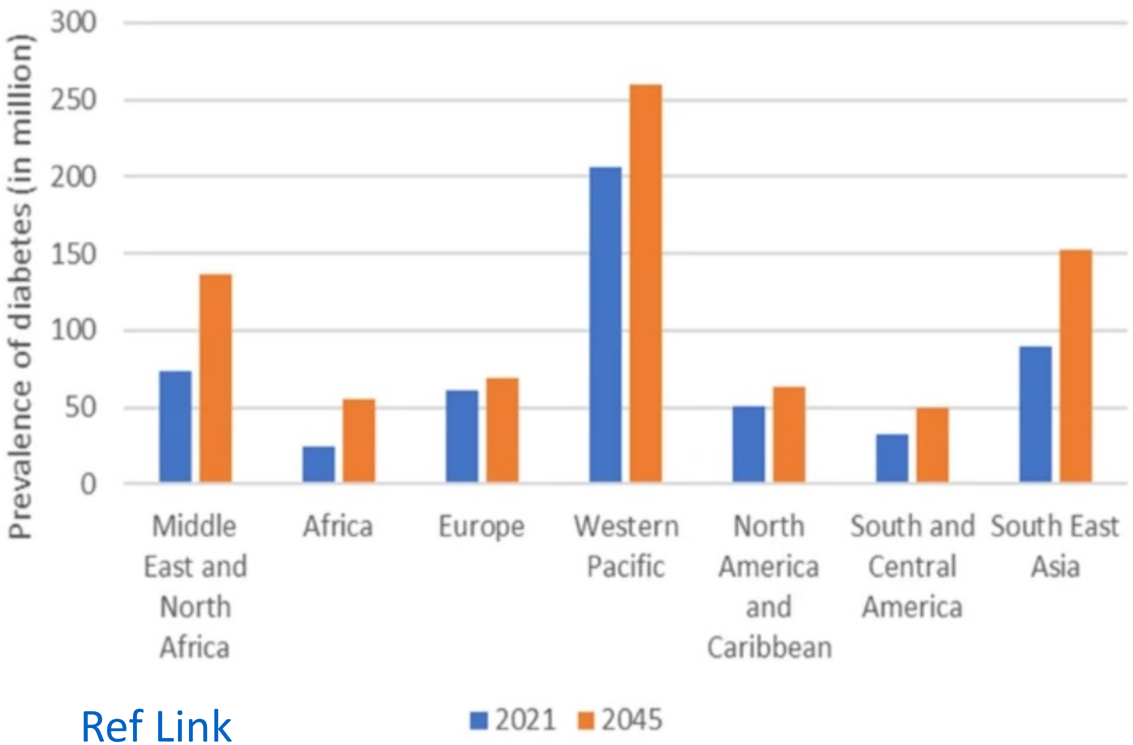
Early Detection and Severity Classification of Diabetic Retinopathy Through Image Processing and Deep Learning

Nelson Nishio[1], Jason Nishio[2], Alexander Nishio[3]

[1,2]BASIS Independent Silicon Valley, San José, California

[3]Department of Computer Science, Purdue University, West Lafayette, USA

Introduction



Ref Link

- Diabetic retinopathy (DR): leading cause of blindness among working-aged adults
- **537 million people globally** live with diabetes and 103 million (22%) of them suffer from DR. In US, 1 in 10 people have diabetes and 1 in 5 of those with diabetes are unaware.
- The increasing prevalence of DR presents a substantial challenge especially in resource-constrained regions where ophthalmologists are burdened with heavy screening workloads
- **DR can be prevented if detected early and systematically treated.**

