INFORMATION EXTRACTION FROM NATIONAL IDENTITY CARD (NIC) AND DRIVING LICENSE

Group Members

S. Sabesan – 160538L

R. Sivaram – 160596K

P. Thineshan – 160633A

Supervisor

Dr. Uthayashankar Thayasivam

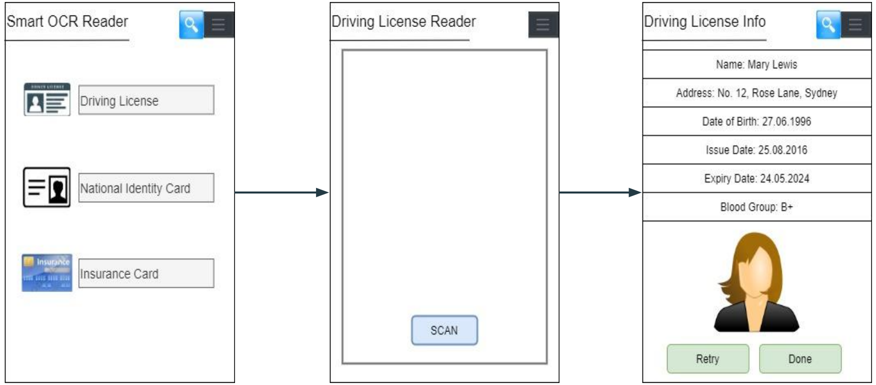
Pannal Members

Dr. Surangika Ranathunga

Dr. Sapumal Ahangama

Introduction

- In Sri Lanka, we have to use our NIC or Driving License in many situations (ex: obtaining SIM card). These situations establishments keep our NIC or Driving License hard copy documents.
- We are going to develop an end to end information extraction system from structural documents such as NIC and Driving License.



Information in Driving License and NIC

- Information in Driving License
  - License number
  - Full name
  - Address
  - NIC number
  - Date of Birth
  - Date of Issue
  - Date of Expiry
  - Blood group
  - Restriction in code form (ex: Specs)

- Information in NIC
  - NIC number
  - Full name
  - Date of Birth
  - Gender
  - Birth place
  - Address
  - Job

Problem Statement

Develop an improved end to end information extraction system which can extract the information from images containing written text (typed, handwritten or printed) of the National Identity Card (NIC) and Driving License of Sri Lanka and deliver the information as an interoperable format.

Motivation

- Sri Lankan Companies and other sectors face problems when they verify the customer with NIC or Driving licence because they use the hard copy of these documents.
- NIC have Sri Lankan native Languages.
- We have good business motivation because we will improve our system for other documents like passports.

Research Motivation

- To develop a model which can improve OCR accuracy.
- To develop a system which can extract whole information from NIC or Driving License.

Objectives

- OCR analysis and selection.
  - To compare different OCR accuracies and select suitable OCR service.
- Develop a NIC type analyzer.
  - To train a model with YOLOv3 using NIC front images.
- Develop best cropping, glare removal and chunking preprocessing approaches to improve OCR accuracy.
- Entity annotation.
  - To develop a regex based approach to improve OCR accuracy.
- Deliver the extracted information in JSON format.

Literature Review - Glare Removal

Document	Methods	Key points
Efficient and robust specular highlight removal[23]	Bilateral filter is used	Needed bilateral filter parameters.
A Global optimization method to remove highlights[21]	Dichromatic based model	Diffuse chromaticity is estimated by correcting hue & saturation
Inpainting with context encoder[2]	Encoder decoder is used as generator to generate the missing parts of the image	Basic architecture design for inpainting with GAN

Literature Review - Glare Removal

Document	Methods	Key points
Image Inpainting Based on Patch-GANs[1]	Local and Global discriminators are used edge preservation loss function	Works well on human faces inpainting
Single Image Reflection Removal Using Deep Encoder-Decoder Network[3]	encoder decoder network supervised learning	Needed pair of glare and non glare images to implement. Takes 2 seconds to process a 512 x 512 image.
Single Image Reflection Removal based on GAN with Gradient Constraint[26]	Generator is UNet++ architecture model	Four losses are used including gradient constraint UNet++ Architecture

Literature Review - Entity Extraction

Document	Methods
Medical records information extraction system [19]	Dictionary based approach, regex based approach, slot-based approach
Information extraction from ID card via computer vision techniques [20]	Templating traversal technique. (using templates of ID card)

Methodology - Dataset

- Pickme gave us 1472 images that contain driving license and NIC front, back parts without actual text
- 21 Images are not driving license or NIC

