Ecole Nationale Supérieure d'Ingénieurs de Tunis

الجمهورية التونسية
وزارة التعليم العالي والبحث العلمي
جامعة تونس
المدرسة الوطنية العليا للمهندسين بتونس

# 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)** | **President** |
| **Mrs. Fadoua Bouafifi Samoud (MA at FSN )** | **Reporter** |
| **Mrs. Afef Kacem Echi (MC at ENSIT)** | **Supervisor** |
| **Mrs. Takwa Ben Aïcha Gader (PhD)** | **Co-supervisor** |

**Academic Year: 2022-2023**

5, Avenue Taha Hussein – Tunis
5 B. P. 56, Bab Menara 1008

Tel.: 496 71 066 : الهات ف
Fax: 71 391 166 : فاكس

5 شارع طه حسين - تونس
ص. ب.: 56 باب منارة 1008

# 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 avions 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 كتّاب 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 بفعاليتهما في ميدان رؤية الحاسوب. يتميز Swin Transformer بقدرته على التعامل مع التباينات ونمذجة التبعيات 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% و٥٠.٨٩٠% و١٦.٦٩٩%. باستخدام عينات الصور من ٥٠ كاتبًا بعدد كبير من الصور، بلغت دقة ResNeSt50 ٤٠.٩٨% وبلغت دقة ResNet50 ٩٠.٩٠% 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 الكتّاب 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 الكتّاب وتحليل النص اليدوي.

---

**كلمات مفاتيح** :كلمات مفتاحية: تعرف الكتّاب، 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

# List of Tables

# List of Figures