International Journal of Multicultural and Multireligious Understanding

http://ijmmu.com
editor@ijmmu.com
ISSN 2364-5369
Volume 11, Issue 11
November, 2024
Pages: 497-506

A Comprehensive Study of Supervised Learning Models on the OCTID Dataset for Retinal Disease Detection with Future Directions

Arman Mohammad Nakib[1]; A K M Jayed[2]; Shah Anisur Rahim Scene[3]; Md Badsha[4]

[1] Artificial Intelligence, Nanjing University of Information Science & Technology, Nanjing, Jiangsu, China

[2] Computer Science and Technology, Nanjing University of Information Science & Technology, Nanjing, Jiangsu, China

[3] Software Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu, China

[4] Computer Science and Technology, Nanjing University of Posts and Telecommunication, Nanjing, Jiangsu, China

http://dx.doi.org/10.18415/ijmmu.v11i11.6403

Abstract

This paper presents a critical analysis of recent approaches to supervised learning techniques on the OCTID database that are important in the early diagnosis of retinal diseases such as AMD, CSR, DME, and MH. The review concentrates on methodologies, experimental configurations, and assessments of various machine learning and deep learning techniques. The limitations observed in the current work are in terms of computational complexity, possible limitation in labeling, and lack of extensive experimentation using videos with OCT. For future work, we suggest using semi-supervised learning as they are more accurate and less costly in medical applications.

Keywords: *OCTID Dataset; Retinal Disease Detection; Semi-Supervised Learning; Future Directions*

1. Introduction

Some of the retinal diseases include: Age-related Macular Degeneration; Diabetic Macular Edema; and Central Serous Retinopathy. Through application of the Optical Coherence Tomography (OCT technology that does not require invasive procedures) detection of these diseases is possible at an early stage. However, the problem is that manual diagnosis based on the OCT images is time-consuming and has a large measure of subjectivity. To overcome these difficulties, the use of machine learning and deep learning has come into the foreground as a natural solution for automated diagnosis.

The OCT Image Database (OCTID) is considered as a well-known dataset for creating classification models aimed at diagnosing pathological changes in the retinal tissue. This review covers supervised learning models that have been used with OCTID and the differences in experiments conducted, their performance, and drawbacks. In addition, it outlines limitations of existing approaches, for example, high computational time and lack of studies on applying semi-supervised learning.

2. OCTID Dataset Details

The Optical Coherence Tomography Image Database (OCTID) is a dataset obtained especially for the classification of macular diseases based on OCT images. It has a pre-eminent function in training various algorithms of Machine learning and Deep learning to diagnose automatically. The dataset is relatively small, consisting of 572 images, divided into five main classes: AMD, CSR, DR, MH, and Normal retina (Gholami, Peyman, et al, 2020).

2.1 Class Breakdown

a) AMD (Age-related Macular Degeneration)

It has a total of 55 images. AMD is a progressive eye condition that affects the macula, leading to the loss of central vision. In the dataset, OCT images show typical signs of AMD such as drusen deposits and retinal atrophy. This class is one of the more challenging to detect due to the subtlety of the early-stage symptoms in OCT images.

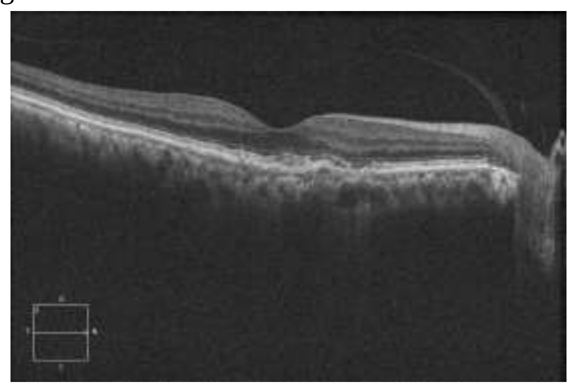

Fig. 1. AMD

b) CSR (Central Serous Retinopathy)

It has a total of 102 images. CSR is characterized by the accumulation of fluid under the retina, leading to visual impairment. The OCT images in this class show the typical signs of subretinal fluid accumulation, with a distinct detachment of the retinal layers.

