



*Division of Computing Science and Mathematics
Faculty of Natural Sciences
University of Stirling*

Project Title

Chiara Cagliari

September 2022

**Dissertation submitted in partial fulfillment for the degree of
Master of Science in Artificial Intelligence**

September 2022

Abstract

Summary of the dissertation **within one page**.

This template starts the page numbering at the foot of this page. While you are printing drafts, you might find it useful to add the printing date and time into the footer – to help you, and your supervisor, tell which version is most current.

It is suggested that the abstract be structured as follows:

- Problem: What you tackled, and why this needed a solution
- Objectives: What you set out to achieve, and how this addressed the problem
- Methodology: How you went about solving the problem
- Achievements: What you managed to achieve, and how far it meets your objectives.

Attestation

I understand the nature of plagiarism, and I am aware of the University's policy on this.

I certify that this dissertation reports original work by me during my University project except for the following (*adjust according to the circumstances*):

- The technology review in Section 2.5 was largely taken from [17].
- The code discussed in Section 3.1 was created by Acme Corporation (www.acme-corp.com/JavaExpert) and was used in accordance with the licence supplied.
- The code discussed in Section 3.5 was written by my supervisor.
- The code discussed in Section 4.2 was developed by me during a vacation placement with the collaborating company. In addition, this used ideas I had already developed in my own time.

Signature:

(you must delete this, then sign and date this page)

Date

Acknowledgements

Acknowledge anyone that you wish to thank who has helped you in your work or supported you in any way: such as your supervisor, technical support staff, fellow students, external organisations or family. Acknowledge the source of any work that is not your own.

How to Create Table of Contents and List of Figures

The table of contents on the following page is automatically generated from the \part, \section, \subsection, etc., commands in the .tex file. Depending on your configuration, you may have to typeset your document twice in order to update this section (the first typesetting will create a .toc file, which is then used during the second typesetting).

Similarly you can automatically generate a list of figures from the \begin{figure} and \end{figure} commands in this document. Again, you may need to typeset your document twice in order for this list to update.

In your document, remove this entire section and only leave the lines

```
\tableofcontents  
\listoffigures
```

These automatically begin new pages that look like Chapters in the finished pdf.

Contents

Abstract	i
Attestation	ii
Acknowledgements	iii
How to Create Table of Contents and List of Figures	iv
1 Introduction	1
1.1 Background and Context	1
1.1.1 Diabetic Retinopathy	1
1.1.2 Explainability	3
1.2 Scope and Objectives	4
1.3 Achievements	4
1.4 Overview of Dissertation	4
2 State-of-the-Art	6
2.1 Diabetic Retinopathy Detection	6
2.1.1 Convolutional Neural Networks (CNNs)	6
2.1.2 Vision Transformers (ViTs)	6
2.2 Explainable AI	7
2.2.1 GradCAM	7
2.3 Available Datasets	7
3 Technical Chapters (change this to something appropriate)	10
3.1 First Section	12
3.1.1 First Subsection	12
3.1.2 Second Subsection	12
4 Conclusion	13
4.1 Summary	13

4.2 Evaluation	13
4.3 Future Work	13
Appendix 1	16
Appendix 2 – User guide	17
Appendix 3 – Installation guide	18

List of Figures

1.1	Proportion of cases of vision impairment that could have been prevented or addressed. Source: WHO [2]	1
1.2	Different microvascular lesions associated with Diabetic Retinopathy. Source: [6]	2
1.3	Stages of Diabetic Retinopathy. Source: APTOS 2019 Dataset [8]	3
3.1	Side by side confusion matrix for model evaluation	11
3.2	GradCAM applied to a sample image classified by the method proposed by [19], as implemented by [17]	11

1 Introduction

According to the most recent *World Health Organization (WHO)* report on vision [2], at least 2.2 billion people across the world are impacted by some vision impairment to a degree. Not only this, but many of these cases go undetected and therefore undergo no treatment or are addressed in an advanced stage of the disease. As pointed out by the *WHO* [2], certain subsets of the population are more likely to not receive appropriate care, due to factors like income, age or geographical location. The situation is also estimated to get worse over the next decade, due to the combination of population increase and the rise in incidence of the diseases most commonly associated with vision loss.

