UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat d'Informàtica de Barcelona

FIB

Master Thesis

Prediction of Diabetic Retinopathy (DR), DR Progression and Relationship with Clinical Data using Optical Coherence Tomography Angiography (OCTA)

Name Ann Christin Rathert

E-mail ann.christin.rathert@estudiantat.upc.edu

Supervisors
ENRIQUE ROMERO MERINO and ALFREDO VELLIDO ALCACENA

October 2021

Abstract

Oftalmological images have shown promising results in diagnostic imaging in recent years. In this work, 2D OCTA cuts as well as 3D OCTA images and the radiomic features extracted therefrom will be analyzed using machine learning techniques to support diagnostic decision making in diabetes mellitus and diabetic retinopathy. For this purpose, a cross-sectional analysis of these medical images collected in a large scale prospective OCTA trial cohort and their respective radiomic features is performed and evaluated with classification models such as Logistic Regression, LDA, Linear SVC and RBF SVC for the respective classification problem (presence of DM and DR). 808 eyes from 404 patients were studied, of which 660 eyes belonged to diabetes mellitus patients and 148 eyes from healthy control eyes. The diagnosis of DM resulted in an AUC value of almost 0.8 with 2D OCTA cuts and 0.7 with 3D OCTA images by only considering the radiomic features. With the addition of clinical data, these values improve slightly.

For the diagnosis of DR, 461 eyes with DM but without DR and 199 with DR were available. Again, with the addition of clinical data, these values improve slightly. AUC values of just below 0.7 for 2D cuts and slightly lower values for 3D images were obtained. Similar to before, the results improve after adding the clinical data. Now, however, percentally stronger than before.

Radiomic techniques applied to OCTA images have shown accurate classification of DM and DR patients. The combination of retinal images and clinical data superior performance than induvidual tests only.

Contents

List of Figures

List of Tables

1 Introduction

Diabetes Mellitus (DM) is a metabolic disease of inadequate control of blood glucose levels, hyperglycemia [4]. DM is called the epidemic of the century [5] as it affects about 400 million people of working age (20-64 years) worldwide [6] with an increasing trend. While in 2019 Europe has a prevalence of 8,9 %, in 2030 this is already estimated to be 9,8 % and 10,3 % in 2045 [7]. There exist different classes of DM including Type 1, Type 2, gestational diabetes and other types, however, this work will focus exclusively on Type 1 Diabetes (T1D).

Diabetic Retinopathy (DR) is a complication of DM, that causes visual impairment and blindness by damaging the retina of the eye. Clasically the diagnosis of DR is clinical, by direct fundus examination in the slitlamp or by direct ophthalmoscopy. Recently, significant advances have been done in the retinal imaging field, and the detection of DR in fundus retinographies has gained relevance in the context of DR screening programs. Optical Coherence Tomography Angiography (OCTA) is a new, non-invasive imaging technique that generates volumetric angiography images of the retina [2]. The analysis of OCTA images and in particular the quantification of OCTA-derived parameters such as vessel density, perfusion density or foveal avascular zone technique has been proven effective to detect DR in DM patients [8]. This thesis will only be based in the diagnosis of DR on OCTA images.

This paper aims to analyse the diagnosis of DR in Type 1 Diabetic patients using radiomics applied to OCTA 3D images. The dataset being analysed is provided by the Ophthalmology department and the Diabetes Unit of Hospital Clínic and originates from a previous exploratory study of diabetic type 1 patients [8].

1.1 Background

As already mentioned in the beginning, this work will only focus on Type 1 DM patients. This type of DM, previously referred to as "juvenile diabetes", affects younger patients, with greater life expectancy and higher proportion of DR than type 2 DM, that is more common in the general population. Although T1D patients constitute only 5-10 % of of all people having diabetes [9], the increasing prevalence of T1D should not be underestimated. A recently published 25-year study shows that the incidence of type 1 diabetes is increasing by 3,4 % per year across Europe [10].

Patients with chronic hyperglycemia of diabetes suffer long-term damages as macrovascular-related stroke, ischemic heart disease, and peripheral artery disease and/or microvascular-related retinopathy, neuropathy, and nephropathy [11]. Microvascular complications of DM are the leading cause of new-onset blindness and vision loss among working-age adults [12]. DR is the most frequent diabetic eye disease and a major complication of DM. It causes visual impairment and blindness. From 1990 to 2010, DR patients registered an increase of 27 % in blindness and 64 % in visual impairment (10).

DR can be divided into various stages. The international classification [13] considers five stages: 1) No DR, 2) mild non-proliferative diabetic retinopathy

(NPDR), 3) moderate NPDR, 4) severe NPDR, and 5) proliferative diabetic retinopathy (PDR). In NPDR vascular neoplasms do not yet occur. It can be recognised by vascular bulges in the capillaries, deposits of fats from the blood plasma (so-called "hard exudates") or also frequent haemorrhages in the retina, depending on which sub-stage is present here. PDR is characterised by new formation of pathological blood vessels in the retina and vitreous humor (11).

The diagnosis of DR is clinical and is established by direct fundus examination by the eye doctor. Recently, non-invasive retinal imaging techniques such as fundus photographs Figure 4a), Optical Coherence Tomography (OCT) Figure 4b), or OCTA Figure 4c-d) can be used as a rapid and non-invasive screening method.

However, early detection is important. Even if a patient has good vision both subjectively and objectively, DR may already be present. The sooner DR is diagnosed and appropriate therapy is started, the greater the chances of success in maintaining a visual acuity that is still considered worth living with as DR blindness is irreversible. It has been shown that if annual retinal examinations are performed with, for example, fundus digital photography, blindness caused by DR can be reduced by more than 95 % [14].

DR screening can be performed by a variety of health professionals, including ophthalmologists, optometrists, general physicians, screening technicians and clinical photographers [15]. Although there seems to be a wide choice here, it should not be disregarded that there is a growing shortage of retina specialists and trained human graders in many countries. Add to this demographic change and increasing prevalence of DR patients, and image analysis becomes more and more costly and time consuming. For these reasons, research on artificial intelligence and autonomous diagnosis is increasing, since unmanned automated applications systems will be inevitable in the near future.

Deep learning and Machine Learning has already shown promising results in diagnostics in various fields such as image recognition, speech recognition and natural language processing. Especially in the health care sector, the application is growing rapidly. In ophthalmology, the first autonomous and assisted analysis of retinal photos was documented in 1990 [16]. More existing research in computational analysis in DR with radiomics is explored in more detail in the following sub-chapter Related Work.

1.2 Related Work

Most studies focus mostly on one sub-disease such as diabetic retinopathy, diabetic macular edema, cataracts, and glaucoma [12]. Further distinctions are rare, although important, especially for small subgroups that branch off from the main groups. Are there general differences between small groups and the overall result or is it not necessary to examine them separately? As already mentioned in the introduction, there are subgroups of diabetic types. Since only 5-10 % of diabetic patients have Type 1 diabetes, a subdivision is often suppressed in studies and publications or only Type 2 diabetes patients were considered. It is extremely important to study them separately, as the diagnosis of eye diseases in T1D patients is much more difficult than in T2D patients. And because they

make up such a low percentage, their possible poor diagnostic classifications are lost in the shuffle.

Since this thesis specialises in evaluation with radiomics features 3, the following section examines existing research and publications on radiomics in ophthalmology.

It can already be pointed out that radiomics has not yet been applied much in retinal images and there are few publications on it. Radiomics has its origins in radiology and oncology [17], where it has already found numerous applications since 2014 [18]. For retinal images only following papers are available:

In "Automatic Diabetic Foot Prediction Through Fundus Images by Radiomics Features" [19] the application is only directed on the disease Diabetic Foot by examining fundus images of the patients for Diabetic Foot prediction.

In "The development of an ophthalmologic imaging CADe structured report for retinal image radiomics research" [20] as well only fundus images are considered but it is not only focused on a particular disease, rather it addresses ophthalmic images in general.

"Discovery Radiomics With CLEAR-DR: Interpretable Computer Aided Diagnosis of Diabetic Retinopathy" [21] comes closest to this work, focusing on DR, but here as well only fundus images were used. CLEAR-DR stands for CLass-Enhanced Attentive Response Discovery Radiomic which is the basis for their radiomics-driven computer aided diagnosis (CAD). They have developed a system that is not only a decision support system for DR, but also indicates in the retinal images which are the decisive points for those decisions.

1.3 Objective

After giving an overview of the fact that very little research exists in the field of radiomics on DR and that radiomics has so far only been applied to fundus images in the ophthalmic field, the aims of this work will now be outlined.

- Gaining an understanding of OCTA images, their formation and the generation of 2D slices and 3D images
- Introduction to radiomics and its features
- Analyzing the obtained radiomic features with machine learning algorithms and creating a robust feature extraction scheme for DM and DR classification. The same as well for the combination of radiomic features and clinical data
- Understanding and comparing of results from OCTA 2D cuts and 3D as well as gaining insights in the relationship with Clinical Data
- Comparing these results with previous work using radiomics applied to fundus retinographies, single OCT sections and single OCTA images

1.4 Roadmap

The structure of the document will be as follows: after providing context in the introduction and presenting the aims of this work, the second chapter 2 will go into more detail about the given dataset and provide details about the retinal images. In order to analyse the data images, certain radiomics values are extracted out of them. These are explained in chapter 3. This is followed by the fundamentals (chapter 4) and the methods (chapter 5) and finally the results (chapter 6) and their conclusions (chapter 7) are presented in the very last part.

2 Retinal Images

Since the analysis in this paper is based exclusively on OCTA images, these will be considered in more detail in the present chapter.

OCTA images are a relatively new technology that has only been commercially available since 2014 [22]. Although OCTA has been analysed in many publications, there are still very few studies that evaluate its clinical utility [23]. This may be either because this imaging technique is one of the newest, or because of disadvantages such as susceptibility to artefacts that can occur during image acquisition [24]. Also, as will be explained below, the imaging time is longer, resulting in slower processing and storage times, and OCTA cannot visualise dye leakage, which is an indicator of inflammatory vascular pathologies.

Nevertheless, artefacts in the image capture can be prevented [22] and significant advantages bring new perspectives to conventional methods. Among these advantages, a better visualisation of the deep capillary plexus and choroid stands out [25], a depth-resolved in-vivo visualisation of the retinal microvessels. In addition, no time-consuming dye injection is necessary. This is not only in favour of the patient, who saves the medication of contrast dye and possible allergic reactions, but also provides higher contrast images. This leads to generally more defined vessels in the image.

