Explainable AI for Diagnosis of Malignant Lymphoma

Hidekata Hontani (Nagoya Institute of Technology, JAPAN)

Abstract— **This presentation discusses the explainability of computer aided subtype diagnosis for malignant lymphoma. While implementing an AI that uses neural networks to classify subtypes by inputting pathological microscope images can be technically feasible, it often presents challenges in providing supervised signals. In a clinical context, what pathologists seek in subtype-diagnosis machines is *explainability*, which in this context refers to the machine's ability to quantitatively elucidate diagnostic evidence using interpretable features for physicians. We will discuss efforts to improve the explainability of subtype classification and its relationship to the WHO diagnostic criteria.**

I. Background of this presentation

Malignant lymphoma is a type of blood cancer, broadly classified into three main categories: Hodgkin lymphoma, B-cell type lymphoma, and T-cell type lymphoma. Each of these categories is further subdivided into numerous subtypes, with over 70 different subtypes identified. Each subtype has distinct treatment strategies and prognoses. One of the primary objectives of pathological diagnosis is the identification of the specific subtype. The classification of subtypes relies on the use of pathological microscope images. Initially, candidate subtypes are estimated through the observation of cellular tissue morphology in images of H&E-stained tissue thin sections. Subsequently, immunostaining is employed to confirm the subtype. For certain subtypes, an assessment of malignancy grade is also conducted.

The classification of subtypes and the assessment of malignancy grade are often performed empirically and qualitatively through visual examination by pathologists. The differentiation degree of malignant lymphoma is relatively high compared to other types of cancer, it is challenging to describe the morphological changes. Achieving a quantitative assessment of these morphological changes can contribute to improving the accuracy and reliability of the diagnoses of malignant lymphoma.

Follicular lymphoma is a subtype of B-cell lymphoma. In its diagnosis, the morphology of cellular structures called follicles is observed. Follicles are spherical structures composed of numerous cells, and in pathological images, their cross-sectional disc-like structure can be observed. Within normal follicles, a variety of cell types are present in significant numbers. As the malignancy grade of follicular lymphoma increases, the composition ratio of cell types within the follicles changes. Generally, as cancer cells proliferate, the diversity of cell types decreases.

The World Health Organization (WHO) defines the assessment criteria for malignancy grade in follicular lymphoma based on the number of cells known as centrocyte and centroblast within the field of view when observing the interior of follicles under high magnification using a microscope. However, the WHO does not mandate the assessment of malignancy grade. This is because visually identifying the cell type of each cell and counting the number of centrocytes and that of centroblasts among the numerous cells present in the field of view is not practical. However, it can become practical by using image processing.

The reason why the WHO did not require evaluation of malignancy was because it was assumed that diagnosis would be made visually. If it becomes possible to accurately identify cell types, it will also be possible to quantitatively evaluate malignancy in accordance with WHO criteria. However, it is not easy to create an AI that accurately identifies cell types. Although pathologists have knowledge about the types of cells within follicles, they do not identify cell types in daily diagnosis. For this reason, it is not easy even for pathologists to identify the types of individual cells, and it takes a great deal of time to identify and label many cell nuclei. Therefore, to construct an AI that identifies cell types, it is necessary to use engineering techniques to select cells to be labeled and to take into account the ambiguity that the labels may contain.

If you are given a large amount of unlabeled data and it is difficult to label all of it, we need to select and label the data that can improve the target task as efficiently as possible. Active learning provides a method for such data selection. A typical technique is a method that utilizes the ambiguity of the classifier. That is, a classifier is firstly constructed based on the few labels currently available, the variance of the estimated posterior probability is evaluated, and then data with a large variance is selected and labeled. Unfortunately, it is known that this popular method is not always the best method, and various other methods have been proposed. A method we employed can improve the efficiency of labeling by selecting data based on the distribution of the large amount of unlabeled data.

The ambiguity information contained in the labels can be used to improve the classification performance when constructing a classifier. When assigning a class label, a pathologist may, for example, takes long time before deciding one final label between two candidate classes. By recording all such confusions during labeling, we can improve the performance of the classifier by using such the confusion records.

We believe these techniques described here are all important to constructing an explainable AI that accomplishes tasks that are not easy even for pathologists themselves.