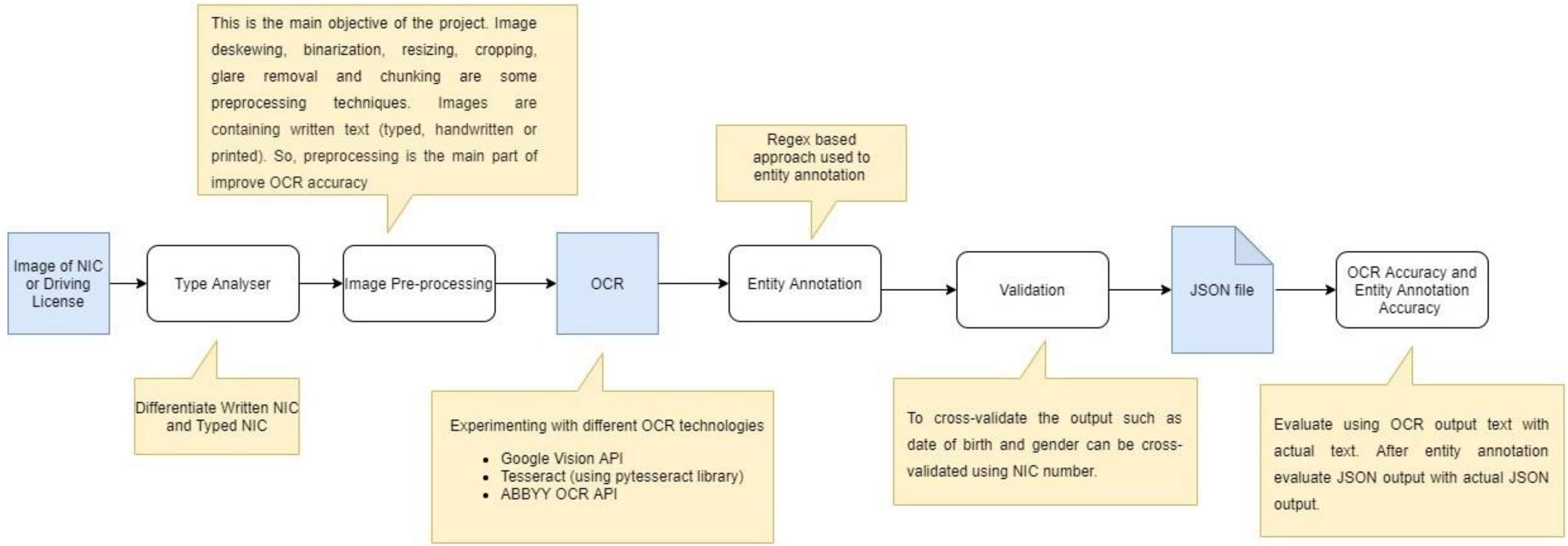
	Driving License	NIC	Total
Front	154	444	598
Back	215	638	853
Total	369	1082	1451

Methodology - OCR Analysis and Selection

- After analyzing OCR result, we came to following conclusions.
  - Tesseract OCR is very low accuracy.
  - Google OCR is better than ABBYY OCR because ABBYY OCR gave some empty results at the same time Google OCR gave us some results for some images.
  - So, we dropped Tesseract and ABBYY OCR.
- Finally we selected Google OCR service.

OCR Service	Accuracy
Cloud Vision API	83.83%
ABBYY OCR API	47.46%

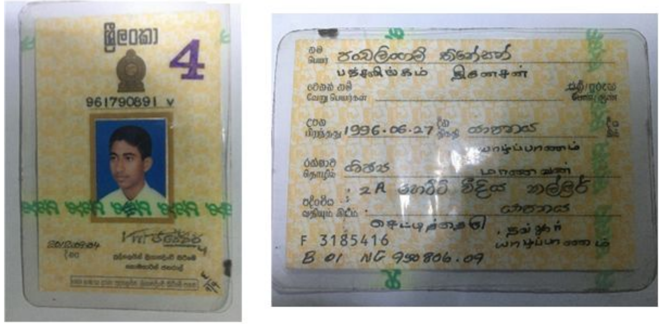
Methodology - Overall Architecture



Methodology - NIC Type Analyzer using YOLOv3

- You Only Look Once is a faster object detection algorithm that uses convolutional neural networks for object detection.
- We trained a model with YOLOv3 to categorize our NICs.
- The model will categorize written NIC as 0 and typed NIC as 1.

