

Master thesis

In order to obtain

Research Master's Degree in Computer Science, Intelligent Systems

Presented and publicly defended in October 2023

By

MEHER BOULAABI

A Deep Learning System for Diabetic Retinopathy, Lesion Segmentation and Disease Grading

Composition of the jury

Mrs. Ines Bayoudh Saadi (MC at ENSIT)

Mrs. Fadoua Bouafifi Samoud (MA at FSN)

Mrs. Afef Kacem Echi (MC at ENSIT)

Mrs. Takwa Ben Aïcha Gader (PhD)

President

Reporter

Supervisor

Co-supervisor

Academic Year: 2022-2023

Abstract

In this project, we tackled two significant challenges using the IDRiD dataset. Our first challenge involved segmenting lesions, which we approached as a binary classification problem for the classes MA, EX, SE, and HE. We performed exceptionally by implementing rigorous data augmentation and enhancing class histograms through CLAHE. Our metrics, including Average IoU, Dice Coefficient, Accuracy, Precision, F1 Score, and Sensitivity, reached an impressive 99%. Additionally, since we had superior annotations for the HE, EX, and SE classes, we adopted Deeplabv3+, which led to outstanding segmentation results.

In our second challenge, we focus on classification, specifically grading images into five grades ranging from 0 to 4. We use the advanced Swin Transformer model to achieve superior results, carefully preprocess and enhance the data, and fine-tune the augmentation parameters. Our approach achieves a Training Accuracy of 95% and a Test Accuracy of 92%, which exceeds the latest research articles. Our innovation lies in utilizing the state-of-the-art Swin Transformer architecture, significantly contributing to our exceptional performance.

To summarize, we thoroughly test our methods in real-life situations to ensure they are practical and effective. Our groundbreaking results in segmentation and classification are achieved through a combination of advanced models, careful data preprocessing, and innovative augmentation strategies. This work highlights the significance of new approaches and systematic experimentation, providing valuable insights and paving the way for future advancements in medical image analysis.

Keywords: Diabetic Retinopathy, Swin Transformer, Deeplabv3+, Disease grading, Lesion segmentation.

Résumé

Dans ce projet, nous avons abordé deux défis majeurs en utilisant l'ensemble de données IDRiD. Notre premier défi concernait la segmentation des lésions, que nous avons abordée comme un problème de classification binaire pour les classes MA, EX, SE et HE. Nous avons obtenu des performances exceptionnelles en mettant en œuvre une augmentation rigoureuse des données et en améliorant les histogrammes de classe grâce à CLAHE. Nos mesures, y compris l'IoU moyenne, le coefficient de disque, l'exactitude, la précision, le score F1 et la sensibilité, ont atteint un niveau impressionnant de 99%. En outre, étant donné que nous avons des annotations supérieures pour les classes HE, EX et SE, nous avons adopté Deeplabv3+, ce qui a conduit à des résultats de segmentation exceptionnels.

Dans notre deuxième défi, nous nous sommes concentrés sur la classification, en particulier le classement des images en cinq niveaux allant de 0 à 4. Pour obtenir des résultats supérieurs, nous avons utilisé le modèle Swin Transformer avancé, prétraité et amélioré soigneusement les données, et ajusté les paramètres d'augmentation. Notre approche a permis d'obtenir une précision d'apprentissage de 95% et d'une précision de test de 92%, ce qui dépasse les derniers travaux de recherche. Notre innovation réside dans l'utilisation de l'architecture Swin Transformer de pointe, ce qui a contribué considérablement à nos performances exceptionnelles.

Pour résumer, nous avons testé en profondeur nos méthodes dans des situations réelles pour nous assurer qu'elles sont pratiques et efficaces. Nos résultats novateurs en segmentation et en classification sont obtenus grâce à une combinaison de modèles avancés, à un pré-traitement minutieux des données et à des stratégies d'augmentation novatrices. Ce travail souligne l'importance des nouvelles approches et de l'expérimentation systématique, fournissant des informations précieuses et ouvrant la voie aux progrès futurs en matière d'analyse d'images médicales.

Mots clés : Rétinopathie Diabétique, Swin Transformer, Deeplabv3+, Classement de la maladie, Segmentation de lésion.

