Research Challenges for Augmenting Endoscopy Image Datasets using Image Combination Methodologies

Alexandre Neto[1,2]
alexandre.h.neto@inesctec.pt
Diogo Libânio[3]
diogo.monteiro@ipoporto.min-saude.pt
Mário Dinis Ribeiro[3]
mdinisribeiro@gmail.com
Miguel Coimbra[2,4]
mcoimbra@fc.up.pt
António Cunha[1,2]
acunha@utad.pt

[1]Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal
[2]Instituto de Engenharia de Sistemas e Computadores, Tecnologia e Ciência, Porto, Portugal
[3]Faculdade de Medicina, Universidade do Porto, Porto, Portugal
[4]Faculdade de Ciências, Universidade do Porto, Porto, Portugal

Abstract

Metaplasia detection in upper gastrointestinal endoscopy is crucial to identify patients at higher risk of gastric cancer. Deep learning algorithms can be useful for detecting and localising these lesions during an endoscopy exam. However, a lot of annotated data is needed to train these models, which can be a problem in the medical field. To overcome this, data augmentation techniques are commonly applied to increase the dataset's variability but must be adapted to the specificities of the application scenario. In this study, we discuss the potential benefits and identify four key research challenges of a promising data augmentation approach: image combination methodologies, such as CutMix, for metaplasia detection and localisation in gastric endoscopy imaging modalities.

1 Introduction

An upper gastrointestinal endoscopy (UGIE) is a vital medical procedure used for diagnosing and treating conditions affecting the esophagus, stomach, and duodenum. It plays a crucial role in identifying and managing gastric cancer and helping select patients for surveillance. H. pylori infection can lead to stomach inflammation and intestinal metaplasia (IM), increasing gastric cancer risk. Detecting IM during UGIE is challenging due to subtle endoscopic features. Narrow-band imaging improves accuracy, but it remains suboptimal, motivating the development of artificial intelligence algorithms [1], [2].

Deep learning aids in diagnosing gastrointestinal lesions in UGIE [3]. However, acquiring and annotating medical images is costly and challenging, limiting data availability. Data augmentation techniques, like CutMix [4], artificially expand training data. Yet, their effectiveness varies across datasets, like IM endoscopic images. Given IM's significance in preventing gastric cancer, this study explores the potential of CutMix-based data augmentation techniques and addresses four key challenges specific to IM detection and localization. Early experiments and real clinical IM images are provided as visual examples. The central research question is: "What are the primary research challenges associated with using CutMix methods to augment endoscopic data, particularly for intestinal metaplasia detection and localization?" By answering this question, this work contributes to improving established data augmentation strategies, potentially making substantial enhancements tailored to endoscopy image datasets' unique characteristics.

2 Data augmentation techniques

Image data augmentation enhances deep learning model performance and robustness by generating additional similar but non-identical examples from original images. It proves especially valuable in scenarios with limited training data, like medical imaging, where obtaining large, annotated datasets is challenging. This approach mitigates the impact of scarce training data, improving a model's ability to generalize to unseen data [5], [6].

Data augmentation techniques fall into three categories: model-free, model-based, and policy-based algorithms. Model-free methods don't rely on pre-trained models and can use one or multiple images. Model-based algorithms employ trained models to generate augmented images, either unconditionally, label-dependent, or image-based. Policy-based algorithms optimize augmentation hyperparameters for ideal results [5], [6].

This work primarily focuses on data augmentation techniques involving image combination, like MixUp and CutMix, due to their success in ImageNet classification [4]. MixUp combines two random images by interpolating their pixel values, generating a new example with a weighted average of pixel values and labels, enabling learning from in-between class examples [7]. CutMix selects two images and overlays a cutout from one onto the other, creating new examples that merge features from multiple images with labels weighted by the cutout area [4].

