Machine Learning for Glaucoma Assessment using Fundus Images

Andres Diaz-Pinto
a.diaz-pinto@leeds.ac.uk

Outline

1. Introduction
2. Segmentation Methods
3. Classification Methods
4. Image synthesis

Outline

1. **Introduction**
   - **Motivation**
   - Anatomy of the retina
   - Types of glaucoma
   - Imaging technology
2. Segmentation Methods
3. Classification Methods
4. Image synthesis

Motivation

- Glaucoma is the second most common cause of blindness worldwide, according to the World Health Organization (United Nations agency).

- It affects more than 60 million people.

Motivation

Early detection and treatment is important to prevent vision loss.

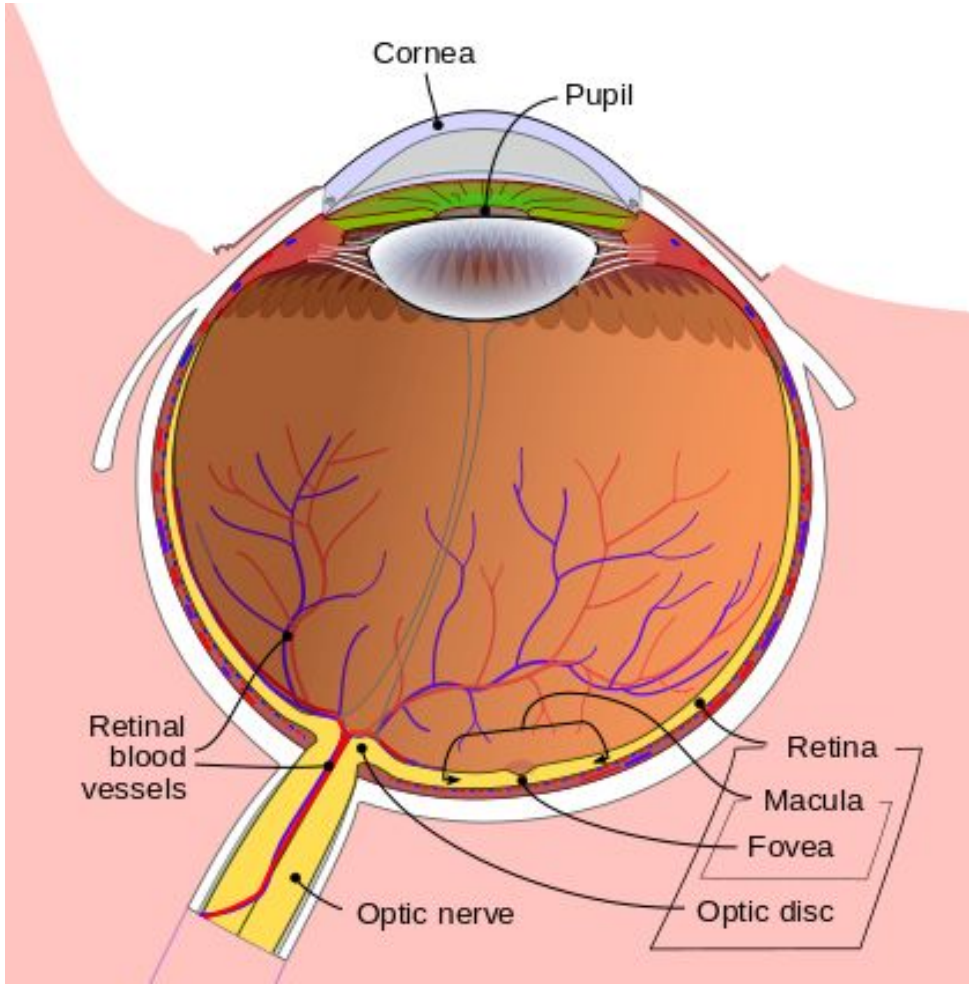
HOWEVER, screening to large population is expensive.

FOR THAT REASON, the development of automatic glaucoma assessment algorithms is of great interest.

Outline

1. **Introduction**
   - Motivation
   - **Anatomy of the retina**
   - Types of glaucoma
   - Imaging technology
2. Segmentation Methods
3. Classification Methods
4. Image synthesis

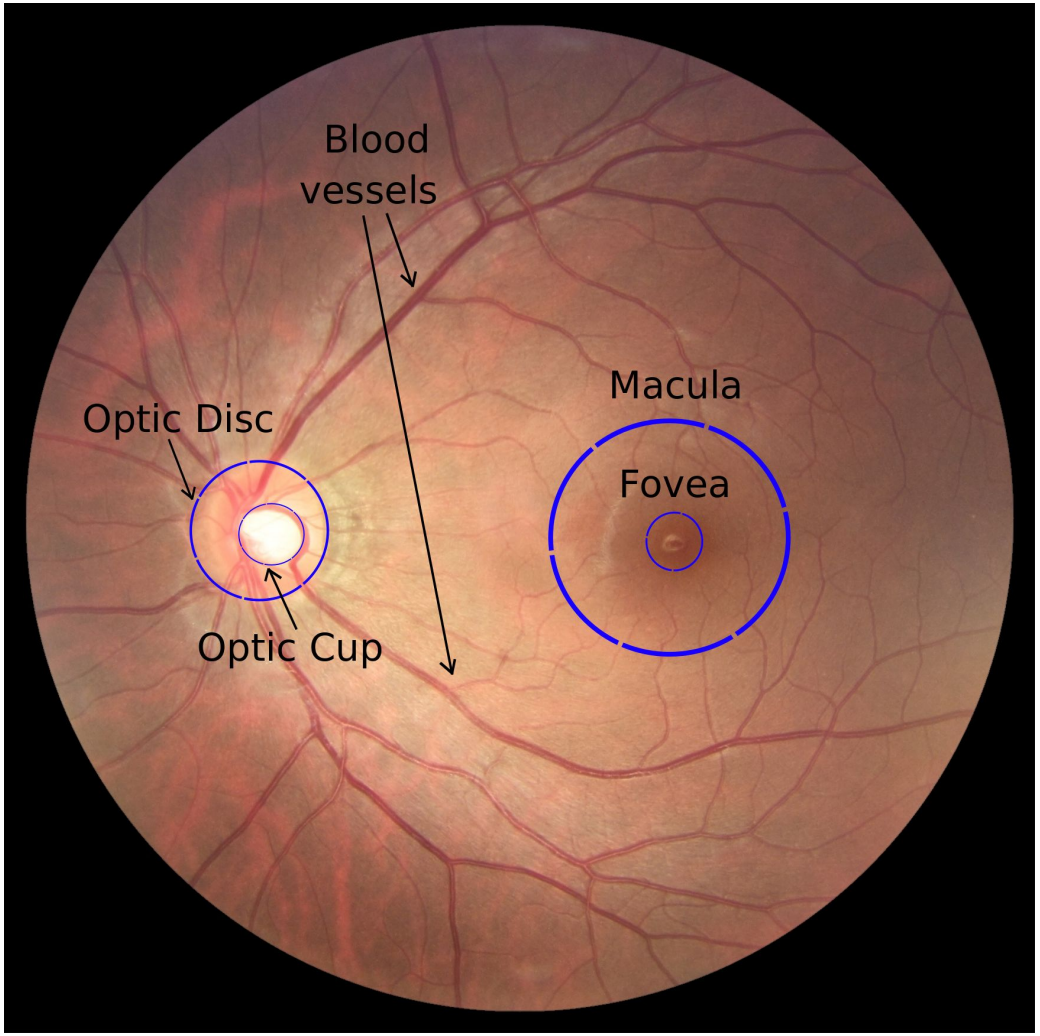
Anatomy of the retina



Three main structures:

- The optic disc
- Retinal blood vessels
- The macula

Anatomy of the retina



Outline

1. **Introduction**
   - Motivation
   - Anatomy of the retina
   - **Types of glaucoma**
   - Imaging technology
2. Segmentation Methods
3. Classification Methods
4. Image synthesis

Types of glaucoma

Glaucoma refers to a deepening or excavation of the optic nerve head.

And there are three main forms of glaucoma:

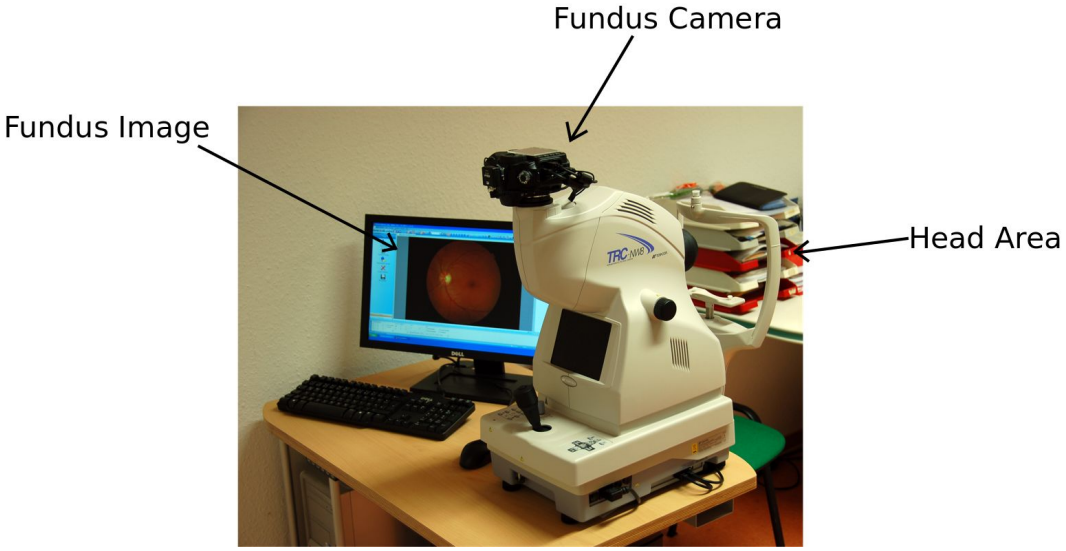
1. Open-angle glaucoma
2. Angle closure glaucoma and
3. Congenital glaucoma

Outline

1. **Introduction**
   - Motivation
   - Anatomy of the retina
   - Types of glaucoma
   - **Imaging technology**
2. Segmentation Methods
3. Classification Methods
4. Image synthesis

Imaging technology

1. Fundus photograph



Imaging technology

Main differences:

Fundus Photograph	OCT
RGB image	Tomography (up to 3D image)
Less accurate measurements	High accurate measurements
Low cost	Prohibitively expensive for mass screening

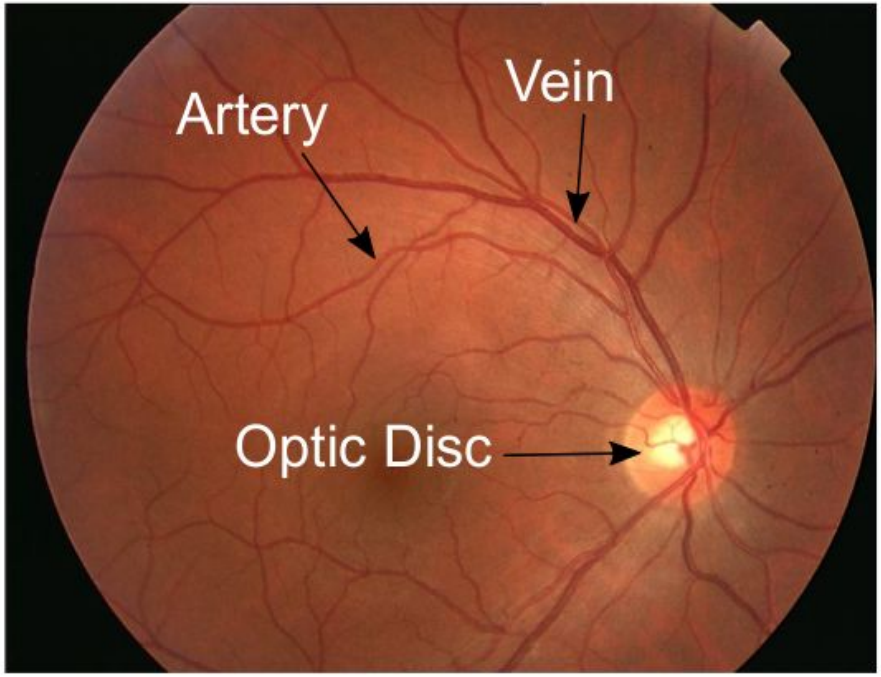
Outline

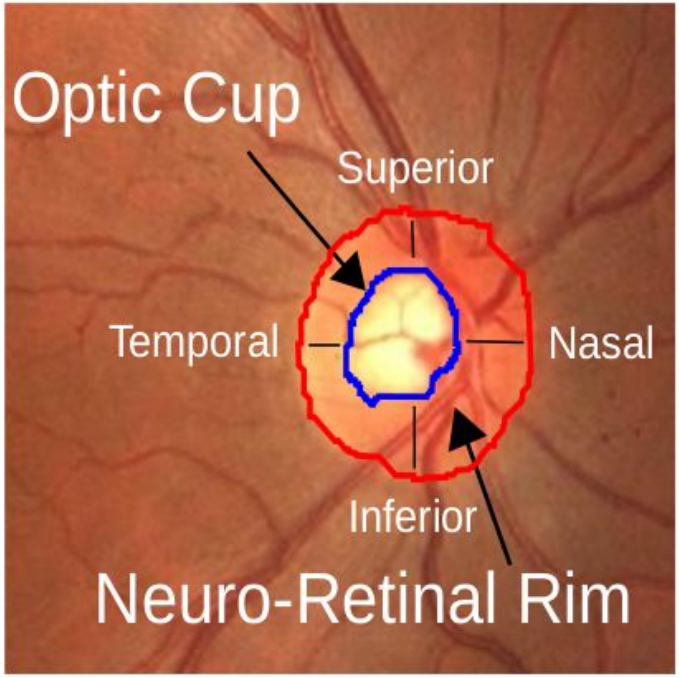
1. Introduction
2. **Segmentation Methods**
   - **Stochastic-Watershed-based method**
   - U-Net-based approach
3. Classification Methods
4. Image synthesis