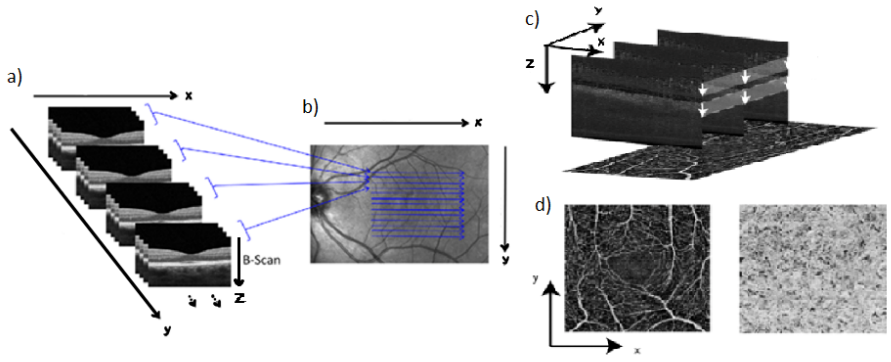
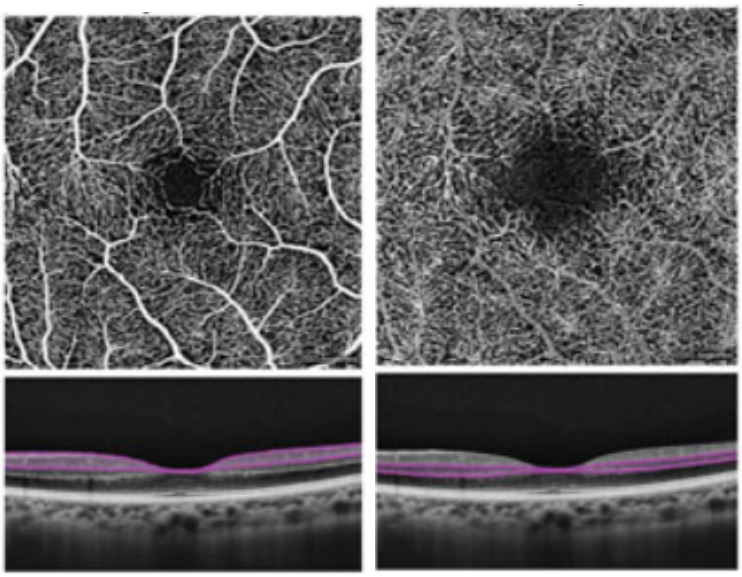


Figure 1: Scanning protocol for OCTA [1, 2]. a) OCT-scans on x- and y-axis to determine depth of vascular structure, b) top view, in which OCTA is marked by the blue arrows. The background is a fundus image, c)

Prior to OCTA, as briefly mentioned in the introduction, fundus imaging and OCT were the primary image technologies for visualising retinal vessels. In Figure 3 these are listed for comparison. Fundus images are two-dimensional views of the retinal vasculature, where Fundus literally stands for the back-part, so central retina, optic disc and macula are imaged. OCT is a three-dimensional cross-sectional imaging of the macula. In general, it can be said that OCTA is an extension of OCT. This is visualised in a straightforward manner in Figure Figure 1. Initially, OCT scans are made on the x- and y-axis, whereby several images are taken at the same point. In OCTAs, the z-axis is included, which



(a) Superficial (b) Deep

Figure 2: Different layers of OCTAs indicated by the purple outline of the corresponding OCT section: a) Superficial b) Deep [3]

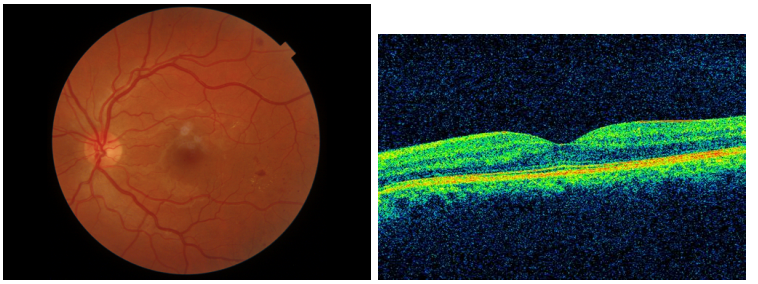
enables clinicians to determine the depth of vascular structure. The innovation in OCTA therefore is the use of motion contrast to detect blood flow within a few snapshots. Similar to a flip book, a sequence of individual images is strung together as a continuous movement. With in this work's datasets available 3x3-mm scans, there are 4 images strung together, with the 6x6-mm images there are two 4. In general 3x3-mm images capture an area of 3x3mm, while the 6x6-mm images capture a wider area of the macula.

OCTA thus takes so-called A-scans at the same point of the retina, which succeed at short intervals, called interscan times, in order to detect the changes in the OCT reflection signal due to the flow through the blood vessels [26]. The superficial and deep capillary plexus of the retina can be visualized separately (see Figure 2) and these settings can also be modified to further segment the retinal vasculature to focus on other layers where previously invisible pathologies may also be present.

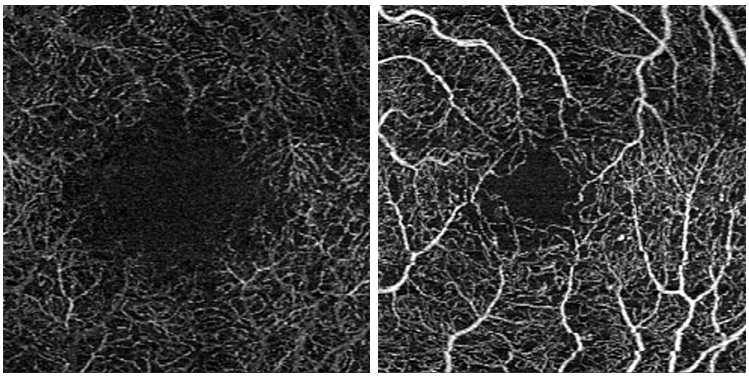
2.1 OCTA scans

Since in the previous work on this matter presented by Laura Carrera Escalé in collaboration with Dr. Javier Zarranz-Ventura, Anass Benali, Ruben Martin, Enrique Romero Merino and Alfredo Vellido Alcacena, the complete set of retinal images including fundus, OCT and enface 2D frontal OCTA images in the superficial and deep capillary plexi, divided into 3x3-mm and 6x6-mm respectively were studied, this work is specifically dedicated to OCTA using the flow OCTA cube data and structural OCTA cube data.

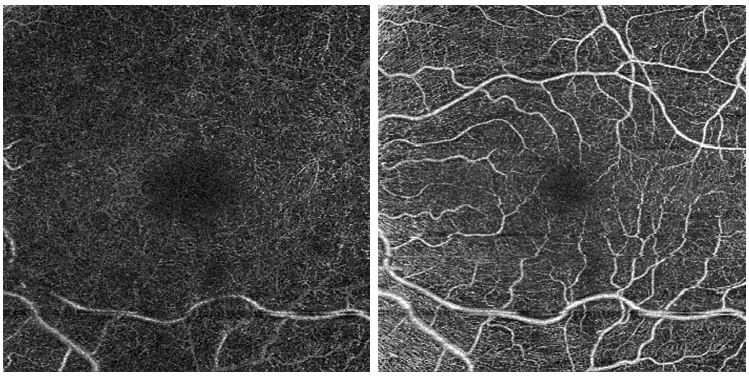
However, the retinal images used in this work are not the standard enface 2D



(a) Fundus (b) OCT



(c) OCTA 3x3-mm Deep (d) OCTA 3x3-mm Superficial



(e) OCTA 6x6-mm Deep (f) OCTA 6x6-mm Superficial

Figure 3: Image types: a) Fundus photography: provides an image that shows the "back" (fundus) of the eye, which includes the central retina, the macula and the optic nerve head, b) OCT: a non-invasive, rapid imaging technique that can depict cross-sectional structures of the retina to picture optical scattering of different retinal layers, c), d), e), f) Enface 2D OCTA images: Motion contrast images with those obtained from multiple B-scans at the same location to obtain a deep-resolution image of the retinal vessels. Images of the superficial and deep retinal layers can be obtained

frontal OCTA images as shown in Figure 3, instead cuts or sections of the 3D OCTAs cubes, displayed as flow OCTA cube data and structural OCTA cube data are used. For this purpose, the raw data storage file was converted in such a way that 2D slices are obtained from the binary file of the raws disk image. To do this, the OCT file converter by Mark Graham [27] was rewritten and adapted to OCTA.

The present data set of OCTAs is extracted from the AngioPlex (TM) system with a Cirrus HD-OCT5000 device. This is important to note as there are differences in acquisition output depending on which device and associated device hardware and software is used [23].

With the OMAG algorithm of the Cirrus device, a scan speed of 68000 A-scan/s is measured. The imaging depth is 2mm and as mentioned previously, there is a scan area of 3x3-mm and 6x6-mm [28]. For the A-scan count, there are 245 scans for 3x3-mm and 350 for 6x6-mm. As can also be seen in Figure 4, there are four repeated B-scan counts for 3x3-mm and two for 6x6-mm. Figure 4 shows one slice out of 245 scans for the 3x3-mm images and one out of 350 for the 6x6-mm images. This means that per patient and per eye there are 245 images each of superficial and deep captures with a volume of 3x3-mm and 350 images each for superficial and deep captures with a 6x6-mm frame. In total, 1190 images per patient and per eye are considered for 2D imaging.

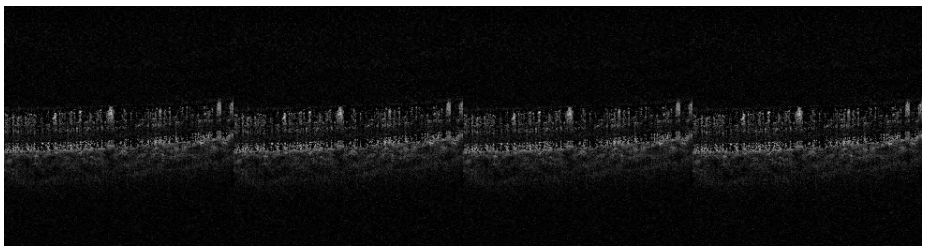
How this large amount of data is processed will be explained in more detail in the following chapters.

2.2 OCTA 3D

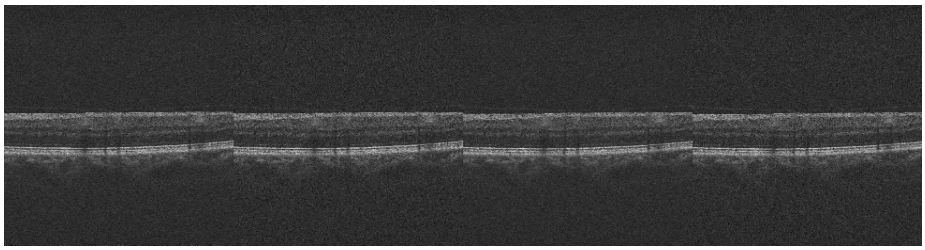
Before the methodology of the data analysis is presented, there is the second image acquisition to introduce. With the 2D slices obtained with the OCT-Converter, it was possible to convert them a second time with the help of a specialized program. Another time this program is initially designed for radiological and oncological medical images, but this can just as well be supplied with retinal images.

3D Slicer is a free and open source software package commonly used for medical, biomedical, and related imaging research [29]. 3D Slicer has a large amount of functionalities and also of additional extensions. However, these will not be considered further here. For this work, 3D Slicer was used exclusively for the creation of 3-dimensional Nrrd files. Nrrd stands for "nearly raw raster data" and was created for scientific visualization and image processing applications [30].

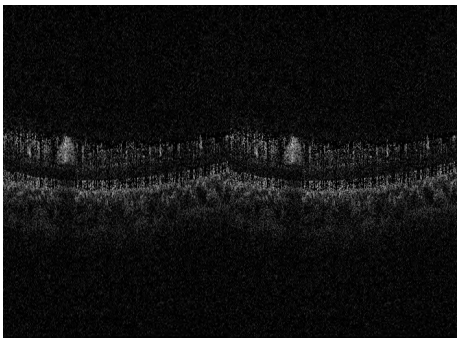
Figure 6 shows the display of the 3D Slicer program for the left eye of patient one, with the resulting Nrrd file which is depicted in the upper right corner. These are stored and created for each eye of the patients, for later analysis.



