MingleNet: A Novel Dual Stacking Approach for Medical Image Segmentation

姜竣維、詹堡崴、許漢強、何天勝、曾漢盛

指導教授:張光遠

NTOU CSE

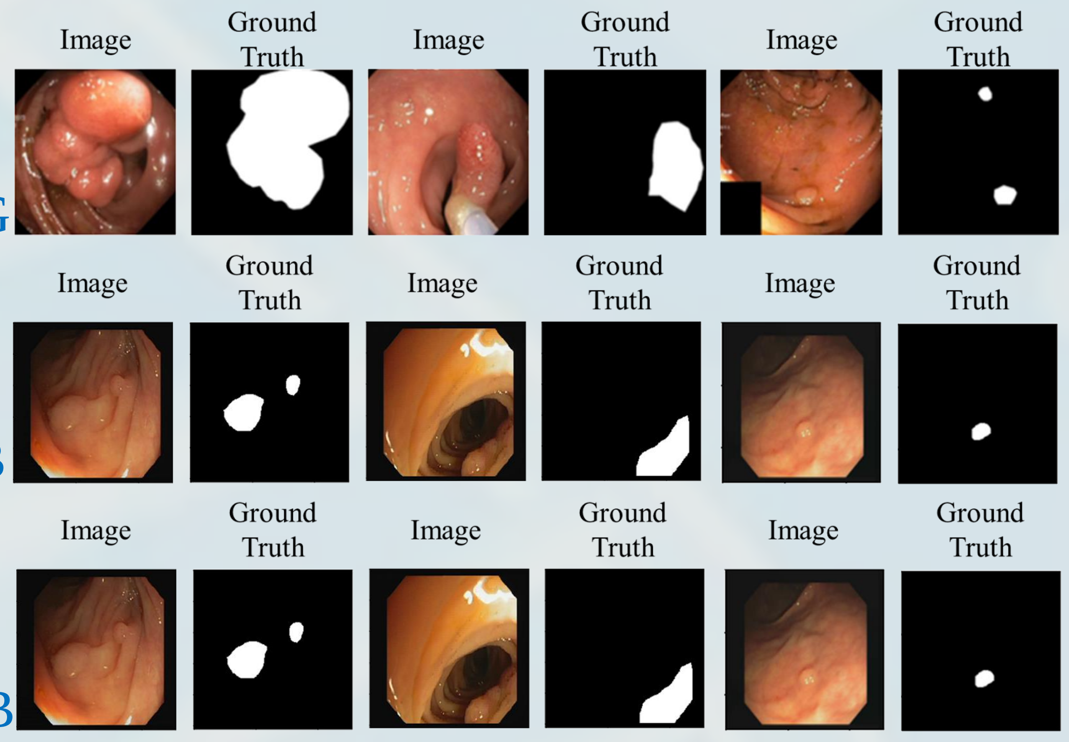
Department of Computer Science and Engineering