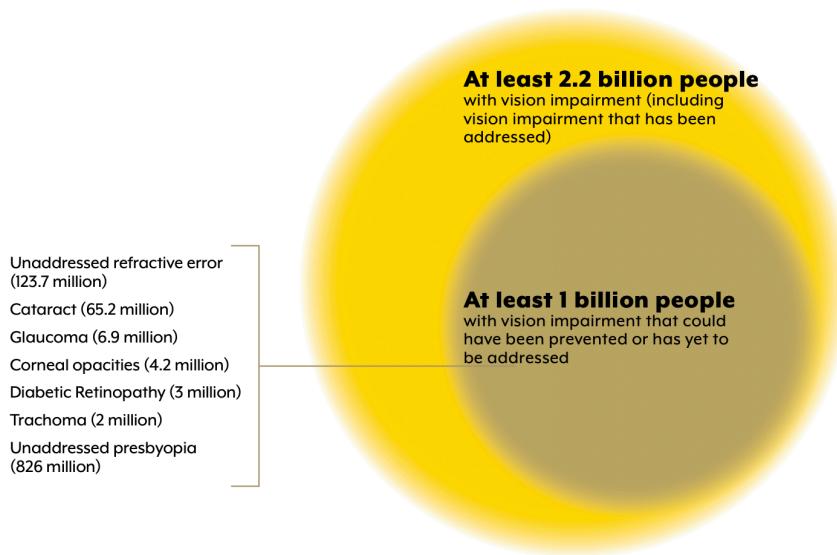


Figure 1.1: Proportion of cases of vision impairment that could have been prevented or addressed. Source: WHO [2]

As shown in *Figure 1.1*, one of the leading causes of vision impairment and blindness is *Diabetic Retinopathy*. This condition, which is a consequence of the incidence of *Diabetes Mellitus*, is destined to increase in incidence due to the rise in number of diabetic patients, reaching 191 million by 2030 [5]. The next section expands on the medical definition of *Diabetic Retinopathy*, and illustrates the challenges related to the integration of automated detection systems in clinical settings.

1.1 Background and Context

1.1.1 Diabetic Retinopathy

Diabetic Retinopathy is defined as a complication of the incidence of *Diabetes Mellitus*, and it is in many cases responsible for severe vision loss. Blood vessels in the eye suffer

ruptures caused by exceptionally high blood sugar levels, therefore preventing the retina from receiving an adequate blood supply to function properly. Since the retina is the structure within the eye responsible for the conversion of light into what we perceive as images, any damage to it will result in vision loss.

Early detection of *Diabetic Retinopathy* is crucial, since starting an appropriate treatment can prevent the condition from spreading and causing ulterior damage. As with *Diabetes Mellitus*, Retinopathy has no cure, but can be treated in a variety of ways, including through laser therapy and injections.

Diabetic Retinopathy is classed as a *microvascular disease* [3], and as such its diagnosis originates from the detection of a variety of *microvascular lesions*. The lesions traditionally associated with *Diabetic Retinopathy* are the following:

- *Microaneurysms*: small ruptures of blood vessels within the eye, usually detected through a dilated eye exam. Microaneurysms are considered to be the most prominent symptom of Retinopathy.
- *Hemorrhages*: can be the result of different complications, including vein occlusion caused by high blood sugar. Within the context of *Diabetic Retinopathy*, we look for *Intraretinal Hemorrhages*, which appear as dense, dark red, and sharply outlined [1].
- *Exudates*: lipid residues originating from leaking damaged capillaries.
- *Hard exudates*: exudates composed of extracellular lipid and usually found in the outer layer of the retina.