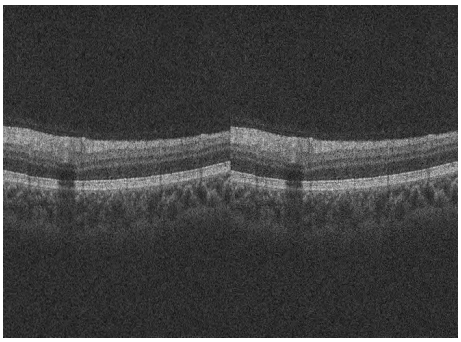
(a) Flow cube OCTA 3x3-mm



(b) Structural OCTA 3x3-mm



(c) Flow cube OCTA 3x3-mm



(d) Structural OCTA 6x6-mm

Figure 4: Image types of Dataset

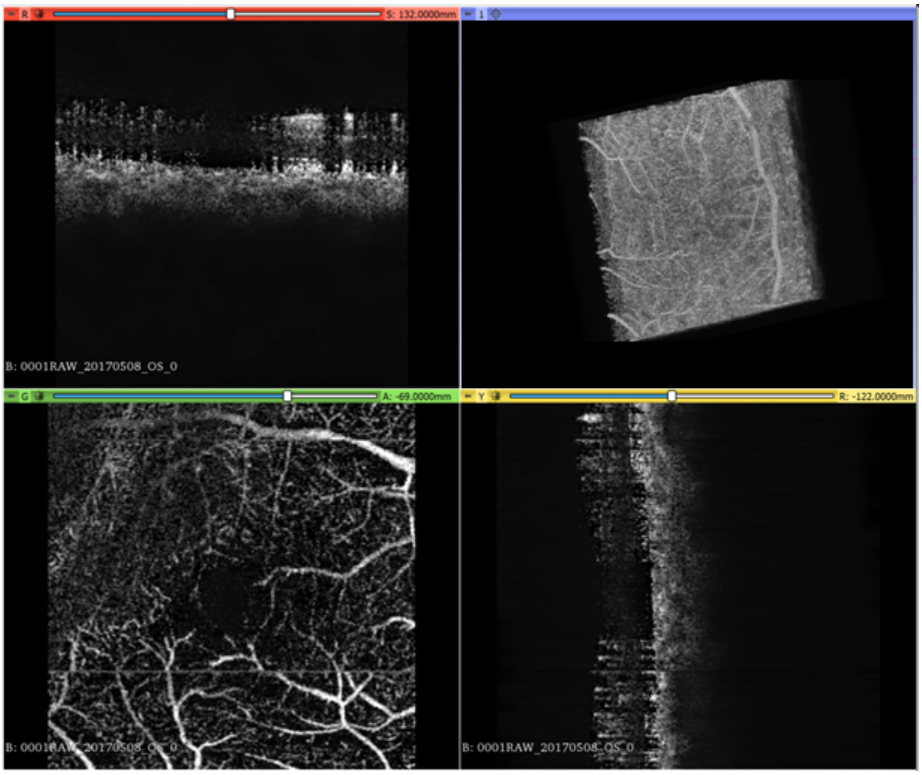


Figure 5: 3D Slicer Output with the original input in the upper left corner. Here there is a slider that makes it possible to scroll through either 245 or 350 images, depending on which area size is selected. At the bottom left, this creates the on top view, while at the bottom right, a side view is displayed. The Volume Rendering option results in the 3D view, as shown above right.

3 Radiomics

As already briefly touched upon in the previous chapters, the method of radiomics comes from the radiological and oncological medical fields and is also most widely used in these [17]. Magnetic resonance imaging (MRI), computed tomography (CT) and positron emission tomography (PET) are most commonly used in the latter [31]. It is generally assumed that there is far more information in medical images than can be seen by experts in image analysis with the bare eye, even if they are specially trained to observe such images [32]. This will be achieved by extracting mathematical extractions from spatial distributions of signal intensities and pixel relationships. The data obtained is then used as information in analysis methods from artificial intelligence [31]. Thus, an improvement of image analysis is created by an automated extraction of large amounts of quantitative features. These features are subdivided into 7 main categories [33]:

- First Order Features

- Shape Features 2D/3D
- Gray Level Co-occurrence Matrix (GLCM) Features
- Gray Level Size Zone Matrix (GLSZM) Features
- Gray Level Run Length Matrix (GLRLM) Features
- Neighbouring Gray Tone Difference Matrix (NGTDM) Features
- Gray Level Dependence Matrix (GLDM) Features

The respective extracted features of this work are listed in Table 3. More information can be found on the Radiomics documentation website [33].

Table 1: Radiomic Features

First-order Features (18)	GLCM (24)	GLZSM (16)	GLRLM (16)
1. Energy	1. Autocorrelation	1. Small Area Emphasis (SAE)	1. Short Run Emphasis (SRE)
2. Total Energy	2. Joint Average	2. Large Area Emphasis (LAE)	2. Long Run Emphasis (LRE)
3. Entropy	3. Cluster Prominence	3. Gray Level Non-Uniformity (GLN)	3. Gray Level Non-Uniformity (GLN)
4. Minimum	4. Cluster Shade	4. Gray Level Non-Uniformity Normalized (GLNN)	4. Gray Level Non-Uniformity Normalized (GLNN)
5. 10th percentile	5. Cluster Tendency	5. Size-Zone Non-Uniformity (SZN)	5. Run Length Non-Uniformity (RLN)
6. 90th percentile	6. Contrast	6. Size-Zone Non-Uniformity Normalized (SZNN)	6. Run Length Non-Uniformity Normalized (RLNN)
7. Maximum	7. Correlation	7. Zone Percentage (ZP)	7. Run Percentage (RP)
8. Mean	8. Difference Average	8. Gray Level Variance (GLV)	8. Gray Level Variance (GLV)
9. Median	9. Difference Entropy	9. Zone Variance (ZV)	9. Run Variance (RV)
10. Interquartile Range	10. Difference Variance	10. Zone Entropy (ZE)	10. Run Entropy (RE)
11. Range	11. Joint Energy	11. Low Gray Level Zone Emphasis (LGLZE)	11. Low Gray Level Run Emphasis (LGLRE)
12. Mean Absolute Deviation (MAD)	12. Joint Entropy	12. High Gray Level Zone Emphasis (HGLZE)	12. High Gray Level Run Emphasis (HGLRE)
13. Robust Mean Absolute Deviation (rMAD)	13. Informational Measure of Correlation (IMC) 1	13. Small Area Low Gray Level Emphasis (SALGLE)	13. Short Run Low Gray Level Emphasis (SRLGLE)
14. Root Mean Squared (RMS)	14. Informational Measure of Correlation (IMC) 2	14. Small Area High Gray Level Emphasis (SAHGLE)	14. Short Run High Gray Level Emphasis (SRHGLE)
15. Skewness	15. Inverse Difference Moment (IDM)	15. Large Area Low Gray Level Emphasis (LALGLE)	15. Long Run Low Gray Level Emphasis (LRLGLE)
16. Kurtosis	16. Maximal Correlation Coefficient (MCC)	16. Large Area High Gray Level Emphasis (LAHGLE)	16. Long Run High Gray Level Emphasis (LRHGLE)
17. Variance	17. Inverse Difference Moment Normalized (IDMN)		
18. Uniformity	18. Inverse Difference (ID)		
	19. Inverse Difference Normalized (IDN)		
	20. Inverse Variance		
	21. Maximum Probability		
	22. Sum Average		
	23. Sum Entropy		
	24. Sum of Squares		

2D Shape Features (9)	3D Shape Features (14)	NGTDM (5)	GLDM (14)
1. Mesh Surface	1. Mesh Volume	1. Small Dependence Emphasis (SDE)	1. Small Dependence Emphasis (SDE)
2. Pixel Surface	2. Voxel Volume	2. Large Dependence Emphasis (LDE)	2. Large Dependence Emphasis (LDE)
3. Perimeter	3. Surface Area	3. Gray Level Non-Uniformity (GLN)	3. Gray Level Non-Uniformity (GLN)
4. Perimeter to Surface ratio	4. Surface Area to Volume ratio	4. Dependence Non-Uniformity (DN)	4. Dependence Non-Uniformity (DN)
5. Sphericity	5. Sphericity	5. Dependence Non-Uniformity Normalized (DNN)	5. Dependence Non-Uniformity Normalized (DNN)
7. Maximum 2D diameter	6. Maximum 3D diameter	6. Gray Level Variance (GLV)	6. Gray Level Variance (GLV)
8. Major Axis Length	7. Maximum 2D diameter (Slice)	7. Dependence Variance (DV)	7. Dependence Variance (DV)
9. Minor Axis Length	8. Maximum 2D diameter (Column)	8. Dependence Entropy (DE)	8. Dependence Entropy (DE)
10. Elongation	9. Maximum 2D diameter (Row)	9. Low Gray Level Emphasis (LGLE)	9. Low Gray Level Emphasis (LGLE)
	10. Major Axis Length	10. High Gray Level Emphasis (HGLE)	10. High Gray Level Emphasis (HGLE)
	11. Minor Axis Length	11. Small Dependence Low Gray Level Emphasis (SDLGLE)	11. Small Dependence Low Gray Level Emphasis (SDLGLE)
	12. Least Axis Length	12. Small Dependence High Gray Level Emphasis (SDHGLE)	12. Small Dependence High Gray Level Emphasis (SDHGLE)
	13. Elongation	13. Large Dependence Low Gray Level Emphasis (LDLGLE)	13. Large Dependence Low Gray Level Emphasis (LDLGLE)
	14. Flatness	14. Large Dependence High Gray Level Emphasis (LDHGLE)	14. Large Dependence High Gray Level Emphasis (LDHGLE)

4 Fundamentals

In this chapter, fundamental knowledge is introduced that is used throughout the following chapters and will be essential to understand the methods and evaluations described there.

4.1 Data Splitting

Splitting data into training- and testingdataset is a technique which evaluates the performance of a machine learning algorithm.

A dataset is split into two subsets, the first subset being the training dataset used to fit the model. The second subset, the test dataset, acts as a collection of data points that evaluate whether the previously trained model generalizes well to new, and previously unknown, data. Thus, the goal is to estimate the performance of the machine learning model on new data, that has not been used to train the model. By creating different samples for training and testing, it is possible to evaluate the performance of the model.

A very common problem that occurs when training a model is overfitting. It occurs when a model performs very well with the training data, but does not generalize well to new, unknown data points. This may be due to noise and details of the training data being adopted and their concepts learned to an extent that they negatively affect the model's performance on new data. The more complex a model is, the more likely it is to overfit.

When a dataset is split into one set of training data and one set of validation data, the performance metrics of the models are highly reliant on these two sets, as they are trained and evaluated only once. Thus the performance depends on that one evaluation and may perform very differently when trained and evaluated on different subsets.

A method to avoid this, is presented in the following subsection.

4.2 Cross-Validation

Since there is never enough data to train a model on, k-fold cross validation introduces a method that makes it possible to validate models multiple times with different subsets.

This is done by splitting the data into k subsets along with repeating the so-called holdout method k times. This means that in each round, one of the k subsets is used as the test set and the other $k-1$ subsets are combined to form the training set. Finally, the performance is averaged over the multiple evaluations of all k trials, giving the overall model effectiveness.

A special case is stratified k-fold cross validation. Here, an unbalanced data set is present, so care is taken to include approximately the same percentage of samples of each target class in each fold.

To get more out of models, it is possible to tune hyperparameters, which can

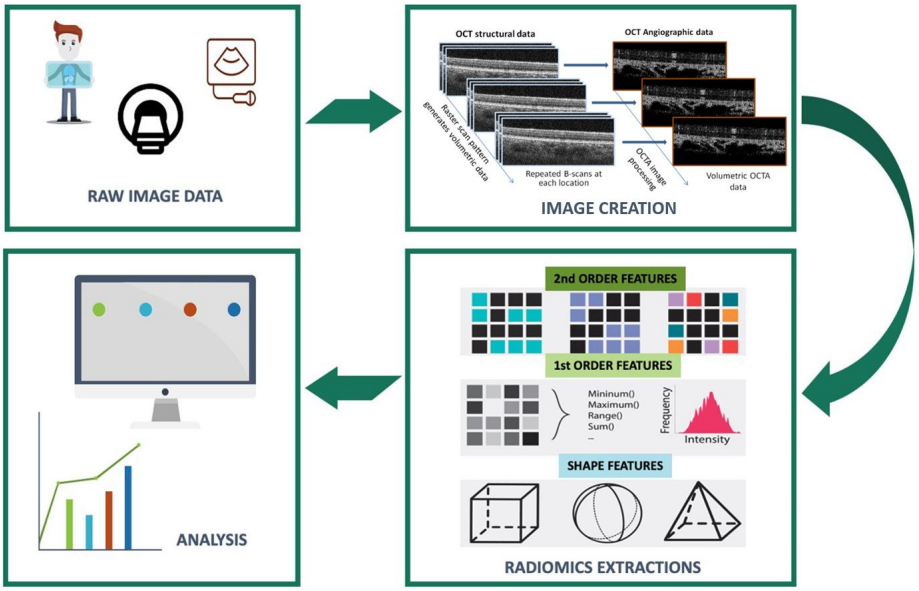


Figure 6: Radiomics procedure

be set before training a Machine Learning model. As there is no one-size-fits-all solution for finding optimum hyperparameters, these must be tested and optimized manually.