Abstract

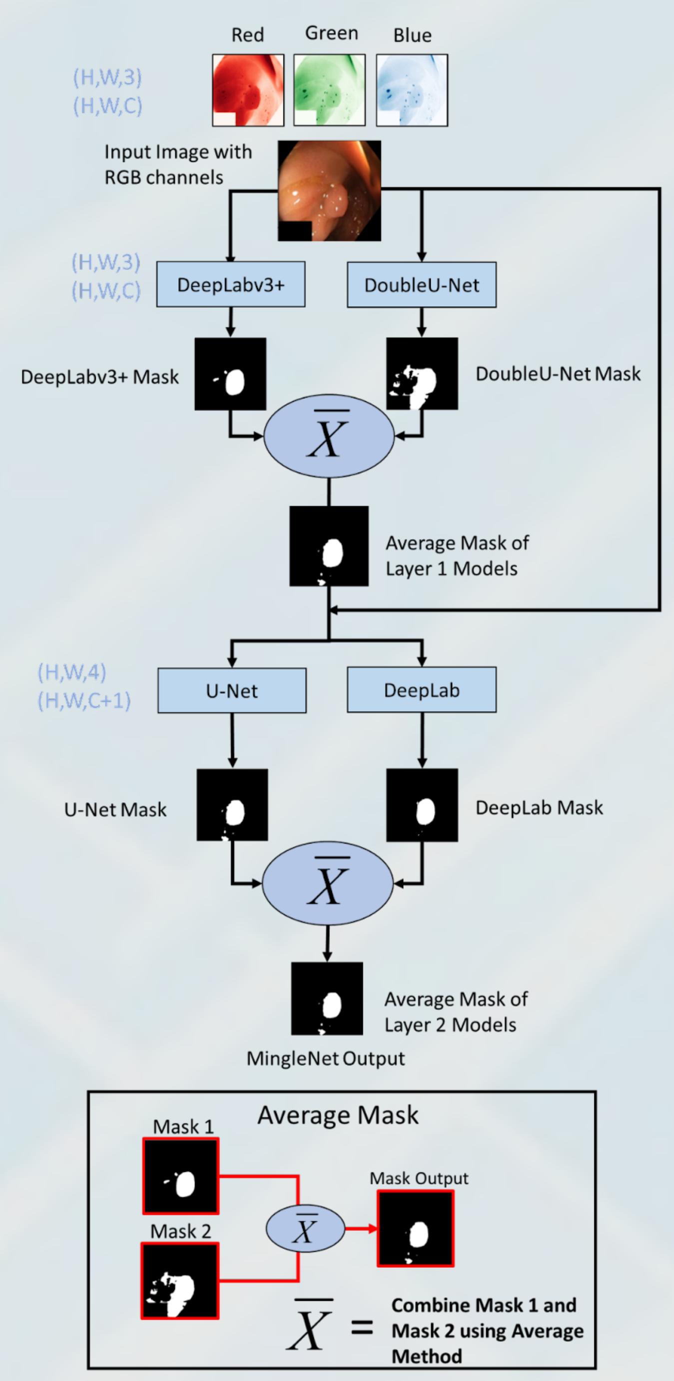
Medical image segmentation is important for disease diagnosis and treatment planning. Ensemble learning, which combines multiple models or predictions, can improve accuracy and performance in medical image segmentation. We propose MingleNet, which uses multiple layers of ensemble learning. MingleNet uses double-stacking of models, such as DoubleU-Net, DeepLabv3+, U-Net, and DeepLab, to produce masks. The first layer's masks are averaged and concatenated with the original images for the second layer. We also apply dynamic data augmentation to enhance model performances.

We evaluate MingleNet on polyp segmentation benchmark datasets: Kvasir-SEG, CVC-ClinicDB, and CVC-ColonDB. On Kvasir-SEG, MingleNet achieves 93.19% Dice, 87.24% IoU, 94.15% precision, 92.25% recall, and 97.87% accuracy. On CVC-ClinicDB, MingleNet achieves 95.99% Dice, 92.29% IoU, 96.08% precision, 95.90% recall, and 99.21% accuracy. On CVC-ColonDB, MingleNet achieves 94.33% Dice, 89.28% IoU, 95.89% precision, 92.83% recall, and 99.10% accuracy. Our proposed method demonstrated competitive performance across the Kvasir-SEG, CVC-ClinicDB, and CVC-ColonDB benchmarks, MingleNet ranks 1st in Dice and IoU. Moreover, MingleNet ranks 8th in Dice and 17th in IoU on the Kvasir-SEG benchmark.