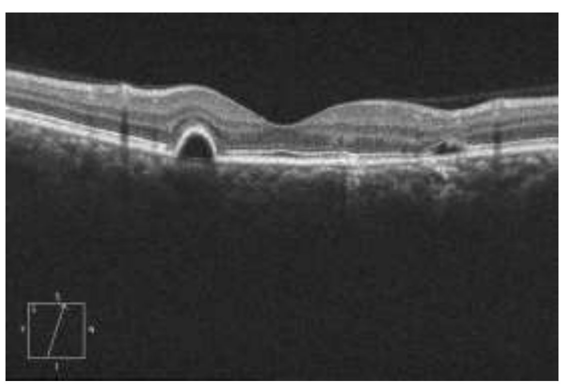

Fig. 2. CSR

c) DR (Diabetic Retinopathy)

It has a total of 107 images. Diabetic Retinopathy involves damage to the retinal blood vessels, which can result in leakage and swelling. The DR images in OCTID exhibit signs such as retinal thickening, microaneurysms, and fluid leakage, which can lead to vision loss if untreated.

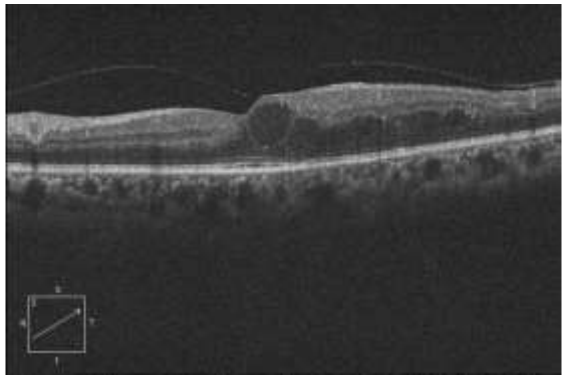

Fig. 3. DR

d) MH (Macular Hole)

It has a total of 102 images. Macular Hole is a condition where a small break in the macula causes blurred or distorted central vision. The OCT images in this class clearly show the full-thickness break in the macula and the surrounding retinal layers, providing a unique signature for model classification.

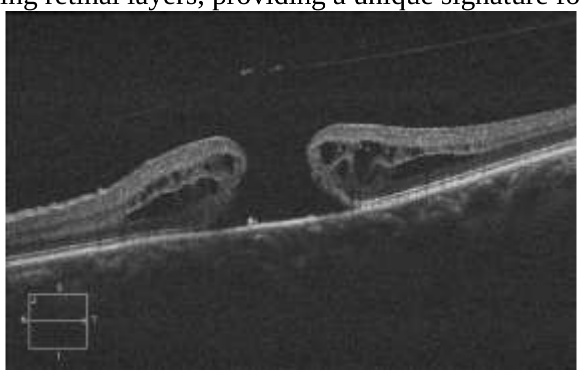

Fig. 4. MH

e) Normal

It has a total of 206 images. This class contains images of healthy retinas with no signs of pathology. These images are vital for training models to differentiate between pathological and healthy retinal structures, ensuring the model can effectively discern between normal and diseased conditions.

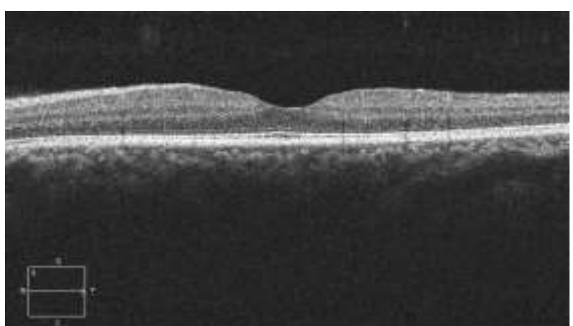

Fig. 5. Normal

2.2 Challenges with OCTID Dataset

a) Small size

The dataset has options of only 572 images, and it is a challenge for deep learning models which work best with large images and data sets. Because of this limitation, this setting is well-suited for analyzing different approaches to semi-supervised learning that are used in cases where both labeled and unlabeled data are available.

b) Class Imbalance

The database is highly unbalanced, which means that the “Normal” group has many more images (206) than the others, for example, AMD has 55 images. However, failure to do so can result in skewed models which may not be desirable.

c) Diversity of conditions

Although the dataset encompasses five different classes, the difficulty arises from the fact that the number of images contributing to each of the classes is small, particularly for diseases like AMD, which would systematically reduce the generality of the model to various forms of those diseases.