Rationale

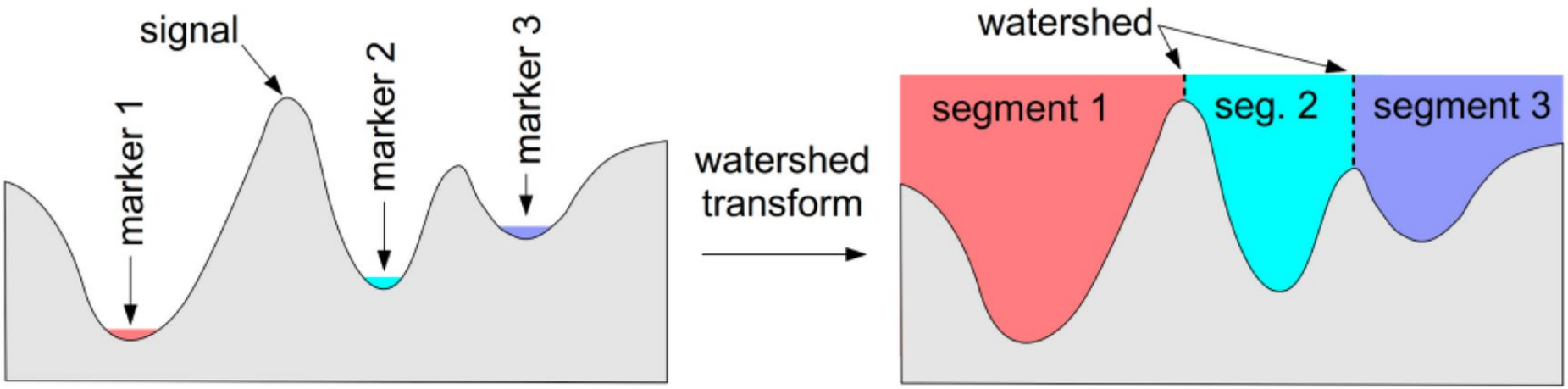
Segment important parts of the retina to measure clinical features.

Classical approach!





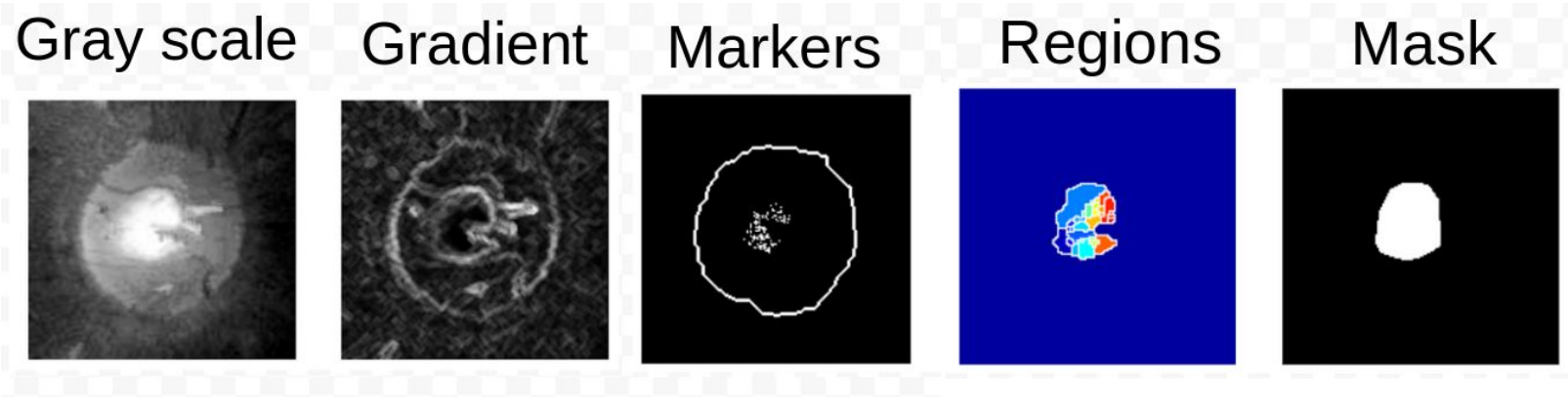
Stochastic-Watershed-based method



Regular Watershed transformation

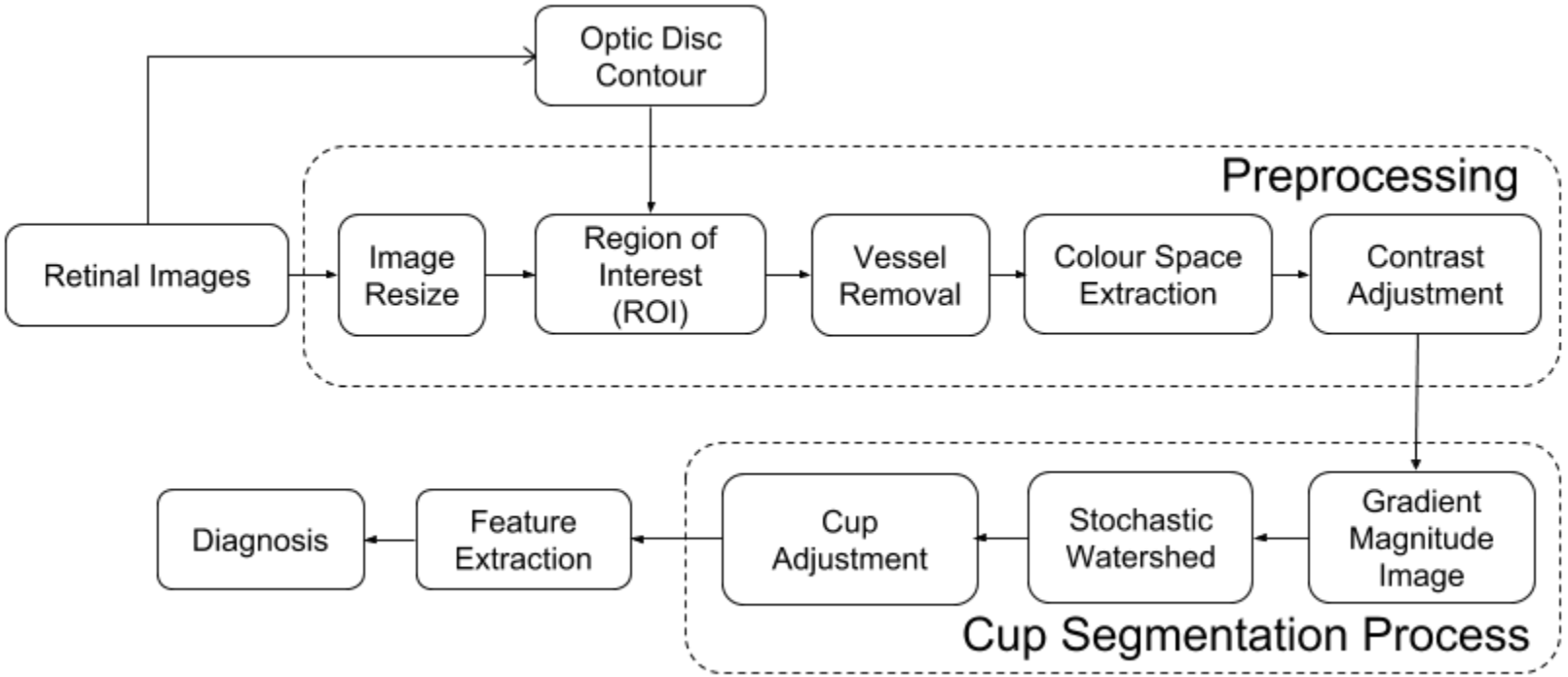
Image taken from https://res.mdpi.com/remotesensing/remotesensing-06-00776/article_deploy/html/images/remotesensing-06-00776f3.png

Stochastic-Watershed-based method

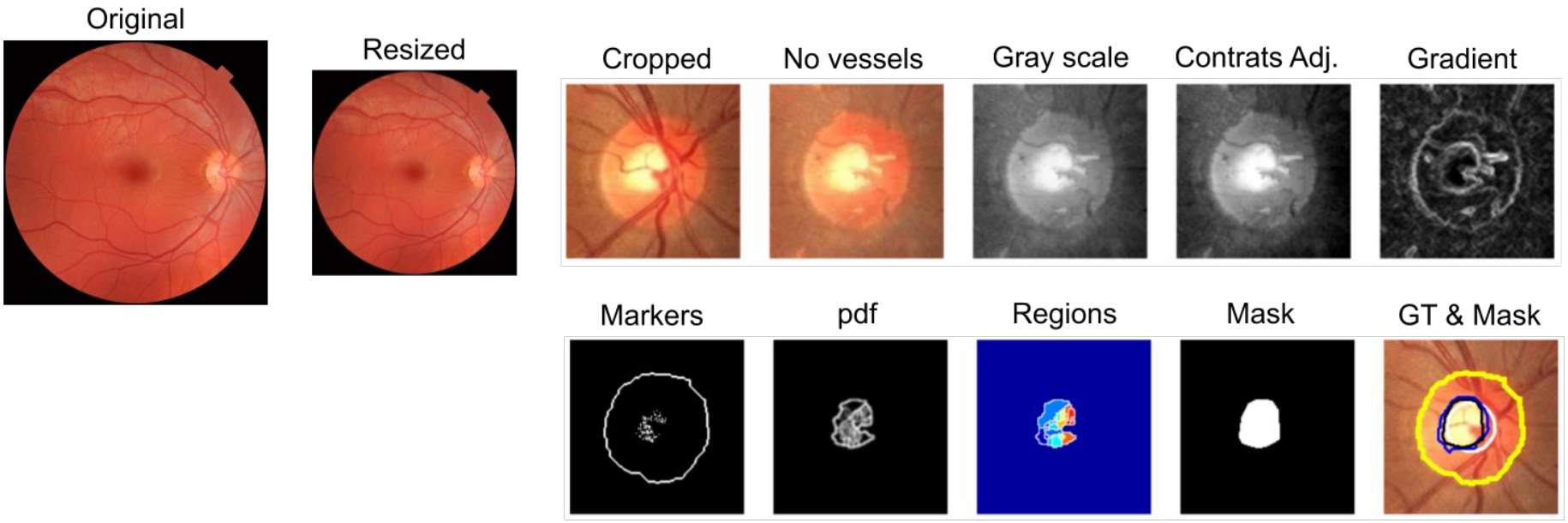


A Dice index of **0.70** was obtained for the optic cup segmentation

Stochastic-Watershed-based method



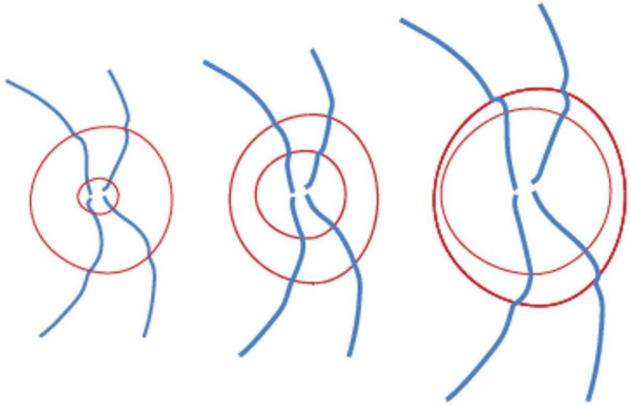
Stochastic-Watershed-based method



Stochastic-Watershed-based method

Clinical features

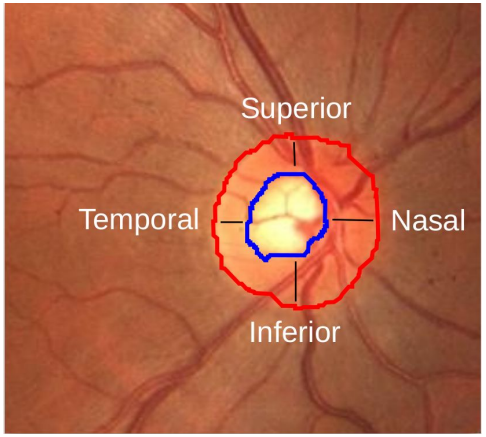
Cup/Disc ratio (**CDR**): Vertical ratio between Cup and Disc



