ポスター

ポスター6
医療データ分析３（画像認識）

2018年11月23日(金) 16:00 〜 17:00 K会場(ポスター、HyperDemo) (2F 多目的ホール)

[2-K-3-3] 医用画像を入力とした Deep-Learningよる放射線画像診断レポート自動生成の検討

○松本 宏[1], 宇宿 功市郎[2,1] （1.熊本大学大学院医学教育部医療情報医学分野, 2.熊本大学医学部附属病院医療情報経営企画部）

In the rapid advancement in informatics, there is a study of generating a caption for an input image. In this study, we investigate the applicability of such method to auto-generate a radiological diagnosis report in Japanese from a medical image. We built a deep-learning report-generating model, consisting of convolutional layers and recurrent layers. For dataset, 5,000 reports previously diagnosed from 2015 through 2018, embedded with key-images are collected. In the collection, pairs of diagnosis statements and first key-image are acquired for training and validation test-set. After running 50 epochs of training, loss and convergence indicator has stayed around 0.46. With variation in diagnosis statement size and vocabulary size the loss has not changed. From our results, it is clear that the training has not converged with 5,000 reports.

Automatic Generation of Radiological Diagnostic Report by Deep-Learning with Medical Image

Hiroshi Matsumoto[1,2] Koichiro Usuku[2] Taishi Nakamura[2] Yoshinori Yamanouchi[2]

Y's Reading inc.[1]

Department of Medical Information Sciences and Administration Planning, Kumamoto University Hospital[2]

Abstract: Since the number of images taken daily basis increases and the burdens of radiologists accumulate, it is essential for radiologists to acquire some writing aid. In the rapid advancement in informatics, there is a study of generating a caption for an input image. In this study, we investigate the applicability of such method to auto-generate a radiological diagnosis report in Japanese solely from a medical image.

Keywords: Radiographs, Natural Language Processing, Natural Language Generating, Image Processing.

1. Introduction

Since the number of images taken daily basis increases and the burdens of radiologists accumulate, it is essential for radiologists to acquire some writing aid. In the rapid advancement in informatics, there is a study of generating a caption for an input image. In this study, we investigate the applicability of such method to auto-generate a radiological diagnosis report in Japanese solely from a medical image.

2. Related Works

In one of daily-emerging deep-learning-based research in informatics, there is a study on auto-generating descriptive captions for the content of a picture. Vinyals et al, 2015, have shown their success in generating captions for various images [2]. SHIN et al, 2015, has shown deep-learning method to extract appropriate tokens from images [3].

3. Methods

In this study, focused on head-image diagnosed reports, the applicability of the deep-learning method of the auto-generating the caption

The dataset used for this experiment is from collected data of 36,669 pairs of images and reports in Y's Reading in between May 9th, 2013 and May 17th 2018, limiting the diagnosis of heads. After cleaning the dataset, limiting the token size of documents to 500 tokens, the total volume came down to 4,900 pairs.

For each report, it bundles with multiple key-images in the order of their importance. Only the first image is used for this experiment.

The deep-learning model used for captioning the radiographs is a joint model of VGGNet for image processing and auto-encoder model for natural language generating. The VGGNet, VGG19, is composed of 16 convolutional layers having small 3x3 filter kernels and 3 fully connected layers, originally designed for ImageNet classification task [1][4]. For this experiment, pre-trained VGGNet model on ImageNet is used. The auto-encoder model consists of LSTM layers in the shape of token size of document, 500, and hidden layer size, 256, for encoder and for the decoder bidirectional LSTM layers are used with same parameters as encoder. For the optimizer and loss function, Adam and categorical cross-entropy are used.

All images are resized to 224 pixels by 224 pixels, applicable size for VGGNet. As for the text data, each document is parsed using MeCab, Japanese morphological analyzer, with default IPADIC and ComeJisyo [6].

In this experiment, 100 report-image pairs are used for training and it was tested for 1,000 training times with batch size of 64.