3. Literature Review

3.1 Deep Learning-Based Automatic Detection of CSR

Hassan et al. introduced AlexNet, ResNet-18, and GoogleNet in their work for the automatic detection of CSR using OCT images derived from the OCTID dataset. The CNN models used in this study were trained with high classification accuracy primarily because imaging was preprocessed to correct image contrast and minimize noise. The highest accuracy of 99.64% had the best model, AlexNet, which indicates the opportunities to apply deep learning in the diagnostics of diseases (Hassan, Syed Al E., et al, 2021).

Model: AlexNet, ResNet-18, GoogleNet

Accuracy: 99.64% (AlexNet)

Strengths: High accuracy with good ability in data preprocessing

3.2 Deep Dictionary Learning and Wavelet Scattering Transform (WST)

The combined method of DDL and WST was presented by Shaker et al. for the classification of OCT. In this paper, I found that the WST model did better than DDL, particularly in multiclass classification. It revealed its stability when extracting features from the OCT images and attained a classification accuracy of 82.5% on five classes. The models were evaluated on the OCTID dataset the results were measured and a comparison between WST and DDL (Shaker, Fariba, et al).

Model: DDL, WST

Accuracy: 82.5% for five-class classification

Strengths: Proper for cases of multiple classes

3.3 A Multipath CNN for AMD Detection

Thomas et al. (2021) followed a multipath convolutional neural network for automated identification of Age-related Macular Degeneration (AMD). The model employs multiple convolutionary paths to generate global and local features and generates high classification rates on different datasets, including OCTID. This high percentage of accuracy specifically 99.61% showed that the model developed possesses a good generality level and can therefore be used in real-time medical application systems in the future (Thomas, Anju, et al, 2021).

Model: Multipath CNN

Accuracy: 99.61% (OCTID)

Strengths: High accuracy for all the datasets

3.4 Model-Based Transformer for OCT Image Classification

Hammou et al. (2023) proposed a Model-Based Transformer (MBT) to predict the type of retinal diseases using OCT images and videos. The authors utilized transformers including Vision Transformer and Swin Transformer to extract outstanding features from OCT data. Compared to prior deep learning models, this model proved to be less complicated whilst also delivering high levels of accuracy on the OCTID data set (Hammou, Badr Ait, et al, 2023).

Model: MBT, Vision Transformer, Swin Transformer

Accuracy: The model determined highly accurate, competitive results based on the F1-score.

Strengths: Considering the transformers at the right places, low computational discussions

3.5 Automatic Segmentation of Macular Holes (MH)

For automatic segmentation of macular holes in OCT images, Mendes et al. (2021)proposed an algorithm. The authors combined two approaches which are Euclidean distance and flood fill algorithms to segment the retinal boundaries for the detection of MH with high accuracy. This approach was tested on OCTID – the segmented images show the possibility of applying the segmentation algorithms for increasing the diagnostic precision (Mendes, Odilon LC, et al, 2021).

Model: Segmentation based on distance and measurement with reference to the Euclidean distance and flood fill.

Accuracy: The MH segmenters demonstrated a high accuracy rate in identifying MH disease.

Strengths: Suitable for the division of particular forms of retinal disease

3.6 Use cases using Deep Learning for AMD Recognition

A detailed literature review of the methods proposed by Koseoglu et al. (2023) for the detection of AMD using OCT images and deep learning approaches has been presented. The paper also highlighted CNN and as well as other new architectures such as Vision Transformers. Elo The paper also explained the benefits of convolutional neural networks (CNNs) and newer networks, such as vision transformers. In particular, the CNN-based models provided reasonably good effectiveness as several of the models showed accuracies of above 90 percent in recognizing AMD in a large dataset like the OCTID. We also first reviewed the limitation of deep learning methods for their high data requirements and inability to explain the results (Koseoglu, Neslihan Dilruba, Andrzej Grzybowski, and TY Alvin Liu, 2023).