Area Cup/Disc ratio (**ACDR**): Ratio between area occupied by the Cup and the Disc

ISNT rule: Inferior > Superior > Nasal > Temporal

A normal eye follows this rule



Stochastic-Watershed-based method

Glaucoma Diagnosis

	CMYK		YIQ		Luv		Lab		PCA		RGB	
	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se
CDR	0,574	0,697	**0,675**	**0,674**	0,650	0,731	0,832	0,563	0,487	0,716	0.545	0.716
ACDR	0,601	0,633	**0,715**	**0,604**	0,688	0,673	0,849	0,509	0,517	0,663	0.574	0.655
ISNT	0,495	0,570	**0,431**	**0,568**	0,422	0,561	0,337	0,609	0,523	0,544	0.499	0.511
Combined	0,545	0,702	**0,730**	**0,602**	0,685	0,635	0,373	0,760	0,376	0,778	0.513	0.742

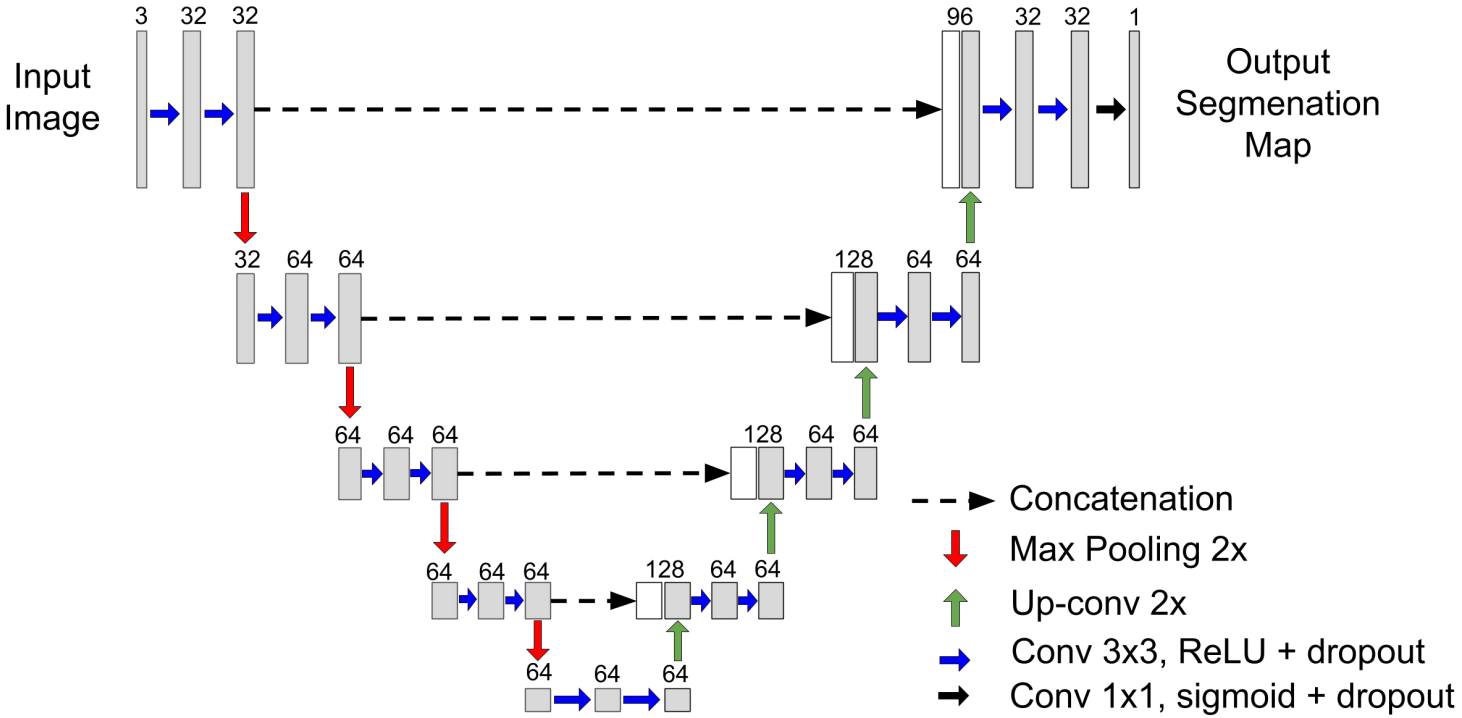
Combined means we used CDR and ISNT rule to assess glaucoma

Outline

1. Introduction
2. **Segmentation Methods**
   - Stochastic-Watershed-based method
   - **U-Net-based approach**
3. Classification Methods
4. Image synthesis

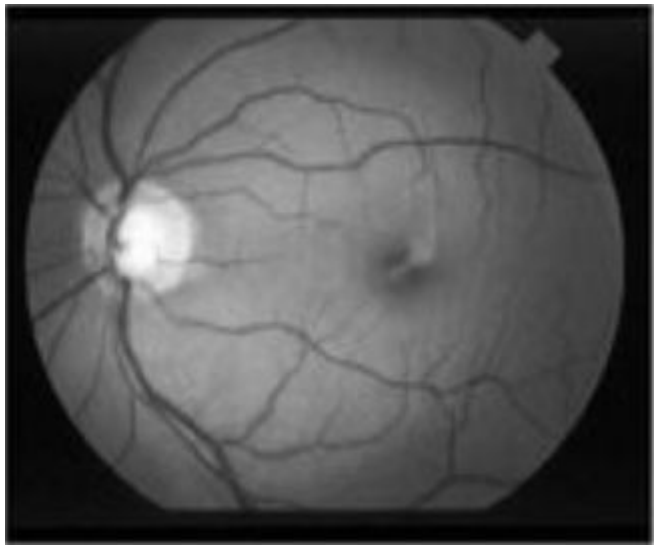
U-Net-based approach

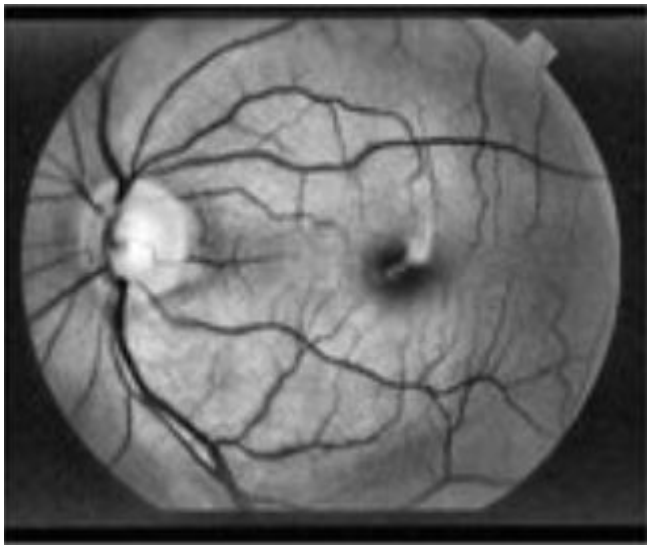
- Fully convolutional network
- The contracting path captures context.
- Symmetric expanding path enables precise localization.



U-Net-based approach

- We used Contrast Limited Adaptive **Histogram Equalization (CLAHE)** as a preprocessing technique





U-Net-based approach

- We applied data augmentation to use the available annotated samples more efficiently.

  Rotation, translation and zoom

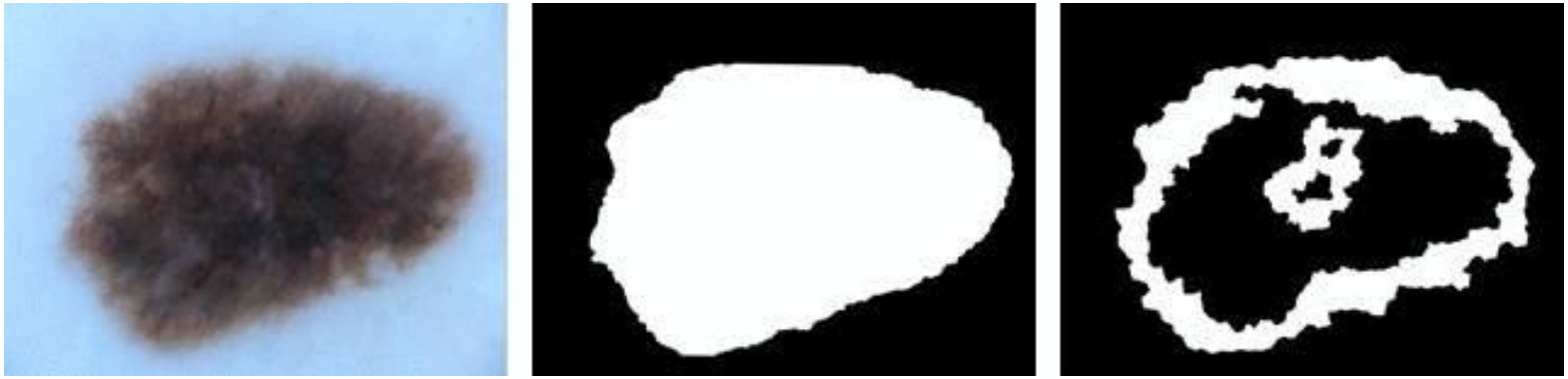
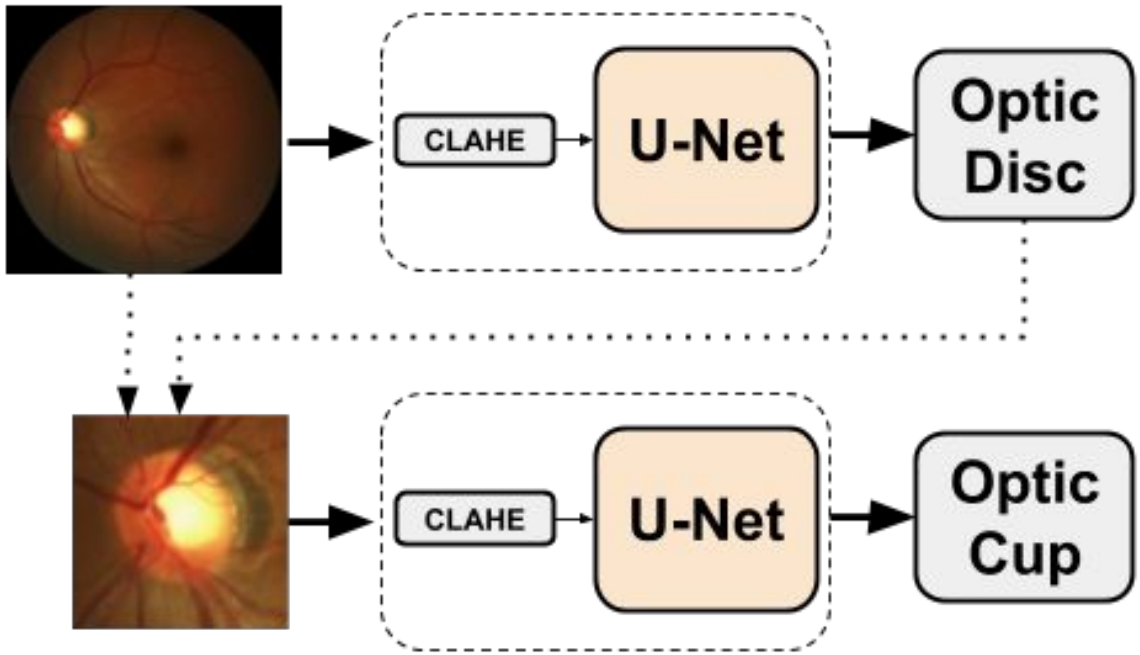


Image taken from: https://github.com/mdbloice/Augmentor

U-Net-based approach

Schema used for Optic Disc and Optic Cup segmentation



A Dice index of **0.91** and **0.78** were obtained for the optic disc and optic cup

Outline

1. Introduction
2. Segmentation Methods
3. **Classification Methods**
   - **ImageNet-trained CNN architectures**
   - Ensemble Setting with CNNs
4. Image synthesis

ImageNet-trained CNN architectures

ImageNet-trained CNN architectures applied for retinal image classification:

VGG16 and VGG19: These CNNs are based on the same model and characterized by their simplicity. Presented by Simonyan in 2014 for the ImageNet challenge