Polyp Image Datasets

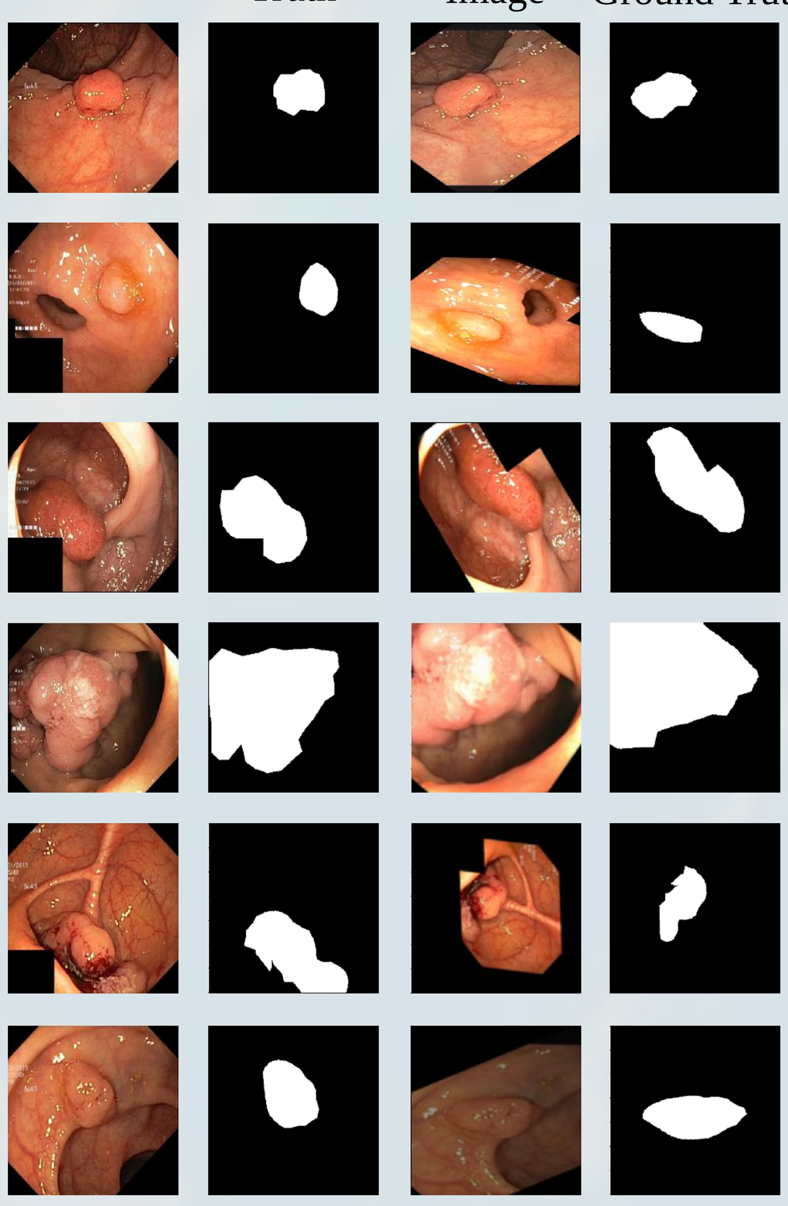


MingleNet Architecture



MingleNet is a convolutional neural network (CNN) architecture that fuses the incorporation of multiple layers of model as an input. Afterward, we average the output masks of the second layer as the final output.

Dynamic Data Augmentation :

