Building An Automated Module for Image Quality Assessing from Narrow-Banding-Imaging Endoscopy Cameras

Students

Pham Khac Long

Nguyen Thuan Thanh

Nguyen The Anh

Supervisor

Bui Van Hieu

FPT University - Hoa Lac Campus

Table of content

INTRODUCTION

Worldwide

~ 40%

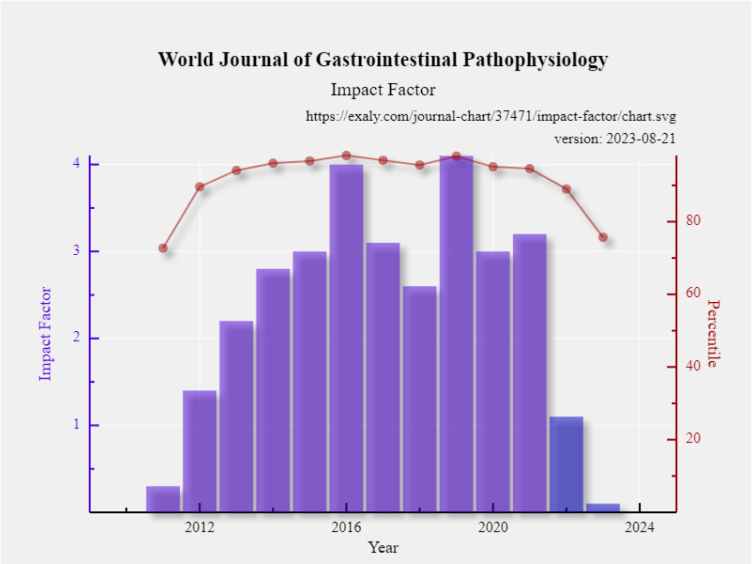
of the population globally affected by gastrointestinal diseases [1]

5th

most diagnosed cancer worldwide [2]

4th

stomach cancer ranked as leading cause of death related to cancer in 2020 [2]



Vietnam

70%

of Vietnamese population is at risk of gastrointestinal diseases [3]

7000/14000

deaths from colorectal cancer in Vietnam [3]

3th

ranked of gastric cancer in the country statistic of most common cancer [3]

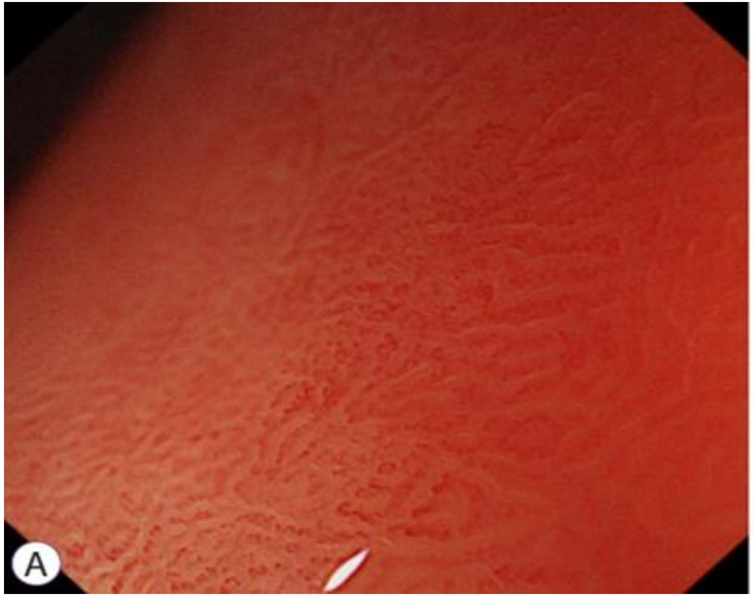
→ **Diagnosis is very important in early detection and treatment !**

Narrow-Banding-Imaging (NBI)

Traditional Endoscopy: White light Endoscopy

- WLE uses light source similarly ordinary daylight
- White light does not penetrate the mucosal layer
- Only detect lesions that have form in surface

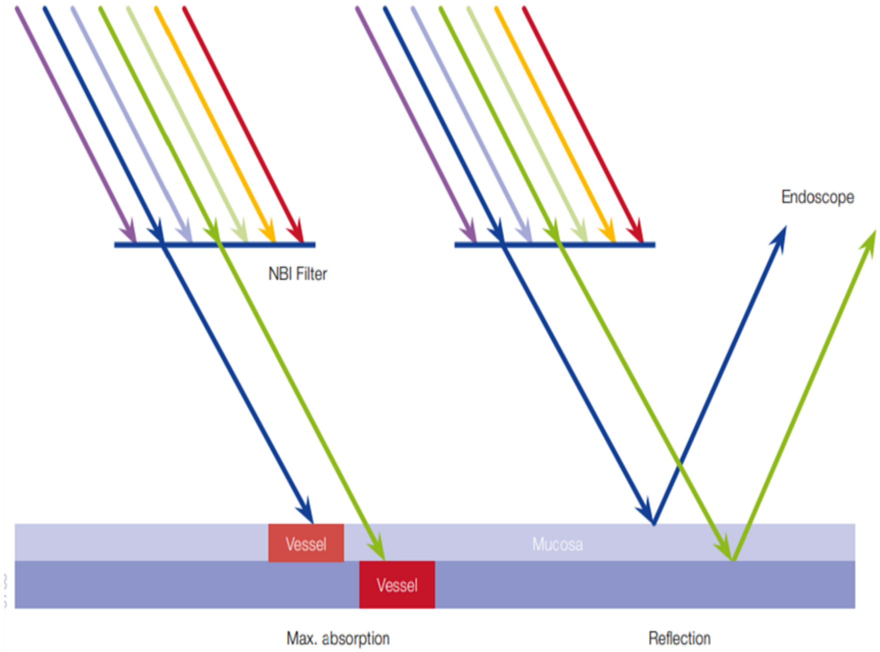
→ Bad result in early detection
Accuracy depends on doctor [4]



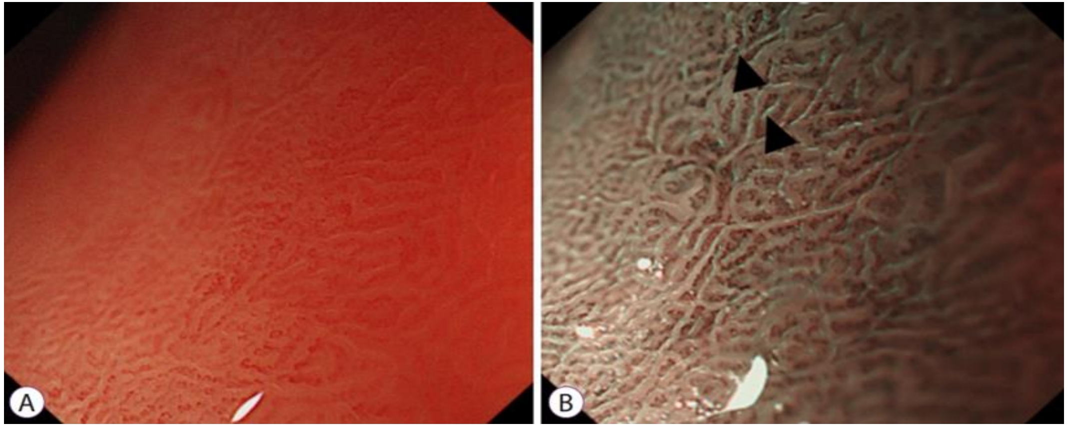
Narrow-Banding-Imaging (NBI)