4.3 Nested Cross-Validation

Nested Cross-Validation, as the name suggests, nests cross-validation and hyper-parameter tuning for trying to overcome the problem of overfitting the training dataset. This is done by running a double loop, an outer and an inner Cross-Validation, where the outer loop will serve the assessing in quality of the model and the inner loop will serve for the parameter selection. If the outer loop is repeated five times, five different test sets will be generated. For each of these iterations, the outer train set will be split into four inner folds, if the inner loop is repeated four times. With having five outer folds and four inner folds, the total number of trained models will be 20.

4.4 Classification Model

A classification model tries to draw some conclusion from the input values given for training. The goal of classification is to accurately predict the target class for each case in the data.

The performance of a classification model is primarily dependent on the nature of the data. Given that the available dataset carries multiple predictors and that the complexity is high, it is difficult to select one single algorithm that would always work out well. Therefore, the usual practice is to try multiple models and figure out the suitable one. For this experiment, three Machine Learning algorithms that are widely known for classification problems were considered.

4.4.1 Logistic Regression

Logistic Regression utilizes the power of regression to do classification to model binary dependent variables. Unlike regression which uses the method of Least Squares, the model uses Maximum Likelihood to fit a sigmoid-curve on the target variable distribution. It has a strong power of explainability of several independent variables on a single outcome variable.

4.4.2 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is a robust, supervised classification method that attempts to model differences among samples assigned to certain groups. A linear discrimination function represents a hyperplane in the feature space to distinguish the classes with the aim to maximize the ratio of the between-group variance and the within-group variance. LDA is a simple and as well effective method for classifiaction.

4.4.3 Support Vector Machine

Support Vector Machine (SVM) is a linear or non-linear classifier based on the kernel which is used. It represents the training data as points in a space separated into categories in such a way that a wide area around the class boundaries remains free of points. It is a so-called Large Marging Classifier. If a linear kernel is used, the classifier and hence the prediction boundary are linear. For separating two classes, a line is drawn in a manner that achieves a maximum marging. This line is drawn at equal distance from both groups. An additional two lines are drawn on both sides, which are called support vectors. SVMs learn from the support vectors, unlike other machine learning models that learn from the true and false data. For non-linear problems, the kernel trick is used. A kernel is a function that takes the original non-linear problem and transforms it into a linear one within the higher-dimensional space. In this work the Radial Basis Function (RBF) Kernel was applied. It is the most generalized form of kernelization and shows similarities to the Gaussian distribution. SVM is very effective in high dimensional spaces.

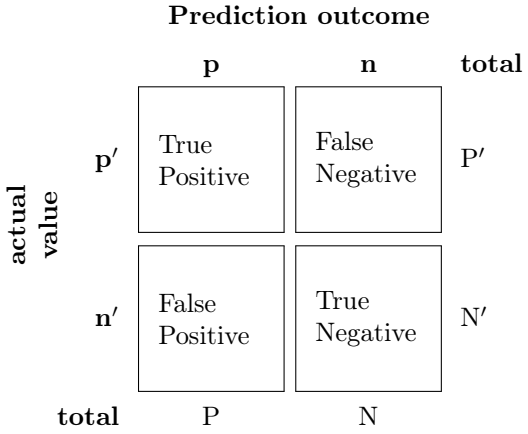
4.5 Evaluation Metrics for Classification

Evaluation metrics are essential when evaluating the performance of a model. Several of these should be considered, as it is possible for a model to perform well on one measurement of an evaluation metric but poorly on another.

4.5.1 Confusion Matrix

For the evaluation of classification task results, a contingency table, or also called confusion matrix, is used to represent the number of correctly and falsely predicted samples per class.

Four possible outcomes could occur while performing classification predictions: true positive (TP), true negative (TN), false positive (FP) and false negative (FN), for which TP means, that the model correctly predicts the positive class and TN the model correctly predicts the negative class. FP is an outcome with



incorrectly predicted positive class and simultaneously TN is an outcome where the model incorrectly predicts the negative class.

4.5.2 ROC Curve

A ROC curve (Receiver Operating Characteristic curve) is a graph showing the performance of a classification model by distinguishing between true positive and negative outcome. This is achieved by plotting sensitivity, the probability of predicting a real positive will be a positive, against 1-specificity, the probability of predicting a real negative will be a positive. Thus it is a way to visualize the tradeoff between these two using different decision thresholds, meaning a threshold for deciding whether a prediction is labeled "true" or "false". Increasing the threshold will generally increase the precision, but a decrease in recall.

The best decision rule is high on sensitivity (true positive rate) and low on 1-specificity (false positive rate).

4.5.3 AUC

AUC is the area under the ROC curve and a metric used to summarize the ROC curve by using a single number. AUC has a range of [0,1]. The greater the value, the better is the performance of the model.

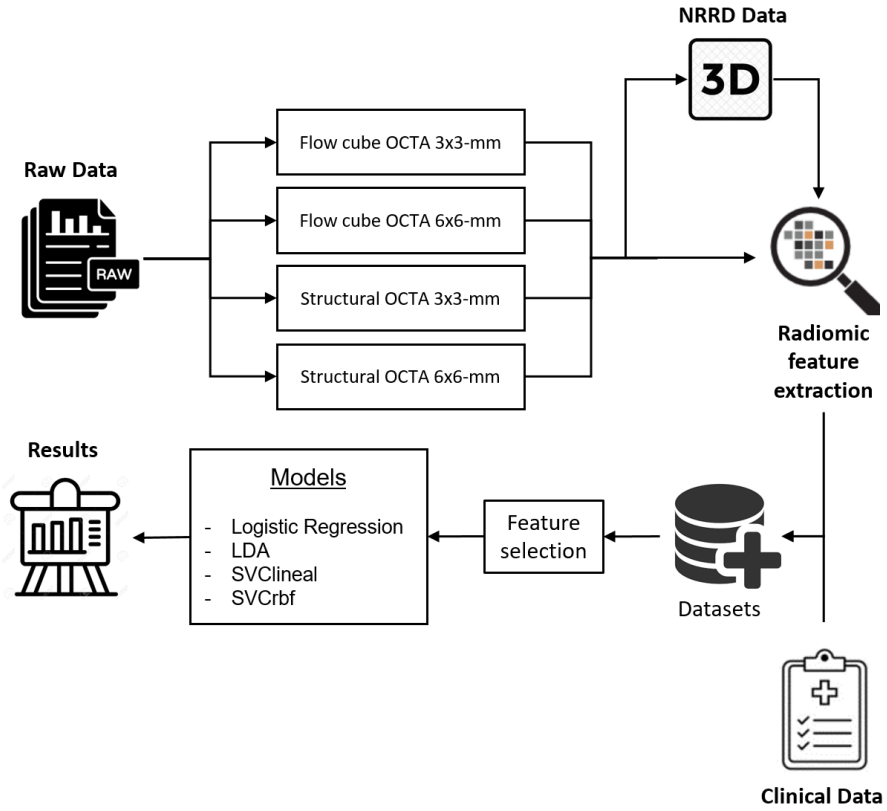


Figure 7: Workflow

5 Methodology

5.1 Experimental Design

Now that the basic concepts of retinal images and radiomics have been introduced, the actual aim of this work can be laid out, which is to analyze the ability of different characteristics to predict DM on the one hand, and to predict DR when DM is present on the other hand.

For this purpose, the data set consisting of clinical data and radiomic features, has to be filtered using the exclusion and inclusion criterion as explained in subsection 5.2.

The radiomic features extracted for 2D and 3D imagery respectively to 3x3-mm and 6x6-mm structural as well as flow cube images have to be aggregated for the 2D imagery. This is done by building the mean from the 245 images of the 3x3-mm images and the 350 6x6-mm images for each eye and each patient.

Using the mutual information technique explained above, a first feature sampling is performed. This is continued with an iterative process in which the bet-

ter parameters are identified. Next, the backward elimination wrapper method is used to eliminate features that are irrelevant or contribute only insignificantly to the model.

Finally, the model is trained, evaluated and tested with a double-cross validation.

The entire process will be further explained in the following by first reviewing the dataset in more detail and then continuing with the implementation and execution.

5.2 Dataset Description

The present data set, generated by the Diabetes Unit of Hospital Clinic, arose from an exploratory study of a cohort of type 1 DM patients [8, 34].

Included are 596 patients who have different pathologies and thus can be subdivided into certain classifications. As exposed in the introduction, there are two subcategories of DR: non-proliferative diabetic retinopathy (NPDR) and proliferative diabetic retinopathy (PDR). The former is further subdivided into mild, moderate and severe. However, also included in this study are patients who may indeed have Type 1 diabetes, but fortunately have not yet been diagnosed with DR. Also, in order not to classify later on only between DR and No DR, control patients are included who do not suffer from T1D and therefore neither from DR. For a better overview see Table 2.

0	Controls
1	No DR
2	Mild
3	Moderat
4	Severe
5	Proliferative DR

Table 2: Group characteristics of study participants

As mentioned previously, not only one classification is considered, but two binary classification problems are analyzed in this work:

- First, the discrimination between the control group [0] against the patients with T1D [1,2,3,4,5].
- Secondly, the discrimination between T1D patients without DR [1] and those with DR [2,3,4,5].

In addition to the retinal images for each eye of the patients, a series of medical records and clinical history were also collected in the dataset generated by the Diabetes Unit of Hospital Clínic. The most interesting data for discrimination are gender, time from T1D diagnosis, age, Body Mass Index (BMI), and smoking status. In this regard, it should be noted that this combination of clinical data and medical data is a positive rarity in existing publications.

Table 3: Distribution of eyes before and after exclusion criteria

Class	Before	After	Excluded
0	227	148	79
1	607	461	146
2	245	169	76
3	42	25	17
4	5	2	3
5	44	3	41
Total	**1170**	**808**	**362**

In consultation with the supervising ophthalmologist, a medical exclusion criterion is to be set as a filter. Patients who have already received treatment or surgery are sorted out, as these could distort the outcome. This filter as a medical exclusion criterion is based on a previous study [8], and examines those variables that are also present in the clinical data. Also a quality exclusion criteria is applied, which excludes images with artifacts, segmentation errors or a signal strength index below seven.

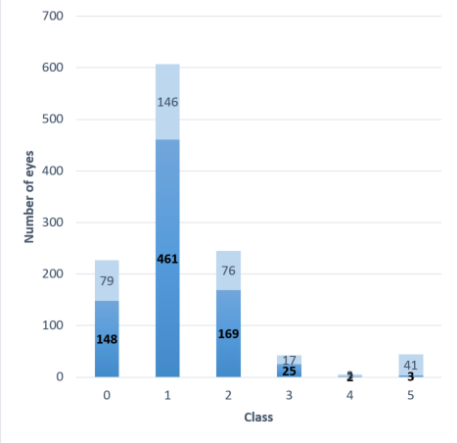


Figure 8: Distribution of Classes

After applying these filters of medical exclusion criteria and quality exclusion criteria, a total of 362 eyes were excluded Figure 8.