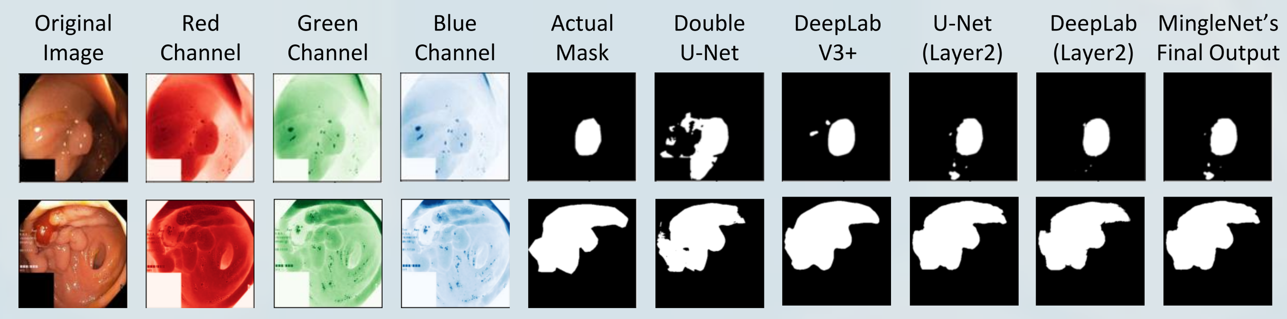


Generate new augmented images before each training epoch. This technique improved the model's generalization.

Performance of MingleNet on the benchmarks

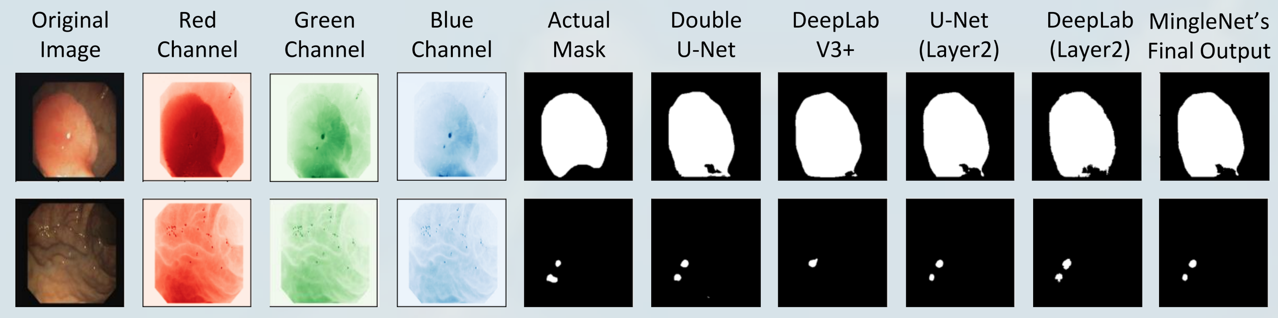
Kvasir-SEG dataset

Model	Dice	IoU	Precision	Recall	Accuracy
DoubleU-Net (Layer 1)	92.58%	86.18%	92.28%	**92.87%**	97.64%
DeepLabv3+ (Layer 1)	91.86%	84.94%	93.76%	90.03%	97.48%
U-Net (Layer 2)	93.13%	87.14%	93.92%	92.35%	97.84%
DeepLab (Layer 2)	92.69%	86.37%	**95.12%**	90.37%	97.74%
MingleNet's Final Output	**93.19%**	**87.24%**	94.15%	92.25%	**97.87%**