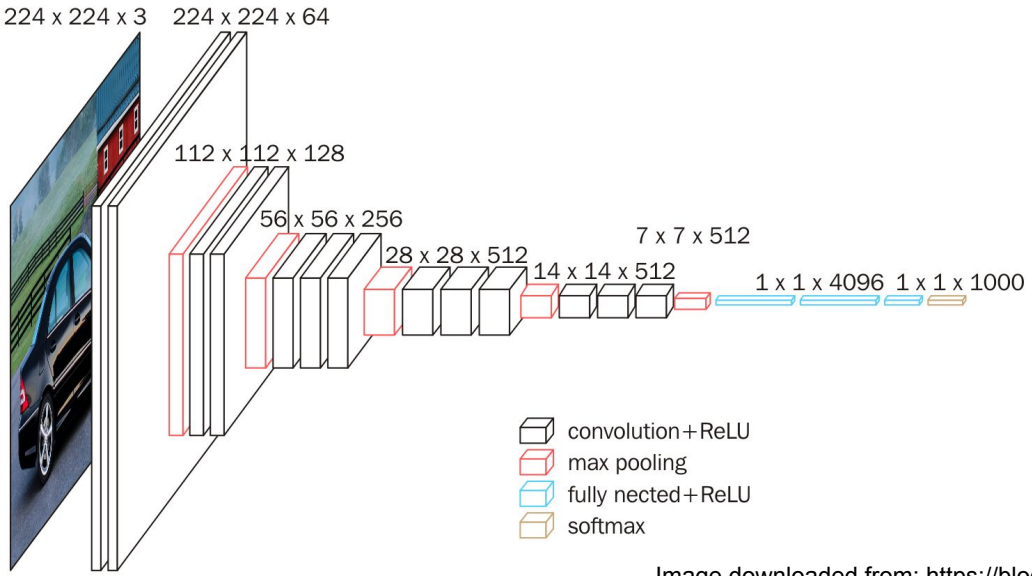
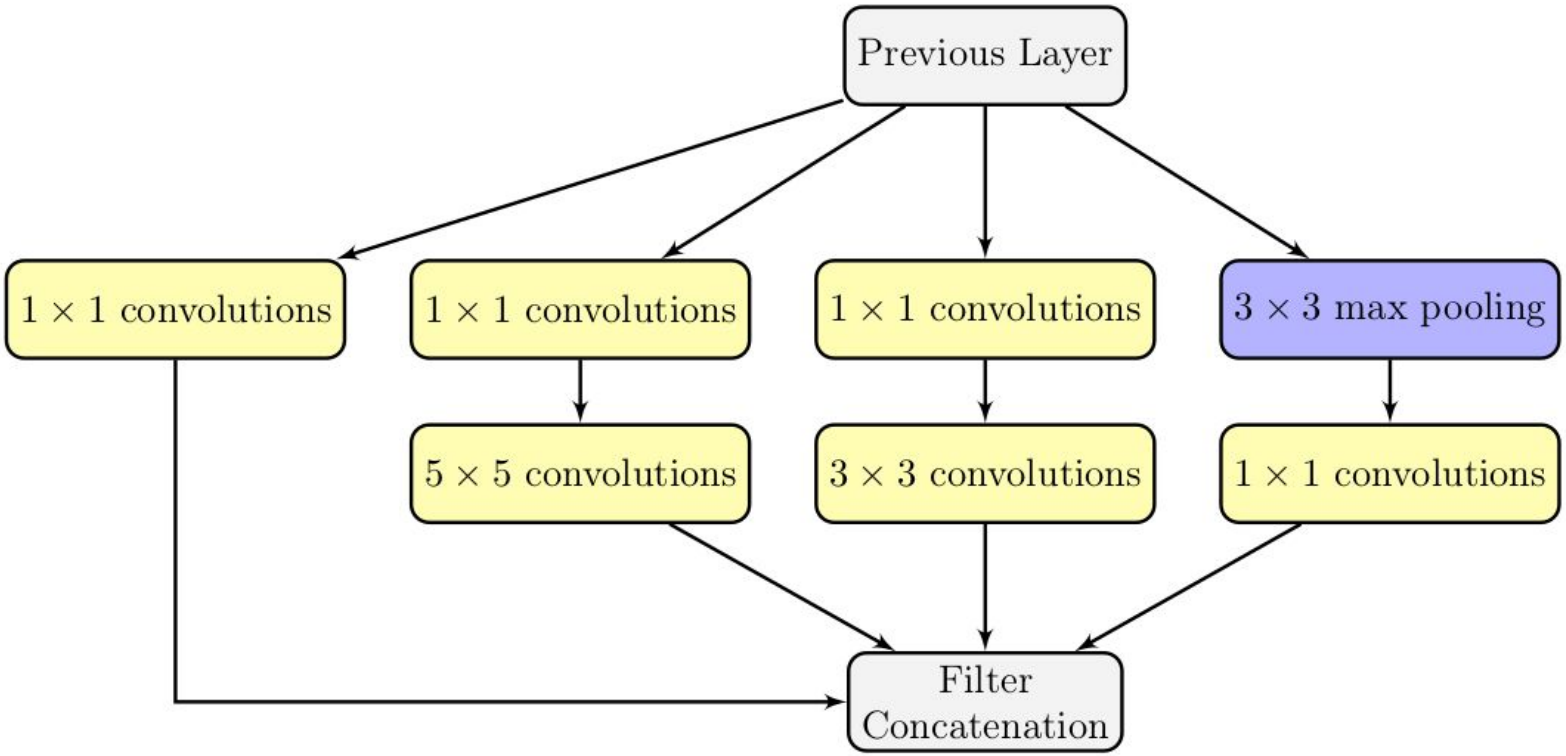


Image downloaded from: https://blog.datawow.io/cnn-models-ef356bc11032

ImageNet-trained CNN architectures

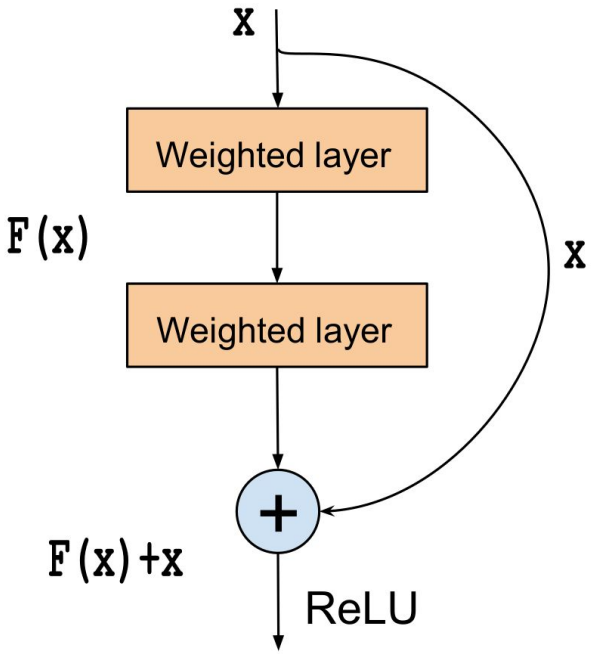
GoogLeNet: It was first introduced by Szegedy et al. in 2015. It is based on the Inception module.



ImageNet-trained CNN architectures

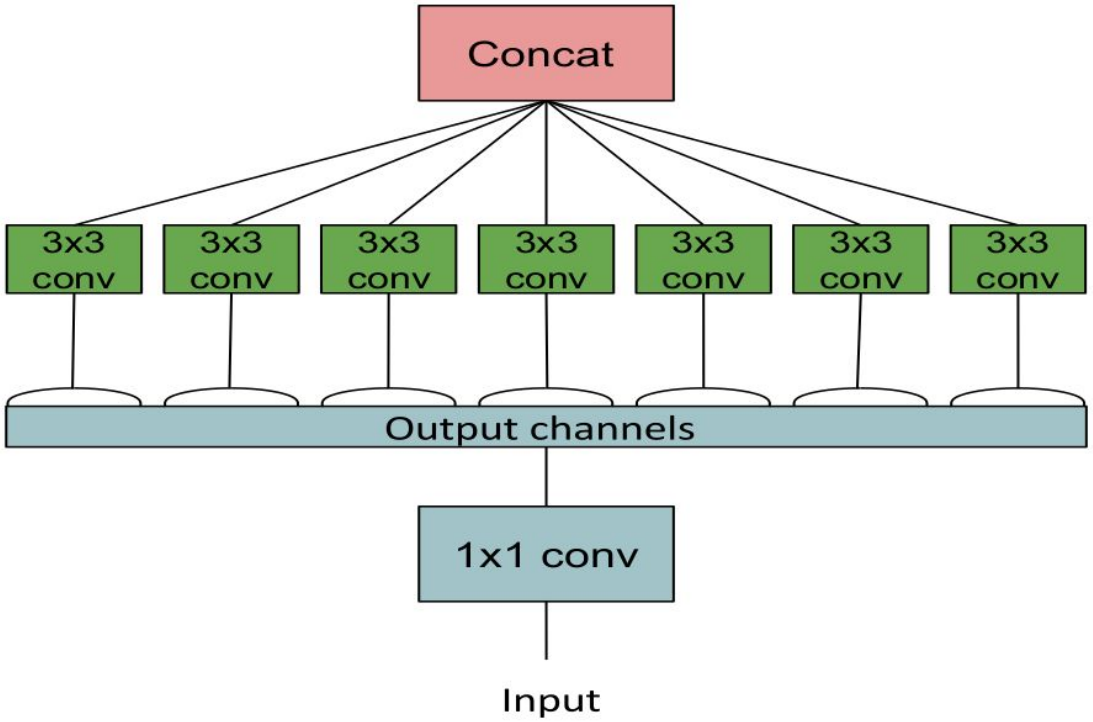
Microsoft ResNet: Proposed by the Microsoft Research Asia team (MSRA) in 2015.

LET'S GO DEEPER!



ImageNet-trained CNN architectures

Xception: or Extreme Inception, it was proposed by F. Chollet in 2016. It is an extension of the Inception architecture.

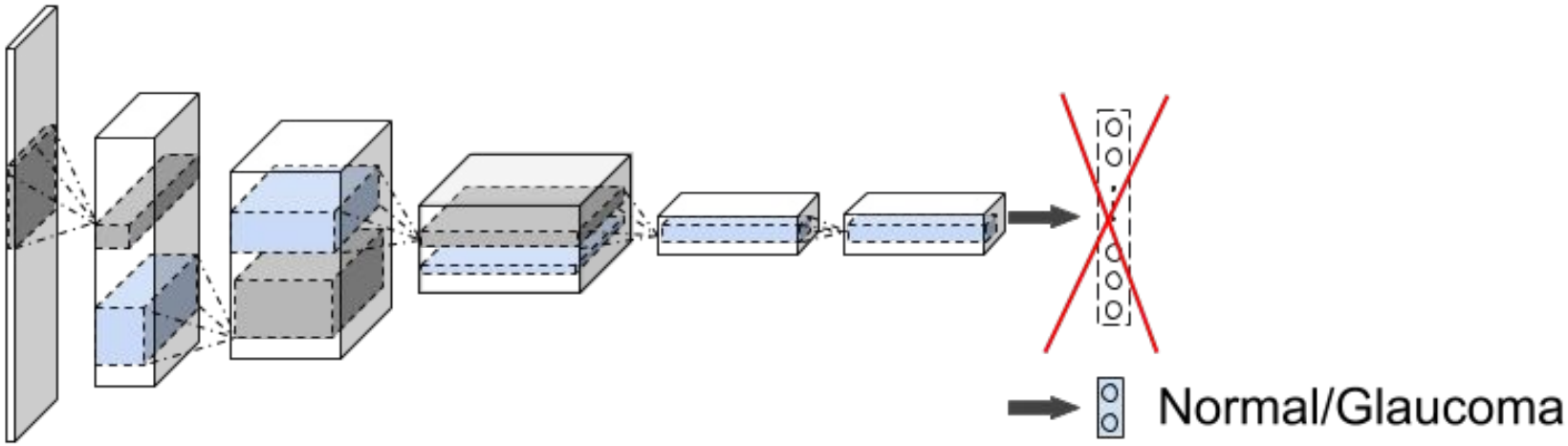


ImageNet-trained CNN architectures

Fine-tuning technique:

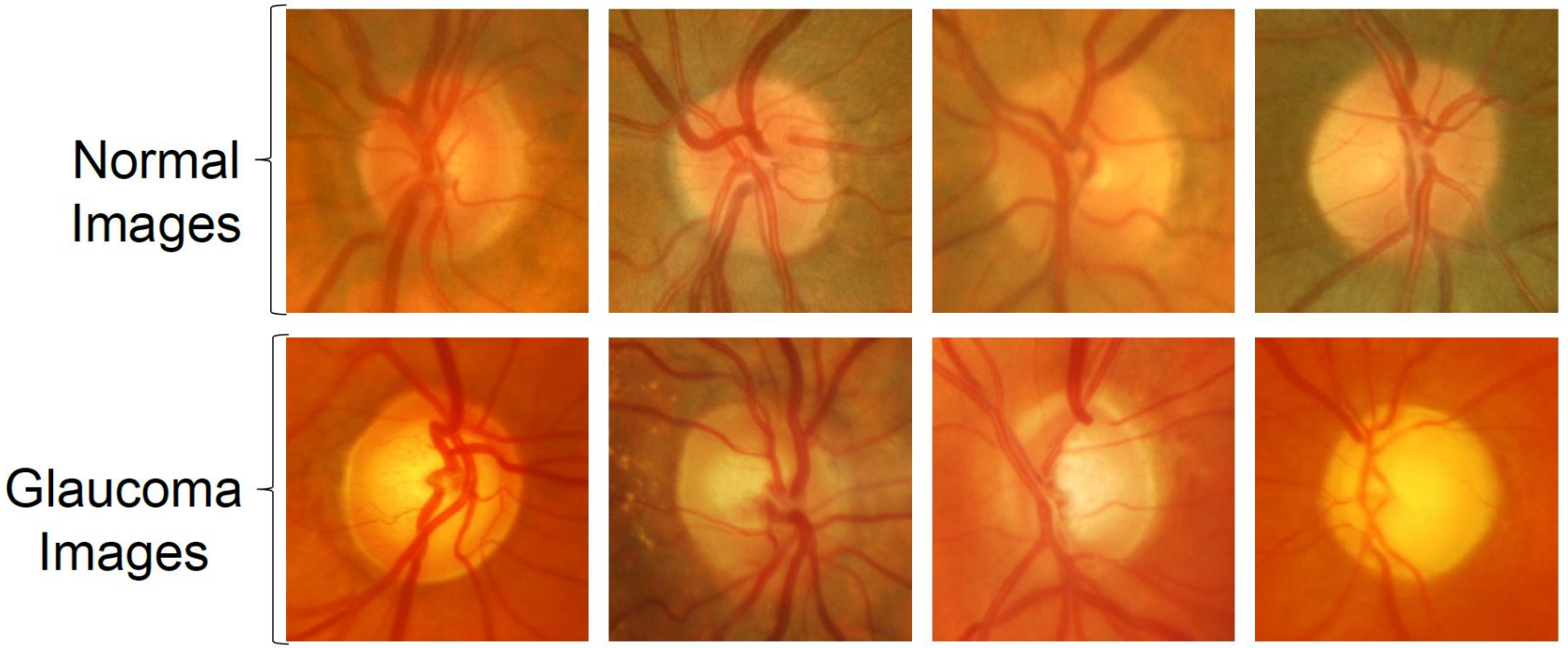
a) The weight initialization of the convolutional layers using the ImageNet weights and

b) The replacement of the classification function or the number of nodes in the last fully connected layer.



ImageNet-trained CNN architectures

All these images were automatically cropped around the optic disc using a deep learning method[1]



1) Xu P, Wan C, Cheng J, Niu D, Liu J. Optic disc detection via deep learning in fundus images. Fetal, infant and ophthalmic medical image analysis.

ImageNet-trained CNN architectures

Results

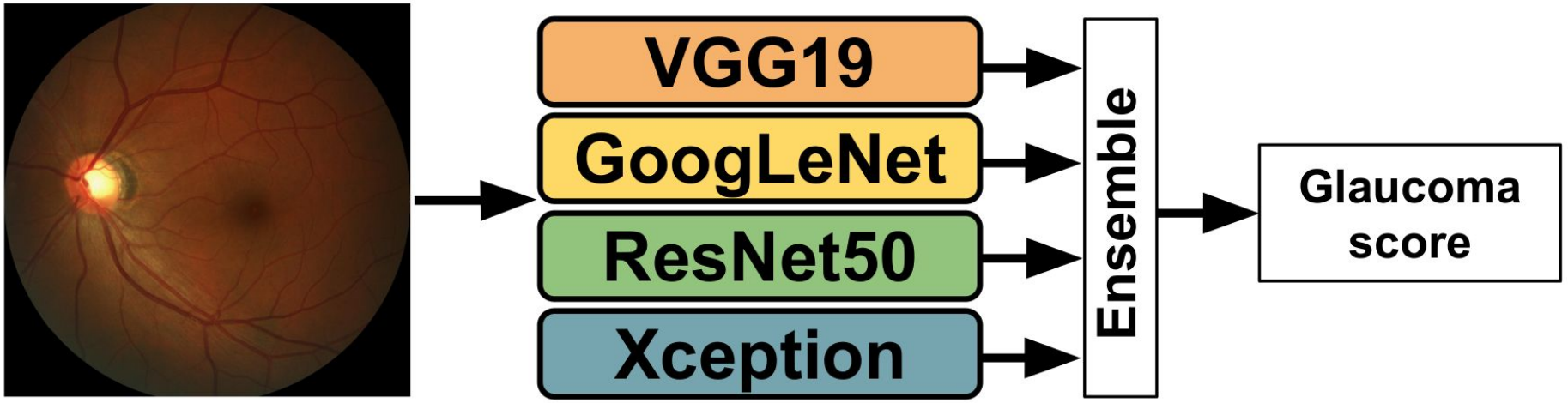
Model Name	AUC	Accuracy	F-score	# parameters (in millions)
VGG16	0.9632	0.8948	0.9005	138
VGG19	0.9686	0.9069	0.9125	144
InceptionV3	0.9653	0.9000	0.9056	23
ResNet50	0.9614	0.9029	0.9076	25
Xception	0.9605	0.8977	0.9051	22

Outline

1. Introduction
2. Segmentation Methods
3. **Classification Methods**
   - ImageNet-trained CNN architectures
   - **Ensemble Setting with CNNs**
4. Image synthesis

Ensemble Setting with CNNs

Ensemble setting reduces the testing error



Common technique used in Kaggle competitions:

https://medium.com/neuralspace/kaggle-1-winning-approach-for-image-classification-challenge-9c1188157a86

https://www.kaggle.com/

Outline

1. Introduction
2. Segmentation Methods
3. Classification Methods
4. **Image synthesis**
   - **Using VAE and DCGAN**
   - Semi-supervised Learning using DCGAN

Image synthesis

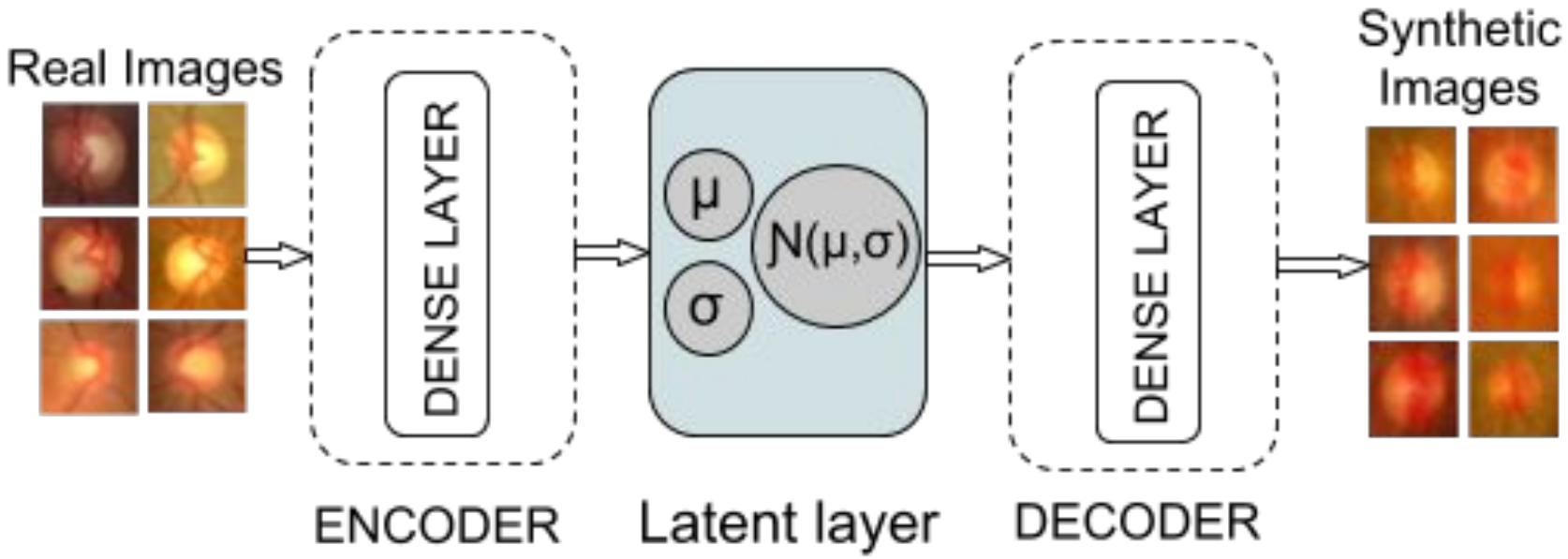
Reasons

- Very limited data
- Use to generalise automatic glaucoma assessment methods

Using VAE and DCGAN

The Variational Autoencoder (VAE)[1] is composed by

- Approximate inference network (or encoder)
- Decoder network



1 **Auto-Encoding Variational Bayes.** Diederik P Kingma, Max Welling - http://kvfrans.com/variational-autoencoders-explained/

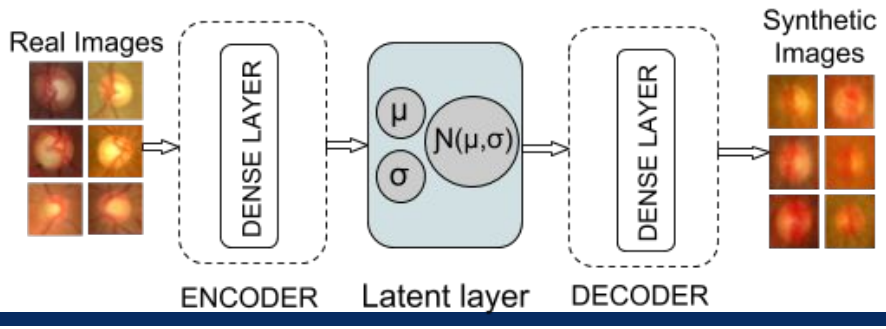
Using VAE and DCGAN

Differences between VAE and standard autoencoder:

- **Latent variables follow a unit gaussian distribution**
- **Loss function composed of separate losses:**

  The generative loss → Mean squared error that measures how accurately the network reconstructed the images

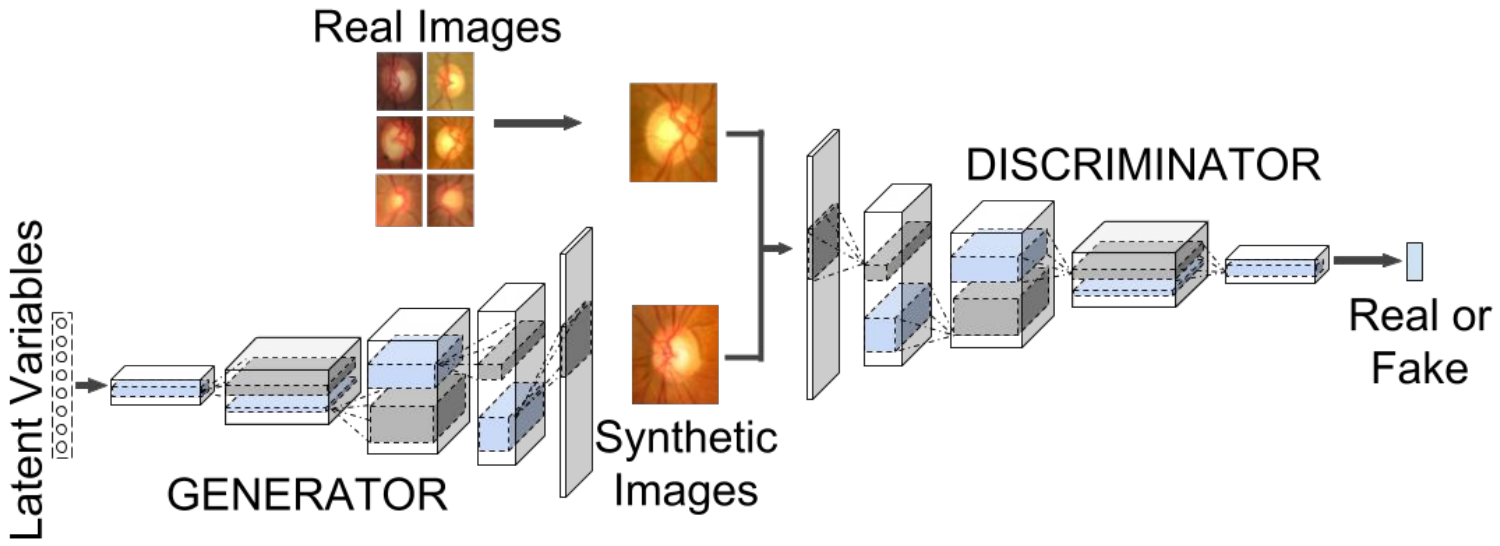
  Latent loss → Kullback Leibler divergence that measures how closely the latent variables match a unit gaussian.