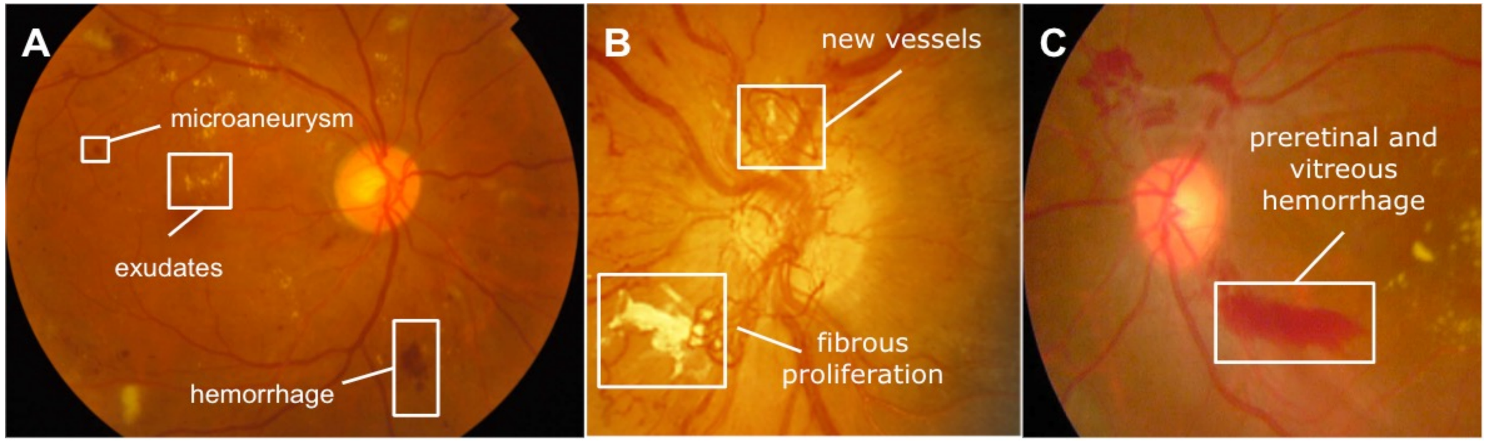
There is an urgent need to develop automated and effective DR diagnosis and prognostic techniques.

Research Gaps and Contributions

- **Research Gaps of DR detection using machine learning**
  1. A lack of systematic studies on applying image processing to DR images to enhance features for better DR detection
  2. Insufficient research on optimizing image processing techniques in alignment with deep learning architectures
  3. A lack of practical applications capable of detecting retinal lesions.

- **Key Contributions of this study**
  1. Carefully selected image processing techniques—CLAHE, Gaussian Blur, and ESRGAN—were systematically experimented and combined to highlight DR lesion features.
  2. The integrated pre-trained ResNet50 and VGG19 deep learning frameworks with proposed image preprocessing methods achieved superior performance.
  3. The proposed image enhancement methods can be integrated into retinal camera applications for self and professional DR screening.

APTOS 2019 Dataset



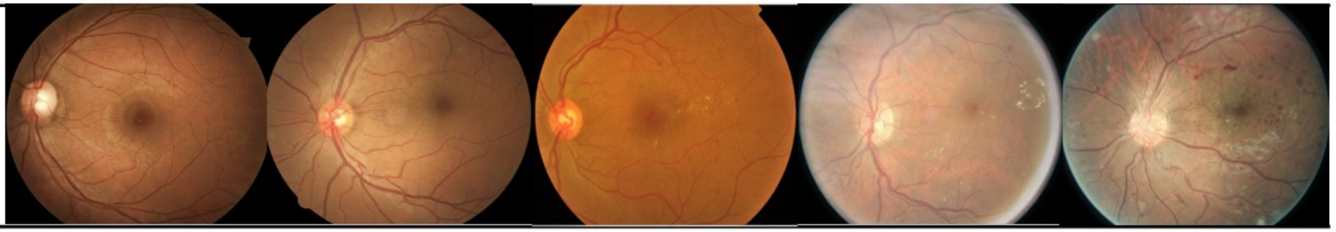
Stage	0 - No DR	1 - Mild	2 - Moderate	3 - Severe	4 - Proliferative
Count	1805	370	999	193	295
Image					

Image Processing: Graham Gaussian Blur

- Proposed by Ben Graham, the winner of the Kaggle 2015 DR grading competition to help remove image variations due to different lighting conditions.
- In this work, sigma is increased from 10 to 20 to make the retinal lesions more visible.

```
def apply_ben(self, image, sigmaX=10):
    image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)
    image=cv2.addWeighted ( image,4, cv2.GaussianBlur( image , (0,0) , sigmaX) ,-4 ,128)
    return image
```

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

x is the distance from the origin in the horizontal axis
y is the distance from the origin in the vertical axis
σ^2 is the sigma (standard deviation square)

Original | Graham10 (sigma =10) | Graham20 (sigma =20)