Typed NIC

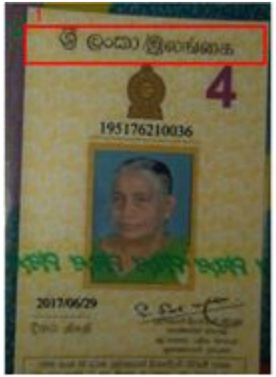


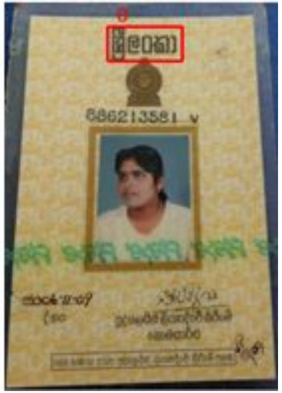
Written NIC

Methodology - NIC Type Analyzer using YOLOv3

- We got 98% accuracy for our trained model

Accuracy = 98%	Written NIC	Typed NIC	Total
Train set	275	110	385
Test set	44	13	57

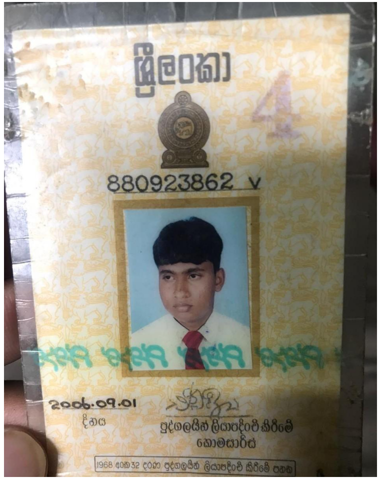


Glare Images

Glares are classified Disability glare, discomfort glare, photostress glare, dazzling glare and dysphotopsia glare[27]