Another 100 sample set aside from the training samples are applied as test set. The result is evaluated using BLEU metrics, widely used for Machine Translation task [7]. It evaluates the matching ratio of reference text to system's output text.

4. Results

The graph below shows the training process with loss value and validation loss.

Figure 1 Training Result

Table 1 shows the output of this trained model with untrained samples as input and corresponding correct text. For the text evaluation, BLEU result is 27.13.

5. Discussion

From training dataset size applied for the training, the training result, and the output text, it is apparent that the training is over-fit. However there are some insights in this result. The output text shows appropriate Japanese sentence constructed even with the small amount training dataset. For the future work, all the dataset will be applied for the training. Furthermore, attention model will be implemented to see the semantics of words in the reports will correspond to the appropriate portion in images.

6. References

[1] SIMONYAN, Karen; ZISSERMAN, Andrew. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

[2] VINYALS, Oriol, et al. Show and tell: A neural image caption generator. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2015. p. 3156-3164.

[3] SHIN, Hoo-Chang, et al. Interleaved text/image deep mining on a very large-scale radiology database. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2015. p. 1090-1099.

[4] RUSSAKOVSKY, Olga, et al. Imagenet large scale visual recognition challenge. International Journal of Computer Vision, 2015, 115.3: 211-252.

[5] Taku Kudo, Kaoru Yamamoto, Yuji Matsumoto: Applying Conditional Random Fields to Japanese Morphological Analysis, Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP-2004), pp.230-237 (2004.)

[6] Kaoru Sagara, ComeJisyo. http://sourceforge.jp/projects/comedic/

[7] PAPINENI, Kishore, et al. BLEU: a method for automatic evaluation of machine translation. In: Proceedings of the 40th annual meeting on association for computational linguistics. Association for Computational Linguistics, 2002. p. 311-318.

	Correct Document	System Output
1	<BOD> 明らか な 頭蓋 内 出血 は 認め られ ませ ん 。 <EOD>	<BOD> 明らか な 頭蓋 内 器質 病変 は 指摘 でき ませ ん 。 <EOD>
2	<BOD> 急性 期 脳 梗塞 は 認め ませ ん 。 <EOD>	<BOD> 明らか な 頭蓋 内 器質 病変 は 指摘 でき ませ ん 。 <EOD>
3	<BOD> 明らか な 頭蓋 内 器質 病変 は 指摘 でき ませ ん 。 <EOD>	<BOD> 明らか な 頭蓋 内 は 同定 は ませ ませ ませ ん 。 <EOD>
4	<BOD> 橋 の 軽度 萎縮 疑い ： 性別 から 正常 下限 レベル の 大き さ かも しれ ませ ん 。 <EOD>	<BOD> 明らか な 頭蓋 内 器質 病変 は 指摘 でき ませ ん 。 <EOD>
5	<BOD> ・ 急性 期 頭蓋 内 病変 を 指摘 でき ませ ん 。 <EOD>	<BOD> 明らか な 頭蓋 内 器質 病変 は 指摘 でき ませ ん 。 <EOD>
6	<BOD> 右 中 前頭 回 の 急性 期 脳 梗塞 ： 中心 部 に 出血 を 伴っ て い ます 。 前年 に 比べ て 大脳 萎縮 が 進行 。 <EOD>	<BOD> 明らか 明らか 頭蓋 内 器質 病変 は 指摘 ませ ませ ん 。 <EOD>
7	<BOD> 症状 の 原因 と なる 病変 は 指摘 でき ませ ん 。 <EOD>	<BOD> 明らか な 頭蓋 内 器質 病変 は 指摘 でき ませ ん 。 <EOD>
8	<BOD> 主 訴 の 原因 は 同定 でき ませ ん 。 <EOD>	<BOD> 明らか な に は は は は ませ ん ん 。 。 <EOD>

Table 1 List of system output corresponding to its correct text