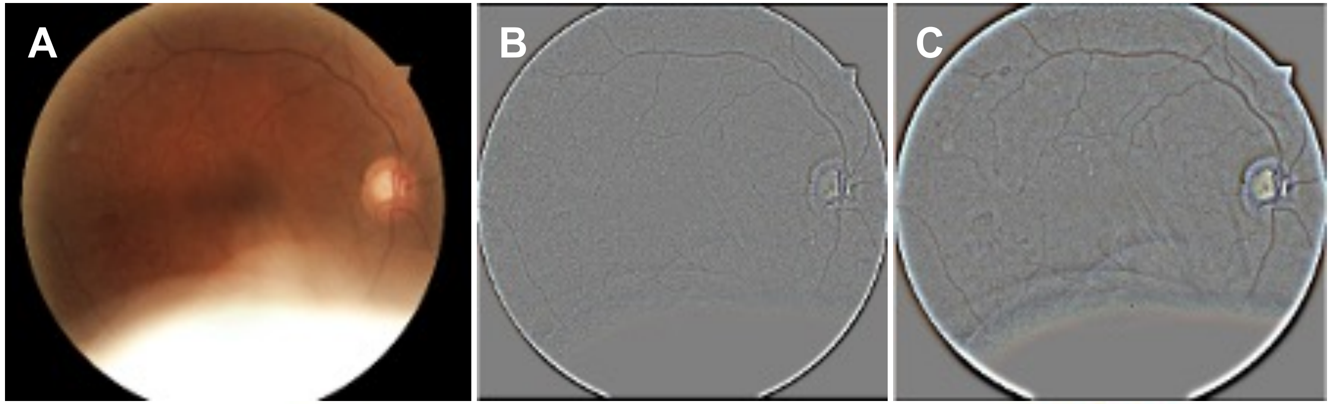
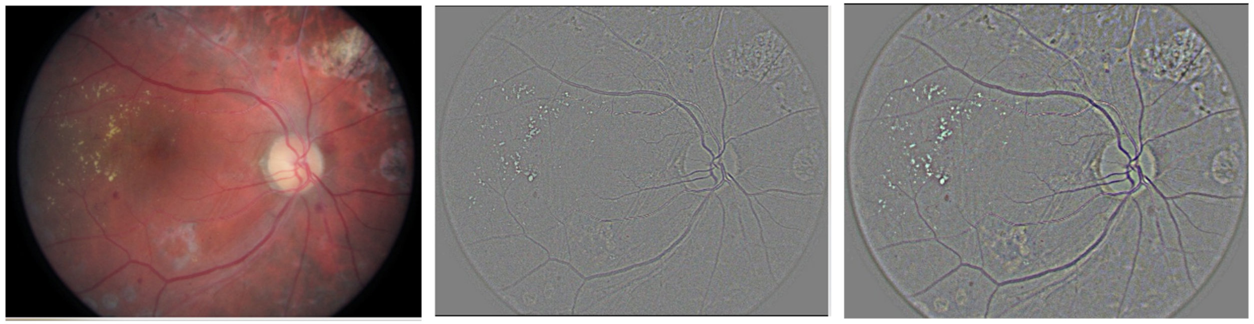


Image Processing: CLAHE

Contrast Limited Adaptive Histogram Equalization

Original | clip_limit = 1.0 | clip_limit = 2.0 (used in this work)

(a) Original | (b) Graham10+CLAHE | (c) Graham20+CLAHE



- CLAHE is a variant of HE designed to **improve image contrast without excessively amplifying noise** by limiting contrast amplification

$$F_m = \frac{I_p\,(x,y) - \mu_m}{\sigma_m}$$

$I_p(x,y)$: image pixels

μ_m : mean

σ_m : standard deviation

Image Processing: ESRGAN

Enhanced Super-Resolution Generative Adversarial Networks

- Trained with synthetic data to enhance details while removing noisy artifacts to restore blurry images and videos in very high resolution
- Selected in this study to denoise and **restore the quality** of **blurry** retinal fundus images

