

ABSTRACT

It is projected that 3.2 million people will suffer from colorectal cancer (CRC), which is one of the most lethal cancers, by 2040. To anticipate this emergency, early screening of CRC via endoscopy or colonoscopy is unequivocally critical. In the last decade, to accurately spot gastrointestinal polyps, which may be an early sign of CRC, deep learning-based methods have been commonly used among researchers in collaboration with clinicians. Nonetheless, deep learning-based methods are oftentimes designed in an overparameterized fashion. Thus, it is nearly impossible to execute the models on low-spec devices, e.g., edge devices, in the limited electricity supply and internet access circumstances. This can be averted by utilizing deep learning-based methods that have fewer parameters or lightweight ones. Even though lightweight deep learning methods have fewer parameters, they can still suffer from high latency due to the inefficiency of the network design, such as the abusive utilization of skip connections. To address the above issues, a novel deep learning-based lightweight gastrointestinal polyp segmentation model is proposed. The proposed model, namely ConvInCEDNet, is built upon novel lightweight low-latency components: the convolution-based CEConv block and the involution-based CInvo block. Structurally, ConvInCEDNet has the form of a U-shape encoder-decoder. The encoder and the decoder contain the CEConv block and the CInvo block, respectively. The downsampling processes are carried out via max pooling, except for the first stage, which utilizes strided convolution. Transposed convolution is used for the upsampling processes. ConvInCEDNet has three variants, i.e., small, medium, and large. The small variant has an extremely tiny number of parameters, i.e., 64.81 k. On the CPU-only embedded device, i.e., the Raspberry Pi 4, the small variant can run fast up to 3.49 FPS. To increase generalization capacity, ConvInCEDNet variants employ instance normalization instead of customarily used batch normalization. To extensively test out the ability of the models to generalize, ConvInCEDNet variants are trained on the PolypGen dataset and tested on six distinct datasets: CVC-ClinicDB, CVC-ColonDB (extended), Endo-tect, ETIS-LaribPolypDB, KvasirCapsule-SEG, and Kvasir-SEG. ConvInCEDNet variants demonstrate competitive computational and segmentation performance in comparison with the prior related models.

Keywords: deep learning, lightweight model, medical image segmentation, polyp segmentation

INTISARI

Diperkirakan 3, 2 juta orang akan menderita kanker kolorektal (CRC), yang merupakan salah satu kanker paling mematikan, pada tahun 2040. Untuk mengantisipasi keadaan darurat ini, skrining dini CRC melalui endoskopi atau kolonoskopi sangatlah penting. Dalam dekade terakhir, untuk mengenali polip gastrointestinal secara akurat, yang mungkin merupakan tanda awal CRC, metode berbasis *deep learning* telah umum digunakan di kalangan peneliti yang bekerja sama dengan dokter. Meskipun demikian, metode berbasis *deep learning* seringkali dirancang dengan parameter yang sangat berlebihan. Sehingga hampir tidak mungkin untuk mengeksekusi model pada perangkat dengan spesifikasi rendah, misalnya *edge device*, dalam kondisi pasokan listrik dan akses internet yang terbatas. Hal ini dapat dicegah dengan memanfaatkan metode berbasis *deep learning* yang parameternya lebih sedikit atau ringan. Meskipun metode *deep learning* yang ringan memiliki parameter yang lebih sedikit, metode ini masih dapat mengalami latensi tinggi karena ketidakefisienan desain jaringan, seperti penggunaan *skip connection* yang sangat berlebihan. Untuk mengatasi masalah di atas, diusulkan model segmentasi polip gastrointestinal ringan berbasis *deep learning* yang baru. Model yang diusulkan, yaitu ConvInCEDNet, dibangun di atas komponen latensi rendah baru yang ringan: blok CEConv berbasis konvolusi dan blok CInvo berbasis involusi. Secara struktural, ConvInCEDNet mempunyai bentuk *encoder-decoder* berbentuk U. *Encoder* dan *decoder* masing-masing berisi blok CEConv dan blok CInvo. Proses *downsampling* dilakukan melalui *max pooling*, kecuali tahap pertama yang menggunakan *strided convolution*. *Transposed convolution* digunakan untuk proses *upsampling*. ConvInCEDNet memiliki tiga varian yaitu kecil, sedang, dan besar. Varian kecil memiliki jumlah parameter yang sangat kecil, yaitu 64, 81 ribu. Pada perangkat tertanam dengan CPU saja, yakni Raspberry Pi 4, varian kecilnya mampu berjalan cepat hingga 3, 49 FPS. Untuk meningkatkan kapasitas generalisasi, varian ConvInCEDNet menggunakan *instance normalization*, bukan *batch normalization* yang biasa digunakan. Untuk menguji kemampuan model dalam melakukan generalisasi secara ekstensif, varian ConvInCEDNet dilatih pada *dataset* PolypGen dan diuji pada enam *dataset* berbeda: CVC-ClinicDB, CVC-ColonDB (*extended*), Endotect, ETIS-LaribPolypDB, KvasirCapsule-SEG, dan Kvasir-SEG. Varian ConvInCEDNet menunjukkan kinerja komputasi dan segmentasi yang kompetitif dibandingkan dengan model terkait sebelumnya.

Kata kunci: *deep learning*, model ringan, segmentasi citra medis, segmentasi polip