Figure 1.2 shows an example of these lesions:

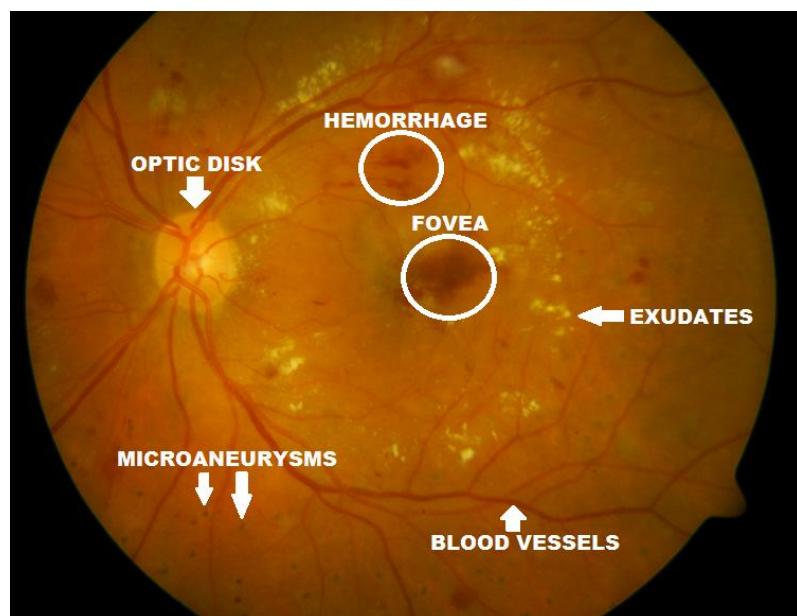


Figure 1.2: Different microvascular lesions associated with Diabetic Retinopathy. Source: [6]

Clinical grading for Diabetic Retinopathy can be provided according to different standards, of which the one proposed by the ETDRS [7] is the most popular. The grading scale used alongside the majority of public datasets resembles that of the ETDRS, although it does not distinguish between Non-High Risk Proliferative DR and High-Risk Proliferative DR. Such grading scale is the *International Clinical Diabetic Retinopathy Disease Severity Scale* [4], and it recognizes Diabetic Retinopathy at five different stages:

- No DR: no fundus alterations attributed to diabetes.
- Mild non-proliferative DR: a few microaneurysms are present.
- Moderate non-proliferative DR: presence of microaneurysms, intraretinal hemorrhages, and non-severe venous beading.
- Severe non-proliferative DR: large hemorrhages, severe venous beading, or severe intraretinal abnormalities.
- Proliferative DR: neovascularization of the structures of the eye, including the retina and the optic disk. When Proliferative DR is diagnosed, it is also necessary to assess the presence of *Macular Edema*.

Such a scale will be used as a reference for automated disease grading.

Figure 1.3 shows examples of each stage on training images from the APTOS 2019 Dataset [8].