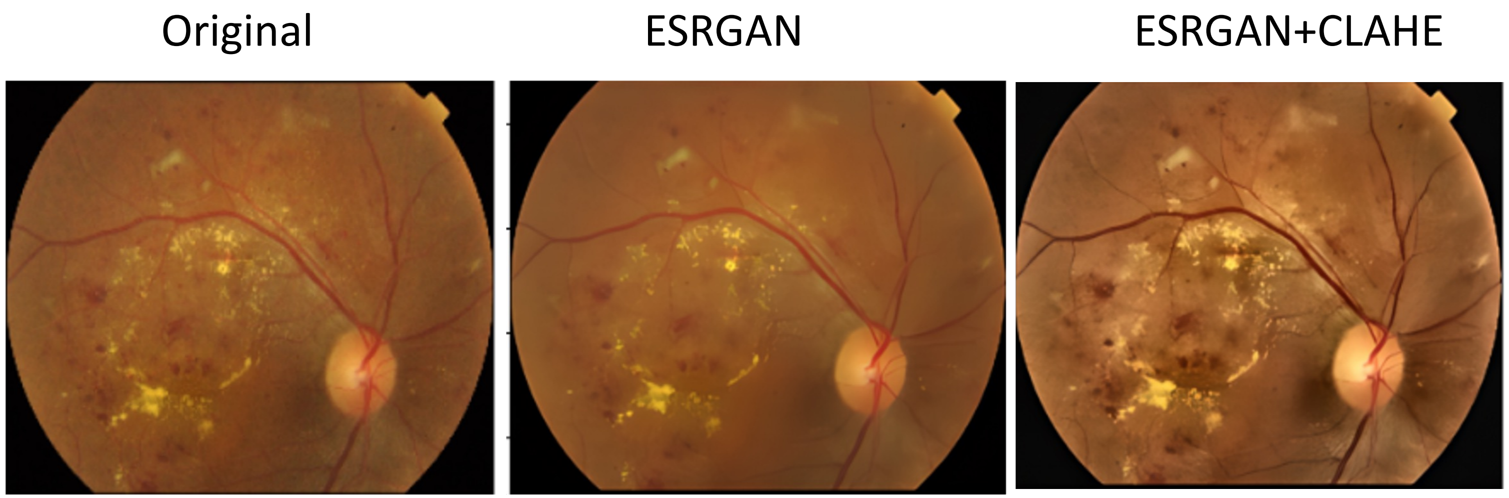
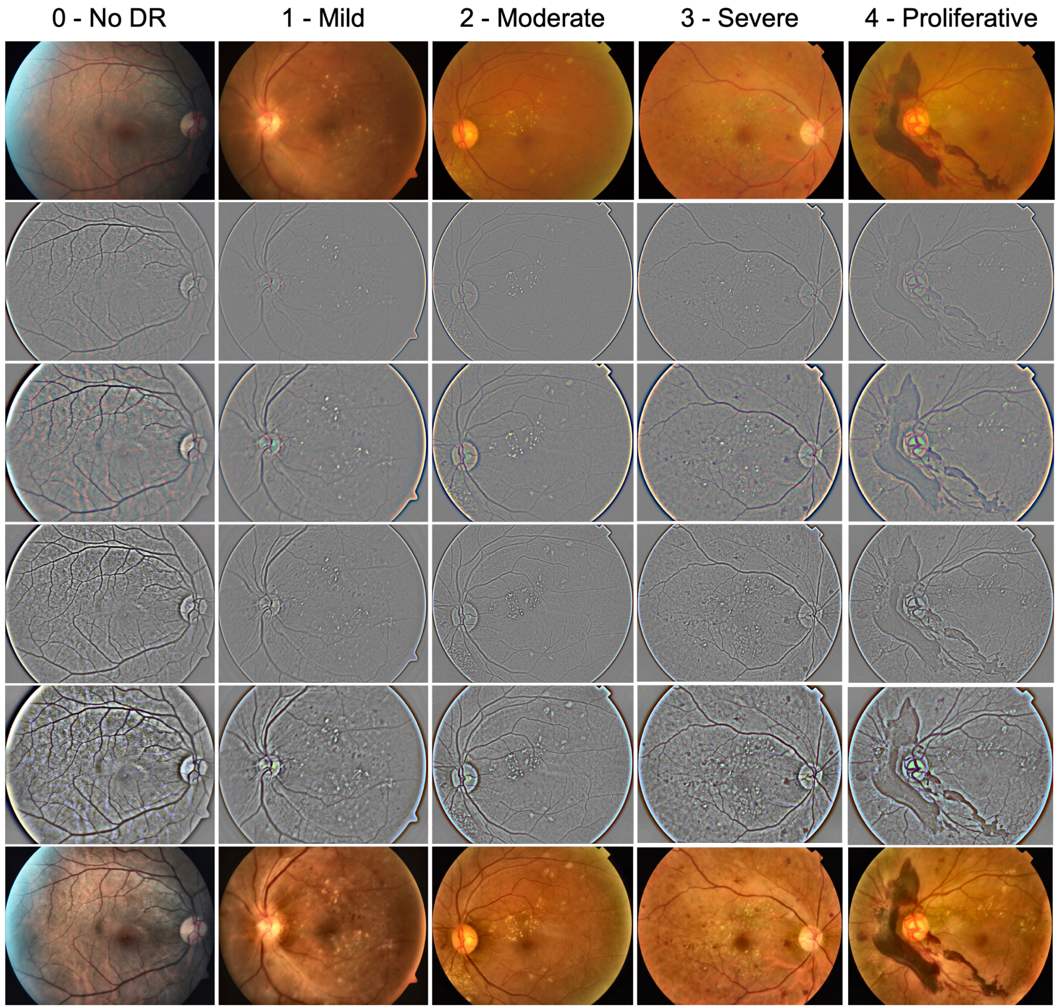


Image Processing: Before vs. After



Original

Graham10 (poor)

Graham20 (Good)