NBI Endoscopy

- Improves the contrast between capillaries and submucosal vessels by manipulating the light source through specialized filters
- Specific wavelengths (415 nm - blue and 540 nm – green) [5,6]



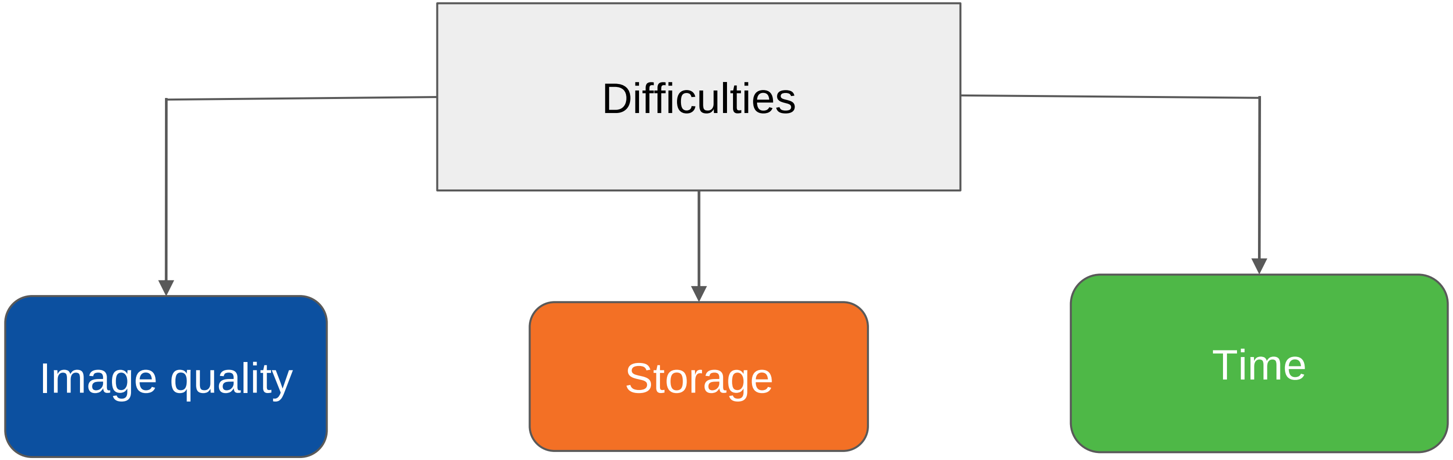
Compare NBI and WLE



(A) White light endoscopy image

(B) NBI endoscopy image

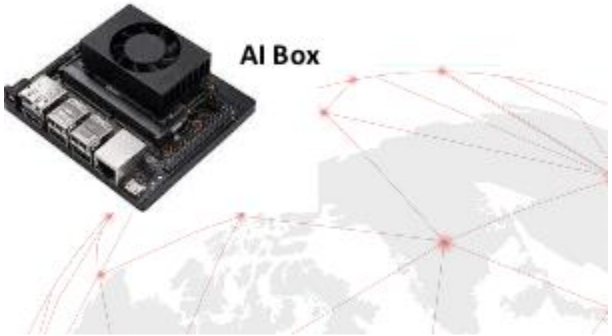
AI in Medical Image Processing



→ **AI plays crucial role in image quality in Medical field**

Our module

❖ A module belongs to AI smart endoscopic system from NBI camera of Viettel Cyberspace (VTCC) with 3 main modules:
   ➢ Automated image quality assessment
   ➢ Automated detect stomach damage
   ➢ Super resolution



RELATED WORK

Full-Reference Image Quality Assessment (FR-IQA)

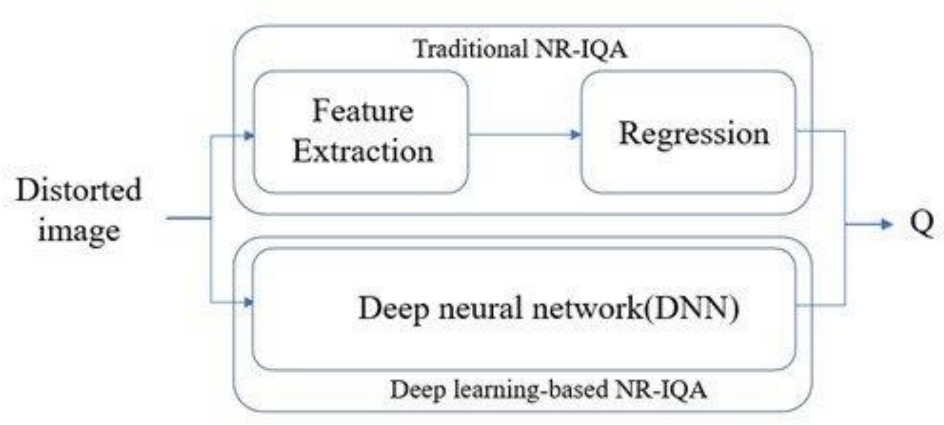
- FR-IQA methods require two types of input: distorted and reference images to estimate their perceptual similarity. [7,8]
- **Non-reference Image Quality Assessment more useful. [9]**

Example of FR-IQA

Non-Reference Image Quality Assessment (NR-IQA)

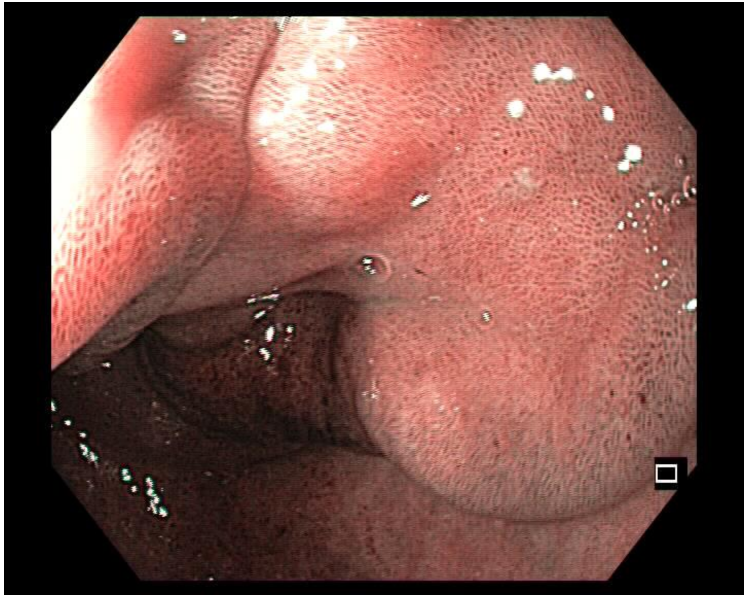
- Providing a solution when a reference image is not available. [10,11]
- In endoscopic domain, the quality in different image regions are very different. [12]