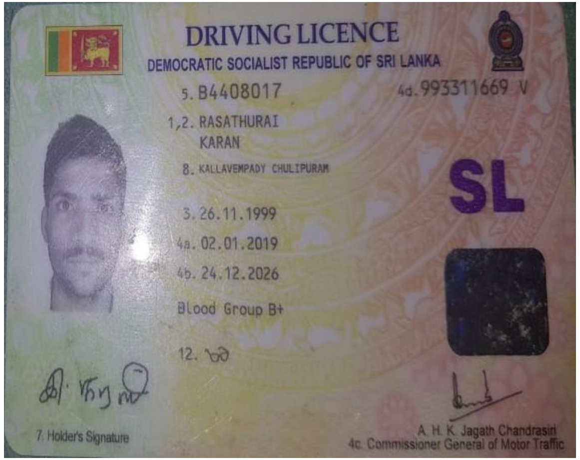
Disability Glare



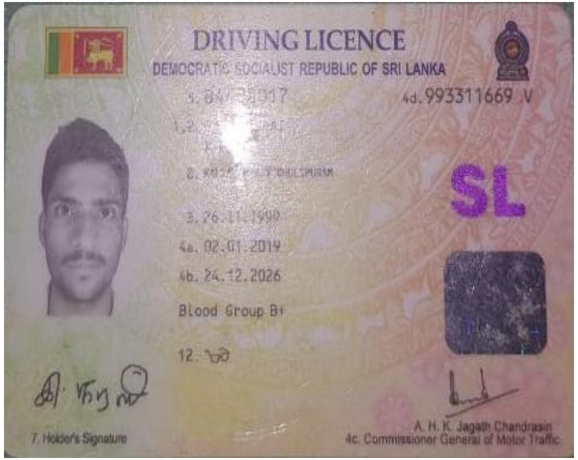
Discomfort glare

Disability Glare Images

The Disability glare can be in the textarea or background area

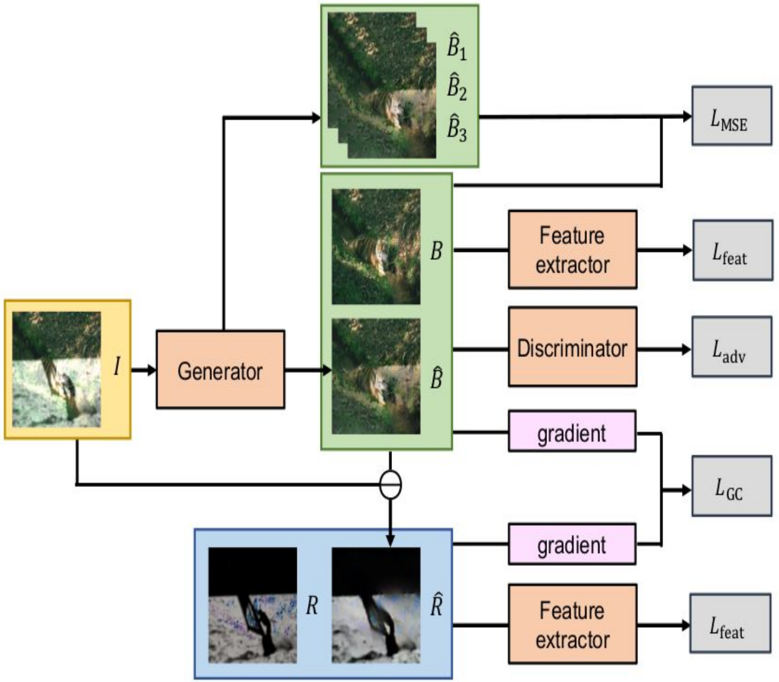


Background disability glare

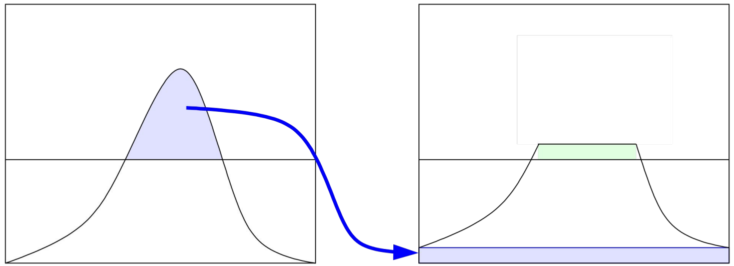


Textarea disability glare

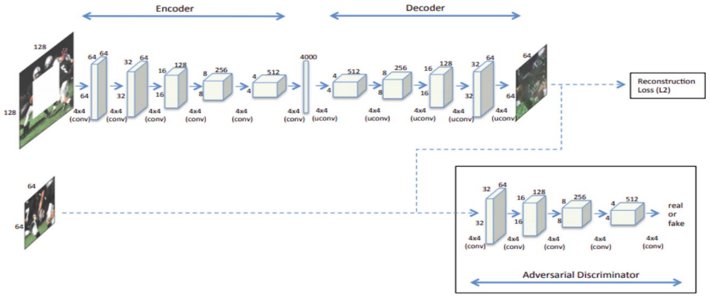
Glare Removal Approach



GAN based Gradient Constraint Network(GCNet)[26]

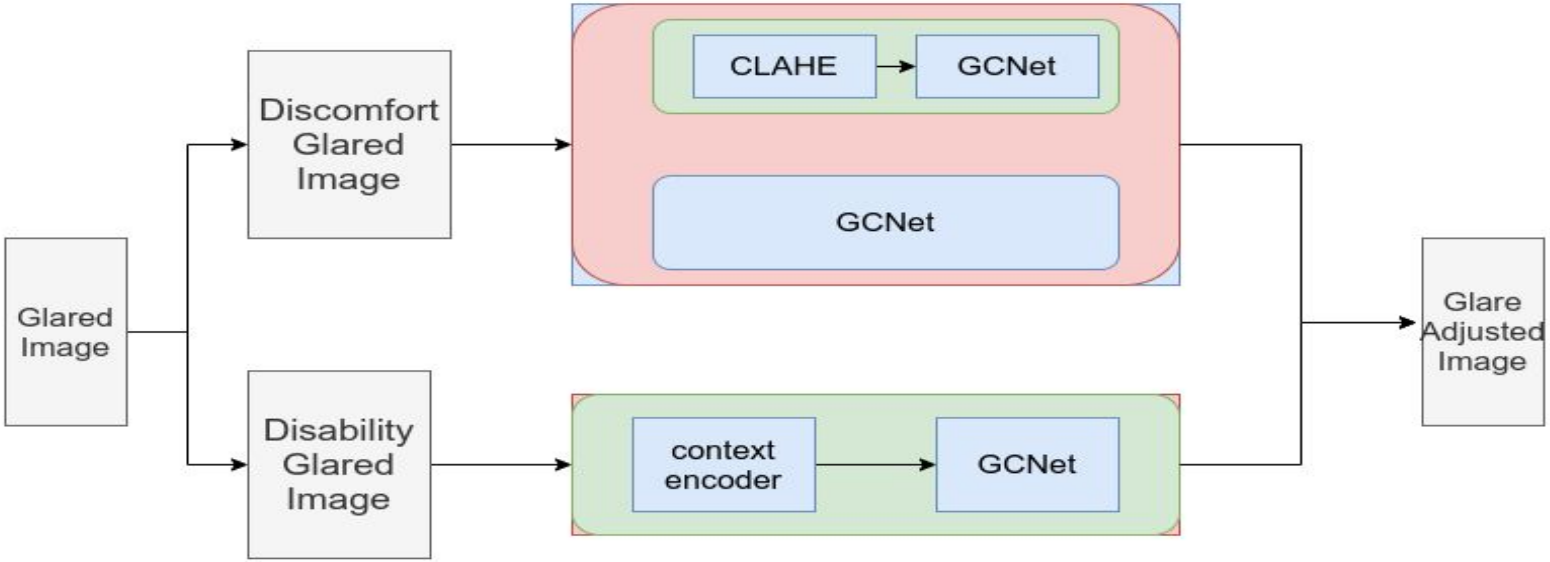


Contrast limited AHE[26]