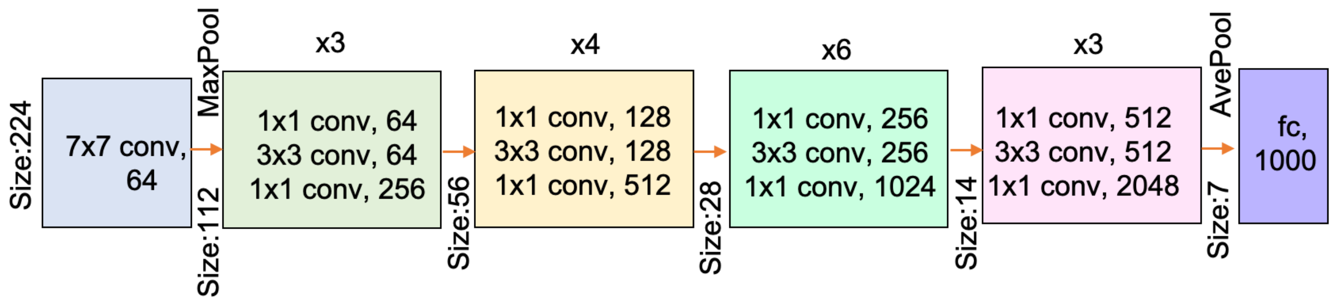
Graham10 + CLAHE

Graham20 + CLAHE (Best)

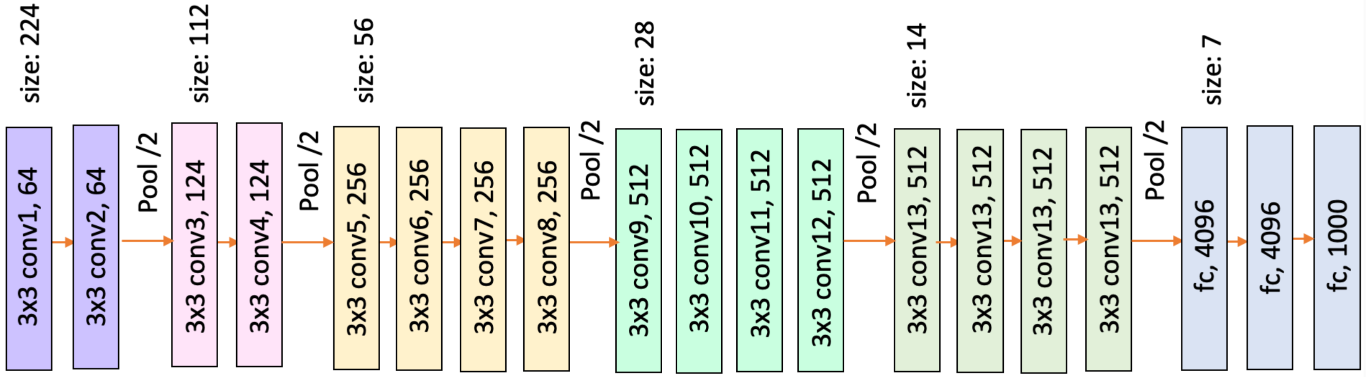
ESRGAN + CLAHE (2nd Best)