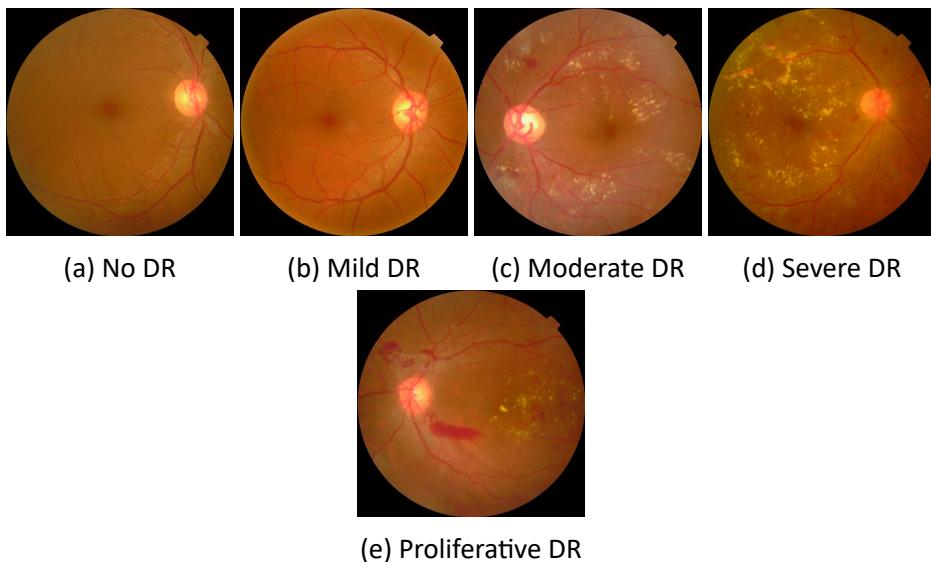


Figure 1.3: Stages of Diabetic Retinopathy. Source: APTOS 2019 Dataset [8]

1.1.2 Explainability

A well-known issue related to the employment of Machine Learning Classification methods is the fact that models are usually regarded as "black-box". This is because although they often achieve remarkable results, it is not always clear which features contribute

more or less to the final classification. If this is less of an issue in the majority of domains, in the context of a medical diagnosis being able to corroborate the output of the model is a prerogative.

In order for a Machine Learning model to be employed in a clinical setting, it is necessary for the medical equip to reach a high confidence level in its performance. Explainable AI comprises a set of tools that can help move in this direction.

That of **Explainable AI (XAI)** is a newborn field within Artificial Intelligence and it aims to address and solve the issue of result interpretability.

Usually, the more complex the model, the least explainable it is. At the same time, in most cases the more complex the model, the most accurate it is [12]. There is definitely a balance to strike between these two factors, especially in high-risk fields such as that of medical diagnosis.

1.2 Scope and Objectives

The aim of this project is to provide an evaluation of explainable

1.3 Achievements

Summarise what you have achieved.

1.4 Overview of Dissertation

The dissertation is organised in the following sections:

- Section 1: **Introduction**: an overview of Diabetic Retinopathy and retinal imaging analysis is provided. This section touches upon the medical context surrounding this project, and it discusses the objectives of the dissertation.
- Section 2: **State Of the Art**: a survey of current literature on Diabetic Retinopathy detection is provided. Both well-established and innovative methods are analysed, and an evaluation of current developments in Explainability is carried out.
- Section 3: **Methodology**: this section illustrates the process of implementing and/or replicating State Of the Art methods in Diabetic Retinopathy detection. Furthermore, it explains the implementation and integration of Explainability techniques.
- Section 4: **Results and Evaluation**: a review of the achieved results is provided in this section. Algorithm performance is analysed and compared against benchmarks, and visual results are presented.

- Section 5: **Conclusion and Future Work:** a final discussion is provided in this section, alongside an evaluation of possible future developments and improvements of the work.

2 State-of-the-Art

This section presents a review of the current State-of-the-Art methods for Diabetic Retinopathy detection and grading, and it gives an overview of recent developments in Explainable AI applied to the medical field. The final section of this chapter will touch upon the different datasets available to the public.

2.1 Diabetic Retinopathy Detection

Several Deep Learning based approaches have been proposed in recent years for the detection and grading of Diabetic Retinopathy. A large part of the available literature focuses on the employment of Convolutional Neural Networks (CNNs) for classification of eye fundus images, which have been proven to yield optimal results. However, more complex and innovative solutions have been developed, together with methods to address low variability in the data, overfitting, image enhancement, and gradability. Some popular approaches include:

- Transfer learning:
- Ensemble methods:
- Ensemble methods:

2.1.1 Convolutional Neural Networks (CNNs)

A simple classification model is proposed by [18], and it aims to employ the Inception-v3 architecture for classification of preprocessed macula-centered retinal fundus images. In this case, the approach taken is that of transfer learning, in which the network is pre-instantiated with weights originating from training on a different dataset, ImageNet in this case. The final algorithm consisted of an ensemble of 10 networks trained on the same data and it was able to provide grading on 4 different levels of disease, including an indication of referable diabetic macular edema (ME). The model was trained on the EyePACS dataset and validated on both EyePACS and Messidor-2, achieving an average sensitivity of 88.3%, and an average specificity of 98.6%.