The dataset is further looked at when classifying DR with two different test scenarios. Once, all eyes are included in the model so that training can be done with a larger data set. Another time, only one eye is included in the model. This is done according to the following principle: if a patient has eyes with 2 different class labels, i.e. if he has DR on one eye and not on the other, the eye with the worse classification is always included in the data set, the better eye is thrown out. However, if the patient's eyes have the same classification, one of the eyes is randomly eliminated.

5.3 Implementation

In the implementation section, the tools required are presented and more insight is given into Data Partitioning, Data Preparation and Data Selection.

5.3.1 Tools

The main programming language of this work is Python 3.8. Python offers the ability to import third-party packages to use implementations.

The main libraries used in this work are listed as follows:

- sklearn - to create and train models
- pandas and numpy - for simple data manipulation
- Pyradiomics - to extract radiomic features
- sklearn - to build machine learning models
- matplotlib and seaborn - for visualization tasks

As already mentioned in section 2 the OCT converter as well as 3D-Slicer were used for image generation of structural and flow cube OCTA images as well as for the creation of 3D Nrrd files.

5.3.2 Data Partitioning

After the presented exclusion and inclusion criteria in subsection 5.2 have been applied, it is necessary to divide the data set into training and test data, preventing overfitting. This is done for each classification problem, model and feature subset. It ensures that the test data is used exclusively for the evaluation of the model performance.

5.3.3 Data Preparation

Following the Data Partitioning, the preprocessing of the data is applied. Missing data entries for the clinical data occur which are either filled with the median if less than 50 % are affected by the entries, or the respective column is taken out if more than 50 % of the data are impacted. Categorical variables are coded with one-hot encoding by transforming the contained n values into n variables and numerical attributes are normalized with the MinMaxScaler to values between zero and one.

5.3.4 Data Selection

Several filter feature-selection methods are performed to ensure the use of the most significant variables for the model. These are performed for each classification task (DM and DR) as well as for each image type.

First, a feature sampling is performed using the Mutual Information Metric. Mutual information is an entropy measure between two random variables and

quantifies the amount of information that one variable gains from the other. If it is equal to 0, then the two variables being tested are independent of each other. With the mutual information, 30 features with the highest MI values are pre-selected. This is done by using a 10-fold cross validation for keeping robustness.

This is followed by an iterative finding of the best fit parameters for the given dataset, on which a backward elimination method is applied to remove irrelevant features for the future model. A stratified double cross validation with 5 outer and 4 inner Cross Validations is carried out with a total of 20 iterations.

Hyperparameters used for Logistic Regression varies from $C = 10^{-3}$ to $C = 10^{3}$, for Linear SVC and RBF SFC from $C = 10^{1}$ to $C = 10^{4}$ and $\gamma = 10^{-4}$ to $\gamma = 10^{1}$.

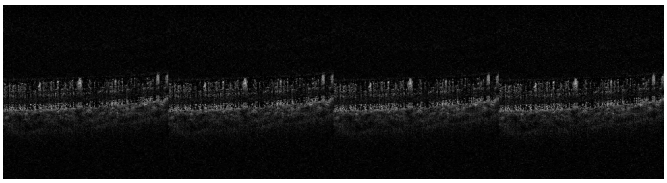
Also with the stratified double cross validation a derivation of the Area under the Curce (AUC) value for each split is created and to compare the different models and feature subset performance.

5.3.5 Image Data Preprocessing

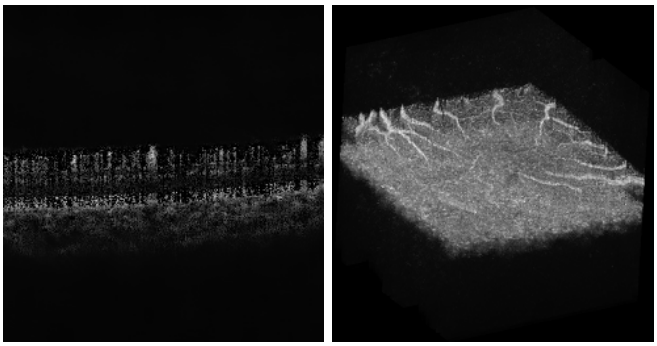
After examining and pre-processing the clinical data, the image data is also treated. A distinction is made in each case for the different image types.

Since in the timelapse images of Flow Cube OCTA and Structural OCTA, attention must be paid to even the smallest differences between the scans of the overall image. Firstly the image size is retained and secondly further image preprocessing is dispensed with. The images are already in grayscale and random noise may be difficult to distinguish from captured point clouds in the OCTA images.

For the 3-dimensional images, the respective cuts were already cropped (Figure 9) before the conversion to 3D, for the simple reason that the amount of data was already very large and the computational effort of the feature extractions would otherwise have been immense.



(a) Flow cube OCTA 3x3-mm - Original



(b) Cropped (c) 3D output

Figure 9: Image preprocessing for 3D creation

6 Results

The following chapter will be divided into the two classification problems, which have already been defined. Considered is the discrimination between the control group [0] against the patients with T1D [1,2,3,4,5] and the discrimination between T1D patients without DR [1] and those with DR [2,3,4,5].

For this purpose, the AUC values has been extracted from each double cross validation split and were plotted for the respective image types which are then compared against the different classification models including Logistic Regression, LDA, Linear SVM, RBF SVC.

6.0.1 Discriminating DM with 2D OCTA

Starting with the 2D OCTA images (Figure 10), which are divided into Flow Cube and Structural, each available for the section of a 3x3-mm area and 6x6-mm area and analysed with radiomic features, there are first of all quiet similar results between the different classification models.

First of all, it is clear to see that for all 4 classification problems, the 3x3-mm image of the flow cube OCTA is the worst, showing an AUC value just above 0.7. For the same image type, but with a 6x6-mm shot, an AUC value of 0.8 is almost achieved and thus stands out strongly from the 3x3-mm. The 6x6-mm image of the flow cube OCTA also achieves a better result for the classification problems Linear SVM and RBF SVC compared to the structural images. In general, these differ only slightly from each other, but the similarities are stronger for Linear SVM and RBF SVC than for Logistic Regression and LDA.

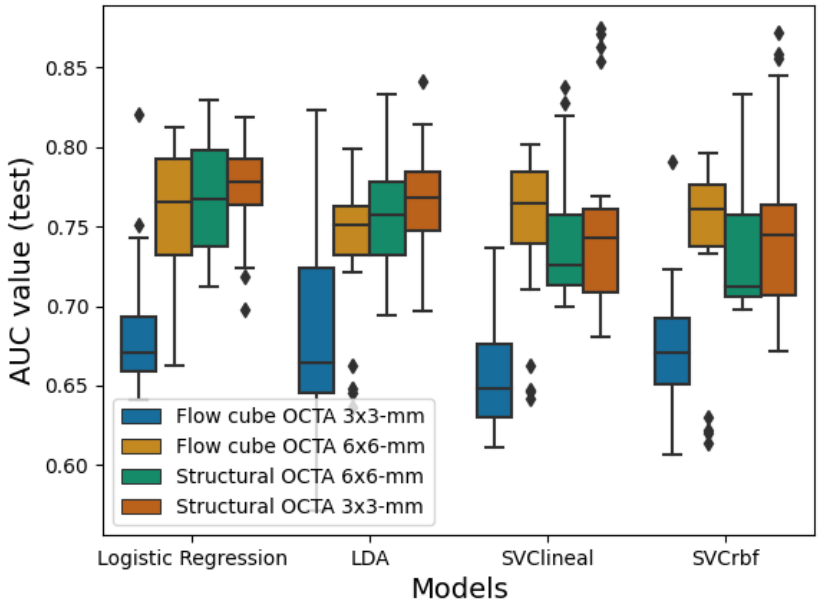


Figure 10: 2D: Discriminating DM with radiomic features

With the clinical data now considered in addition to the radiomic features, a model is built from both. The results seen in Figure 11 are obtained.

As was the case with the radiomic features, Flow Cube OCTA 3x3-mm also performs the worst here, although this time the difference from the other image types is not too great. At least the lower quartile is above an AUC value of 0.7. Whereas the 6x6-mm image of the same image type could still shine in the radiomic features for the classification problems Linear SVM and RBF SVC, these now perform second worst for all the classification problems. The first and second place are shared by the structural image types.

In general, it can be said that Flow Cube images offer stronger performance in discriminating DM than the Structural images for all the models considered. In the Logistic Regression and LDA models, the closer 3x3-mm images seem to perform better than the wider images of the 6x6-mm area images for both Flow cube and Structural images. However, this cannot be confirmed with the Linear SVM and RBF SVC models, where there are no fixed cutoffs. The 6x6-mm image performs better in the Linear SVM and RBF SVC models and conversely 3x3-mm performs better in Logistic Regression and LDA. Structural OCTA 6x6-mm almost reaches an AUC value of 0.88 with the upper quartile.

6.0.2 Discriminating DR with 2D OCTA

For the discrimination of DR, both the dataset containing both eyes of the patient and the dataset containing the worse eye of the patient were included.

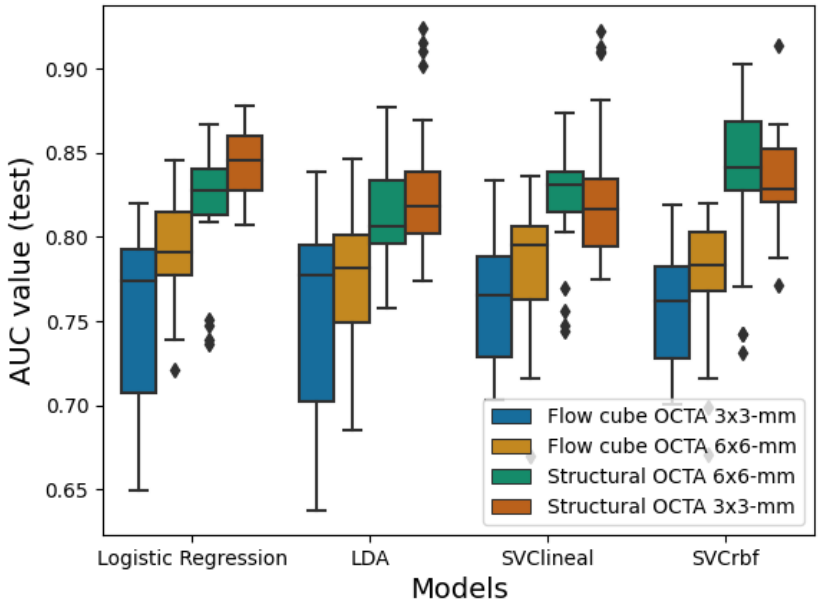


Figure 11: 2D: Discriminating DM with radiomic features and clinical data

Starting with the larger dataset containing all eyes with the corresponding radiomic features, a contrasting result is obtained as with the discrimination of DM.

While in the discrimination of DM the flow cube OCTA 3x3-mm performed best, the discrimination of DR shows the exact opposite with regard to the flow cube OCTA 3x3-mm images. These perform best in all models. With logistic regression and LDA, these AUC values can also be significantly differentiated from the other image types.

However, the maximum value is below 0.8 with the upper quartile of the OCTA 3x3-mm images. Overall, the results are all somewhat higher than an AUC value of 0.6. It is striking that the other image types are all very close to each other and, with the exception of Structural OCTA 6x6-mm, have a very low interquartile range.

If clinical data is now added to the radiomic features, improvements in the AUC values are obtained for all image types, but especially for the structural OCTA images (Figure 13). These previously performed worst, but now show the best results, so that an AUC value of just under 0.875 can be achieved.