Model: CNN, Vision Transformers

Strengths: General overview of models; high recall rate for identifying AMD

3.7 Automatic Diagnosis of Retinal Diseases Based on OCT B-scan

Another AI-based system Marciniak and Stankiewicz (2023) uses OCT B-scans to classify the retinal diseases. This study tested VGG16 and Inception V3 models trained on the LOCT dataset, which contains over 84,000 images across four classes: normal, CNV, DME, and Drusen. The authors applied transfer learning and compared the performances of these models and gained a classification accuracy of 96.28% using the VGG16 model optimized with RMSprop (Marciniak, Tomasz, and Agnieszka Stankiewicz, 2023).

Model: VGG16, Inception V3

Accuracy: 96.28% (VGG16)

Strengths: Very accurate results on big data; transfer learning is also very useful in reducing the amount of retraining needed

3.8 Wavelet Scattering Transform in Classification of Retinal Abnormality

Baharlouei et al. (2023) used the Wavelet Scattering Transform (WST) for the classification of retinal pathology in OCT images. The study showed that the application of WST could sail high from the regular CNNs based on the less computational hurdle. The findings of this study showed that the model attained 100% classification for binary and multiclass problems on the OCTID dataset. Translation-invariant property of the WST model along with its ability to handle slight deformations was particularly advantageous for medical image classification (Baharlouei, Zahra, Hossein Rabbani, and Gerlind Plonka, 2023).

Model: Wavelet Scattering Transform

Accuracy: 100% for binary, classification and 82.5% for multiclass

Strengths: Computational complexity decreased, invariant to transformations in orientation

3.9 An Application of Machine Learning in the Diagnosis of Retinal Diseases Using OCT Scan

The author Aggarwal in his study from 2019, applied deep convolutional neural networks (CNNs) for the automated detection of retinal diseases using the OCT images. This paper also explored data augmentation enhancement techniques, including rotation and flipping in improving the performance of the model. The augmented CNN kept an average of 97% accuracy, which is consistently higher compared to the accuracies of the non-augmented version. This approach demonstrated that using overview data in combination with the original dataset could enhance model classification efficiency by up to three (Aggarwal, Pushkar, 2019).

Model: CNN with data augmentation

Accuracy: 97% (OCTID)

Strengths: Another area of successful application of data augmentation aimed at increasing performance.

3.10 Adaptive Window-based Feature Extraction for Dry AMD Detection

Sahoo et al. (2023) proposed a new adaptive window-based feature extraction process for diagnosing Dry Age-related Macular Degeneration (Dry-AMD). The method accompanied by the weighted ensemble classifier had a high classification accuracy of the OCTID and SD-OCT Noor databases. The advanced method applied a superposition of median filters to enhance the contrast of the RPE layer and performed segmentation on the same layer to classify medical images using adaptive window-based approaches (Sahoo, Moumita, Madhuchhanda Mitra, and Saurabh Pal, 2023).

Model: Weighted Majority Voting Ensemble Classifier New.

Accuracy: 96.94% (OCTID)

Strengths: High accuracy with a new method of feature extraction

4. Research Gap

Despite the promising results of the reviewed models, several gaps remain:

a) Computational Complexity

Most deep learning models like the CNNs and the transformers have high computational demand and this might not work well for real-time use in clinical operations. In some cases, additional data preprocessing and feature extraction are required in some of the models, which more or less contributes to the computational cost.

b) Dependence on Labeled Data

Almost all the existing approaches of supervised learning involving big, labeled datasets such as OCTID entail huge annotation costs. The problem of labeling medical images is a time and cost-intensive process which hinders the scalability of these models.

c) Limited Exploration of Semi-Supervised Learning

However, while big data sets like OCTID exist, only supervised learning models have been considered up to now. Pretraining using a small amount of labeled data with other types of data makes semi-supervised learning a promising method of obtaining high performance for the classification of problems with a relatively small amount of labeled data. This is an obvious area of research gap, which in the future needs to be explored extensively.

d) Generalization to Different Modalities

As for the specific models that have considered video data, there are few, among them transformers at the moment, though the majority of the models are created for static image recognition. To the author's knowledge, there is limited discussion on using OCT video-based diagnostics, even though OCT videos involve more temporal information which in turn increases diagnostic precision.

e) Generalization Across Diseases

Most models address the identification of certain diseases, like for example, AMD, CSR, or DME. Nevertheless, compared to models that are capable of detecting only one disease, models that are bidirectional and can detect multiple diseases are scarce. This lack of generalization can diminish the clinical value of these models as genuine clinical application may involve a choice of several possible diagnoses.

f) Explainability and Interpretability

Despite the relatively strong performance of deep learning models, these models are generally considered "black boxes," which remain a problem in healthcare and clinical practice, where it is necessary to understand the model or algorithm used. Most of the studies lack information on how their models function, which remains a major drawback in the use of the various models in practice.