Using VAE and DCGAN

The Deep Convolutional Generative Adversarial Network (DCGAN)[1].

- It also consists of two networks, the generator and discriminator.
- A major improvement on the first GAN.



1 Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks.
Alec Radford, Luke Metz, Soumith Chintala

Using VAE and DCGAN

- **Analysed resolutions**

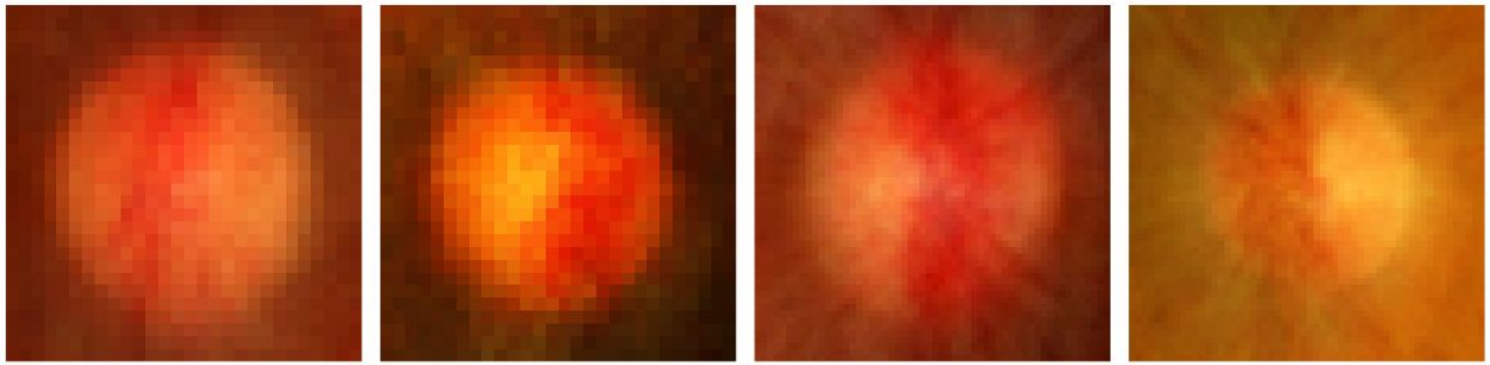
  28x28 pix, 56x56 pix, 112x112 pix and 224x224 pix.

- **Latent space**

  From 32 to 100 latent variables (multivariate Gaussian).

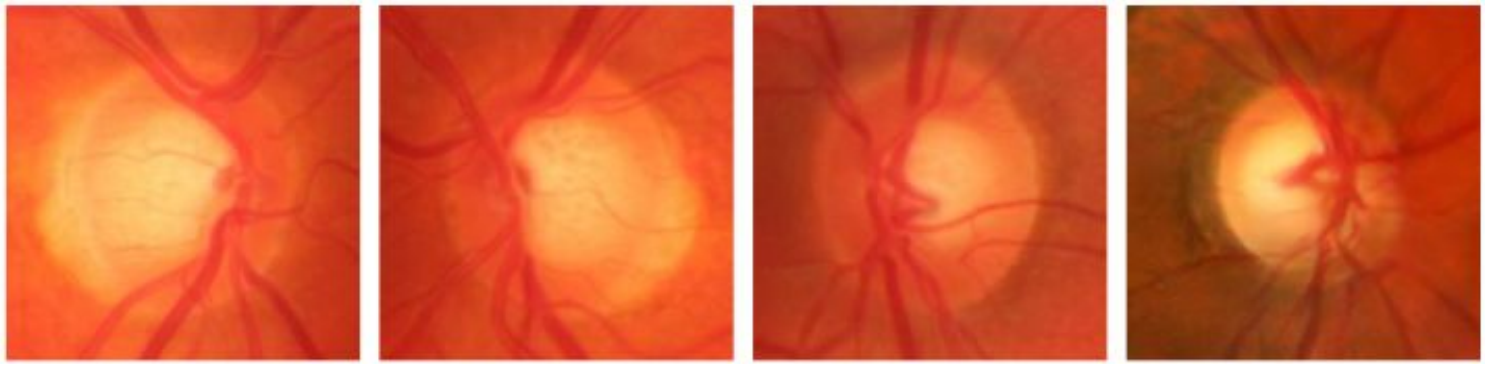
Using VAE and DCGAN

Results VAE: **100 latent variables**



28x28 px 56x56 px

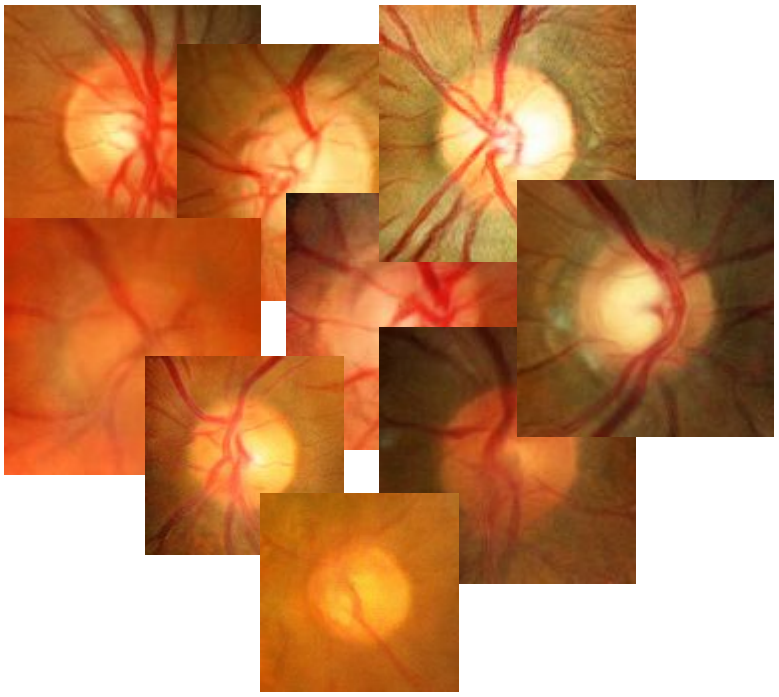
Results DCGAN: **100 latent variables**



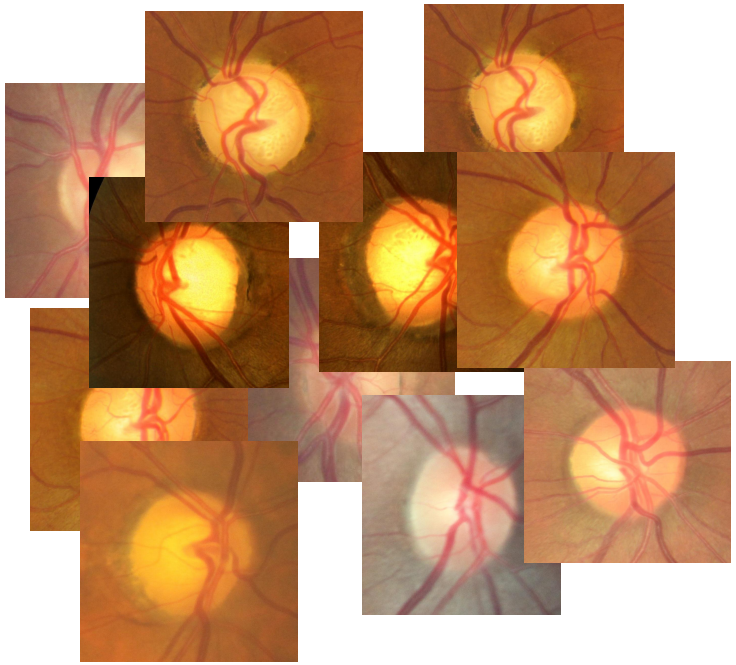
224x224 px

Using VAE and DCGAN

For qualitative and quantitative evaluation, 100 synthetic images and 100 real images were selected