ملخص

في هذا البحث، نقدم تقنيات متقدمة للتعرف على كُتّاب يكتبون بأنماط كتابة متنوعة، باستغلال قوة التعلم العميق. تستخدم الأنظمة المقترحة نماذج رؤية مثل ResNet50 و ResNet101 و Swin Transformer و ResNeSt-50. يعرف ResNet50 و ResNet101 بفعالتهما في ميدان رؤية الحاسوب. يتميز Swin Transformer بقدرته على التعامل مع التباينات ونمذجة التبعية على المدى البعيد، مما يمكنه من التقاط السياق وإجراء تنبؤات قوية. من خلال التدريب المكثف على مجموعات واسعة من البيانات عينات النص اليدوي، يعمل Swin Transformer على تحليل تسلسلات من صفائح الصور، ويتعلم إنشاء تمثيل قوي لأسلوب كل كاتب بشكل فريد. من ناحية أخرى، يتيح ResNeSt50 مع طبقاته المتعددة التعلم من التمثيلات المعقدة لأسلوب فريد للكاتب والتمييز بين الأنماط المختلفة للكتابة بدقة كبيرة. يساعد وحدة SE في ResNeSt النموذج على التركيز على سمات الكتابة اليدوية المميزة مع تقليل الضوضاء. التجارب العملية تظهر أداءً استثنائيًا. على قاعدة CVL (تحتوي على صور أسطر نصية)، وصلت نماذج ResNet101 و Swin Transformer و ResNeSt50 على التوالي إلى دقة تبلغ ٢٣.٩٣% و ٥٠.٨٩% و ١٦.٦٩%. باستخدام عينات الصور من ٥٠ كاتبًا بعدد كبير من الصور، بلغت دقة ResNeSt50 ٤٠.٩٨% وبلغت دقة ResNet50 ٩.٩٠% على قاعدة بيانات IAM. بالإضافة إلى ذلك، بلغت دقة ResNeSt50 ٣٠.٩٧% على مجموعة بيانات IAM الكاملة التي تحتوي على ٦٥٦ كاتبًا. تقدم هذه البحث تقدمًا في مجال التعرف على الكتاب عن طريق إظهار فعالية هذه النماذج. تسلط الدقة المحققة الضوء على إمكانية هذه النماذج في معالجة وفهم الكتابة اليدوية المعقدة بفعالية، مما يفتح الباب أمام تطبيقات مستقبلية في مجال التعرف على الكتاب وتحليل النص اليدوي.

كلمات مفاتيح: كلمات مفتاحية: تعرف الكتاب، التعلم العميق، ResNet50 ، ResNet101 ، Swin Transformer ، ResNeSt50 ، تحليل الكتابة اليدوية.

Dedications

To my dear parents

For all your sacrifices, unwavering love, boundless tenderness, steadfast support, and heartfelt prayers that have guided me through my educational journey..

To my dear brothers,

For your constant encouragement and unwavering moral support, helping me overcome challenges and reach for the stars.

To my dear friends,

For your enduring friendship, support, and motivation that have been my pillars of strength, inspiring me to push my boundaries and achieve my goals.

To all my family,

For standing by my side, offering endless support, and believing in my dreams throughout my university career. Your collective presence has been my greatest source of strength and inspiration.

Acknowledgments

I would like to express my sincere thanks to the people who have helped me and contributed to the elaboration of this project and to the success of this academic year.

I would like to warmly thank ***Dr. Afef Kacem Echi*** who allowed me to benefit from their guidance, the advice they provided us, the confidence they showed us were crucial in the realization of my project.

My thanks also go to ***Dr. Sameh Mbarek*** for his help in understanding the medical concepts in our project and also in studying our data.

I would like to sincerely thank ***Dr. Takwa Ben Aïcha Gader*** for the time they devoted throughout this period to all my questions.

Finally, I would like to take this opportunity to thank the members of the jury who agreed to assess my final project. I hope that they will find in this thesis the insight and motivation that they are looking for.

Contents

Abstract	i
Résumé	ii
General Introduction	1
0.1 Diabetic Retinopathy: Medical Context	1
0.2 Motivation: Early Detection of Diabetic Retinopathy	2
0.3 Objective: Deep learning for Diabetic Retinopathy	3
0.4 Outline for Manuscript	4
1 State of the art in DL for DR Screening	5
1.1 Introduction	5
1.2 Diabetic Retinopathy	5
1.3 Diabetic Retinopathy in Tunisia	7
1.4 Computer-aided Application for Diabetic Retinopathy	7
1.5 Exploring Deep Learning	8
1.5.1 Main Models in Deep Learning	8
1.5.2 Main functions in Deep Learning Models	9
1.5.3 Deep Learning in Computer Vision	9
1.6 Deep Learning for Diabetic Retinopathy Analysis	11
1.6.1 Screening and Recognition	12
1.6.2 Retinal Blood Vessel Segmentation	13
1.6.3 Detection	14
1.6.4 Classification of the Lesions	16
1.7 Overview	17
1.8 Performance measures	17
1.9 Conclusion	20