Table 1. Comparison of the Previous Methods with Limitations

Paper	Methodology	Dataset	Accuracy	Limitations and Future Directions
Hassan et al. (2021)	CNN(AlexNet, ResNet18,GoogleNet) for CSR detection	OCTID	99.64% (AlexNet)	**Limitations**: Focused only on CSR; requires labeled data. **Future Directions**: Integrate semi-supervised learning to reduce reliance on labeled data; extend to multiclass detection.
Shaker et al.	Deep Dictionary Learning (DDL) and Wavelet Scattering Transform (WST)	OCTID	82.5% (5-class, WST)	**Limitations**: High computational complexity, and limited scalability. **Future Directions**: Implement semi-supervised WST to reduce computational load; extend to large datasets with unlabeled images.
Thomas et al. (2021)	Multipath CNN for AMD detection	OCTID,Mendeley, Duke, SD-OCT Noor	99.61% (OCTID)	**Limitations**: No discussion on computational costs. **Future Directions**: Explore real-time applicability; adapt for multiclass detection using semi-supervised learning.
Hammou et al. (2023)	Model-Based Transformer (MBT), Vision Transformer, Swin Transformer	OCTID	High F1-score, high precision	**Limitations**: Focuses on transformer-based models; limited video data. **Future Directions**: Extend to OCT video-based diagnostics with semi-supervised learning; reduce model complexity.
Mendes et al. (2021)	Segmentation using Euclidean Distance, Flood Fill for MH detection	OCTID	High segmentation accuracy	**Limitations**: Focused solely on MH detection; lacks multiclass capability. **Future Directions**: Develop models for general retinal disease detection (e.g., combining segmentation and classification).
Koseoglu et al. (2023)	CNN and Vision Transformers for AMD detection	OCTID	Over 90% (average)	**Limitations**: High computational needs, and reliance on large labeled data. **Future Directions**: Implement semi-supervised CNN to reduce reliance on labels; and enhance explainability for clinical use.
Marciniak and Stankiewicz (2023)	Transfer learning with VGG16 and Inception V3	LOCT dataset	96.28% (VGG16)	**Limitations**: Limited to four classes; does not generalize to small datasets like OCTID. **Future Directions**: Explore semi-supervised learning and expand to additional retinal conditions; test applicability to OCT video data.
Baharlouei et al. (2023)	Wavelet Scattering Transform (WST)	OCTID, TOPCON, Duke, Heidelberg	100% (binary); 82.5% (5-class)	**Limitations**: Limited performance in multiclass tasks, and limited generalization. **Future Directions**: Extend the WST model for multiclass classification using semi-supervised learning; reduce computational overhead.
Aggarwal (2019)	CNN with data augmentation for OCT image classification	OCTID	97%	**Limitations**: Requires large labeled data for training. **Future Directions**: Employ semi-supervised data augmentation; reduce the model's complexity for better scalability.
Sahoo et al. (2023)	Adaptive window-based feature extraction,Weighted Ensemble Classifier	OCTID, SD-OCT Noor	96.94%	**Limitations**: High complexity due to feature extraction and segmentation. **Future Directions**: Implement semi-supervised learning for adaptive window-based methods to reduce computational needs.

5.Future Directions: Semi-Supervised Learning in OCT Classification

Based on the recognized research gaps, it is possible to consider semi-supervised learning as a suitable solution, provided that it proposes a solution to the problem of reliance on a significant amount of labeled data and an increase in computation time. Regarding modeling, semi-supervised learning methods enable learning from a small subset of labeled data with a larger source of unlabeled data which decreases the need for additional labeling.