However, these techniques have limitations. MixUp assumes a linear relationship between inputs and target variables, which may not always hold, potentially generating unrealistic or uninformative data points. It can also produce similar images, lacking diversity. CutMix assumes rectangular lesion instances, not representative of real-world scenarios. It might not suit tasks like object detection, where object location matters. The randomly selected aggregated patch may introduce noise instead of relevant lesion features, as can be seen in figure 1 [4].

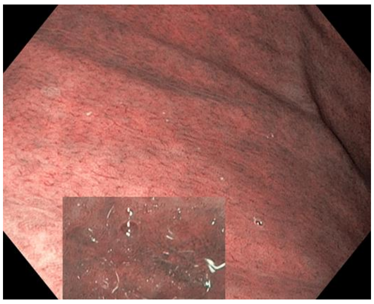

Figure 1: Examples of a noise augmented image resulting from CutMix.

3 Challenges for data augmentation in endoscopy images

In this section, we delve into the challenges associated with employing image combination techniques for endoscopy images, specifically in the context of IM detection. Our focus is on integrating lesion instances from IM samples into normal endoscopy images.

3.1 Image Blending

Seamlessly blending cutout patches into original images is a critical aspect of image augmentation, especially in medical imaging like endoscopy. Realism is paramount to ensure that the augmented image appears as if the lesion genuinely belongs to the endoscopic examination. Techniques like color correction, histogram matching, texture synthesis, or even applying MixUp on the lesion itself can facilitate this blending process. The ultimate goal is to create a natural and realistic augmented image where the cutout lesion seamlessly integrates into the normal mucosa, enhancing its quality for tasks such as classification, segmentation, and detection (Figure 2).

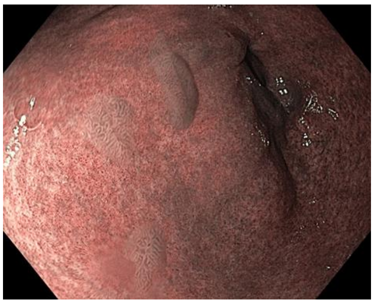

Figure 2: Example of augmented image with lesions blended into the mucosa.

3.2 Multi-Scale Image Combination

Combining endoscopy images requires careful consideration of factors, including the scale of the mixed examples. Ensuring the scaling of the lesion aligns with the target image is crucial. This consideration maintains accuracy and fidelity to the original images. Incorporating images with varying zoom scales introduces variability and allows the model to learn lesion characteristics across different scales, thereby enhancing its robustness and generalization.

3.3 Biologically Viable Images

Preserving the biological realism of endoscopy images with pasted lesions is paramount. Randomly placing lesions into images lacks biological coherence (Figure 3a). Blending techniques help integration, but precise lesion placement is equally critical for clinical and biological consistency. Strategic placement of the lesion within regions where such anomalies would typically occur in real-life scenarios is vital. This ensures the image remains a valid and reliable representation of real-world situations (Figure 3b).

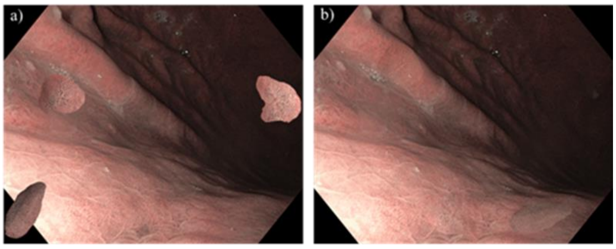

Figure 3: Example of *a)* non-biologically viable images and *b)* biologically viable images.

3.4 Relevant Geometrical Transformations

To further enhance the deep learning model's variability, robustness, and generalization for detecting lesions in endoscopy images, geometrical transformations can augment the instance paste approach. Applying these transformations in a randomized manner, offers insights into effective hyperparameters and policies specific to endoscopy's biological and clinical context. This exploration can lead to improved model performance and a deeper understanding of lesion features and characteristics in endoscopy images. The question arises: Which of these geometric transformations applied to the patched lesion contribute to better-trained models, and which do not? This investigation is crucial for advancing the field of endoscopy image analysis.