Pre-Trained Deep Learning Models

ResNet-50 Architecture

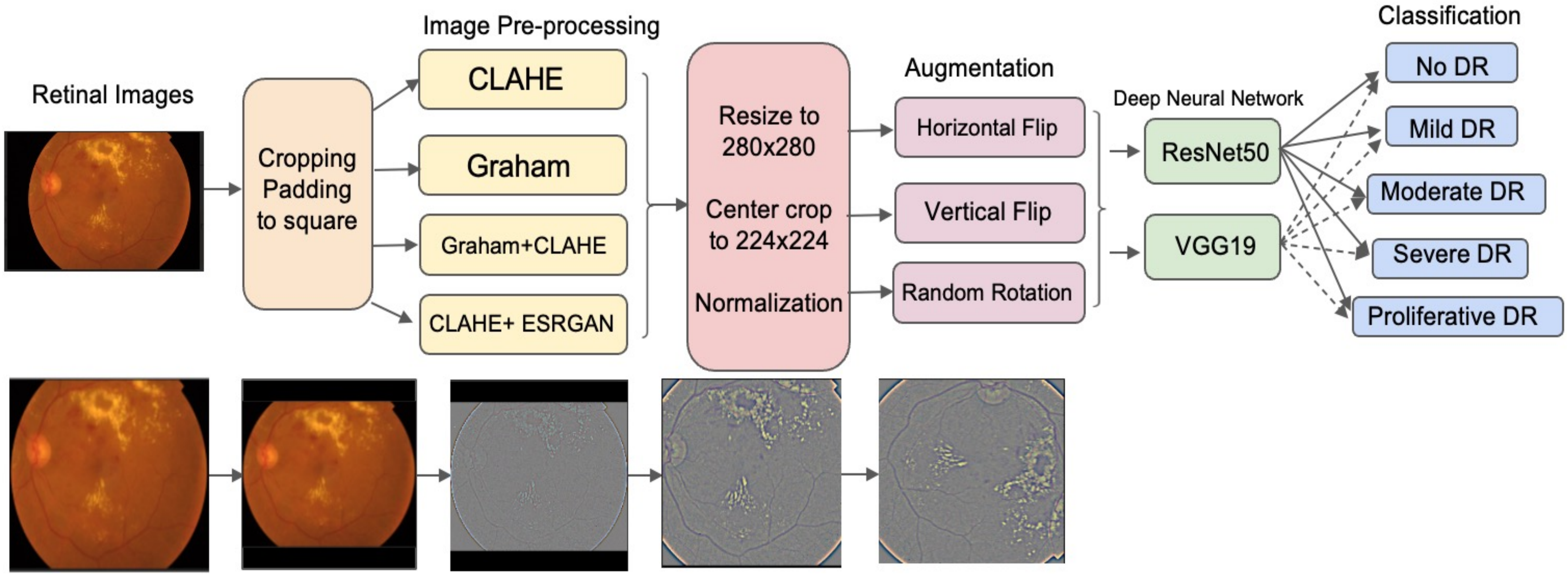


VGG-19 Architecture



- Default weights from Pytorch pre-trained library.
- Initial learning rate = 0.0003 and decreases every 5 epochs by the factor of 0.5
- Adam optimizer performs better than SGD
- Batch size = 32; Max epoch = 50
- Cross entropy loss function
- PyTorch for data loader and normalization
- Scikit-learn for splitting data and performance matrices calculation
- Smote for augmentation
  - Training set before and after over-sampling: 3296 and 8075, respectively
  - Sample count in each class: 1615

Image preprocessing and DR Classification Flow



Results: DR Detection Performance

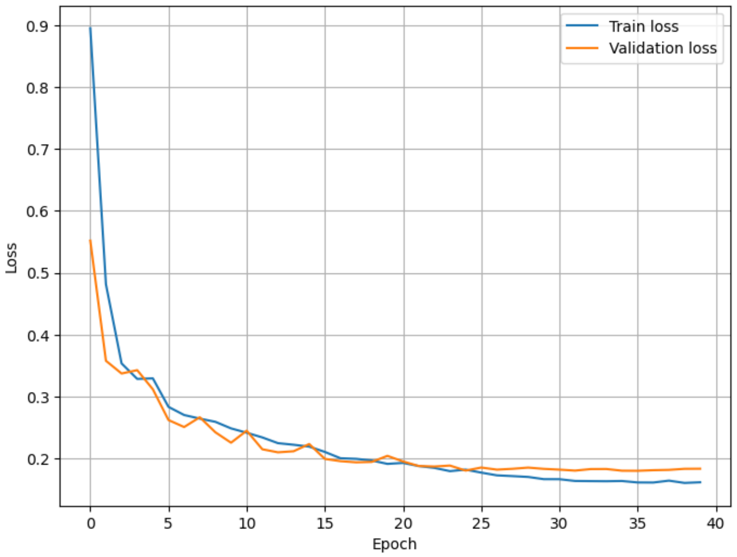
Image Enhancement	Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
CLAHE	ResNet50	97.54	96.89	98.42	97.65
	VGG19	98.09	97.91	98.42	98.16
Graham10	ResNet50	98.09	99.46	96.84	98.13
Graham20	ResNet50	98.36	98.94	97.89	98.41
CLAHE + ESRGAN	ResNet50	97.81	98.92	96.84	97.87
	VGG19	$\boxed{98.91}$	99.47	98.42	98.94
Graham10 + CLAHE	ResNet50	97.36	98.91	95.79	97.33
Graham20 + CLAHE	ResNet50	$\boxed{98.63}$	98.94	98.42	98.68
	VGG19	97.37	97.37	97.37	97.37