5.1 Semi-Supervised Learning Models

a) MixMatch

MixMatch jointly minimizes consistency regularization with pseudo-labels, in order for the model to incorporate the unlabeled data. By further improving its performance, MixMatch can alleviate the need for OCT images that are both labeled and numerous in tasks within the OCT classification.

b) FixMatch

FixMatch also comes up with a simplified pseudo-labeling strategy by employing a confidence threshold that defines pseudo-labels. This method could be especially valuable in medical applications, where the correct labeling is so important.

c) Self-training

Finally, self-training techniques make use of their prediction to assign pseudo labels to the unlabeled data. The above iterative approach can lead to better model performance on large and unlabeled datasets such as OCTID due to the difficulties involved in labeling such datasets.

5.2 Advantages of Semi-Supervised Learning

a) Reduced Annotation Effort

Semi-supervised learning reduces the required amount of labeled data which would make these models applicable where labeling is expensive or can hardly be done due to the size of the data set.

b) Improved Accuracy

Due to that, Semi-supervised learning is able to enhance the generalization of the learned models especially when faced with limited labeled data.

c) Scalability

Apart from this, semi-supervised learning methods can be scaled for large data sets and can allow the building of models encompassing all the overlapping pathological conditions of the retina possible in humans.

d) Applicability to Video Data

Generalizations of the semi-supervised approaches can also be employed to extend OCT video classification and such an approach enables the use of the temporal data available in the video sequences for developing a precise diagnosis.

Conclusion

This review has described the methodologies used in supervised learning models, performance, and limitations that accompany that technique applied to the OCTID dataset in particular and credit risk prediction problems in general. The models reviewed have shown good performance in the identification of retinal diseases namely AMD, CSR, DME, and MH with some of the models having the efficiency of over ninety-nine percent. Of course, there are certain drawbacks: many models are computationally expensive; their functioning is often based on large labeled datasets; there are issues with generalization to video data and multiple disease identification.

To that end, leveraging the availability of the limited labeled data, this work will incorporate semi-supervised learning techniques that reduce the reliance on labeled data, scalability, and increase model accuracy. Further work should involve the enhancement of SSCs that address both image and video data and make minimal or no use of human supervision, as well as the constant report of model explanations to make them more clinically applicable.

This way, the future models will enhance diagnostic reliability while at the same time cutting down on costs and time to practice actual OCT image analysis in real-world medical settings.

References

Aggarwal, Pushkar. "Machine learning of retinal pathology in optical coherence tomography images." *Journal of Medical Artificial Intelligence* 2 (2019).

Baharlouei, Zahra, Hossein Rabbani, and Gerlind Plonka. "Wavelet scattering transform application in classification of retinal abnormalities using OCT images." *Scientific Reports* 13.1 (2023): 19013.

Gholami, Peyman, et al. "OCTID: Optical coherence tomography image database." *Computers & Electrical Engineering* 81 (2020): 106532.

Hassan, Syed Al E., et al. "Deep learning-based automatic detection of central serous retinopathy using optical coherence tomographic images." *2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)*. IEEE, 2021.

Hammou, Badr Ait, et al. "MBT: Model-Based Transformer for retinal optical coherence tomography image and video multi-classification." *International Journal of Medical Informatics* 178 (2023): 105178.

Koseoglu, Neslihan Dilruba, Andrzej Grzybowski, and TY Alvin Liu. "Deep learning applications to classification and detection of age-related macular degeneration on optical coherence tomography imaging: a review." *Ophthalmology and Therapy* 12.5 (2023): 2347-2359.

Mendes, Odilon LC, et al. "Automatic Segmentation of Macular Holes in Optical Coherence Tomography Images." *IEEE Access* 9 (2021): 96487-96500.

Marciniak, Tomasz, and Agnieszka Stankiewicz. "Automatic diagnosis of selected retinal diseases based on OCT B-scan." *Klinika Oczna/Acta Ophthalmologica Polonica* 126.1 (2023): 8-14.

Shaker, Fariba, et al. "Application of Deep Dictionary Learning and Predefined Filters for Classification of Retinal Optical Coherence Tomography Images.".

Sahoo, Moumita, Madhuchhanda Mitra, and Saurabh Pal. "Improved detection of dry age-related macular degeneration from optical coherence tomography images using adaptive window based feature extraction and weighted ensemble based classification approach." *Photodiagnosis and Photodynamic Therapy* 42 (2023): 103629.

Thomas, Anju, et al. "Automated Detection of Age-Related Macular Degeneration from OCT. Images Using Multipath CNN." *J. Comput. Sci. Eng.* 15.1 (2021): 34-46.

Copyrights