Diagnosis for Tubal Patency with Contrast Medium in Hysterosalpingography Images Using Asymmetric Contrastive Learning

Siwon Kim[1], Myeongju Kim[1], Jin Byeong Park[1], Haerin Paik[2], Yeon Hee Hong[2] and Jung Ryeol Lee[2]

[1] Center for Artificial Intelligence in Healthcare, Seoul National University Bundang Hospital, Seongnam-si, Republic of Korea, {kimww42, myeongju.j.kim, parkjinb2}@gmail.com

[2] Department of Obstetrics and Gynecology, Seoul National University Bundang Hospital, Seongnam-si, Republic of Korea, haerin.paik@gmail.com, lotusown@empas.com, leejrmd@snu.ac.kr

A. *Problem Setup*

The number of infertility patients is steadily increasing worldwide due to rising age at marriage, lifestyle changes, and various social factors. In 85% of infertility cases, the cause can be identified, with tubal infertility being one of the significant factors. To evaluate tubal patency as part of infertility work-up, Hysterosalpingography (HSG) using X-ray imaging is commonly performed. In order to receive infertility treatment in South Korea, examinations is required. Therefore, accurate interpretation of HSG images is crucial for determining effective treatment plans for patients. In tertiary hospitals, the interpretation of results is performed by radiology specialists. However, in general infertility clinics, the requesting gynecologists often perform the readings themselves, which can lead to a relatively inaccurate and time-consuming. AI-assisted X-ray image reading is a global trend that can reduce labor costs and automate HSG interpretation. It can also aid beginners who are interpreting these images for the first time. Therefore, this study explores a deep learning approach to assist in the interpretation of HSG images for infertility examinations.

B. *Novelty*

Hysterosalpingography (HSG). HSG is an examination in where a contrast medium is injected into the uterus and X-ray images are continuously taken over time. The diagnosis can be confirmed by observing the normal spillage of the contrast medium from the uterine tubes into the peritoneal cavity on the X-ray. Although X-ray-based HSG is a common examination in gynecology, interpretation using artificial intelligence has not been attempted. Liu and Ren (2021) attempted to interpret MRI-based HSG images using Convolutional Neural Network (CNN), but MRI-based HSG examinations are less frequently used than X-ray-based ones due to their high cost and limited accessibility as a primary diagnostic modality. Thus, there are limitations to its actual clinical application. The AI-assisted interpretation of HSG is a method being attempted for the first time in this study, and this approach is more cost-effective and patent-friendly.

Supervised Contrastive Learning. Khosla et al. (2020) proposed a supervised contrastive learning method that addresses the limitations of self-supervised contrastive learning. This method is training the model to bring data points from the same class closer in the representation space, even if they are not augmented versions of the same anchor. This approach enhances class separation, even with small datasets, making it well-suited for the characteristics of medical data. Along with this method, we applied a 2-stage learning method with effective Asymmetric Loss on asymmetric classification tasks to enhance discrimination on imbalanced datasets.

The main contributions of this paper are summarized as follows. First, we demonstrate the pioneering feasibility of a deep learning approach to classify contrast medium spillage into uterine tubes in X-ray-based HSG images. Second, we demonstrate that the 2-stage learning method with Supervised Contrastive loss and Asymmetric loss improves the ability to discriminate spillage in uterine tubes. Third, we confirm the effectiveness of using Supervised Contrastive Learning and Asymmetric Loss on medical data with asymmetric characteristics.

C. *Algorithm*

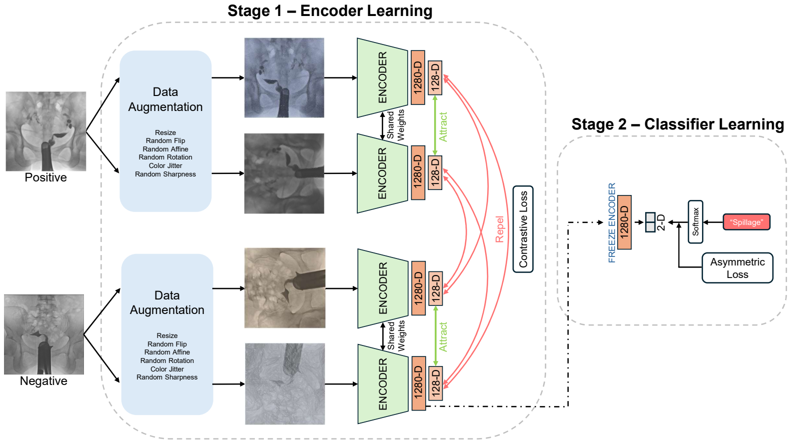

Fig. 1. An overview of the proposed method.