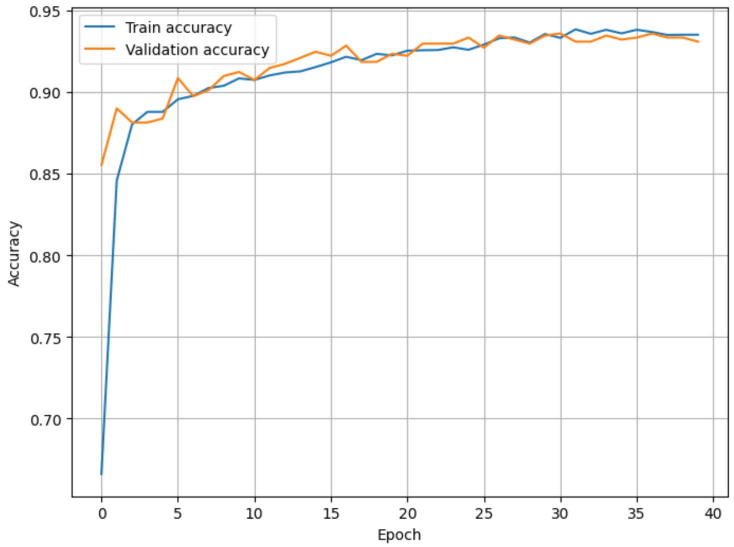
Results: DR Severity Classification

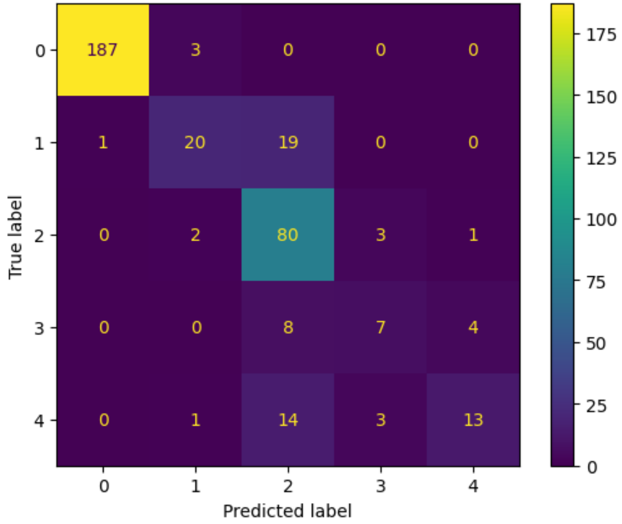
Image Enhancement	Model	Optimizer	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
CLAHE	ResNet50	Adam	83.88	83.21	83.88	83.11
	VGG19	Adam	83.88	83.64	83.88	83.50
Graham10	ResNet50	Adam	83.61	84.16	83.16	83.57
Graham20	ResNet50	Adam	82.79	82.77	82.79	82.67
CLAHE + ESRGAN	ResNet50	Adam	$\boxed{84.15}$	84.40	84.15	81.13
	VGG19	Adam	83.88	84.49	83.88	82.91
Graham10 + CLAHE	ResNet50	Adam	80.87	81.07	81.87	80.69
Graham20 + CLAHE	ResNet50	Adam	$\boxed{86.07}$	85.89	86.07	85.89
	ResNet50	SGD	77.23	76.93	77.23	73.12
	VGG19	Adam	81.69	83.18	81.69	79.64

Best Result of VGG19 Model

w/ CLAHE + ESRGAN preprocessing: 83.88% Accuracy







	precision	recall	f1-score	support
0	0.9947	0.9842	0.9894	190
1	0.7692	0.5000	0.6061	40
2	0.6612	0.9302	0.7729	86
3	0.5385	0.3684	0.4375	19
4	0.7222	0.4194	0.5306	31
accuracy			0.8388	366
macro avg	0.7372	0.6404	0.6673	366
weighted avg	0.8449	0.8388	0.8291	366

Accuracy for 0: 98.91%
Accuracy for 1: 92.9%
Accuracy for 2: 87.16%
Accuracy for 3: 95.08%
Accuracy for 4: 93.72%