2	DeepLabV3+ based System for Lesion Segmentation	21
2.1	Introduction	21
2.2	IDRiD : The Used Database	22
2.3	Proposed DeepLabV3+ for DR Segmentation	23
2.4	Scenario 1: Integration of multiple classes into a single mask	24
2.4.1	Combining multiple classes into a single mask	25
2.4.2	Analysis and comparison of the segmentation results at each stage of the preprocessing	30
2.4.3	Overview of Scenario 01:	30
2.5	Scenario 2: Enhancement of Class-specific binary segmentation	31
2.5.1	Process for scenario 2: Class-specific binary segmentation	32
2.5.2	EXUDATES (EX)	33
2.5.3	HEMORRHAGES (HE)	38
2.5.4	MICROANEURYSMS (MA)	43
2.5.5	SOFT EXUDATES (SE)	47
2.6	Comparison with Related Works	52
2.7	Conclusion	55
3	Swin Transformer-based System for Diabetic Retinopathy Grading	56
3.1	Introduction	56
3.2	Proposed Swin Transformer for DR Grading	57
3.3	Scenario 01: Investigating the Untouched Terrain of Original Images	57
3.3.1	Data preprocessing steps	58
3.3.2	Model selection and training	61
3.3.3	Result analysis	62
3.4	Scenario 02	68
3.4.1	Preprocessing Steps:	68
3.4.2	Data Augmentation: Enhancing Model Robustness in the Face of Limited Data	71
3.5	Analyzing the Enhanced Grade Distribution in the Training Set	72
3.5.1	Result analysis	74
3.5.2	Discussion of the differences in performance between Scenarios 1 and 2	79
3.6	Comparison to Related Works	80
3.6.1	Classification of DR Lesions using Deep Features	80
3.7	Conclusion	83
	Conclusion and perspectives	84
	Bibliography	90

List of Tables

1.1	Description of Metrics.	19
2.1	Evaluation Metrics for the Model.	28
2.2	Evaluation Metrics of EX.	34
2.3	Evaluation Metrics of EX after Data augmentation.	36
2.4	Comparison of Evaluation Metrics of EX.	36
2.5	Evaluation Metrics of HE.	39
2.6	Evaluation Metrics of HE after data augmentation.	40
2.7	Comparison of Evaluation Metrics of HE.	41
2.8	Evaluation Metrics of MA.	43
2.9	Evaluation Metrics of MA after Data augmentation.	44
2.10	Comparison of Evaluation Metrics of MA.	45
2.11	Evaluation Metrics of SE.	48
2.12	Evaluation Metrics of SE after data augmentation.	49
2.13	Comparison of Evaluation Metrics of SE.	50
2.14	Overview of reported techniques for DR lesion segmentation.	53
3.1	Classification Report for the Five Grades.	64
3.2	Training and Test Metrics.	65
3.3	Classification Report for the Five Grades.	76
3.4	Training and Test Metrics for the Model.	76
3.5	Reported deep feature extraction methods and classifiers used in various studies.	81

List of Figures

1.1	(a) the eye structure of a non-DR patient, (b) the eye structure of a patient with DR.	6
1.2	Comparison of semantic segmentation, classification and localization, object detection, and instance segmentation [22].	11
2.1	Color fundus photograph showing various retinal lesions associated with diabetic retinopathy, including microaneurysms, soft exudates, hemorrhages, and hard exudates [34].	22
2.2	Image-Mask example from the database.	25
2.3	DeepLabV3+ Architecture.	26
2.4	Model Evaluation.	27
2.5	Predicted images for each preprocessing.	29
2.6	Predicted and Ground-truth images.	34
2.7	Training and validation Loss and accuracy curves.	35
2.8	predicted images for each preprocessing.	37
2.9	Predicted and Ground-truth images of HE.	39
2.10	Predicted and Ground-truth images of HE.	40
2.11	Predicted images for each preprocessing.	42
2.12	Predicted and Ground-truth images of MA	44
2.13	Predicted and Ground-truth images of MA.	45
2.14	Predicted images for each preprocessing.	46
2.15	Predicted and Ground-truth images of SE.	48
2.16	Predicted and Ground-truth images of SE.	49
2.17	Predicted images for each preprocessing.	51
3.1	Overall Swin Transformer architecture, Adapted from [24].	61
3.2	The Swin Transformer block, with 2 sub-units. The first sub-unit applies W-MSA, and the second applies SW-MSA, Adapted from [24].	62
3.3	Evaluation of the model.	63
3.4	Validation loss and validation accuracy.	63
3.5	Confusion matrix.	66

3.6	Precision, Recall, and F1-score for each class.	67
3.7	All data: distribution of samples across all five grades.	69
3.8	Training Data: distribution of samples across all five grades.	69
3.9	Validation Data: distribution of samples across all five grades.	70
3.10	Testing Data: distribution of samples across all five grades.	70
3.11	Class distribution in the balanced training data.	72
3.12	Train and Validation losses.	74
3.13	Validation loss and validation accuracy.	74
3.14	The confusion matrix of scenario 2.	77
3.15	Precision, recall, and F1-score for each class.	78