→ **Inputting the whole image into IQA model may not be optimal.**



Example of NR-IQA

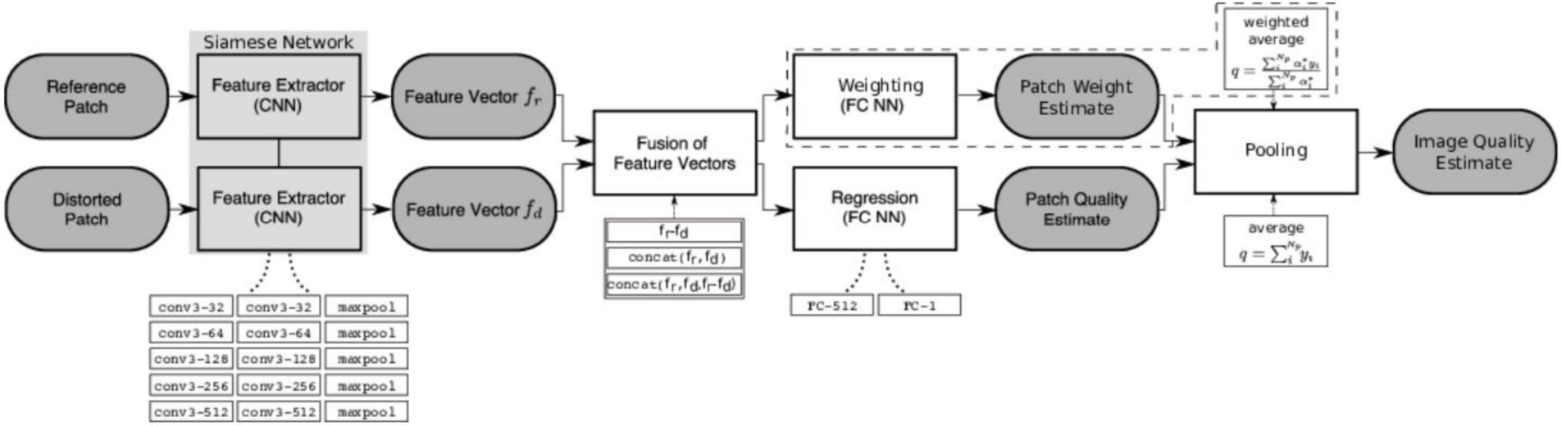
Non-Reference Image Quality Assessment (NR-IQA)



Describing uneven image quality across an endoscopic image.

Patch-based classification

- Dividing the image randomly or consecutively into small patches [13,14]
- Aggregate the features of those patches.



Example of patch-based classification

Contribution

- Two-stage NR-IQA framework for quality image from NBI endoscopy cameras:
  - **First stage**: Patch-based classification model extract from multi-layer features of Convolutional Neural Network (CNN).
  - **Second stage**: Aggregation process based on statistical method.
- Using **Feature magnitude loss** [15] in endoscopy IQA to clearly classification patch quality.
- **Inference speed improvement** method.

METHODOLOGY

General idea

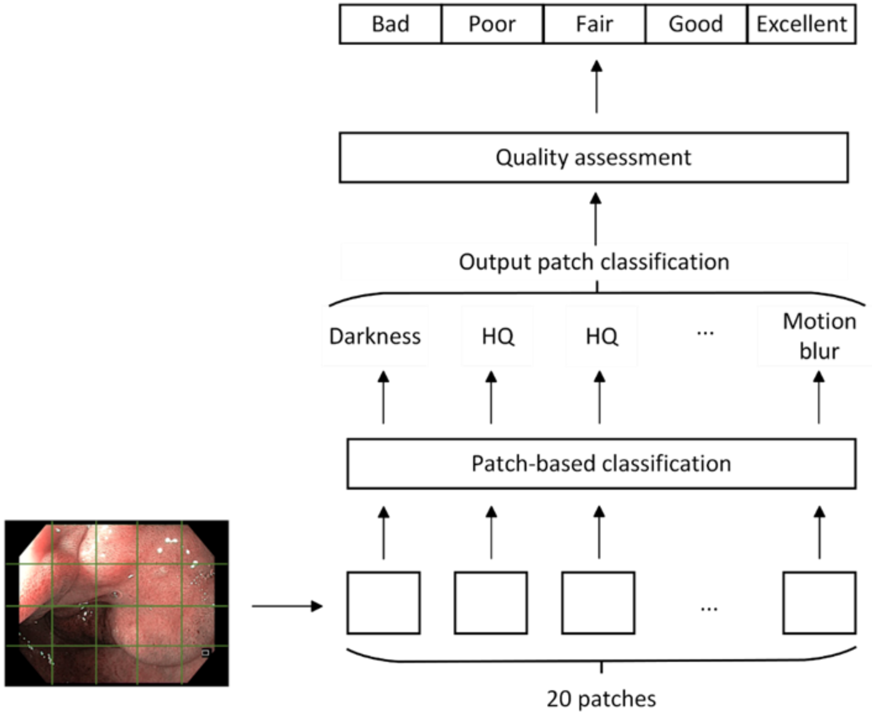
Patch-based classification

- Size of image: 512x640
- Size of patch: 128x128

→ **20 consecutive image patches**

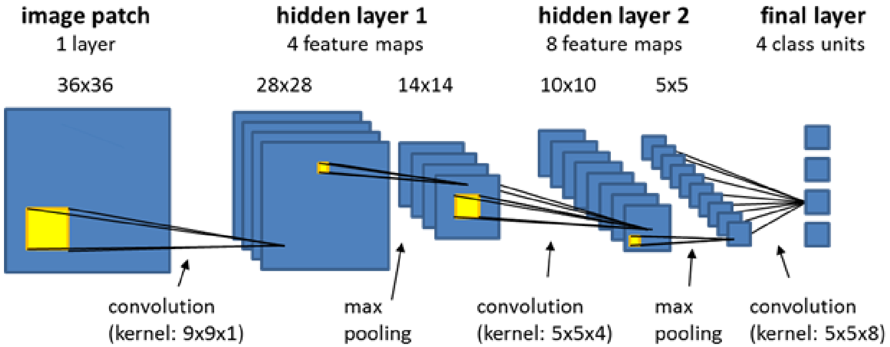
Quality assessment

- Aggregation process
- Statistical method



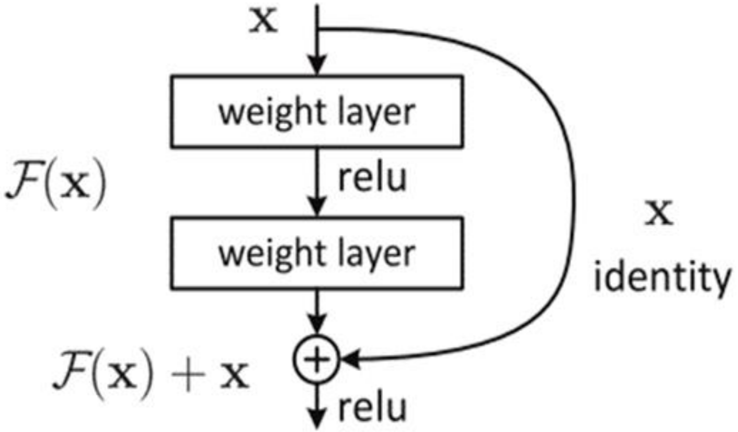
The diagram of the Endoscope Image Quality Assessment model