2.1.2 Vision Transformers (ViTs)

A more recent approach to image classification comes with the rise in popularity of **Vision Transformers**. A Transformer is a deep neural network which is heavily reliant on its self-attention mechanism [23]. This type of architecture is capable of achieving a similar or better performance compared to traditional CNNs, although the two are quite often used in conjunction. Within the context of retinal image classification, however, the model of

interest is ViT [24]. This specific architecture is regarded as a *Pure Transformer*, as it directly applies a Transformer to sequences of image patches.

The method proposed by Wu *et. al* [25] achieves an accuracy of 91.4%, specificity of 0.977, sensitivity of 0.926 and area under curve (AUC) of 0.986. The model (ViT) was trained to distinguish between the 5 traditional disease grades, and its structure is adapted from the work of Dosovitskiy *et al.* [24]. The retinal fundus images are split into non-overlapping patches and then converted into sequences through flattening and embedding operations. Such resulting sequences are then input to a series of multi-head attention layers. The model is trained on the EyePACS dataset, which undergoes preprocessing and data augmentation as usual.

2.2 Explainable AI

The recent literature on Explainable AI, specifically when applied in a medical context, highlights two main popular tools: GradCAM [16] and SIDU.

2.2.1 GradCAM

Gradient-weighted Class Activation Mapping (GradCAM) aims to highlight important features for the classification process by using class-specific gradient information [16]. [13] [14]
[15]

2.3 Available Datasets

There are quite a few publicly available datasets which can be used to train DR detection models, some of which also provide grading for the disease and a risk index for developing Macular Edema. *Table 2.1* shows a summary of the characteristics of each dataset. In order for the grading information to be reliable, all images have been subject to evaluation by multiple ophthalmologists.

The first two datasets worth of mention are the *Kaggle EyePACS* [9] and the *Kaggle APTOS 2019* [8] datasets, which were both assembled in the context of challenges held by Kaggle for the detection of Diabetic Retinopathy.

Both the Kaggle datasets provide high-quality data for the model to be trained on, with the only limitation being the absence of labels for the test sets. The *EyePACS* dataset is the largest available, with over 80.000 samples distributed across 5 different disease grading levels. Uniquely from the rest, this dataset provides specimens of both eyes for each

patient. Additionally, different devices were used to capture the images, contributing to the variability of the dataset. This has the potential to enhance the flexibility of the final model, allowing it to perform well on a wide variety of inputs. In order to be used for training, this dataset requires some attentive preprocessing.

Similarly, the *Aptos 2019* dataset provides DR grading across 5 different levels. The dataset was put together by Aravind Eye Hospital in India, with the hope to aid detection of Diabetic Retinopathy in rural areas. Like in the EyePACS dataset, the images also present noise and variability due to the different devices used for acquisition, but this time only one fundus image is available for each patient.

The *Messidor* and *Messidor-2* datasets were put together by aggregating acquisitions from different sources within the Messidor Consortium in France, within the context of a project funded by the French Ministry of Research and Defense. Specifically, the images within the dataset come from three different ophthalmologic departments, each using different equipment for the acquisition., which ensures variability in the data. Differently from the previous ones, the Messidor dataset only provides grading on four levels of disease, but it does provide the additional risk index for Macular Edema. However, grading on 5 levels was provided by a third party for the Messidor-2 dataset, alongside the risk index for ME.

The *DDR* dataset provides fundus images with image-level, pixel-level and bounding-box-level annotations [10]. Additionally tho the previous ones, this dataset also provides an annotation for images that were deemed ungradable.