In addition to the full data set with all eyes, it is now investigated whether there is a difference if only one eye of the respective patient is included (Figure 14).

In general, one can say that the results are very similar to the results of the

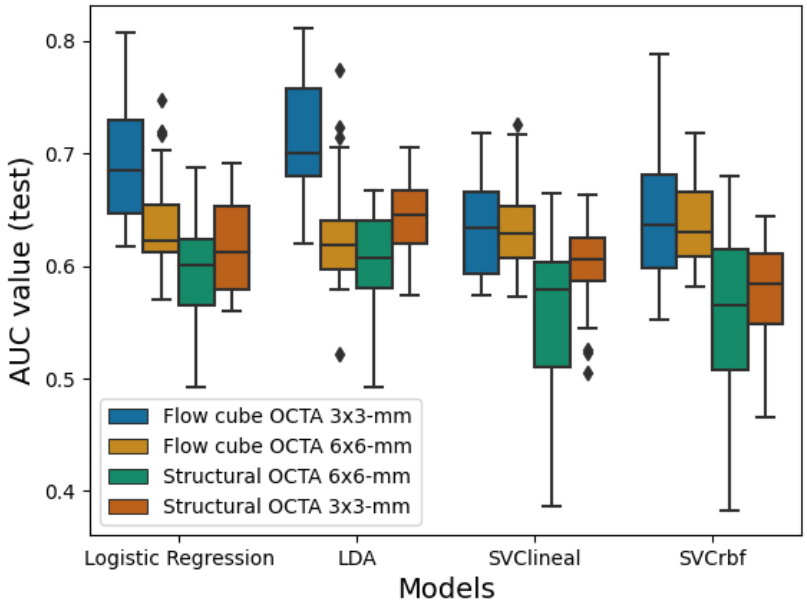


Figure 12: 2D: Discriminating DR with radiomic features

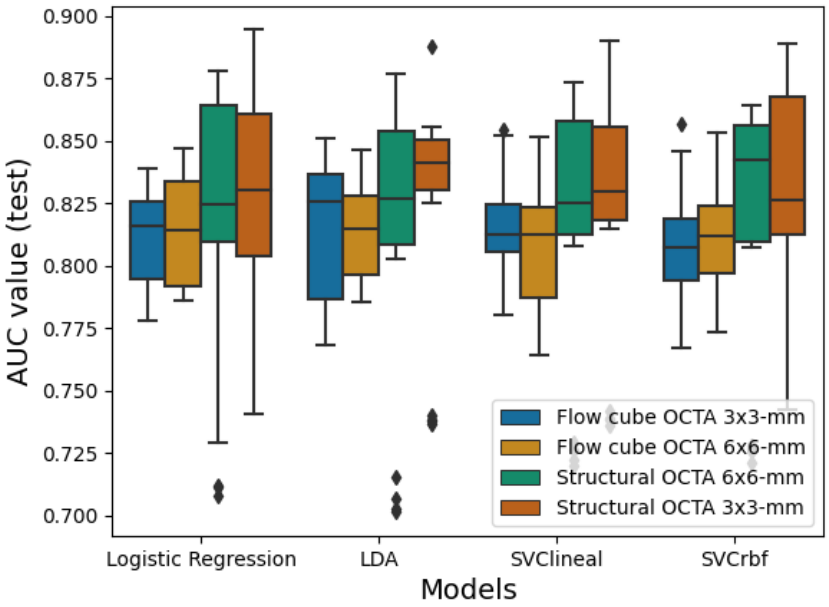


Figure 13: 2D: Discriminating DR with radiomic features and clinical data

entire data set. It is noticeable that the interquartile range is smaller and thus the results are more robust when only one eye is considered. For example, it can be seen that in the RBF SVC model, the image type of structural 6x6-mm performs worse than structural 3x3-mm, where in the larger data set the interquartile range was too wide such that a distinction these was not possible.

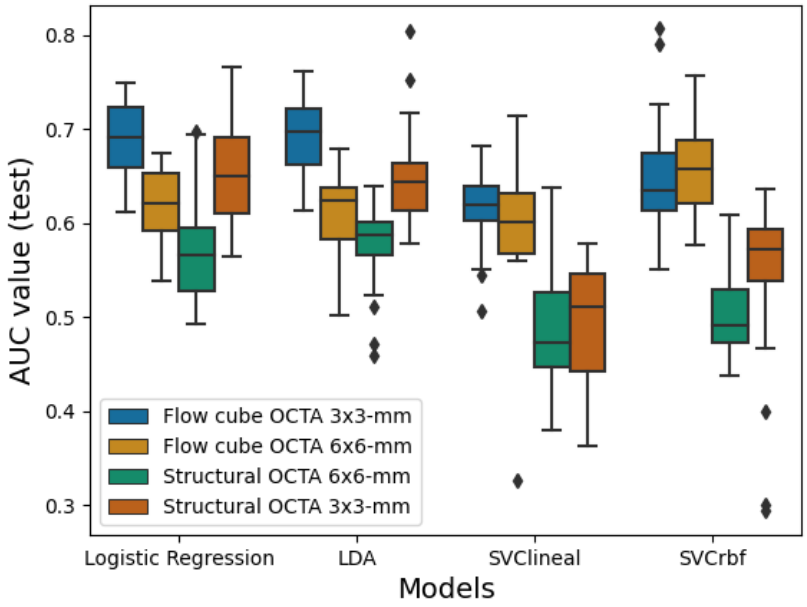


Figure 14: 2D: Discriminating DR with radiomic features with dataset containing only one eye per patient

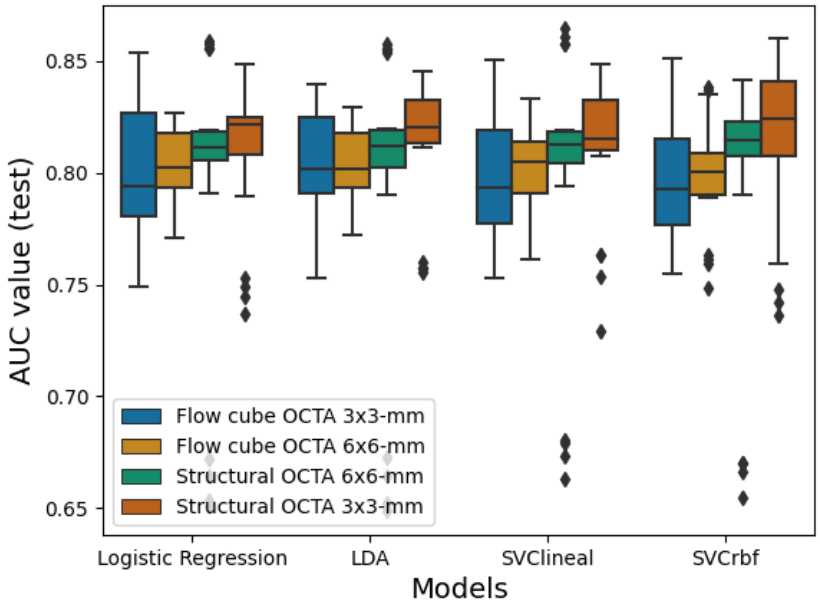


Figure 15: radiomic features and clinical data

6.0.3 Discriminating DM with 3D OCTA

After the 2D OCTA images have been analyzed, next steps are the 3D OCTA images. For this, the 3D-Nrrd images were analyzed and likewise divided into flow cube OCTA 3x3-mm and 6x6-mm as well as structural OCTA 3x3-mm and 6x6-mm.

First considered is again the first classification problem, which deals with the prediction of DM.

In this one, a very strong separation into two groups can be recognized, namely the flow cube and structural OCTA images (Figure 16). The former are clearly set apart from the structural OCTA images and thus achieve an AUC value of almost 0.75 with the 3x3-mm images. The 6x6-mm images reach a value just above 0.7. This is valid for all 4 classification problems. The structural images do not reach the 0.6 limit of AUC value and stay just above an AUC of over 0.5. One can see a slight difference between the first two models Logistic Regression and LDA and the last two models linear SVC and RBF SVC. The latter have slightly worse values than the former.

After including the clinical data (Figure 17), there is also an improvement again for overall models. The flow cube images now perform with an AUC value of just over 0.8, the structural images with an AUC value of 0.75. For the flow cube images, the interquartile ranges have become slightly wider than in the models without the clinical data; for the structural images, they have remained

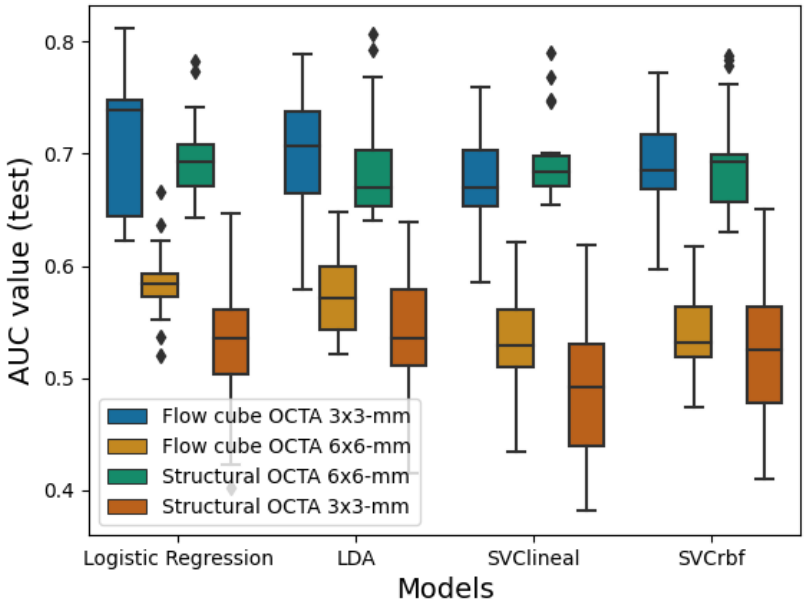


Figure 16: 3D: Discriminating DM with radiomic features

about the same.

6.0.4 Discriminating DR with 3D OCTA

The second classification problem, which deals exclusively with DM patients and which tests whether DR is present or not, shows similar results for all models (Figure 18). Flow cube OCTA 3x3-mm performs best, with AUC values of 0.6-0.7. This is followed by 6x6-mm images, and structural images bring up the rear.

These results are increasing by considering also the clinical data (**??**). A median AUC value of 0.84 is obtained with the Flow Cube OCTA 3x3-mm for the Logistic Regression and LDA models. With linear SVC and RBF SVC, generally wider interquartile ranges are observed. This is true for all image types. The smallest AUC value for the median is approx. 0.8 and is recorded by the Flow Cube OCTA 6x6-mm.

Now looking again at the data set which contains only one eye of each patient.