Resnet



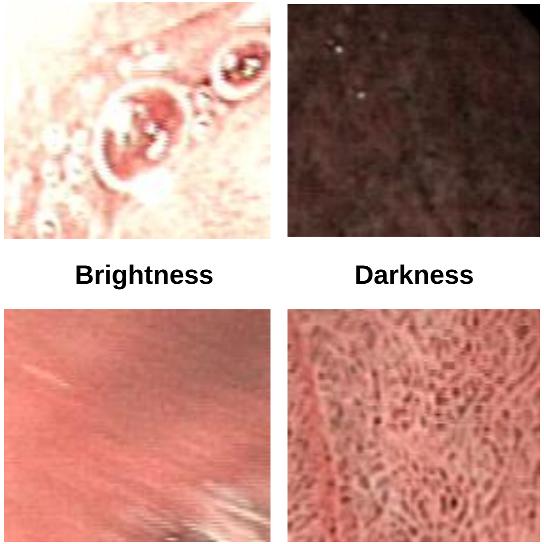
- Image classification, object detection, segmentation
- Hierarchical representations

→ **Using Resnet 18 with several changes in architecture**



Patch-based classification model

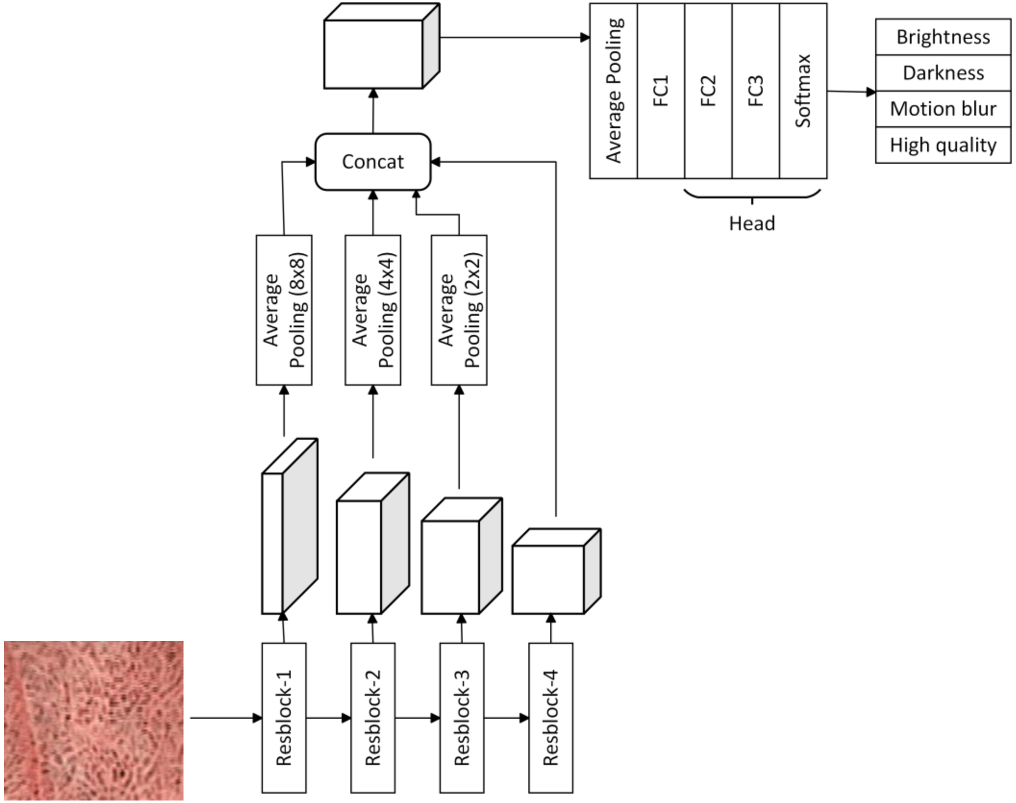
- **Brightness:** The light reflection of gastric juice
- **Darkness:** The light cannot be evenly distributed
- **Motion blur:** The relative motion of camera **(Accounts for a very large proportion)**
- **High-quality:** Sharp areas, relative camera motion and stomach surface is low



Brightness **Darkness**

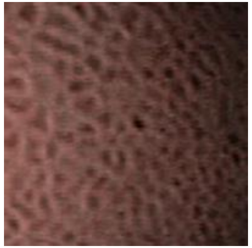
Motion blur **High-quality**

Patch-based classification model

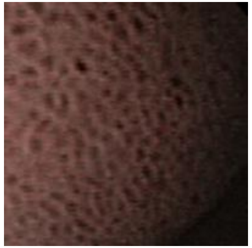


The improved AI model architecture for endoscope image patch quality classification

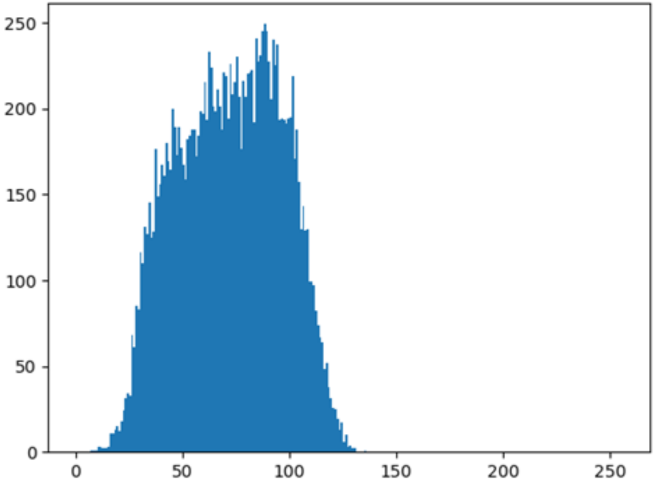
Patch-based classification model



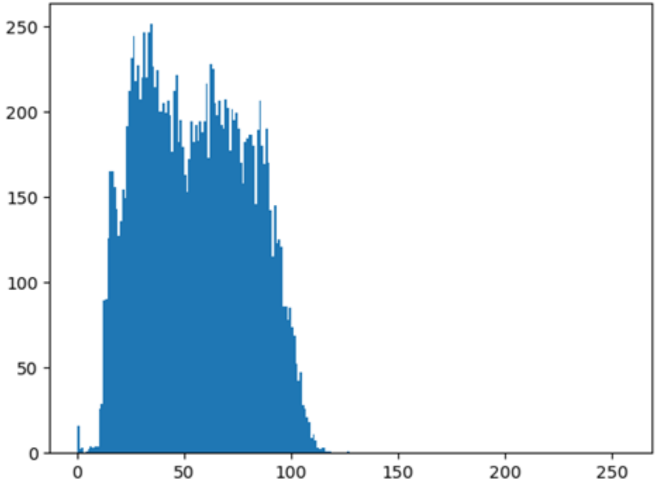
High-quality patch



Darkness patch



The histogram of high-quality patch



The histogram of darkness patch

Feature magnitude loss function

Given the entire image $I = \{(P_i,\ y_i)\}_{i=1}^{N}$

The proposed network $r_{\theta,\emptyset} = f_{\emptyset}(s_{\theta}(P_i))$ $\longrightarrow$ N-dimensional feature $[0,1]^T$

The end-to-end model is training with the total loss:

$$l = min_{\theta,\emptyset} \sum_{i,j=1}^{N} (1-\alpha) l_s\left(s_{\theta}(P_i), s_{\theta}(P_j), y_i, y_j\right) + \alpha l_f\left(f_{\emptyset}\left(s_{\theta}(P_i)\right), y_i\right)$$