Finally, the *ODIR* dataset comes from another Kaggle Competition and it is structured in a different way from the rest. Differently from the rest, it does not exclusively deal with Diabetic Retinopathy, but it contains sample images related to Glaucoma, Cataract, Age related Macular Degeneration, Hypertension, Pathological Myopia, other than a general class of abnormal samples [11]. The dataset is rather small (about 5000 samples), but offers a good starting point for the development of a multi-disease classifier.

Dataset	No. Samples	Features
Kaggle EyePACS	> 80.000	Fundus images for DR detection (graded 0-4)
Kaggle APTOS 2019	5590	Fundus images for DR detection (graded 0-4)
MESSIDOR	1200	Fundus images for DR detection (graded 0-3) and assessment of ME risk (graded 0-2)
MESSIDOR-2	1748	Fundus images for DR detection (graded 0-4) and assessment of ME risk (graded 0-2)
DDR	12522	Fundus images for DR detection (graded 0-5)
ODIR	5000	Fundus images for detection of: Diabetes (D), Glaucoma (G), Cataract (C), Age related Macular Degeneration (A), Hypertension (H), Pathological Myopia (M), Other diseases/abnormalities (O)

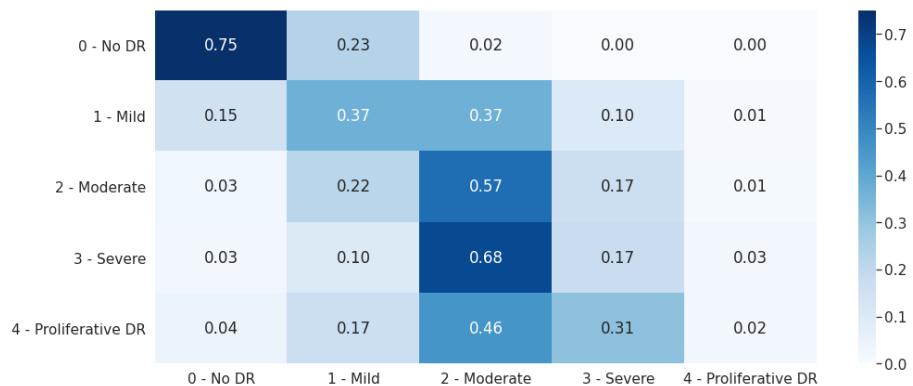
Table 2.1: Publicly available datasets.

3 Technical Chapters (change this to something appropriate)

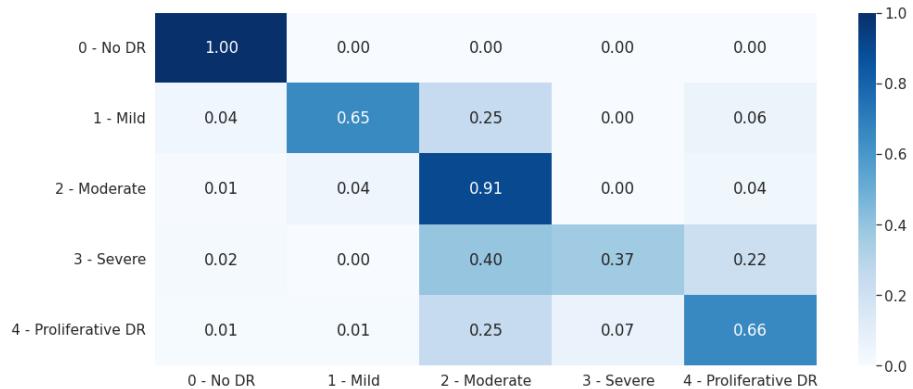
The CNN method proposed by Tymchenko et. al [19] was replicated, as implemented by [17]. The test accuracy score achieved was 0.872, which is considerably lower than the accuracy advertised by the authors, which is 0.928. However, the code implementation used for the replication advertises a Test Accuracy of 0.576.

