compensation and CWT preprocessing," *2023 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, Toronto, ON, Canada, 2023, pp. 1-5, doi: 10.1109/BioCAS58349.2023.10389002.

[11] T. Cao *et al.*, "Edge PoolFormer: Modeling and training of PoolFormer network on RRAM crossbar for edge-AI applications," *IEEE Trans.*

Very Large Scale Integr. (VLSI) Syst., doi: 10.1109/TVLSI.2024.3472270.

[12] Borgli, H., Thambawita, V., Smedsrud, P.H. *et al. HyperKvasir*, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy. *Sci Data* **7**, 283 (2020). https://doi.org/10.1038/s41597-020-00622-y

[13] T. Cao, W.S. Ng, W. L. Goh and Y. Gao., "DWT-PoolFormer: Discrete wavelet Transform-based quantized parallel PoolFormer network implemented in FPGA for wearable ECG monitoring," *2024 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, Xi'an, China, 2024.

[14] C. S. De Jonge, A. M. J. Sprengers, K. L. Van Rijn, A. J. Nederveen, and J. Stoker, "Assessment of fasted and fed gastrointestinal contraction frequencies in healthy subjects using continuously tagged MRI," *Neurogastroenterology Motil*, vol. 32, no. 2, p. e13747, Feb. 2020, doi: 10.1111/nmo.13747.

[15] H. Law and J. Deng, "Cornernet: Detecting objects as paired keypoints," *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018.

[16] T. Cao, Z. Zhang, C. Liu, W. L. Goh, Y. Zheng and Y. Gao., "Neuromorphic FeRAM-Based Co-Design for Imaging Enhancement in Handheld Photoacoustic Systems," *2025 EEE International Symposium on Circuits and Systems (ISCAS)*, London, UK, 2025.

[17] W. Yu and X. Wang, "MambaOut: Do we really need Mamba for vision?" *arXiv preprint arXiv:2405.07992*, 2024.

[18] A. Sahafi, Y. Wang, C. L. M. Rasmussen, *et al.*, "Edge artificial intelligence wireless video capsule endoscopy," *Sci. Rep.*, vol. 12, no. 1, Art. no. 13723, 2022, doi: 10.1038/s41598-022-17502-7.

[19] C. Zhou *et al.*, "An ingestible near-infrared fluorescence capsule endoscopy for specific gastrointestinal diagnoses," *Biosens. Bioelectron.*, vol. 257, Art. no. 116209, Aug. 2024, doi: 10.1016/j.bios.2024.116209.