Context Encoder Decoder[2]

Glare Removal Approach



OCR Accuracy

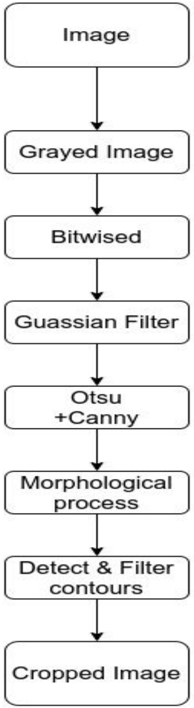
Image	CLAHE	GCNet	Inpainted	CLAHE + GCNet	Inpainted + GCNet
Disability Glared	0.7900	0.8605	0.816	0.8064	0.8668
Discomfort Glare	0.8547	0.8963	0.8741	0.9071	0.8522

Accuracy is improved from **0.8815** to **0.8833**

Good Images

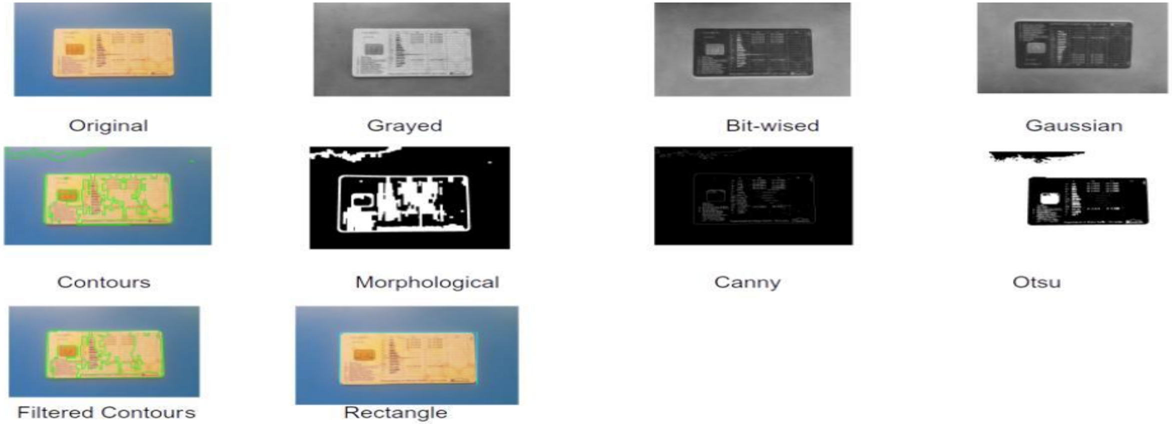
- Textarea size > 300000 px
- DPI > 300
- Clear plan background
- No skewing
- No glare/Reflection

Methodology - Image Cropping pipeline



Constraints for our Cropping pipeline

- Background has more edges.
- Glare or more shadows are occured in the image.
- Image with external text



Methodology - Image Cropping pipeline

Images	Total Images	Correctly cropped	Accuracy
Good Images	110	101	0.91
Glare Images	57	41	0.71
Background problems	65	25	0.38

Background

Background

Glare

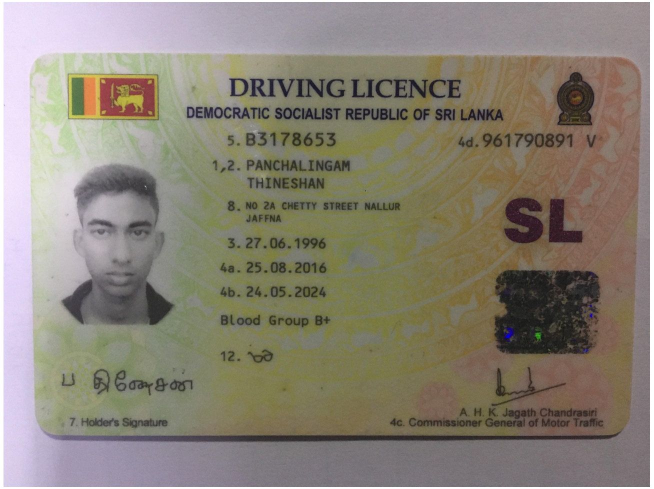
Background

Methodology - Image Face Detection

- Single shot detector framework with a ResNet base network is used to detect the face in the image[28]
- This face detector Confidence level is set to 0.8 and CLAHE is applied to improve the face detection

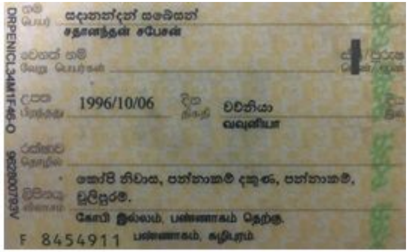
Images	Total	Detected	Accuracy
Image without CLAHE	76	64	0.84
CLAHE applied Image	76	74	0.97