The model was trained on the APTOS 2019 Dataset [8], although the test set was derived from the set of training images due to the lack of labels of the provided test set. This reduced the amount of samples in the training, validation, and test sets. Although in the original paper the main evaluation metric is the Kappa score, since the model is evaluated on the unlabeled test set, we only referred to the accuracy metric for comparison. Before training, the images go through a preprocessing stage which consists of cropping and resizing the samples. In addition, data augmentation was used to lower correlation between features and avoid overfitting.

The method proposed by the authors falls under the category of transfer-learning methods, as it makes use of a CNN architecture pretrained on the ImageNet Dataset [?].



(a) Original model



(b) Replicated model

Figure 3.1: Side by side confusion matrix for model evaluation

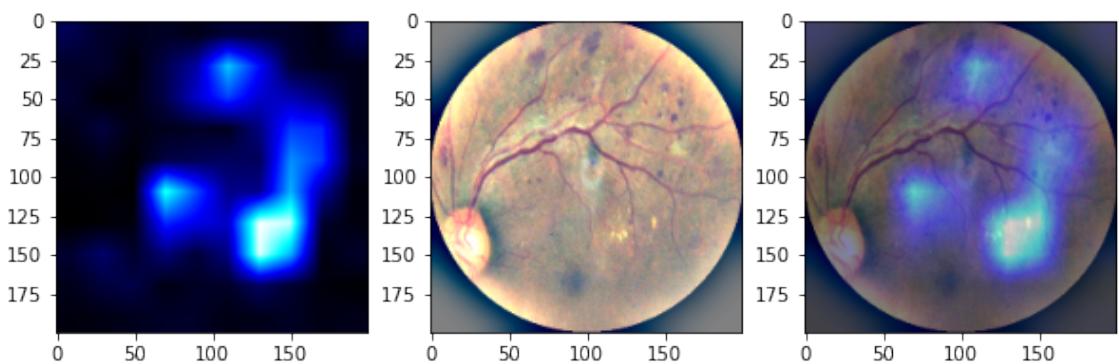


Figure 3.2: GradCAM applied to a sample image classified by the method proposed by [19], as implemented by [17]

3.1 First Section

3.1.1 First Subsection

3.1.2 Second Subsection

4 Conclusion

4.1 Summary

Summarise what you have achieved.

4.2 Evaluation

Stand back and evaluate what you have achieved and how well you have met the objectives. Evaluate your achievements against your objectives in section [1.2](#). Demonstrate that you have tackled the project in a professional manner.

4.3 Future Work

Bibliography

- [1] <https://www.ncbi.nlm.nih.gov/books/NBK560777/> acc: Jul 19
- [2] World report on vision. Geneva: World Health Organization; 2019. Licence: CC BY-NC-SA 3.0 IGO.
- [3] Wang, W.; Lo, A.C.Y. Diabetic Retinopathy: Pathophysiology and Treatments. *Int. J. Mol. Sci.* 2018;19, 1816
- [4] L. Wu, P. Fernandez-Loaiza, J. Sauma, E. Hernandez-Bogantes, E. Masis, *Classification of diabetic retinopathy and diabetic macular edema*, World journal of diabetes, 2013
- [5] Shu Wei Ting D., Chui Ming Cheung G., Wong T. Y., Diabetic retinopathy: global prevalence, major riskfactors, screening practices and public health challenges: a review, *Clinical and Experimental Ophthalmology* 2016; 44: 260–277
- [6] Júnior, Sérgio Welfer, Daniel. (2013). Automatic Detection of Microaneurysms and Hemorrhages in Color Eye Fundus Images. *International Journal of Computer Science and Information Technology*. 5. 21-37.
- [7] Early Treatment Diabetic Retinopathy Study Research Group. "Grading diabetic retinopathy from stereoscopic color fundus photographs—an extension of the modified Airlie House classification: ETDRS report number 10." *Ophthalmology* 98, no. 5 (1991): 786-806
- [8] <https://www.kaggle.com/c/aptos2019-blindness-detection>
- [9] <https://www.kaggle.com/c/diabetic-retinopathy-detection>
- [10] Li, Tao, et al. "Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening." *Information Sciences* 501 (2019): 511-522.
- [11] <https://www.kaggle.com/datasets/andrewmvd/ocular-disease-recognition-odir5k>
- [12] Gunning, D., Stefk, M., Choi, J., Miller, T., Stumpf, S., Yang, G. Z. (2019). XAI—Explainable artificial intelligence.
- [13] Niu, Y., et al., Explainable Diabetic Retinopathy Detection and Retinal Image Generation, *IEEE Journal of Biomedical and Health Informatics* (2021)
- [14] Muddamsetty, S. M., Jahromi, M. N. S., Moeslund, T. B., Expert level evaluations for explainable AI (XAI) methods in the medical domain, *Pattern Recognition. ICPR International Workshops and Challenges* (2021)
- [15] Kind, A., Azzopardi, G.. An explainable AI-based computer aided detection system for diabetic retinopathy using retinal fundus images, *International Conference on Computer Analysis of Images and Patterns*. Springer, Cham, (2019)