Our method is shown in figure 1. In Stage 1, input data is augmented to create two batches. Each batch is transformed into 1280-dimensional embeddings via an encoder, which are then reduced to 128 dimensions. Using Supervised Contrastive Loss, these embedding vectors are trained to place samples place samples belonging to other classes further. This process helps to classify spillage uterine tubes and occlusion uterine tubes in the embedding space.

In Stage 2, the encoder trained in Stage 1 is frozen, and only the classifier is trained. In this process, Asymmetric Loss, is applied to help more effective classification in datasets where spillage and occlusion samples are imbalanced.

D. Experiments

We retrospectively collected data on cases in which HSG was performed for infertility at Seoul National University Bundang Hospital from June 2003 to May 2023 (IRB No. B-2307-839-102). From this dataset, we selected 410 cases for analysis. For each case, we chose two-images taken during the midpoint of the procedure. An obstetrician-gynecologist labeled each case's uterine tubal spillage status as binary. The labeled dataset is imbalanced, with 75% of the cases being spillage and 25% being occlusion. We used 80% of the dataset as the training set and 20% as the evaluation set.

To compare the effects of Supervised Contrastive Loss on each stage separately with the effects of Asymmetric Loss, we conducted experiments using six different methods: training with Cross Entropy Loss, Focal Loss, and Asymmetric Loss in a single-stage approach, and using a 2-stage approach with Supervised Contrastive Loss combined with Cross Entropy Loss, Focal Loss, and Asymmetric Loss. In all experiments, the backbone encoder used efficientnet-b0, maintained the same hyperparameter values, and used the results of the models with the minimal loss value from each experiment. To ensure the reproducibility of each experiment, the experiments were conducted with three fixed seeds (2023, 2024, 2025).

Methods	Accuracy (%)	Precision	Recall	F1	AUPRC
CE	84.388 ± 2.223	0.813 ± 0.107	0.508 ± 0.029	0.623 ± 0.037	0.761 ± 0.100
Focal	85.564 ± 1.593	0.750 ± 0.023	0.650 ± 0.087	0.695 ± 0.049	0.758 ± 0.025
ASL	85.564 ± 0.967	**0.832 ± 0.105**	0.558 ± 0.063	0.663 ± 0.020	**0.827 ± 0.048**
SupCon+CE	86.709 ± 2.282	0.792 ± 0.112	**0.675 ± 0.090**	0.720 ± 0.032	0.766 ± 0.038
SupCon+Focal	87.342 ± 1.675	0.820 ± 0.008	0.641 ± 0.095	0.717 ± 0.050	0.776 ± 0.045
SupCon+ASL	**87.764 ± 0.731**	0.827 ± 0.057	0.650 ± 0.066	**0.724 ± 0.030**	0.769 ± 0.036

Table 1. Results on comparing each methods. The best and second-best results are bolded and underlined, respectively. (**CE**: Cross-Entropy Loss, **Focal**: Focal Loss, **ASL**: Asymmetric Loss, **SupCon**: Supervised Contrastive Loss)

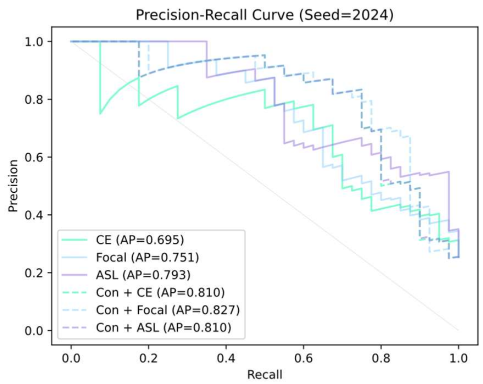

Fig. 2. Results of comparing each method with PR CURVE.

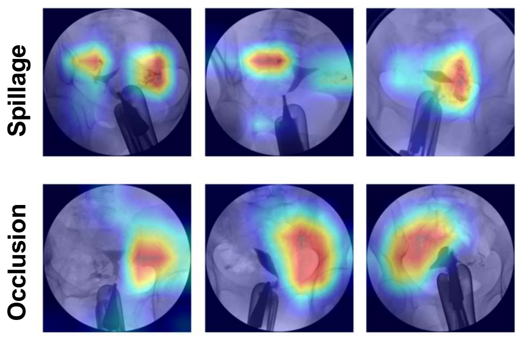

Fig. 3. Visualization with Grad-CAM

In these comparisons, the method combining Supervised Contrastive Loss and Asymmetric Loss achieved the highest performance in accuracy and F1 score, and the second-best performance in precision and recall. This approach proved effective in identifying uterine tube spillage in HSG images and is expected to improve performance when using imbalanced medical data.

E. Acknowledgement

This work was partly supported by Institute of Information communications Technology Planning Evaluation (IITP) grant funded by the Korea government(MSIT) (No.RS-2022-II220184, Development and Study of AI Technologies to Inexpensively Conform to Evolving Policy on Ethics) and Seoul National University Bundang Hospital Research Fund (Grant number 2023-03501).