

Title	視覚情報の少ない物体の検出
Author(s)	PHO, NGOC DANG KHOA
Citation	
Issue Date	2022-09
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/18140
Rights	
Description	Supervisor: 吉高 淳夫, 先端科学技術研究科, 博士

氏 名	PHO, Ngoc Dang Khoa
学 位 の 種 類	博士 (情報科学)
学 位 記 番 号	博情第 484 号
学 位 授 与 年 月 日	令和 4 年 9 月 22 日
論 文 題 目	Detection of Object with Less Visual Information
論 文 審 査 委 員	主査 吉高 淳夫 北陸先端科学技術大学院大学 准教授
	鵜木 祐史 同 教授
	岡田 将吾 同 准教授
	赤木 正人 同 名誉教授
	戸田 真志 熊本大学 教授

論文の内容の要旨

Object Detection plays an essential role in many practical applications such as video analysis, image understanding, security, etc. Recent years have witnessed the breakthroughs of Deep Learning based methods in various applications in daily life images. Deep Learning-based methods achieve human-level performances on classification tasks for identifying successfully more than tens of thousands of categories such as animal species, vehicles, household objects, etc. However, deep neural networks are still limited in specific medical image domains. The neural networks often fail for reasons such as a small number of training data, difficulty characterizing the target objects, or a small number of differences among the target objects.

This study is motivated by the need for a comprehensive method for medical image domains to accelerate the diagnosis and treatment processes. However, the target objects of medical image domains often (i) have less visual information, (ii) have a small number of training data, and (iii) have various appearances. This dissertation detects objects with a few features appearing in medical images. This study assumes that each target category has enough distinctive features to identify even with a small number of features. Finding those distinctive features is essential to detecting objects with less visual information.

In the case of small training data, the problem becomes more difficult since the method must find the correct distinctive features within a few samples. A particular surrounding object may have a high probability of appearing along with the target object. With a small number of training samples, the deep neural network detectors view the background information to identify the objects. This study focuses on finding features that characterize the target objects rather than the unique background features to overcome the problem. The segmentation-driven mechanism is proposed to guide the detector to focus only on the regions of the target objects. The mechanism is integrated into a neural network detector to form Segmentation-driven RetinaNet to filter out the background by the segmentation mask and then detect and identify the objects in the filtered image.

While the characteristic features are efficient for detection tasks, distinctive features are essential for identifying the objects. Detecting objects in grayscale images is also a challenging problem. Grayscale images with only one color channel have much smaller feature spaces than general color images. The objects in grayscale images are characterized only by outer shapes, connectivity, and the intensity of the pixels. The attention-driven mechanism is proposed by replacing the segmentation with the attention mask to guide the deep network in focussing on the distinctive features of the target objects.

Finally, this study explores the relationship among the various appearances of a target object category. An object category may have several appearances, and each appearance only shows some category features. In many cases, an appearance of one category may have more similarities to that of another category than its intraclass appearances. Motivated by the taxonomy of animals, this study investigates the hierarchical multi-label classifier and the category hierarchy structure. Training samples of each category are clustered concerning the appearances. Multiple labels following the category hierarchy structure are assigned to training samples. The hierarchical classifier is integrated with the Segmentation-driven RetinaNet to form a unified network for detection.

Experiments are conducted on realistic datasets from the protozoa and DNA Profiling domains as examples of objects with less visual information in color and grayscale images, respectively. Experiments show that the Attention-driven mechanism effectively guides the neural network detectors to find the distinctive characteristic features of the target objects. Even with at most five samples per subcategory for training, this study successfully trained the proposed method for detecting the protozoa in the micrographs. With 16 training samples, the proposed method achieves the highest performance on the DNA Profiling image dataset. Besides, the integrated hierarchical multi-label classifier boosts the detection performance for the polymorphism problems.

Keywords: Detection, Segmentation, Identification, Protozoa, Genome Profiling

論文審査の結果の要旨

深層学習を画像処理に応用した研究は近年活発に行われており、物体検出や識別の問題に対して従来手法よりも飛躍的に精度を向上させることを可能とした。人や動物、自動車等の人工物を検出する一般物体検出に関する研究例は多いが、それと比較して、医用画像等の視覚的特徴がより乏しい画像を対象とした物体等の検出、識別はより困難である。また、深層学習により十分な精度を得るためには一般に大量の学習データが必要であるが、検出、識別対象によっては、学習データを多数収集することが人的、時間的コストの点から容易でないような状況も多く見受けられる。

本学位論文は、上記一般物体と比較して種の識別をする手がかりとなる視覚的特徴がより乏しい対象で、かつ、少数の画像データしか取得できない状況を問題として設定し、そのような状況においてもより精度高く対象物を検出、或いは識別する深層学習手法について取り組んだものである。具体的には、人に対して疾患の原因となる寄生虫を含む試料の顕微鏡写真画像から寄生虫体を検出し、その種を識別する問題や DNA 解析に関する 1 手法である TGGE（温度勾配電気泳動）画像を入力とし、その画像から種識別の特徴点となる変曲点を検出する問題を取り上げ、これらの画像が限定的な数しか取得できない状況においても、検出、識別精度を向上させる深層学習手法を確立したものである。

深層学習による寄生虫体の識別においては、識別に寄与した部位を可視化する手法により分析を進めたところ、対象となる虫体領域だけでなく、それ以外の背景領域の特徴も判定に寄与していることが判明したため、まず虫体領域を検出するネットワークを適用してそれ以外の領域を除外し、その画像に対して識別処理をするネットワークを適用する手法を提案した。さらには特定の種であっても画像取得条件や種によっては複数の様態を持つことを考慮し、この問題に対応するために複数階層から成る分類機を構成することにより対処することも提案した。これらにより、従来法と比較して検出、識別精度を向上させられたことを示した。また、グレースケール画像で軌跡の幅や濃淡が一定しないという特徴を持つ TGGE 画像の変曲点検出問題に対しては、アテンションを導入することにより、より適切に特徴的な部分を反映した検出処理を可能とすることを示した。

以上、本論文は、視覚的特徴の乏しい対象物に対するより高精度な検出、識別手法を確立したものであり、学術的に貢献するところが大きい。よって博士（情報科学）の学位論文として十分価値あるものと認めた。