Methodology - Image Chunking

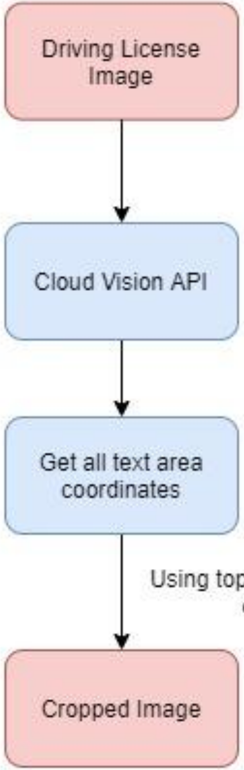


- Purpose: Getting the chunks of images row wise (separating the needed information as chunks)
- First, we need to crop the images and get all bounding boxes.

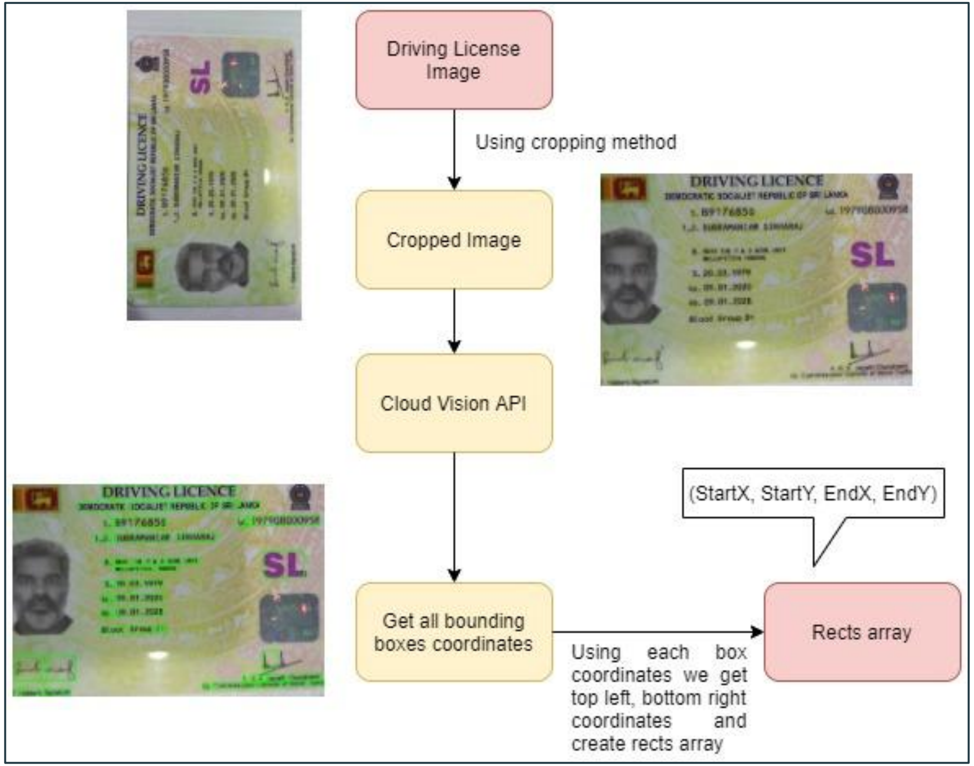
Methodology - Image Cropping using Cloud Vision API Coordinates

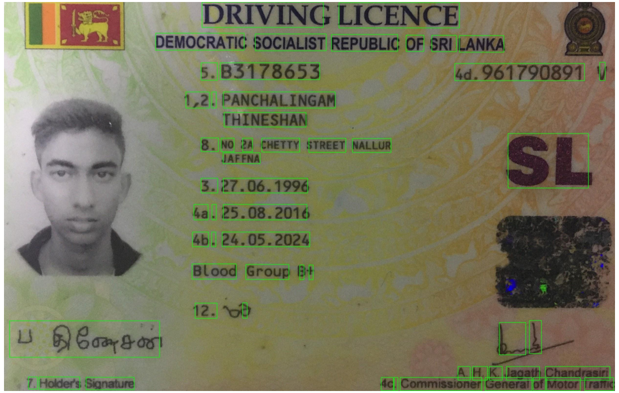


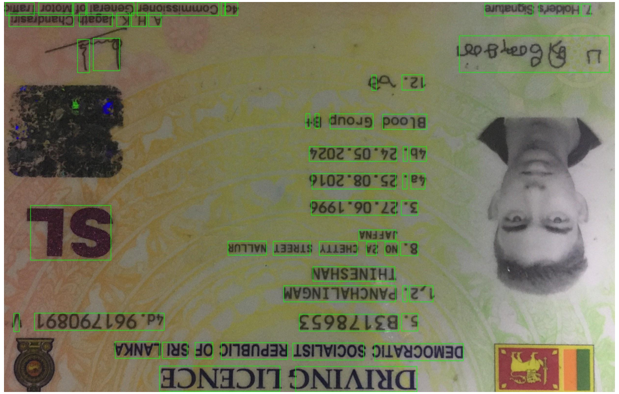

```
[x: 53
y: 366
, x: 473
y: 366
, x: 473
y: 649
, x: 53
y: 649
]
```