Where: $s_{\theta}: P \rightarrow X$ is the patch feature extractor
$f_{\emptyset}: X \rightarrow [0,1]^T$ is the patch classifier
α: weight for each term
$l_s(.)$: loss function that maximizes the separability between darkness and high-quality
l_f: loss function to train the patch classifier

Feature magnitude loss function

The feature magnitude loss function can be further defined as:

$$l_s(s_\theta(P_i), s_\theta(P_j), y_i, y_j) = \max\left(0, m - d\left(g_\theta(X_i), g_\theta(X_j)\right)\right)$$

$$if\ y_i, y_j \in \{Darkness, High - quality\}$$

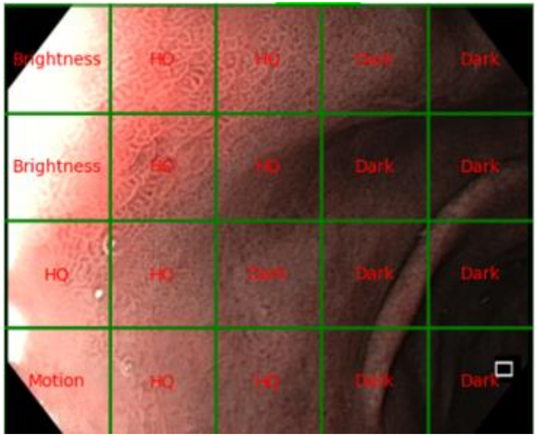
Where: m is pre-defined margin

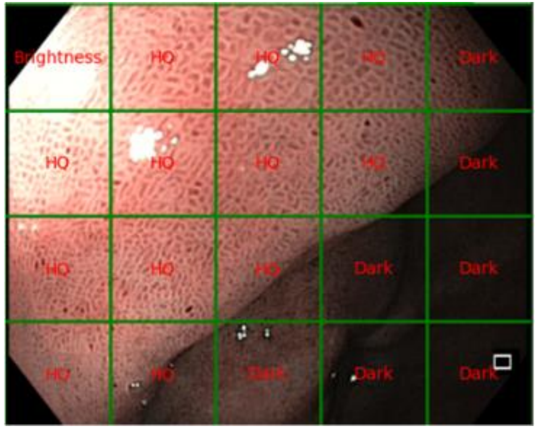
$X_i = s_\theta(P_i)$ is the patch feature

g_θ calculates the feature magnitude of the patch feature

d represents separability function that computes the difference between two feature magnitudes

Quality assessment





The high-quality patches are located **adjacent horizontally or vertically**

→ **Apply Breadth-first search (BFS) algorithm**

$$n = \frac{N}{total\ number\ of\ patches\ per\ image},$$

n: the percentage of adjacent high-quality patches
N: total number of adjacent high-quality patches

Quality assessment

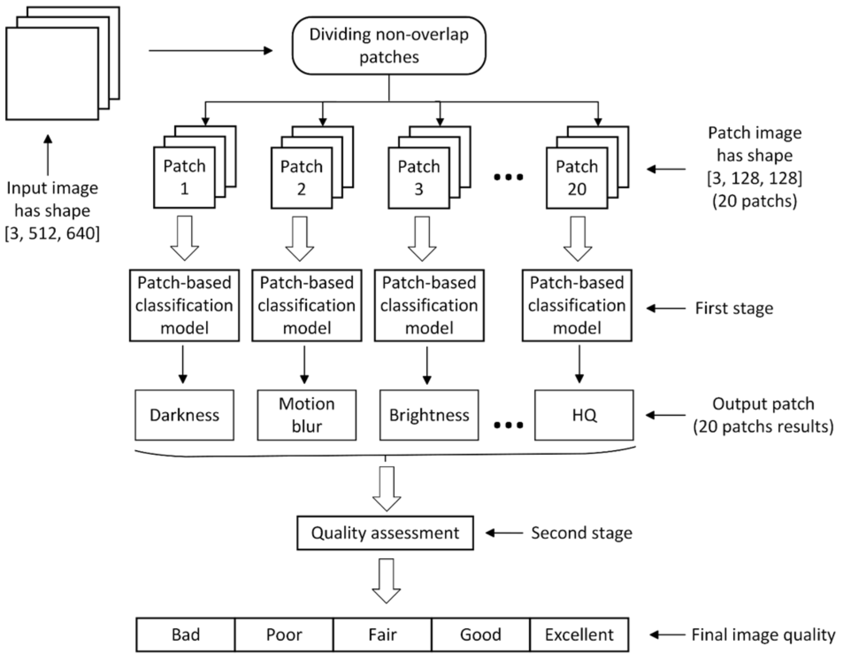
Brightness	Darkness	Motion blur	High-quality (adjacent patches)
a%	b%	c%	n%

The percentage of each patch type in an image

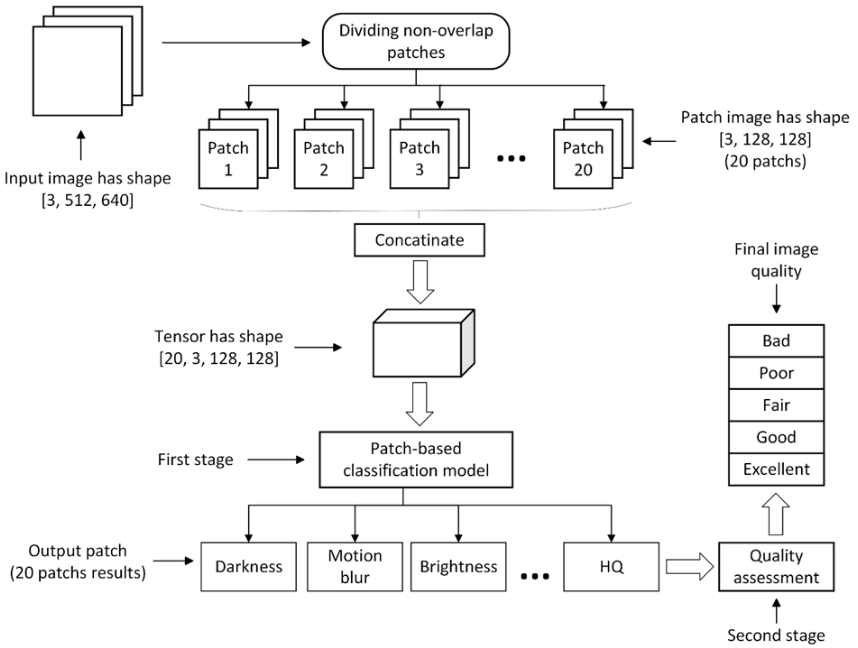
Bad	Poor	Fair	Good	Excellent
$c > 45\%$	$35\% \leq c \leq 45\%$	$c < 35\%$ $n < 35\%$	$35\% \leq n \leq 55\%$ $c < 35\%$	$n > 55\%$ $c < 35\%$

Quality assessment for the entire frame

Inference process



The original way of implementing the inference pipeline.

