【Grant-in-Aid for Scientific Research (S)】

Joint optimization of physical encoder with object recognition model and its pathological application

Principal Investigator	Osaka University, Institute for Datability Science, Professor NAGAHARA Hajime	Researcher Number:80362648
Project Information	Project Number : 23H05490 Keywords : Computational Photography, Machine learning, Pathology	Project Period (FY) : 2023-2027

Purpose and Background of the Research

●Outline of the Research

An object recognition model by regular deep learning is only optimized in the digital layer after capturing RGB digital image as an input. Light is an electromagnetic wave with multiple modality information. Still, a regular camera can only measure the part of optical modalities because it measures RGB images assuming human visual characteristics. The possible solution utilizing such multiple optical modalities is to scan high-dimensional multi-modal images and feed them into a digital encoder to learn the features for recognition, as shown in Figure 1a. However, this straightforward approach would not be feasible due to the huge scanning time and data volume. In this study, we propose a framework, deep physical sensing that a part of the encoder is replaced by a physical encoder, as shown in Fig. 1b, and optimize the hardware parameters by learning to achieve highly accurate recognition with a single coded image. We will apply this framework to pathological diagnosis to show the feasibility as an example of an actual use case.

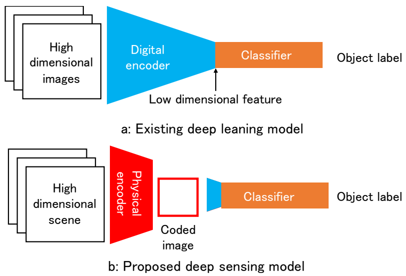

Figure 1. Comparison of existing and proposed approaches

●Multi-modality of light and physical encoder

An RGB image, as an input of regular deep learning, loses information by image capturing, which has limited recognition performance. As shown in Figure 2, light has multiple modalities; amplitude, wavelength, phase, polarization, and time. These optical modalities contain physical information about the target objects, such as the object's reflectance and transmittance, thickness, refraction, etc. There is a lot of bioimaging research that has utilized hyper-spectrum, interference, polarization, etc. However, it is not obvious which modality or combination of modalities will improve recognition performance for a specific task such as pathological diagnosis. It is also difficult to train the model because of the huge scanning time, training time and memory requirements, as shown in Fig. 4 when we use those multimodal images as input. In this research, a physical encoder assumes the imaging hardware, as shown in Fig. 3, and optimizes it to capture the coded image. We will achieve highly accurate object recognition from a single captured image since the optimized coded image effectively combines the modalities for task-specific object recognition.

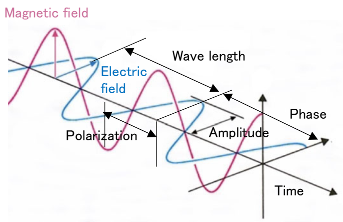

Figure 2. Multimodality of light

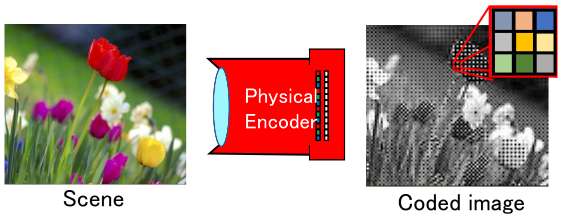

Figure 3. Physical encoder

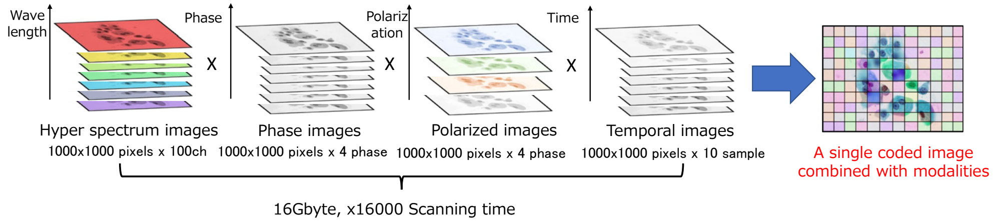

Figure 4. Multi-modality images and a proposed coded image

Expected Research Achievements

●Proposing deep physical sensing（Academic impact）

We propose a framework called "Deep physical sensing," which models the imaging hardware and jointly optimizes the parameters and recognition model to enhance the accuracy of the specific task.

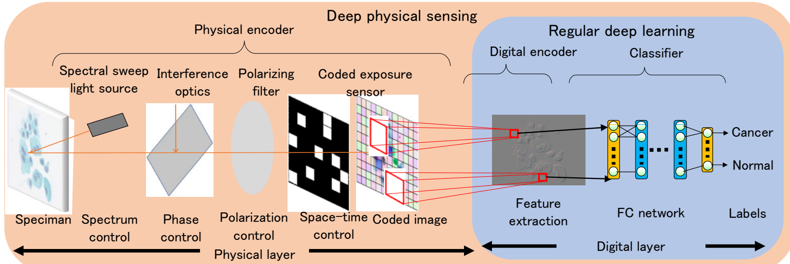

Figure 5. Deep physical sensing

●Applying to pathology（Social impact）

We will apply deep physical sensing to the automated pathological diagnosis to validate that the concept works for the application. There is a lot of work for automatic pathological diagnosis using an RGB image. It is the approach that AI mimics a human visual system for recognizing cancer cells. However, our approach is that AI decides what kind of image is the best input for the diagnosis task by itself. This concept is the new direction of AI research and will be affected by any applications.

●Developing multimodal microscopy（Prototyping of physical encoder）

We will implement the microscopy hardware shown in Fig. 6, which is programmable to control light modalities and get the coded image combined with them. We will use the prototype for the verification of pathology applications.

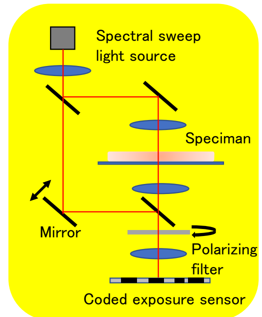

Figure 6. Multi-modal microscopy

Homepage Address, etc.	https://www.is.ids.osaka-u.ac.jp/project/kibans-23h05490/ E-mail: nagahara@ids.osaka-u.ac.jp