For the models that only contain the radiomic features (Figure 20), the images are again set off as flow cube OCTA 3x3-mm and achieve a slightly better value than the data set with 2 eyes. With flow cube OCTA 3x3-mm, a very robust value close to the 0.8 AUC limit is achieved. However, this robustness only applies to flow cube OCTA 3x3-mm. The other image types are all at the same values at first and their robustness is no longer particularly stable as it was the

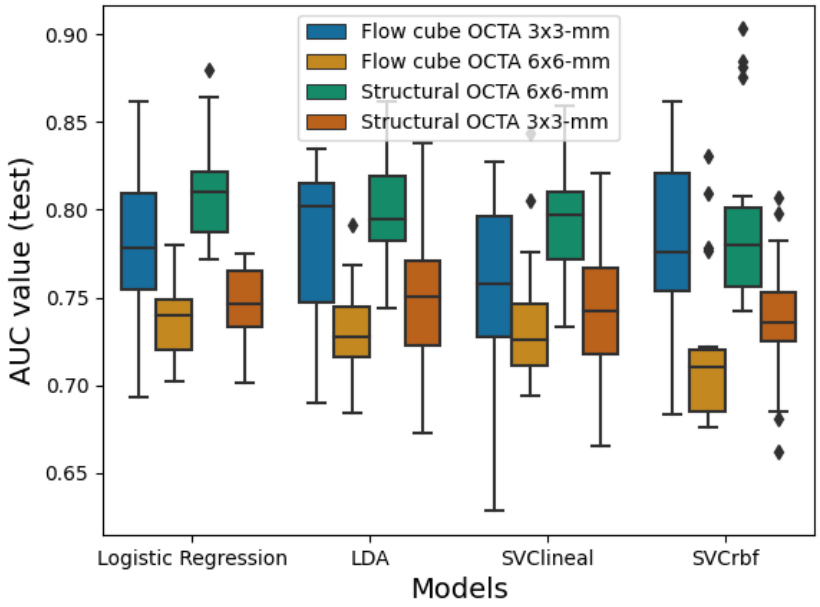


Figure 17: 3D: Discriminating DM with radiomic features and clinical data

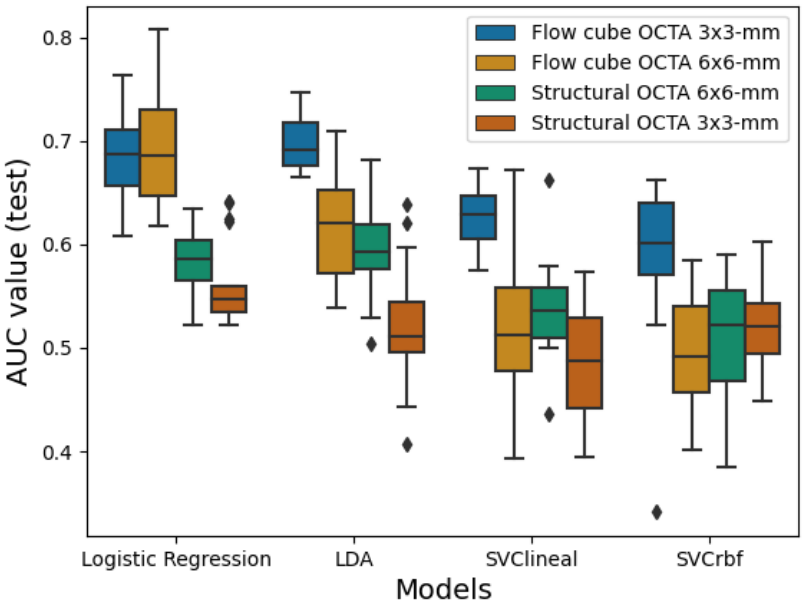


Figure 18: 3D: Discriminating DR with radiomic features

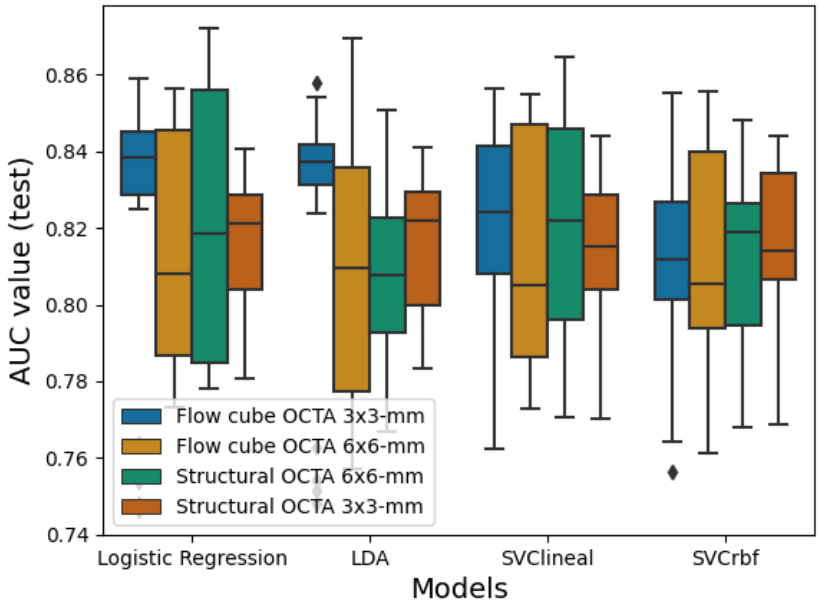


Figure 19: 3D: Discriminating DR with radiomic features and clinical data

case for the structural OCTA images.

For the results where the clinical data have been added, it is first noticeable that all models are very close to each other (Figure 21). However, this could also be visually deceptive, as the interquartile ranges are generally very wide. The medians of the models with only one eye are all above 0.8, but do not exceed the 0.85 limit, as is the case with the models with both eyes. Here, a median of 0.84 is obtained for the Flow Cube OCTA 3x3-mm images.

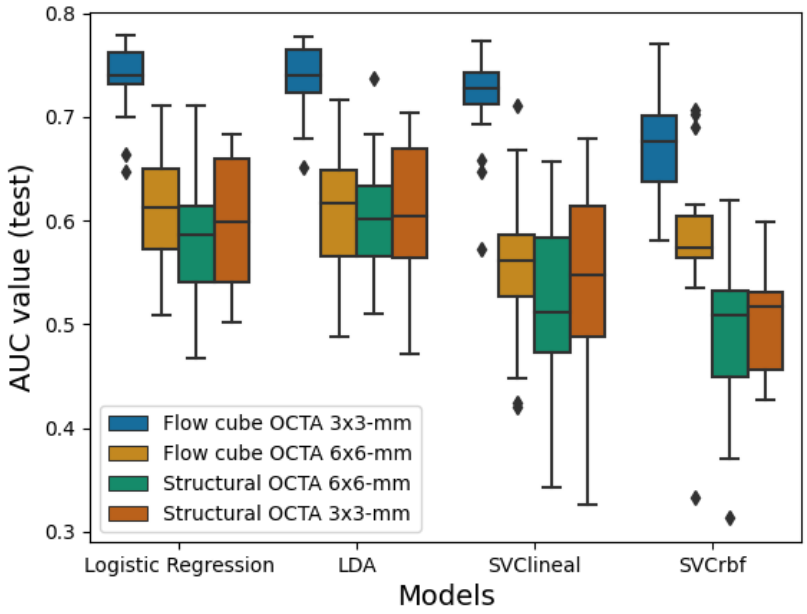


Figure 20: 3D: Discriminating DR with radiomic features with dataset containing just one eye

6.1 Evaluation

Overall, it can be said that surprisingly good results have been achieved for the classification of DM, as well as for DR. This can also be confirmed for the different image types of 2D and 3D images.

For the discrimination of DM, a median of 0.85 AUC value was achieved with the 2D cuts of the OCTA images. Here, the structural images performed particularly well. The flow cube images have an AUC value of just under 0.8.

With the 3D images, an AUC of slightly above 0.8 is achieved. These are paired with the flow cube OCTA 3x3-mm images and the structural OCTA 6x6-mm images. In contrast, the Flow cube OCTA 6x6-mm and Structural OCTA 3x3-mm images bring up the rear with values up to 0.75.

In the second classification problem, it is not quite so easy to make a clear distinction between the 2D and 3D images. A median AUC value of 0.825 is recorded for the 2D images. This was, for example, 0.84 for the 3D images with Flow Cube 3x3-mm, but also just over 0.8 with Flow Cube 6x6-mm. All examinations are close to each other.

Overall, it can be said that both 2D and 3D OCTA images achieved better results with the addition of the clinical data, even if the robustness decreased somewhat.

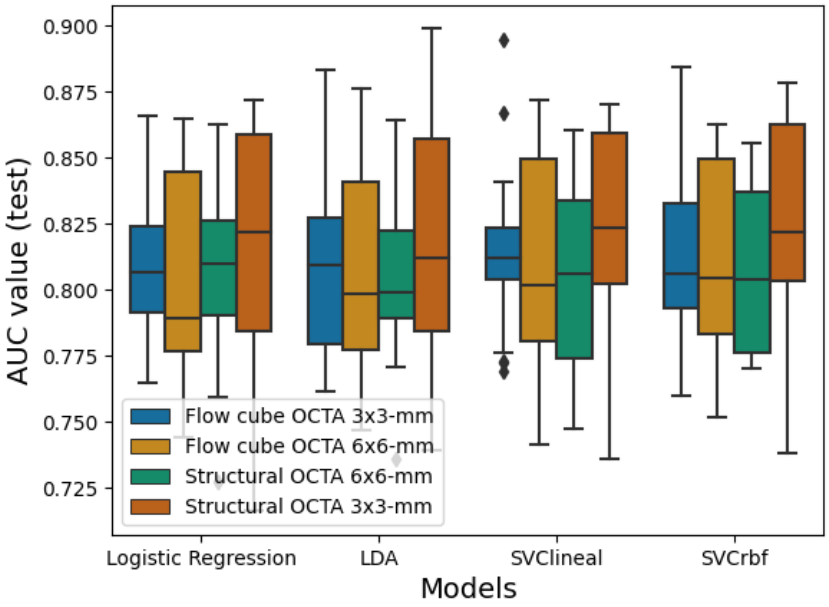


Figure 21: 3D: Discriminating DR with radiomic features and clinical data with dataset containing just one eye

Without the clinical data, however, quite high values are still observed. For example, AUC values of over 0.7 are obtained for discrimination of DM and values just below 0.7 for discrimination of DR.

Regarding the respective image types, it can be said that for the classification problem of DM the flow cube OCTA 3x3-mm performed by far, as well as structural 6x6-mm, for DR it is the flow cube OCTAs 3x3-mm and 6x6-mm.

Furthermore, when examining two different data sets, once with only one eye per patient, the other time with both, a small deterioration can be found in the smaller data set.

7 Conclusion

It can only be repeated that the 2D images and 3D images showed remarkably good results.

The fact that 2D images perform better than 3D images is astonishing, since the mean was ultimately formed via the 245 or 350 2D OCTA images and it was to be assumed that important information could be narrowed down. However, depending on the image area, the 2D images contain either 2 or 4 images that were taken at the same location in a few seconds one after the other. This means that the blood flow can also be mapped in one image.

When comparing the discriminations of DM and DR, it was also noticeable that different image types grouped together performed better than others. Thus, the flow cube images with the respective areas of 3x3-mm and 6x6-mm were able to find the best predictions for the classification problem of DR. For the classification problem of DM, however, this cannot be grouped according to the recorded area or image type, rather they cross each other. Flow cube OCTA 3x3-mm and structural OCTA 6x6-mm achieved the best predictions.

Finally, with the addition of variables such as gender, BMI index, smoking status, etc., all results improve a little. The combination of retinal images and clinical data is the key to achieving this.

7.1 Related Work

Since the previous work on this matter presented by Laura Carrera Escalé in collaboration with Dr. Javier Zarranz-Ventura, Anass Benali, Ruben Martin, Enrique Romero Merino and Alfredo Vellido Alcacena, focused on the set of retinal images including fundus, OCT and enface 2D frontal OCTA images in the superficial and deep capillary plexi, a comparison to these is given in the following. Primarily the results of the enface 2D frontal OCTA images are considered, since these have the same origin as the images studied in this work.

Again, a distinction is made between the classification problems of DM and DR. For the discrimination of DM, the related work showed an AUC of just below 0.7 for all 4 models, varying between 0.7 and 0.6 for logistic regression, LDA and Linear SVM, but an interquartile range of 0.35 to 0.65 for RBF SVC. These values are exceeded once with the 2D cuts and their values up to 0.8 easily and also for the corresponding 3D images this is the case, even though the values here were somewhat lower.

For the classification problem of DR, the enface 2D frontal OCTA images achieve surprisingly high and robust values with AUC values up to 0.8.