Confusion Matrices of ResNet50

Graham10 + CLAHE
Accuracy: 80.87%

Graham20
Accuracy: 82.79%

Graham20 + CLAHE
Accuracy: 86.07%

CLAHE + ESRGAN
Accuracy: 84.15%

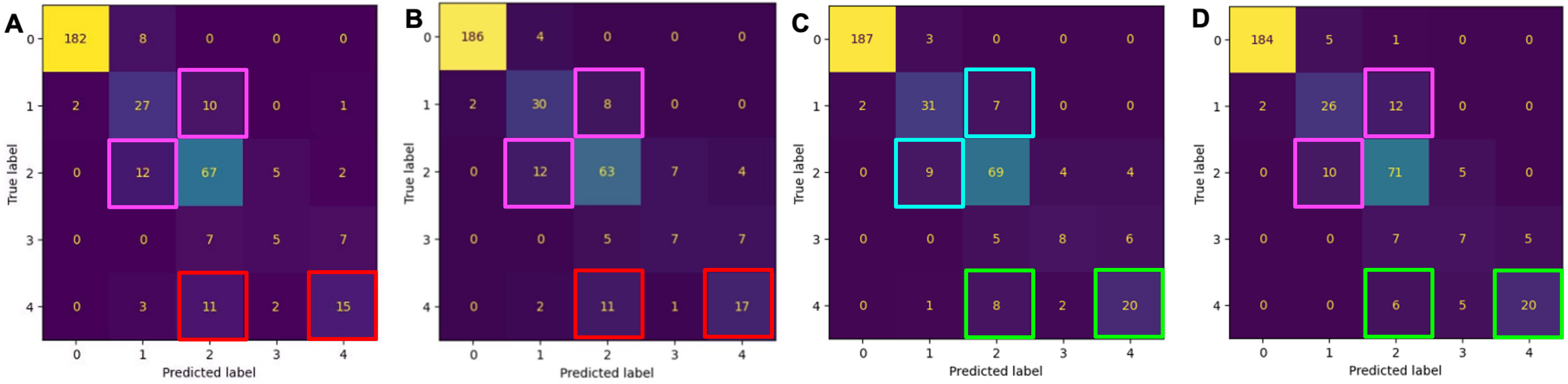


Image Processing Customization by DR Stage

- Users can select the optimal signal processing method to highlight the DR lesions.
- CLAHE + Graham20 is the overall best option for all DR stages.

DR Stage	CLAHE (%)	Graham10 (%)	Graham20 (%)	CLAHE + Graham10 (%)	CLAHE + Graham20 (%)	CLAHE + ESRGAN (%)
0: No DR	98.54	98.09	98.36	98.36	**98.63**	97.81
1: Mild DR	93.08	91.80	92.35	92.44	**93.99**	92.08
2: Moderate DR	86.89	87.70	87.16	86.61	**89.89**	88.80
3: Severe DR	94.17	**95.63**	94.54	93.99	95.36	93.99
4: Proliferative DR	93.81	93.99	93.17	93.99	94.26	**95.63**

Comparison with Previous Research

Ref	Year	Method	Preprocessing	Accuracy (%)	Precision (%)	Recall (%)
Binary Classification						
[7]	2021	Hybrid of VGG16 + capsule networks	None	95.5		
[8]	2022	VGG16	Grad-CAM	73.04	-	-
		DenseNet121	Grad-CAM	72.95	-	-
[11]	2022	Supervised contrastive learning	CLAHE	98.36	98.36	98.37
[14]	2024	Hybrid ResNet50 + EfficientNetB0	CLAHE	97.95	97.84	98.11
This Work	**2024**	**ResNet50**	**Graham20+CLAHE**	**98.63**	**98.94**	**98.42**
		VGG19	**CLAHE+ESRGAN**	**98.91**	**99.47**	**98.42**
Multi-class classification						
[7]	2021	Hybrid of VGG16 + capsule networks	None	75.81		
[18]	2020	NASNet + ν-SVM	None	77.90	76.00	77.00
[9]	2019	ResNet50	Graham10	74.64	-	56.52
		Modified Xception	Graham10	83.09	-	88.24
[8]	2022	VGG16	Grad-CAM	48.43	-	-
[11]	2022	Supervised contrastive learning	CLAHE	84.36	70.51	73.84
[14]	2024	Hybrid ResNet50 + EfficientNetB0	CLAHE	83.17	82.66	83.17
This Work	**2024**	**ResNet50**	**Graham20+CLAHE**	**86.07**	**85.89**	**86.07**
		ResNet50	**CLAHE+ESRGAN**	**84.15**	**84.40**	**84.15**

Conclusion & Future work

- Achievements: The optimal combinations of three signal processing methods are explored and proposed to achieve the high accuracies:
  - DR detection: VGG19 model with CLAHE + ESRGAN achieved an accuracy of 98.91%.
  - DR severity classification: ResNet50 with Graham20 + CLAHE achieved the highest accuracy of 86.07%.
  - Not only APTOS dataset, but we also extended the established framework to histopathology images for breast cancer classification successfully.
- Contributions to Society:
  - The proposed image enhancement methods can be integrated into Apps, allowing users at home or ophthalmologists in clinics to select different image enhancement options to clearly visualize DR lesions for self or professional DR screening.
  - The established ResNet50 deep learning framework can be integrated with electronic health systems, helping diagnose DR stages more efficiently and enable timely treatment for patients.
- Limitations and Future Work:
  - 566 out of 3662 APTOS images are noisy or duplicated: Data clean-up is needed.
  - Imbalanced and insufficient APTOS dataset to build a robust classifier: Increasing the training set size by merging with another dataset or finding a larger and balanced dataset
  - Only ResNet50 and VGG19 used: Exploring other CNN models such as EfficeintNetV2

Thanks and QA