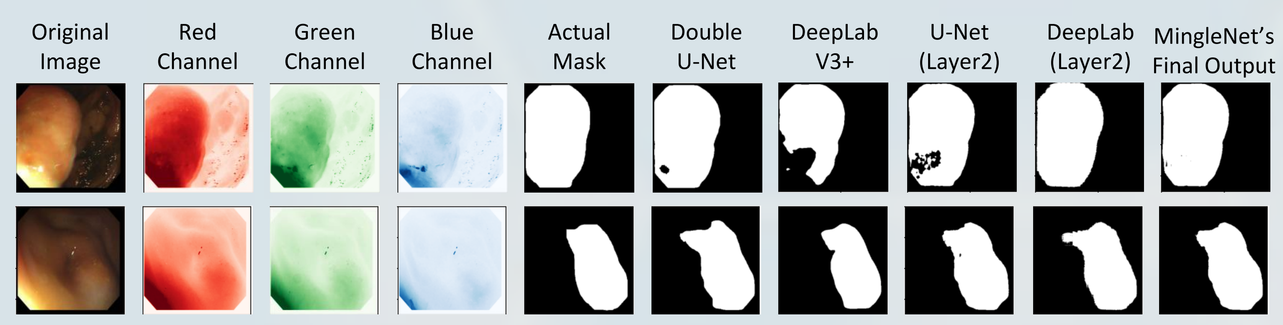
CVC-ClinicDB dataset

Model	Dice	IoU	Precision	Recall	Accuracy
DoubleU-Net (Layer 1)	95.89%	92.11%	95.61%	96.17%	99.18%
DeepLabv3+ (Layer 1)	95.19%	90.83%	95.72%	94.67%	99.05%
U-Net (Layer 2)	95.94%	92.20%	**96.17%**	95.71%	99.20%
DeepLab (Layer 2)	95.56%	91.51%	94.69%	**96.46%**	99.11%
MingleNet's Final Output	**95.99%**	**92.29%**	96.08%	95.90%	**99.21%**



CVC-ColonDB dataset

Model	Dice	IoU	Precision	Recall	Accuracy
DoubleU-Net (Layer 1)	93.74%	88.21%	94.94%	92.56%	99.00%
DeepLabv3+ (Layer 1)	90.37%	82.44%	**96.29%**	85.14%	98.53%
U-Net (Layer 2)	93.60%	87.97%	95.88%	91.42%	98.99%
DeepLab (Layer 2)	93.91%	88.52%	94.74%	**93.08%**	99.02%
MingleNet's Final Output	**94.33%**	**89.28%**	95.89%	92.83%	**99.10%**



MingleNet performance compared to other Deep Learning Models

Kvasir-SEG

MingleNet ranks **8th** in Dice (93.19%) and **17th** in IoU (87.24%) on the Kvasir-SEG benchmark.*

Model	Dice	IoU	Precision	Recall	Accuracy
U-Net (without our augmentation)	83.59%	74.28%	86.45%	76.14%	93.94%
U-Net (with our augmentation)	86.73%	76.57%	90.04%	83.65%	95.95%
FCN-Transformer(pre-trained)	92.20%	85.54%	92.38%	92.03%	97.49%
OURS MingleNet (with our augmentation)	**93.19%**	**87.24%**	**94.15%**	**92.25%**	**97.87%**

*https://paperswithcode.com/sota/medical-image-segmentation-on-kvasir-seg

CVC-ClinicDB

MingleNet ranks **1st** in both Dice (95.99%) and IoU (92.29%) on the CVC-ClinicDB benchmark.**

Model	Dice	IoU	Precision	Recall	Accuracy
U-Net (without our augmentation)	91.06%	81.51%	95.26%	87.22%	97.77%
U-Net (with our augmentation)	91.42%	84.20%	93.29%	89.63%	98.34%
FCN-Transformer(pre-trained)	88.00%	78.58%	**96.59%**	80.82%	96.45%
OURS MingleNet (with our augmentation)	**95.99%**	**92.29%**	96.08%	**95.90%**	**99.21%**

**https://paperswithcode.com/sota/medical-image-segmentation-on-cvc-clinicdb

CVC-ColonDB

MingleNet ranks **1st** in both Dice (94.34%) and IoU (89.28%) on the CVC-ColonDB benchmark.***

Model	Dice	IoU	Precision	Recall	Accuracy
U-Net (without our augmentation)	86.50%	76.2%	94.34%	79.94%	97.99%
U-Net (with our augmentation)	82.12%	69.67%	92.12%	74.09%	97.39%
FCN-Transformer(pre-trained)	90.73%	83.04%	91.07%	90.40%	98.99%
OURS MingleNet (with our augmentation)	**94.34%**	**89.28%**	95.89%	**92.84%**	**99.10%**

***https://paperswithcode.com/sota/medical-image-segmentation-on-cvc-colondb