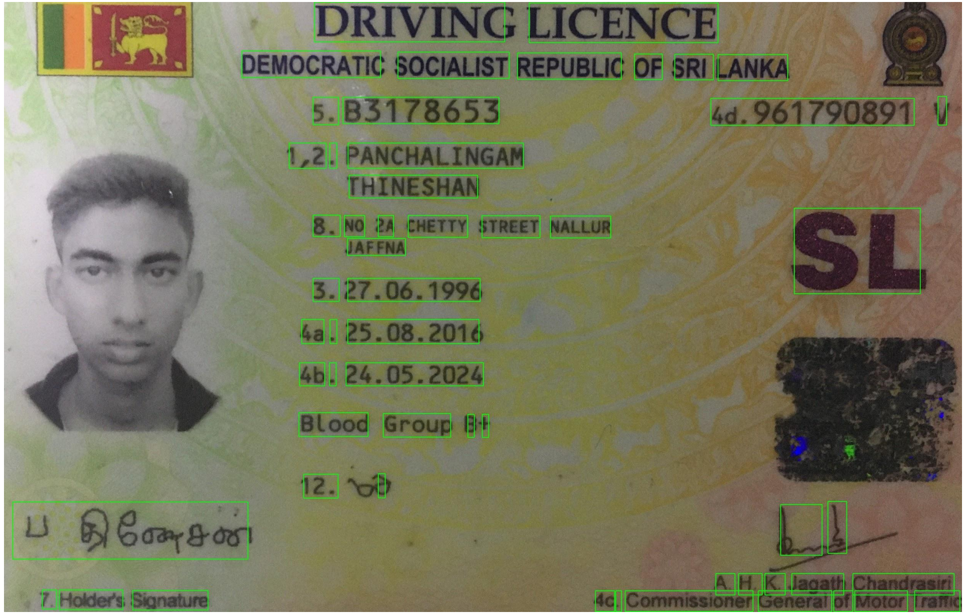
Methodology - Image Bounding Boxes using Cloud Vision API