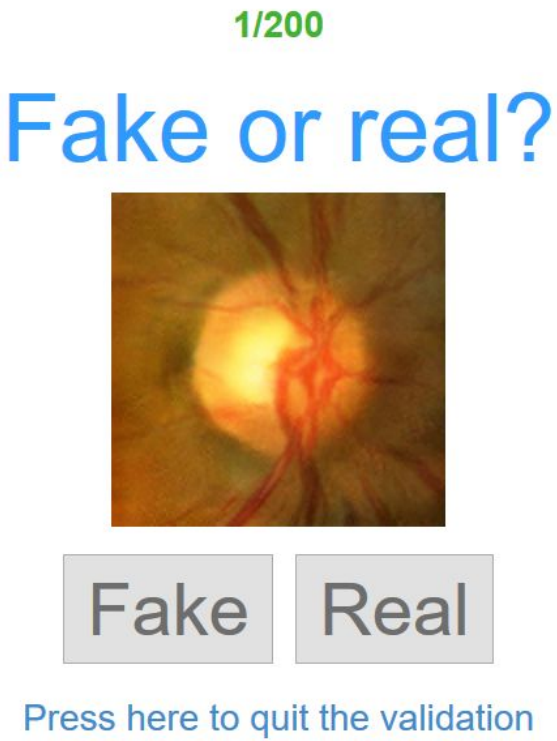
Synthetic images



Real images

Using VAE and DCGAN

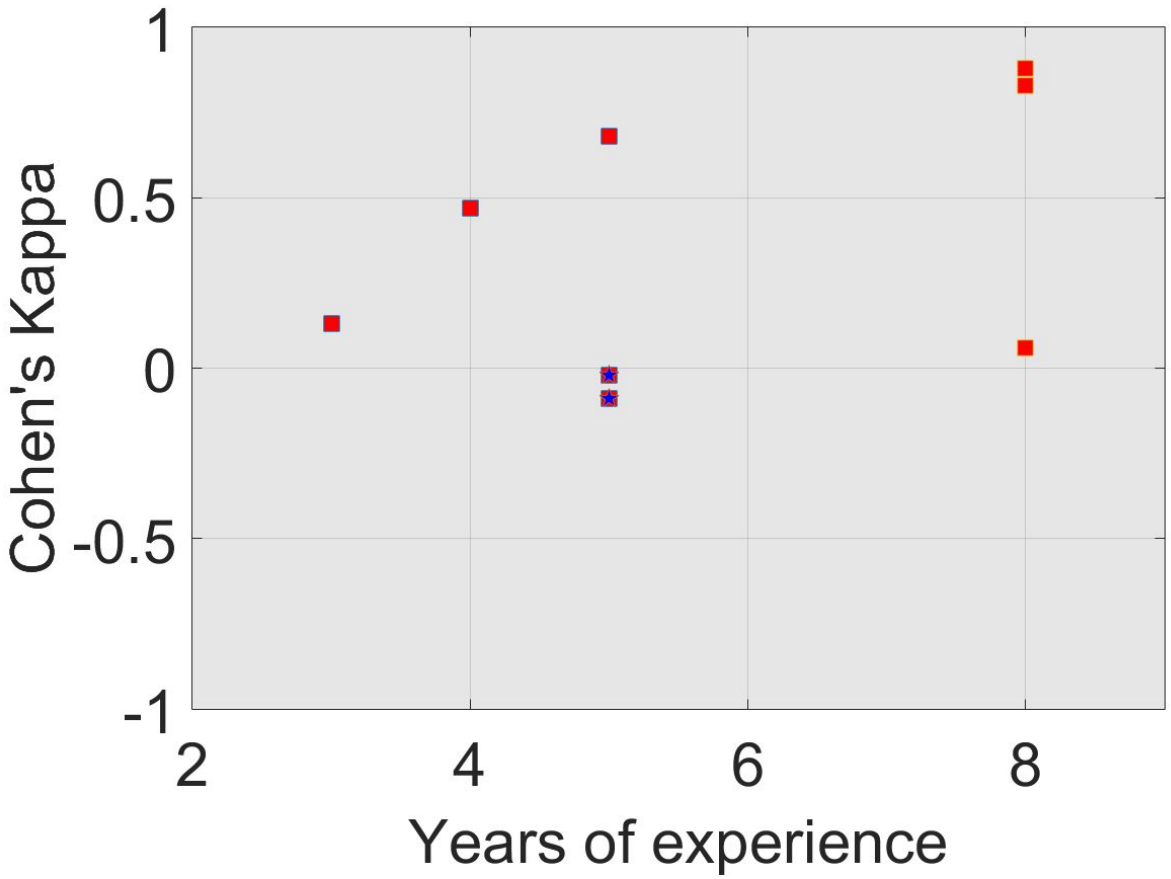
Qualitative evaluation



Web App: https://cvblab.synology.me/ganval/index.php

Using VAE and DCGAN

Results qualitative evaluation DCGAN images



Ten experts with 3 to 8 years of experience

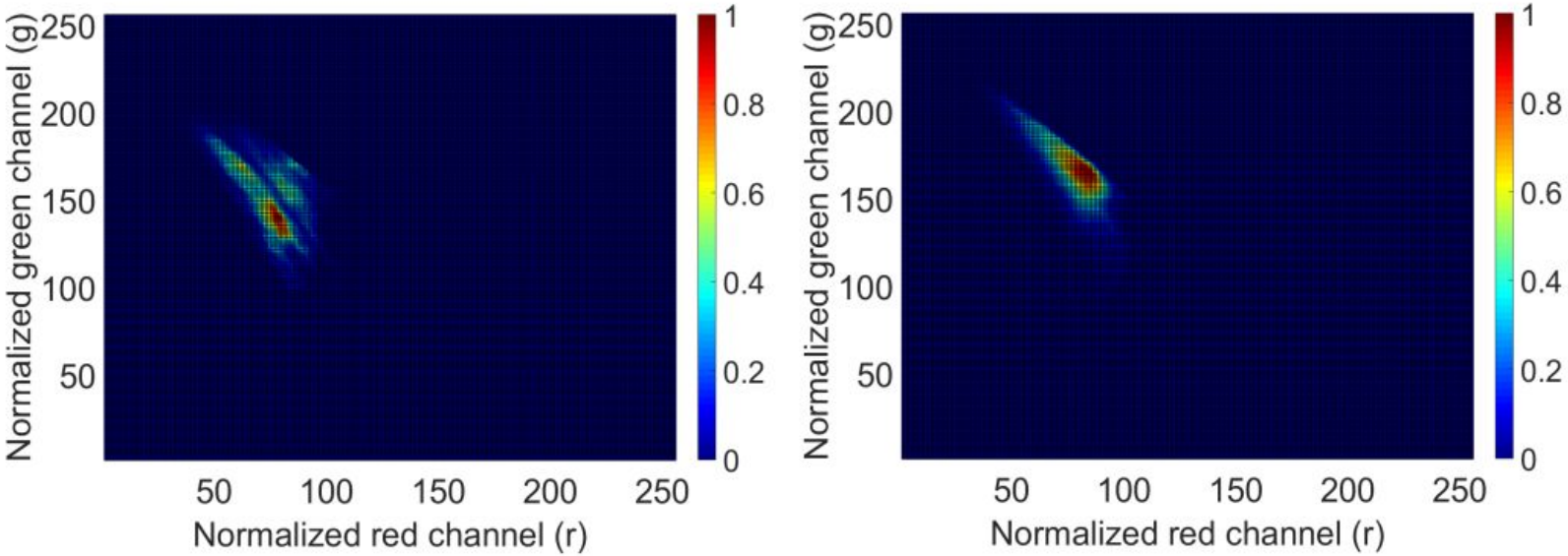
Cohen's kappa

- 0 represents random chance
- 1 represents a perfect agreement between the ground-truth and the expert.

Using VAE and DCGAN

Results quantitative evaluation DCGAN and Real images: 2D-histograms[1]

Average 2D-histograms: RGB channels normalized by the luminance



DCGAN images ***Real images***

1 Adrian Colomer et al., Colour normalization of fundus images based on geometric transformations applied to their chromatic histogram.

Using VAE and DCGAN

Results quantitative evaluation DCGAN and Real images:

Average and standard deviation of the mean-squared error

Average 2D-histogram	**Real Images**	**Synthetic Images**
Real	0.0028 (0.000325)	0.0036 (0.000543)
Synthetic	**0.0031 (0.000461)**	0.0022 (0.000562)

Using VAE and DCGAN

Results quantitative evaluation DCGAN and Real images:

Average vessel, Optic Disc and Background proportion

	Synthetic Images	**Real Images**
Vessel proportion[1]	0.1431 (0.0306)	0.1519 (0.0306)
Optic Disc proportion	0.1776 (0.0339)	0.2456 (0.0722)
Background	0.6792 (0.0428)	0.6025 (0.0795)

Quality evaluation of synthetic images should be specific for each application[2]!

1 Sandra Morales et al., Computer-Aided Diagnosis Software for Hypertensive Risk Determination Through Fundus Image Processing.
2 L Theis et al., A note on the evaluation of generative models.

Outline

1. Introduction
2. Segmentation Methods
3. Classification Methods
4. **Image synthesis**
   - Using VAE and DCGAN
   - **Semi-supervised Learning using DCGAN**

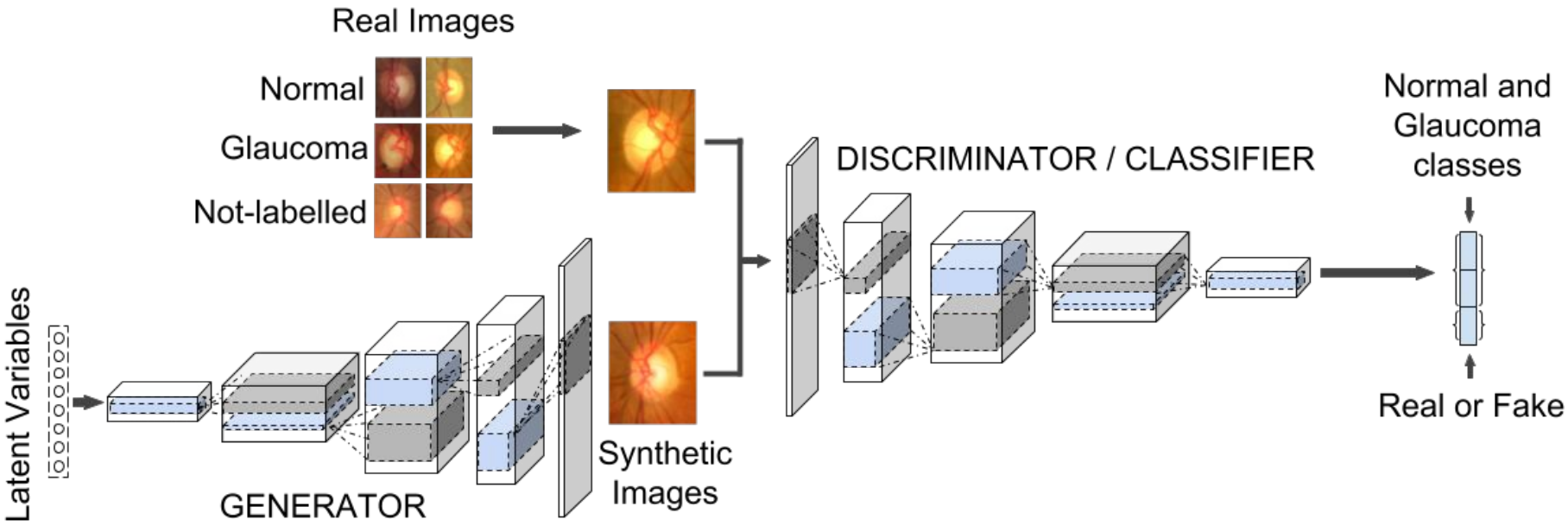
Semi-supervised Learning using DCGAN

- **We trained the DCGAN as image synthesizer and as semi-supervised learning method**

- **Using semi-supervised learning better classifier can be built with a large amount of unlabelled data and small set of labelled data**

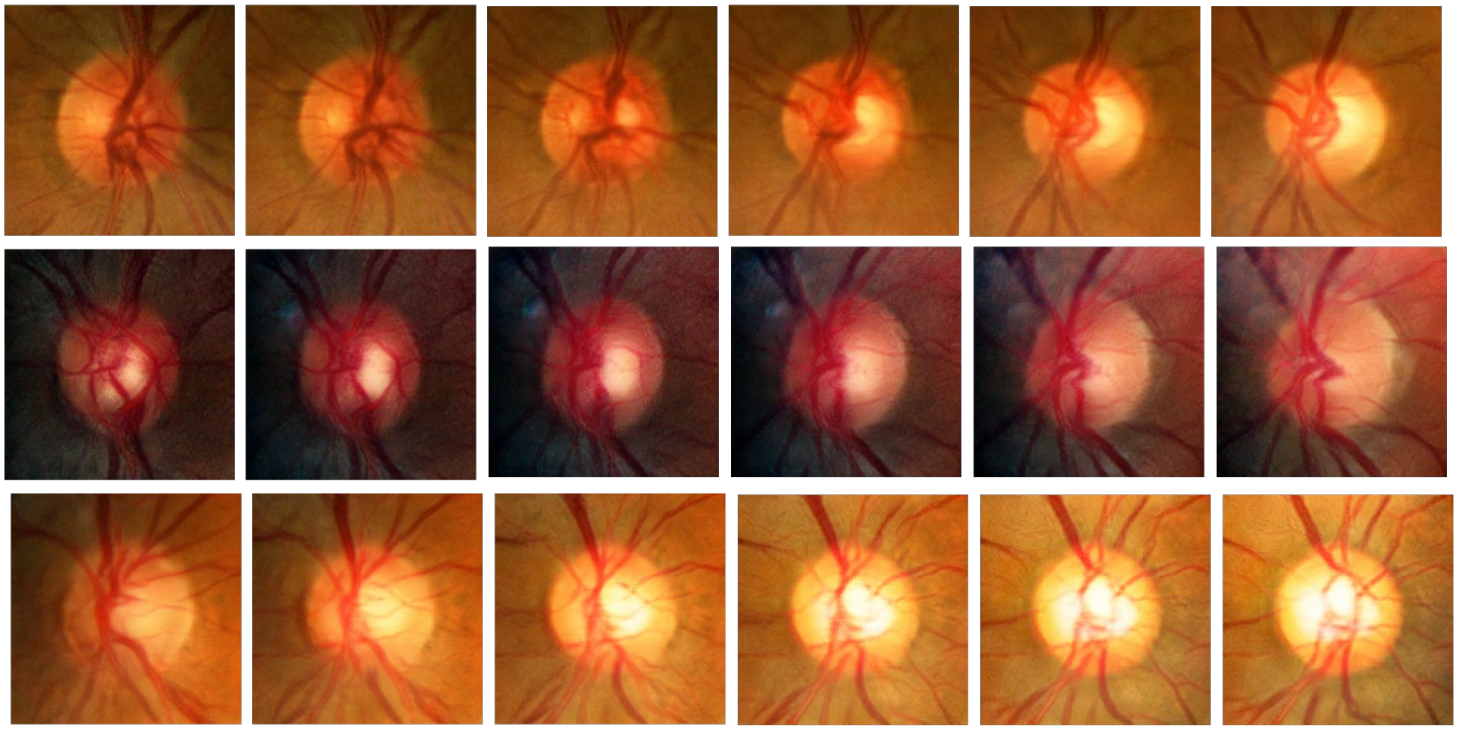
Semi-supervised Learning using DCGAN

SS-DCGAN architecture



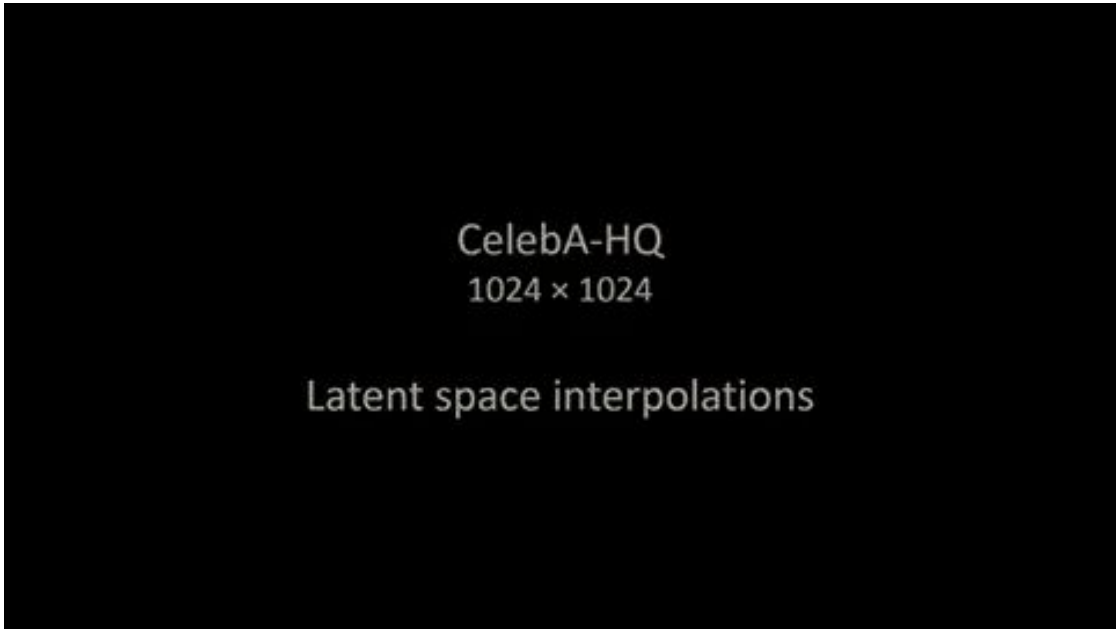
Semi-supervised Learning using DCGAN

Examples of the DCGAN



Semi-supervised Learning using DCGAN

Examples of the Progressive Grown GANs by NVIDIA

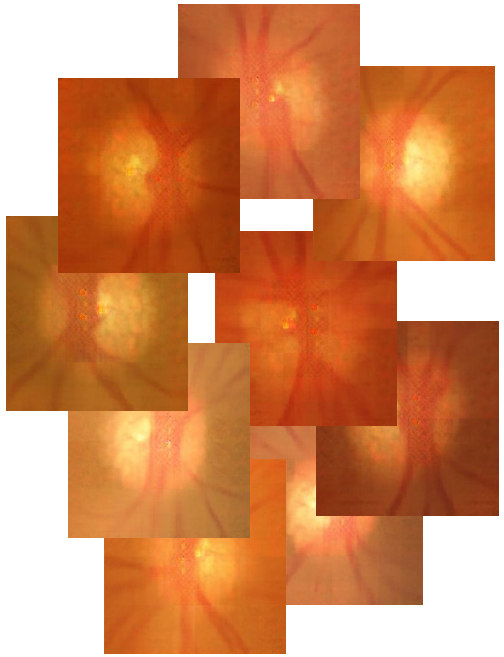


Tero Karras et al., Progressive Growing of GANs for Improved Quality, Stability, and Variation. ICLR 2018