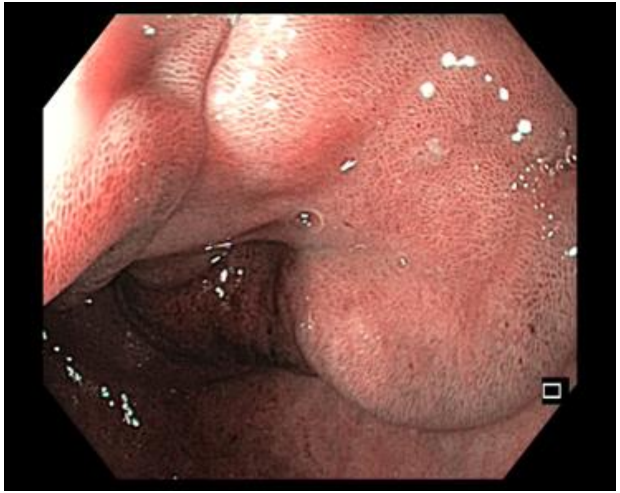


The proposed way of executing the inference process.

IMPLEMENTATION DETAIL

Dataset

- Private dataset from the medical image database of Viettel Cyberspace Center.
- Endoscopic images extracted from specialized NBI cameras in real-world endoscopy cases
- Original size of 720 x 576 (width x height)



Example of image with black border surrounded

Data Preparation

- Cut off most of the surrounding black border
- Size of images for training or testing process became 640x512.

Endoscopic image after cut off most of the black border.

Data Preparation

The original number of 4 categories:

	Brightness	Darkness	High-quality	Motion blur
Train	205	592	3278	918
Val	107	212	699	639

Dataset used for training patches-based classification model

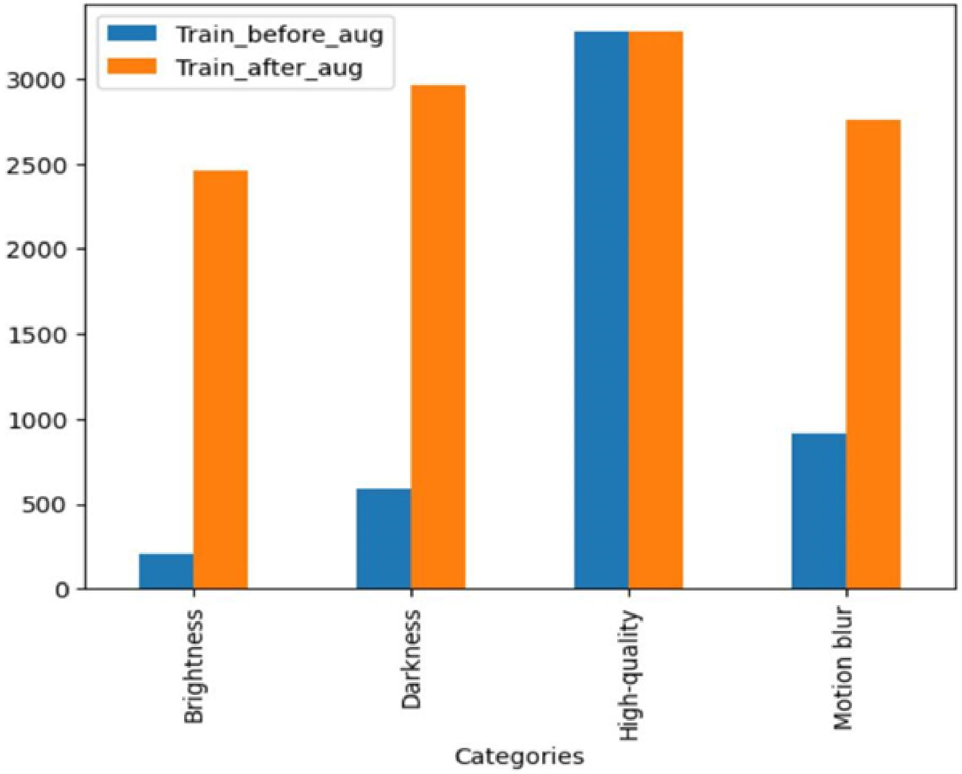
The number of 4 categories after applied data augmentation techniques:

	Brightness(x12)	Darkness(x5)	High-quality	Motion blur(x3)
Train	2460	2960	3278	2754
Val	107	212	699	639

Dataset after using the data augmentation techniques

Data Preparation

Datasetset is much more balanced compared to the original



Quantity of training dataset before and after utilizing augmentation techniques

EXPERIMENTAL RESULT

Results

Experimental results of baseline patch-based classification baseline model

Class	Precision	Recall	F1-score	Accuracy
Brightness	88.23	98.13	92.92	95.65
Darkness	87.61	93.40	90.41	
Motion blur	97.53	98.75	98.13	
High-quality	97.89	93.13	95.45	

Experimental results of improved patch-based classification model

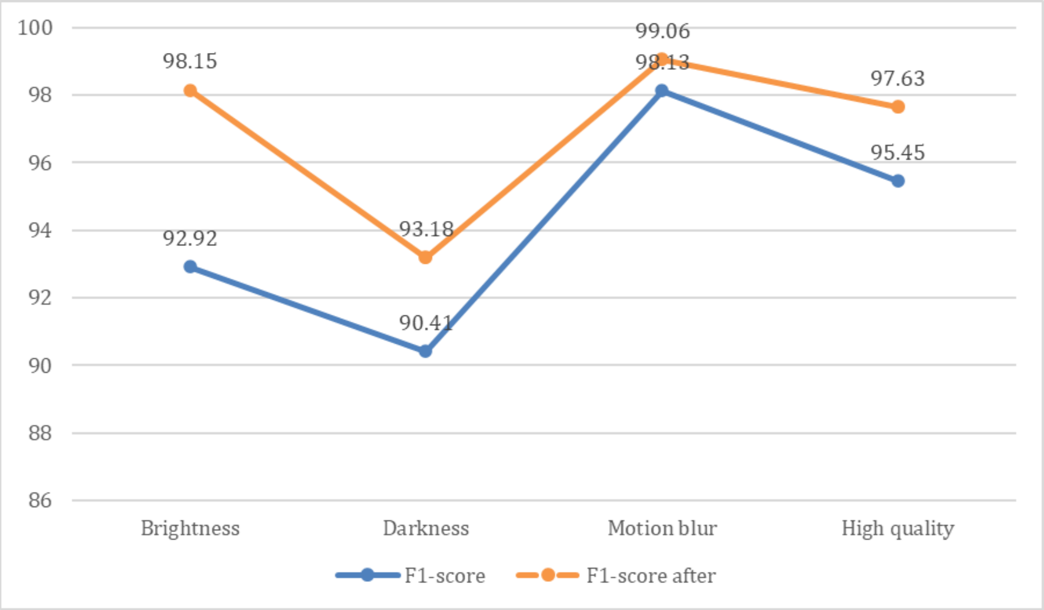
Class	Precision	Recall	F1-score	Accuracy
Brightness	97.25	99.07	98.15	97.7
Darkness	92.96	93.40	93.18	
Motion blur	99.06	99.06	99.06	
High-quality	97.84	97.42	97.63	

Results

By using:

- Data augmentation techniques
- Multi-feature fusion strategy
- Feature magnitude learning