4 Discussion and conclusions

Traditional data augmentation techniques like CutMix and MixUp, while effective for general datasets, may not suit endoscopy imaging. Blindly pasting lesions into endoscopy images poses challenges related to visual coherence, biological realism, and clinical viability. These challenges lead to four key research areas for optimizing this approach in endoscopy:

Optimal Blending: How can we seamlessly integrate lesion instances into normal endoscopy images using blending methods to enhance visual realism? Achieving harmonious edges and a natural lesion insertion in the mucosa is crucial.

Geometric Transformations: How do geometric transformations contribute to improving model training? Exploring transformations that vary scale, rotation, texture, and colour can add valuable variability beyond CutMix.

Scale Variation: Can combining magnifying and non-magnifying endoscopy images enhance model robustness? This approach could enable lesion identification across a broader range of scales.

Biological Placement: To maintain biological coherence, how can we intelligently place lesions only in regions that make sense from a biological and clinical perspective? Random placement is insufficient.

Future research aims to conquer these challenges by customizing and implementing data augmentation algorithms designed explicitly for endoscopy image datasets, with the ultimate goal of substantially improving IM detection.

Acknowledgements

This work is financed by National Funds through the Portuguese funding agency, FCT—Fundacão para a Ciência e a Tecnologia, within project PTDC/EEI-EEE/5557/2020. Co-funded by the European Union (grant number 101095359) and supported by the UK Research and Innovation (grant number 10058099). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Health and Digital Executive Agency (HaDEA). Neither the European Union nor the granting authority can be held responsible for them. Alexandre Neto's work is also funded by the indivual FCT grant reference number 2022.13770.BD.

References

[1] D. S. Early et al., 'Appropriate use of GI endoscopy', Gastrointest. Endosc., vol. 75, no. 6, pp. 1127–1131, Jun. 2012, doi: 10.1016/j.gie.2012.01.011.

[2] P. Pimentel-Nunes et al., 'Management of epithelial precancerous conditions and lesions in the stomach (MAPS II): European Society of Gastrointestinal Endoscopy (ESGE), European Helicobacter and Microbiota Study Group (EHMSG), European Society of Pathology (ESP), and Sociedade Portuguesa de Endoscopia Digestiva (SPED) guideline update 2019', Endoscopy, vol. 51, no. 04, pp. 365–388, Apr. 2019, doi: 10.1055/a-0859-1883.

[3] F. Renna et al., 'Artificial Intelligence for Upper Gastrointestinal Endoscopy: A Roadmap from Technology Development to Clinical Practice.', Diagn. Basel Switz., vol. 12, no. 5, May 2022, doi: 10.3390/diagnostics12051278.

[4] S. Yun, D. Han, S. Chun, S. J. Oh, Y. Yoo, and J. Choe, 'CutMix: Regularization Strategy to Train Strong Classifiers With Localizable Features', in 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South): IEEE, Oct. 2019, pp. 6022–6031. doi: 10.1109/ICCV.2019.00612.

[5] M. Xu, S. Yoon, A. Fuentes, and D. S. Park, 'A Comprehensive Survey of Image Augmentation Techniques for Deep Learning'. arXiv, Nov. 23, 2022. Accessed: Jan. 20, 2023. [Online]. Available: http://arxiv.org/abs/2205.01491

[6] C. Shorten and T. M. Khoshgoftaar, 'A survey on Image Data Augmentation for Deep Learning', J. Big Data, vol. 6, no. 1, p. 60, Dec. 2019, doi: 10.1186/s40537-019-0197-0.

[7] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, 'mixup: Beyond Empirical Risk Minimization'. arXiv, Apr. 27, 2018. Accessed: Jan. 20, 2023. [Online]. Available: http://arxiv.org/abs/1710.09412