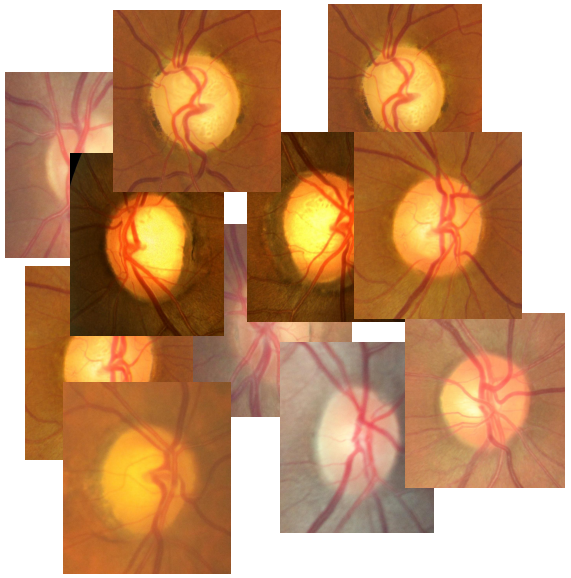
Using VAE and DCGAN

For qualitative and quantitative evaluation, 100 DCGAN images, 100 Costa's images and 100 real images were selected

DCGAN images



Costa's images

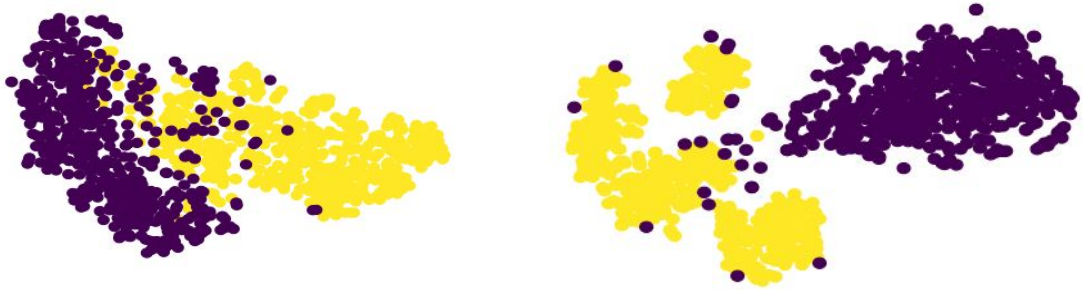


Real images

Semi-supervised Learning using DCGAN

Qualitative evaluation using t-SNE

100 features extracted from the ResNet50 trained on ORIGA-light



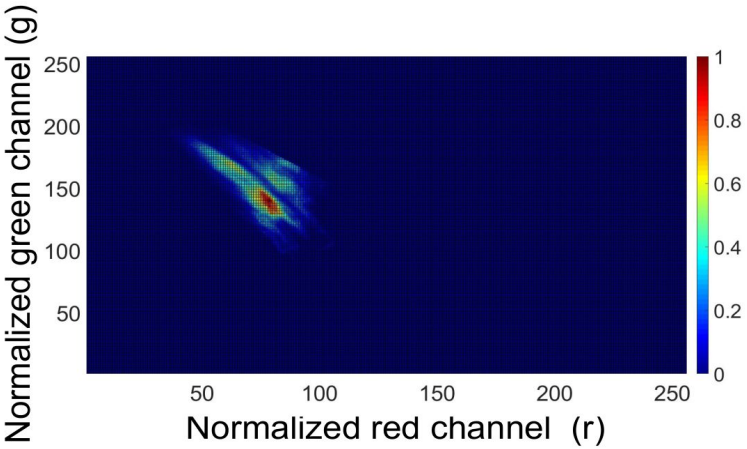
DCGAN images **Costa's images**

Yellow dots represent the features extracted from the real images

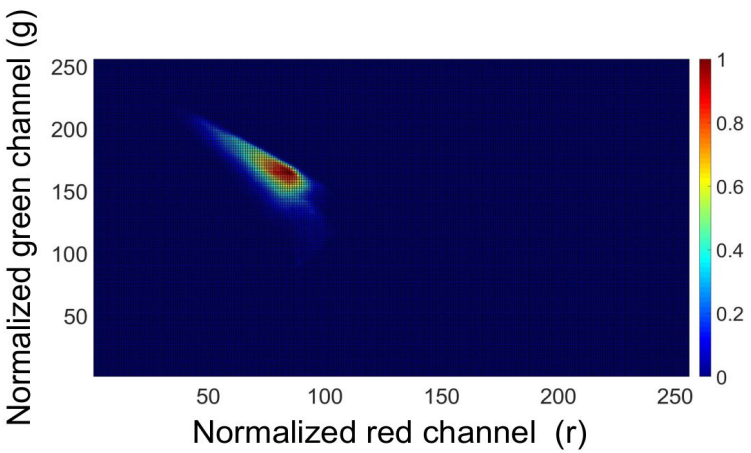
t-SNE stands for **t-Distributed Stochastic Neighbor Embedding**. It is a technique for dimensionality reduction that is particularly well suited for the visualization of high-dimensional datasets

Semi-supervised Learning using DCGAN

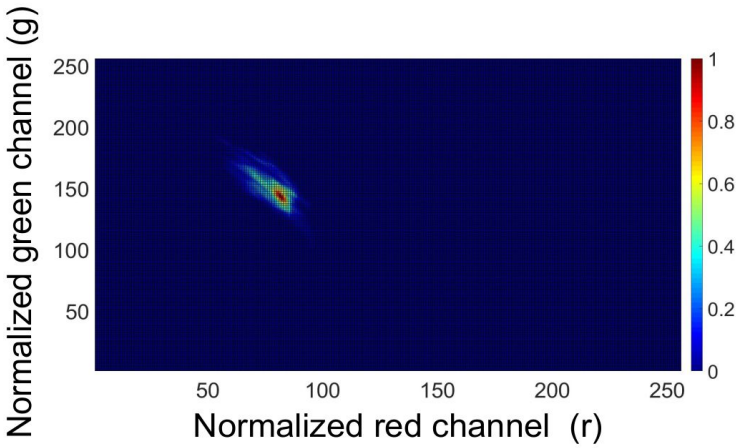
Quantitative evaluation: Average 2D-histograms



DCGAN images



Real images



Costa's images

Semi-supervised Learning using DCGAN

Quantitative evaluation

Average and standard deviation of the mean-squared error

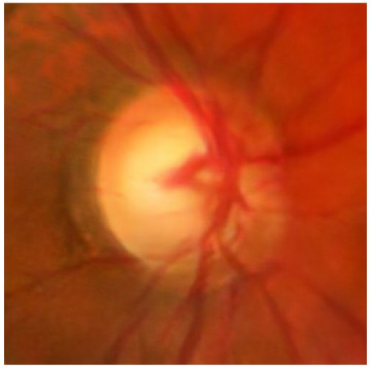
Average 2D-histogram	**Real Images**	**DCGAN method**	**Costa's method**
Real	0.0028 (0.000325)	0.0036 (0.000543)	0.0013 (0.000262)
DCGAN method	**0.0031 (0.000461)**	0.0022 (0.000562)	0.0016 (0.000439)
Costa's method	**0.0031 (0.000126)**	0.0035 (0.000178)	0.0010 (0.000163)

Semi-supervised Learning using DCGAN

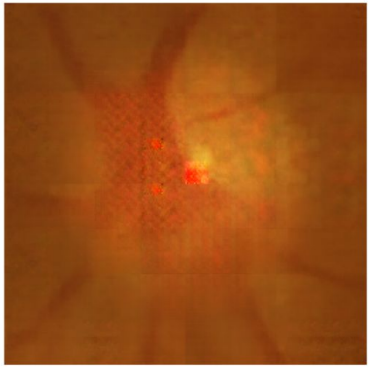
Quantitative evaluation: *Vessel, Optic Disc and background proportion*

	Real Images	DCGAN Images	Costa's method
Vessel proportion	**0.1519 (0.0306)**	**0.1431 (0.0306)**	**0.1026 (0.0195)**
Optic Disc proportion	0.2456 (0.0722)	0.1776 (0.0339)	0.1851 (0.0396)
Background	0.6025 (0.0795)	0.6792 (0.0428)	0.7122 (0.0437)

Real sample



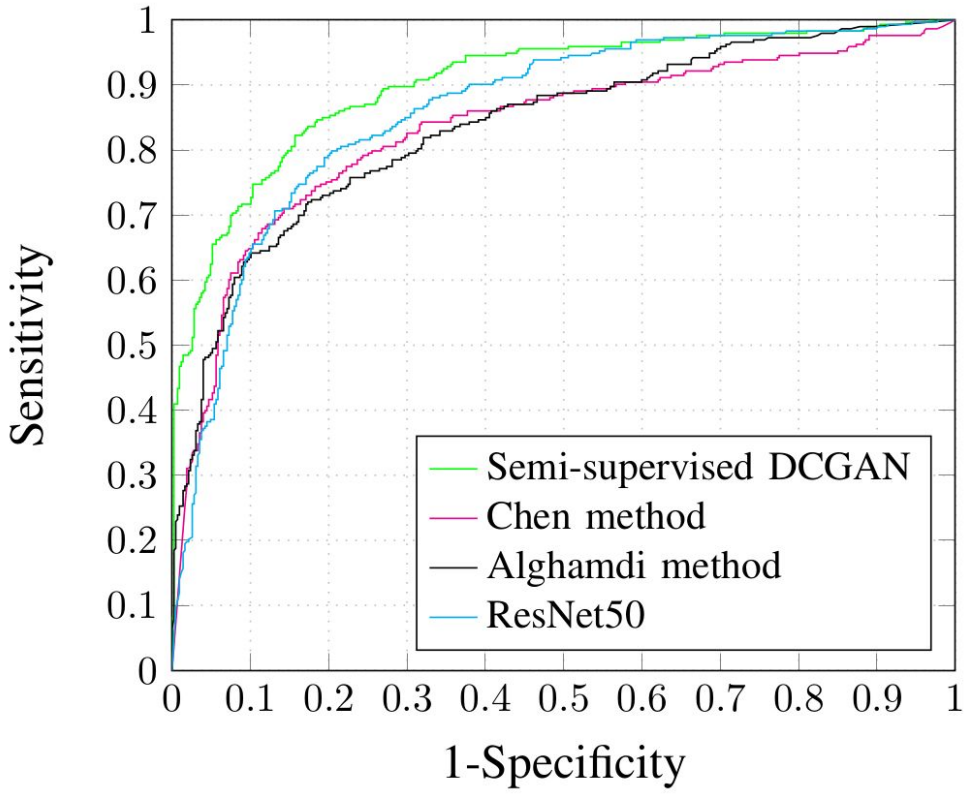
DCGAN sample



Costa's sample

Semi-supervised Learning using DCGAN

Results glaucoma classifier using SS-DCGAN



ROC curve illustrates the diagnostic ability of a binary classifier system

Thank you!

Andres Diaz-Pinto
a.diaz-pinto@leeds.ac.uk