This cannot be matched by the 2D cuts, which are almost all below 0.7, with the exception of flow cube OCTA 3x3-mm, which achieve a higher value of 0.75. For 3D images, the value of 0.7 cannot be surpassed either.

7.2 Improvements

It certainly would be of interest to see if there is an improvement in the values if image cropping were not applied to the 3D OCTA images. Since time was limited and the computing power had already been high enough with the preprocessed images, the image cropping step could not be avoided of this work.

As already explained, this work arose from a project in which several students and supervisors worked on the same data set. Different techniques were applied to this data set. Anass Benali, for example, worked thoroughly on image preprocessing [35]. The images used were all adjusted by segmentation, rotation, curvature and skewness. Here again, it would not have been possible in the time available to apply these preprocessing steps to the 245 or 350 images per eye.

7.3 Future work

As briefly noted in the Introduction, an interpretable Computer Aided Diagnosis of Diabetic Retinopathy already exists. In [21], however, this has so far only been verified with fundus images. In the work of Laura Carrera Escalé it is confirmed that fundus images are by far the worst performers compared to OCT and OCTA images. For this reason it would be interesting to create an interpretable Computer Aided Diagnosis of Diabetic Retinopathy also for OCTA images and their variants. This will also make it easier to introduce Computer Aided Diagnosis into practical use when it comes to medical professionals using computer aided diagnosis in the future. It can be used to understand why the program predicts a particular diagnosis. As already experimented in [21], the area that is to justify the made diagnosis is directly marked in the input itself.

8 Glossary

DM	Diabetes Mellitus.
DR	Diabetic Retinopathy.
LDA	Linear Discriminant Analysis.
NPDR	Non-proliferative diabetic retinopathy.
Nrrd	Nearly raw raster data.
OCT	Optical Coherence Tomography.
OCTA	Optical Coherence Tomography Angiography.
PDR	Proliferative diabetic retinopathy.
RBF	Radial Basis Function.
SVM	Support Vector Machine.
T1D	Type 1 Diabetes.

Appendices

8.1 Acknowledgements

I would like to thank the following people, without whom I would not have been able to complete this thesis. My supervisor Enrique Romero Merino and co-supervisor Alfredo Vellido Alcacena, as well as Javier Zarranz-Ventura, whose insight and knowledge into the subject matter steered me through this research. Thanks for including me in the team together with Ruben Martin, Laura Carrera Escalé and Anass Benali and for the many meetings we had together. A special thank you to Ruben, who helped me with the engineering of the OCTA images, and to Laura, who made a lot of preparatory steps with her master thesis. Anass, who helped with his knowledge in image preprocessing, even if this could not be applied in the end. I would also like to mention my gratitude to casa del moviment, whose yoga classes always brought me back down to earth. But my biggest thanks to Nil for all the unconditional support you have shown me through this process.

References

[1] L. Wang, O. Murphy, N. G. Caldito, P. A. Calabresi, and S. Saidha, "Emerging applications of optical coherence tomography angiography (octa) in neurological research," *Eye and Vision*, vol. 5, no. 1, pp. 1–11, 2018.

[2] W. Drexler, M. Liu, A. Kumar, T. Kamali, A. Unterhuber, and R. A. Leitgeb, "Optical coherence tomography today: speed, contrast, and multimodality," *Journal of biomedical optics*, vol. 19, no. 7, p. 071412, 2014.

[3] Y. M. Chan, E. Ng, V. Jahmunah, J. E. W. Koh, O. S. Lih, L. Y. W. Leon, and U. R. Acharya, "Automated detection of glaucoma using optical coherence tomography angiogram images," *Computers in biology and medicine*, vol. 115, p. 103483, 2019.

[4] A. Sapra and P. Bhandari, "Diabetes mellitus," *StatPearls Publishing*, 2019.

[5] A. T. Kharroubi and H. M. Darwish, "Diabetes mellitus: The epidemic of the century," *World journal of diabetes*, vol. 6, no. 6, p. 850, 2015.

[6] W. R. Rowley, C. Bezold, Y. Arikan, E. Byrne, and S. Krohe, "Diabetes 2030: insights from yesterday, today, and future trends," *Population health management*, vol. 20, no. 1, pp. 6–12, 2017.

[7] R. Williams, S. Karuranga, B. Malanda, P. Saeedi, A. Basit, S. Besançon, C. Bommer, A. Esteghamati, K. Ogurtsova, P. Zhang, *et al.*, "Global and regional estimates and projections of diabetes-related health expenditure: Results from the international diabetes federation diabetes atlas," *Diabetes research and clinical practice*, vol. 162, p. 108072, 2020.

[8] M. Barraso, A. Alé-Chilet, T. Hernández, C. Oliva, I. Vinagre, E. Ortega, M. Figueras-Roca, A. Sala-Puigdollers, C. Esquinas, E. Esmatjes, *et al.*, "Optical coherence tomography angiography in type 1 diabetes mellitus. report 1: diabetic retinopathy," *Translational vision science & technology*, vol. 9, no. 10, pp. 34–34, 2020.

[9] L. Kahanovitz, P. M. Sluss, and S. J. Russell, "Type 1 diabetes–a clinical perspective," *Point of care*, vol. 16, no. 1, p. 37, 2017.

[10] C. C. Patterson, V. Harjutsalo, J. Rosenbauer, A. Neu, O. Cinek, T. Skrivarhaug, B. Rami-Merhar, G. Soltesz, J. Svensson, R. C. Parslow, *et al.*, "Trends and cyclical variation in the incidence of childhood type 1 diabetes in 26 european centres in the 25 year period 1989–2013: a multicentre prospective registration study," *Diabetologia*, vol. 62, no. 3, pp. 408–417, 2019.

[11] E. J. Duh, J. K. Sun, and A. W. Stitt, "Diabetic retinopathy: current understanding, mechanisms, and treatment strategies," *JCI insight*, vol. 2, no. 14, 2017.

[12] N. E. Institute, "Facts about diabetic eye disease." https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/diabetic-retinopathy/, 2021. [Online; accessed 15-October-2021].

[13] C. Wilkinson, F. L. Ferris III, R. E. Klein, P. P. Lee, C. D. Agardh, M. Davis, D. Dills, A. Kampik, R. Pararajasegaram, J. T. Verdaguer, *et al.*, "Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales," *Ophthalmology*, vol. 110, no. 9, pp. 1677–1682, 2003.

[14] M. E. Hartnett, W. Baehr, and Y. Z. Le, "Diabetic retinopathy, an overview," *Vision research*, vol. 139, pp. 1–6, 2017.

[15] D. S. W. Ting, L. R. Pasquale, L. Peng, J. P. Campbell, A. Y. Lee, R. Raman, G. S. W. Tan, L. Schmetterer, P. A. Keane, and T. Y. Wong, "Artificial intelligence and deep learning in ophthalmology," *British Journal of Ophthalmology*, vol. 103, no. 2, pp. 167–175, 2019.

[16] M. Abramoff, J. Staal, M. Suttorp, B. Polak, M. Viergever, *et al.*, "Low level screening of exudates and hemorrhages in background diabetic retinopathy," *Comp. Assi. Fun. Image Anal*, vol. 15, 2000.

[17] V. S. Parekh and M. A. Jacobs, "Deep learning and radiomics in precision medicine," *Expert review of precision medicine and drug development*, vol. 4, no. 2, pp. 59–72, 2019.

[18] H. J. Aerts, E. R. Velazquez, R. T. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, J. Bussink, R. Monshouwer, B. Haibe-Kains, D. Rietveld, *et al.*, "Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach," *Nature communications*, vol. 5, no. 1, pp. 1–9, 2014.

[19] X. Liang, E. N. Alshemmary, M. Ma, S. Liao, W. Zhou, and Z. Lu, "Automatic diabetic foot prediction through fundus images by radiomics features," *IEEE Access*, vol. 9, pp. 92776–92787, 2021.

[20] H. Wu, J. Liu, S. Zhang, A. Zhu, C. Sulistio, J. Li, A. Sang, J. Dong, and B. J. Liu, "The development of an ophthalmologic imaging cade structured report for retinal image radiomics research," in *Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications*, vol. 10579, p. 1057910, International Society for Optics and Photonics, 2018.

[21] D. Kumar, G. W. Taylor, and A. Wong, "Discovery radiomics with clear-dr: interpretable computer aided diagnosis of diabetic retinopathy," *IEEE Access*, vol. 7, pp. 25891–25896, 2019.

[22] R. F. Spaide, J. G. Fujimoto, N. K. Waheed, S. R. Sadda, and G. Staurenghi, "Optical coherence tomography angiography," *Progress in retinal and eye research*, vol. 64, pp. 1–55, 2018.

[23] E. C. Greig, J. S. Duker, and N. K. Waheed, "A practical guide to optical coherence tomography angiography interpretation," *International Journal of Retina and Vitreous*, vol. 6, no. 1, pp. 1–17, 2020.

[24] Q. Zhang, C. S. Lee, J. Chao, C.-L. Chen, T. Zhang, U. Sharma, A. Zhang, J. Liu, K. Rezaei, K. L. Pepple, *et al.*, "Wide-field optical coherence tomography based microangiography for retinal imaging," *Scientific reports*, vol. 6, no. 1, pp. 1–10, 2016.

[25] E. Borrelli, D. Sarraf, K. B. Freund, and S. R. Sadda, "Oct angiography and evaluation of the choroid and choroidal vascular disorders," *Progress in retinal and eye research*, vol. 67, pp. 30–55, 2018.

[26] K. Y. Tey, K. Teo, A. C. Tan, K. Devarajan, B. Tan, J. Tan, L. Schmetterer, and M. Ang, "Optical coherence tomography angiography in diabetic retinopathy: a review of current applications," *Eye and Vision*, vol. 6, no. 1, pp. 1–10, 2019.

[27] M. Graham, "Oct-converter." https://github.com/marksgraham/OCT-Converter, 2020.

[28] X.-X. Li, W. Wu, H. Zhou, J.-J. Deng, M.-Y. Zhao, T.-W. Qian, C. Yan, X. Xu, and S.-Q. Yu, "A quantitative comparison of five optical coherence tomography angiography systems in clinical performance," *International journal of ophthalmology*, vol. 11, no. 11, p. 1784, 2018.

[29] "3d slicer image computing platform." https://www.slicer.org/. Accessed: 2011-10-20.

[30] "Nrrd - nearly raw raster data." http://teem.sourceforge.net/nrrd/index.html. Accessed: 2011-10-20.

[31] J. E. van Timmeren, D. Cester, S. Tanadini-Lang, H. Alkadhi, and B. Baessler, "Radiomics in medical imaging—"how-to" guide and critical reflection," *Insights into Imaging*, vol. 11, no. 1, pp. 1–16, 2020.

[32] M. Mannil, J. von Spiczak, R. Manka, and H. Alkadhi, "Texture analysis and machine learning for detecting myocardial infarction in noncontrast low-dose computed tomography: unveiling the invisible," *Investigative radiology*, vol. 53, no. 6, pp. 338–343, 2018.

[33] "Radiomics documentation - radiomic features." https://pyradiomics.readthedocs.io/en/latest/features.html. Accessed: 2011-10-20.

[34] J. Zarranz-Ventura, M. Barraso, A. Alé-Chilet, T. Hernandez, C. Oliva, J. Gascón, A. Sala-Puigdollers, M. Figueras-Roca, I. Vinagre, E. Ortega, *et al.*, "Evaluation of microvascular changes in the perifoveal vascular network using optical coherence tomography angiography (octa) in type i diabetes mellitus: a large scale prospective trial," *BMC medical imaging*, vol. 19, no. 1, pp. 1–6, 2019.

[35] A. Benali, *Using machine learning on the sources of retinal images for diagnosis by proxy of idabetes mellitzs and diabetic retinopathy.* Master's thesis, Universitat Politècnica de Catalunya, 2021.