- [16] Selvaraju, Ramprasaath R., Abhishek Das, Ramakrishna Vedantam, Michael Cogswell, Devi Parikh, and Dhruv Batra. "Grad-CAM: Why did you say that?." arXiv preprint arXiv:1611.07450 (2016)
- [17] <https://github.com/debayanmitra1993-data/Blindness-Detection-Diabetic-Retinopathy>
- [18] D.S. Ting, C.Y. Cheung, G. Lim, G.S. Tan, N.D. Quang, A. Gan, H. Hamzah, R. Garcia-Franco, I.Y. San Yeo, S.Y. Lee, E.Y. Wong, *Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes*, Jama, 2017
- [19] B. Tymchenko, P. Marchenko, D. Spodarets, *Deep learning approach to diabetic retinopathy detection*, arXiv preprint, 2020
- [20] I. Qureshi, J. Ma, Q. Abbas, *Diabetic retinopathy detection and stage classification in eye fundus images using active deep learning*, Multimedia Tools and Applications, 2021
- [21] K. Shankar, A.R. Sait, D. Gupta, S.K. Lakshmanaprabu, A. Khanna, H.M. Pandey, *Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model*, Pattern Recognition Letters, 2020
- [22] P. Costa, A. Campilho, *Convolutional bag of words for diabetic retinopathy detection from eye fundus images*, IPSJ Transactions on Computer Vision and Applications, 2017
- [23] Han, K., Wang, Y., Chen, H., Chen, X., Guo, J., Liu, Z., ... Tao, D. (2022). A survey on vision transformer. IEEE transactions on pattern analysis and machine intelligence.
- [24] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S. and Uszkoreit, J., 2020. An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929.
- [25] Wu, J., Hu, R., Xiao, Z., Chen, J., Liu, J. (2021). Vision Transformer-based recognition of diabetic retinopathy grade. Medical Physics, 48(12), 7850-7863.

Appendix 1

You may have one or more appendices containing detail, bulky or reference material that is relevant though supplementary to the main text: perhaps additional specifications, tables or diagrams that would distract the reader if placed in the main part of the dissertation. Make sure that you place appropriate cross-references in the main text to direct the reader to the relevant appendices.

*Note that you should **not** include your program listings as an appendix or appendices. You should submit one copy of such bulky text as a separate item, perhaps on a disk.*

Appendix 2 – User guide

If you produced software that is intended for others to use, or that others may wish to extend/improve, then a user guide and an installation guide appendices are **essential**.

Appendix 3 – Installation guide

If you produced software that is intended for others to use, or that others may wish to extend/improve, then a user guide and an installation guide appendices are **essential**.