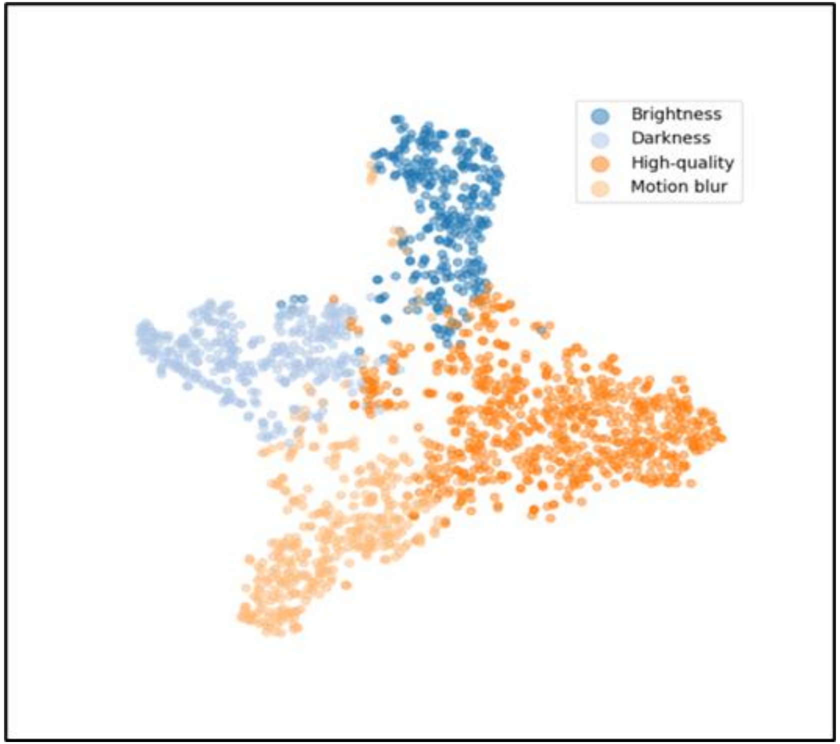
→ Results can be improved significantly



Comparison F1-score between baseline and improved model

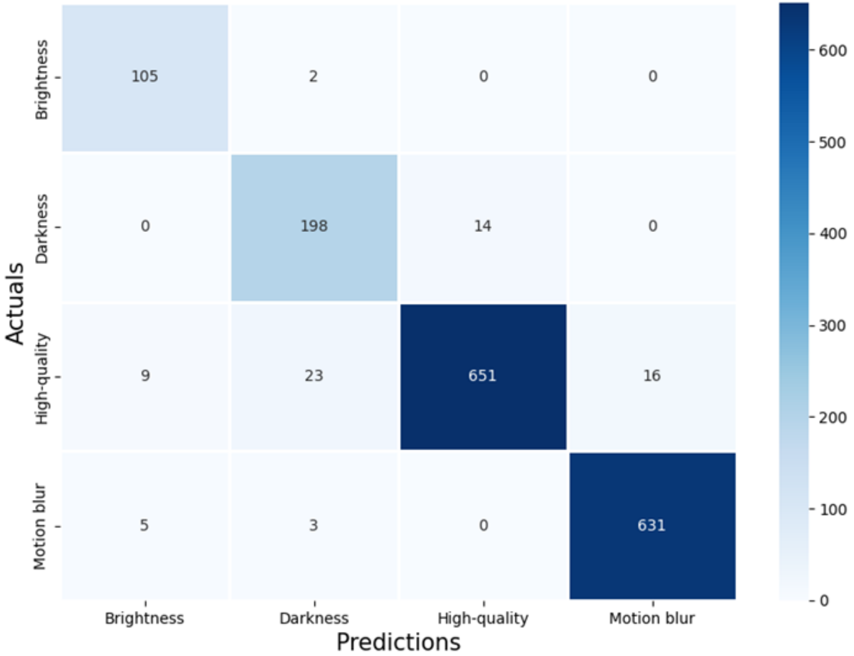
Results

t-SNE of baseline patch-based classification baseline model

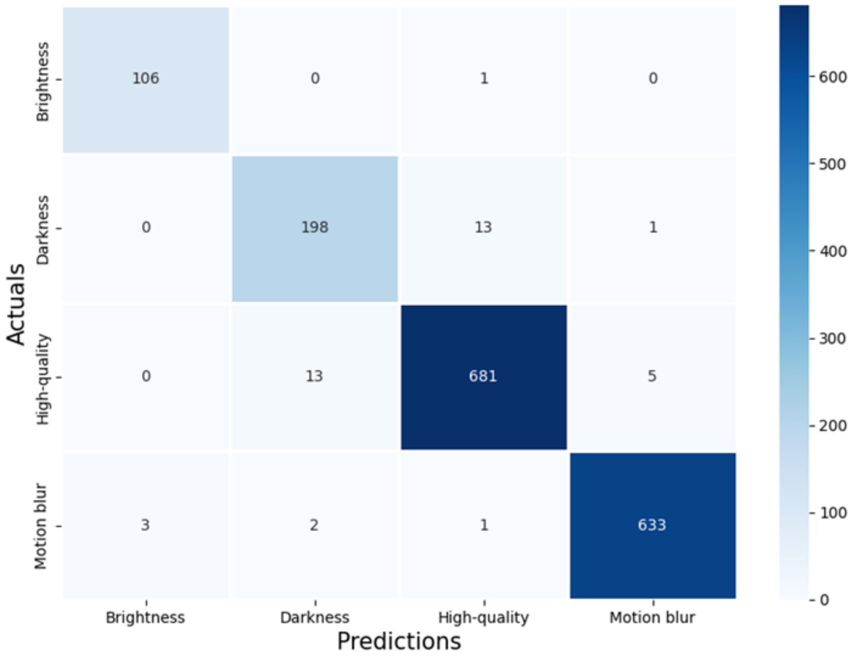


t-SNE of improved patch-based classification improved model

Results



Confusion matrix of baseline patch-based classification model



Confusion matrix of improved patch-based classification model

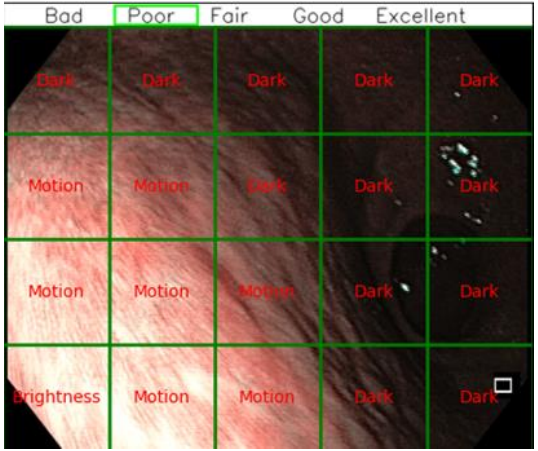
Results in practical environment

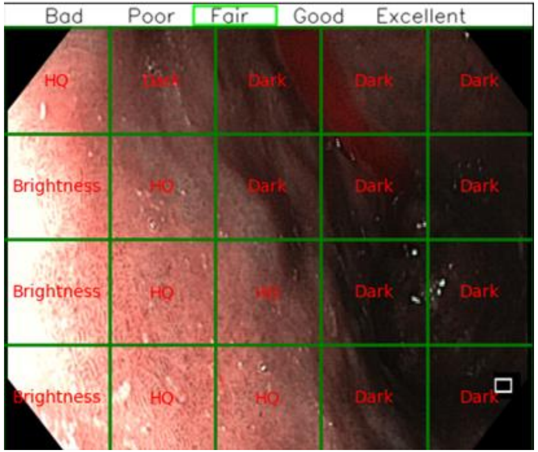
Bad and Poor:

- Most of the patches are motion blur

Fair:

- Majority are brightness and darkness patches
- Small number of adjacent high-quality patches

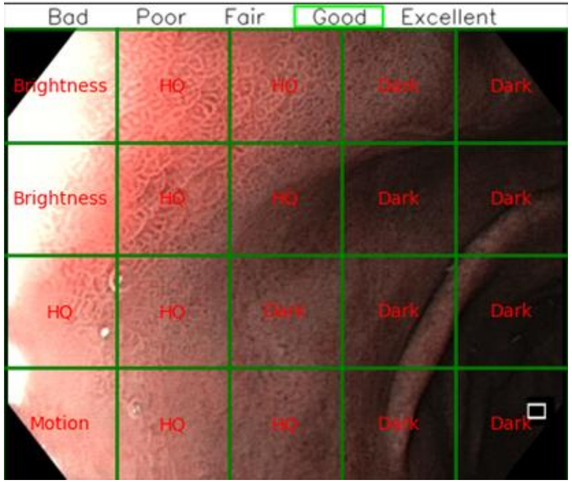


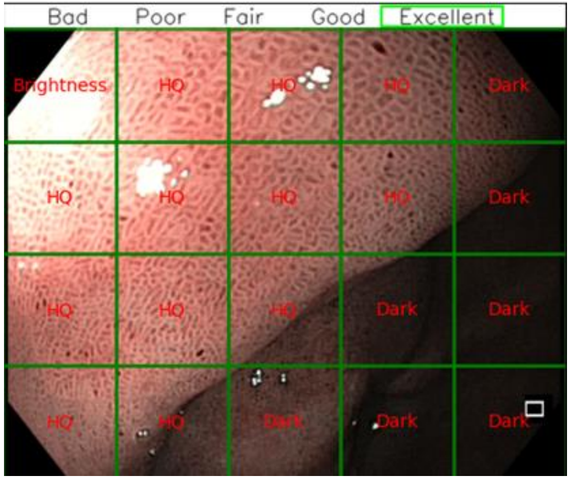


Results in practical environment

Good and Excellent:

- Motion blur patches is almost zero
- High-quality patches is the majority





Inference process result

Results measurement of two ways implementing inference process on GPU NVIDIA QUADRO RTX 4000:

Version	Time processing (mean ± std)	Frames per second (FPS)
Original	0.0848 ± 0.00149	12 FPS
Proposed	**0.0212 ± 0.00138**	**48 FPS**

Time measurement of the original and the proposed way

Storage efficiency

Storage efficiency of the IQA module with F, G, E are the abbreviated image quality levels for Fair, Good and Excellent respectively:

Video	Length	Size (mb)	Total number of frames	Quality threshold	Number of extracted frames	Amount of storage saved (%)
14-2-2018Sequence_15-14-3-228 original.avi	1m29s	123.6 mb	2244	F, G, E	950	57.66%
				G, E	625	72.15%
				E	451	79.90%
Azoulay 28032018.mp4	1m04s	40.8 mb	1623	F, G, E	298	81.64%
				G, E	237	85.40%
				E	164	89.90%

Effectiveness of the IQA module

CONCLUSION AND FUTURE WORKS

Conclusion

Based on multi-layer features of CNN, patch-based classification and feature magnitude loss, we achieve:

- Nearly 98% overall accuracy
- 48 FPS, faster 4 times
- Highly appreciated by medical professionals from Viet Duc and K Tan Trieu hospitals
- Storage saved nearly 90%

Future works

- Finding a more effective dividing patches strategy
- Consult with more experts

REFERENCES

[1] Aziz I, Palsson OS, Tornblom H, et al. The prevalence and impact of overlapping rome iv-diagnosed functional gastrointestinal disorders on somatization, quality of life, and healthcare utilization: a cross-sectional general population study in three countries. Am J Gastroenterol 2018;113:86–96.

[2] Sung, Hyuna, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries, 4 Feb. 2021.

[3] Le Nga, “Medical experts identify the gut as the most vulnerable organ for Vietnamese”, VnExpress, 2019

[4] Dinis-Ribeiro M, da Costa-Pereira A, Lopes C, LaraSantos L, Guilherme M, Moreira-Dias L, Lomba-Viana H, Ribeiro A, Santos C, Soares J, Mesquita N, Silva R, Lomba-Viana R. Magnification chromoendoscopy for the diagnosis of gastric intestinal metaplasia and dysplasia. Gastrointest Endosc 2003.

[5] “Achieve More with NBI", Olympus, 4 May 2017

[6] Gono, Kazuhiro; Obi, Takashi; Yamaguchi, Masahiro; Ohyama, Nagasaki; Machida, Hirohisa; Sano, Yasushi; Yoshida, Shigeaki; Hamamoto, Yasuo; Endo, Takao (May 1, 2004). "Appearance of enhanced tissue features in narrow-band endoscopic imaging". Journal of Biomedical Optics. 9 (3): 568–577.

[7] Sewoong Ahn, Yeji Choi, and Kwangjin Yoon. Deep learning-based distortion sensitivity prediction for full reference image quality assessment. In IEEE Conference on Computer Vision and Pattern Recognition Workshops, pages 344–353, 2021.

[8] Keyan Ding, Kede Ma, Shiqi Wang, and Eero Simoncelli. Image quality assessment: Unifying structure and texture similarity. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020.

REFERENCES

[9] Wang, Z.: Objective image quality assessment: facing the real-world challenges. In: IS&T International Symposium on Electronic Imaging, pp. 1–6 (2016).

[10] S. Su, Q. Yan, Y. Zhu, C. Zhang, X. Ge, J. Sun, and Y. Zhang, “Blindly assess image quality in the wild guided by a self-adaptive hyper network,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3667– 3676, 2020.

[11] Yunhong Li, Huanhuan Zhang, Jinni Chen, Peng Song, Jie Ren, QiuMing Zhang, KaiLi Jia. “Non-reference image quality assessment based on deep clustering”. Signal Processing: Image Communication, Volume 83, 2020.

[12] Wei Tan, Chao Xu et al., An Endoscope Image Enhancement Algorithm Based on Image Decomposition, Electronics 2022, https://doi.org/10.3390/electronics11121909

[13] Xue Z, Angara S, Guo P, Rajaraman S, Jeronimo J, Rodriguez AC, Alfaro K, Charoenkwan K, Mungo C, Domgue JF, Wentzensen N, Desai KT, Ajenifuja KO, Wikström E, Befano B, de Sanjosé S, Schiffman M, Antani S. Image Quality Classification for Automated Visual Evaluation of Cervical Precancer. Milland Workshop (2022). 2022 Sep; 13559:206-217.

[14] Wang Y., Song Y., Wang F., Sun J., Gao X., Han Z., Shi L., Shao G., Fan M., Yang G. A two-step automated quality assessment for liver MR images based on convolutional neural network. Eur. J. Radiol. 2020; 124:108822. doi: 10.1016/j.ejrad.2020.108822

[15] Y. Tian, G. Pang, Y. Chen, R. Singh, J. W. Verjans and G. Carneiro, "Weakly-supervised Video Anomaly Detection with Robust Temporal Feature Magnitude Learning," 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 2021, pp. 4955-4966, doi: 10.1109/ICCV48922.2021.00493.

THANKS YOU FOR LISTENING