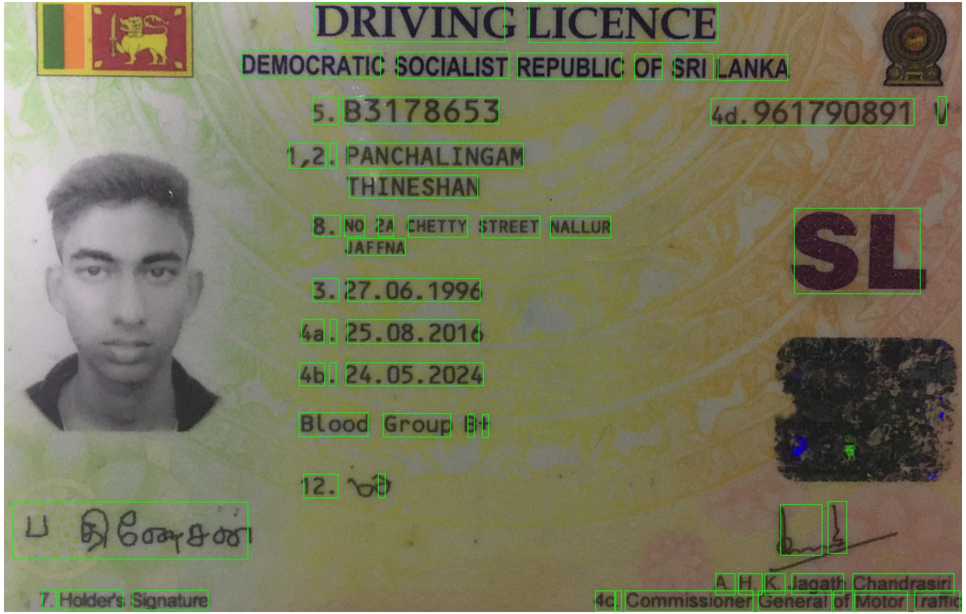
Methodology - Image Chunking Algorithm

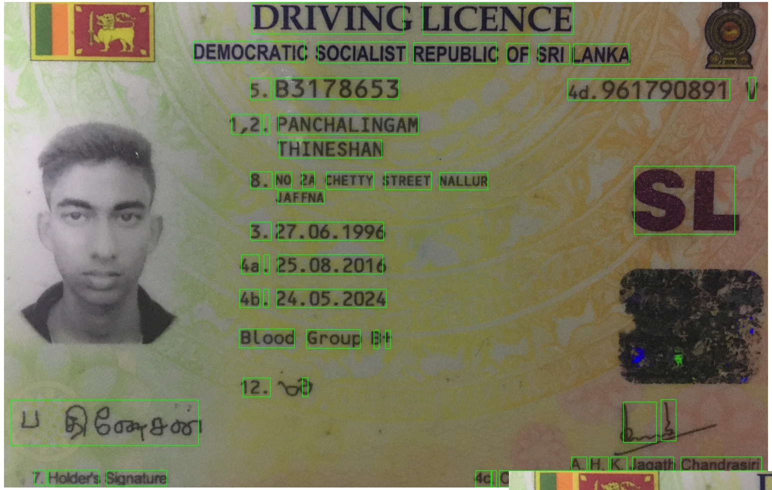


Factors Considered when chunking the images

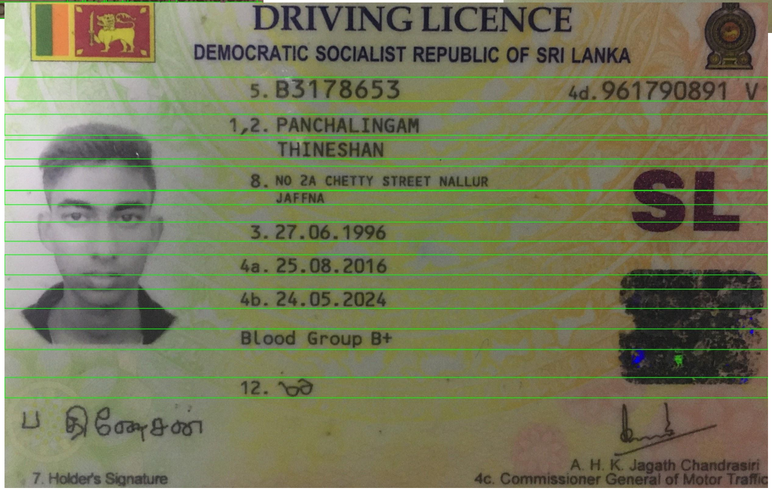
- Chunking the image without resizing the image.
- Consider punctuations (, and .).
- Chunks never have to cut the numbers or characters in the data.
- We need to get the chunks of images with needed information..

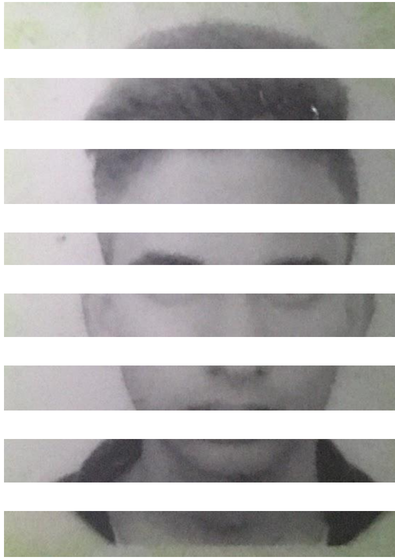
Methodology - Image Chunking Algorithm




DRIVING LICENCE
DEMOCRATIC SOCIALIST REPUBLIC OF SRI LANKA
5. B3178653
4d. 961790891 V
1,2. PANCHALINGAM
THINESHAN
8. NO 2A CHETTY STREET NALLUR
JAFFNA
3. 27.06.1996
4a. 25.08.2016
4b. 24.05.2024
Blood Group B+
SL
12.
A. H. K. Jagath Chandrasiri
7. Holder's Signature


DRIVING LICENCE
DEMOCRATIC SOCIALIST REPUBLIC OF SRI LANKA
5. B3178653
4d. 961790891 V
1,2. PANCHALINGAM
THINESHAN
8. NO 2A CHETTY STREET NALLUR
JAFFNA
3. 27.06.1996
4a. 25.08.2016
4b. 24.05.2024
Blood Group B+
SL
12.
A. H. K. Jagath Chandrasiri
4c. Commissioner General of Motor Traffic


DRIVING LICENCE
DEMOCRATIC SOCIALIST REPUBLIC OF SRI LANKA
5. B3178653
4d. 961790891 V
1,2. PANCHALINGAM
THINESHAN
8. NO 2A CHETTY STREET NALLUR
JAFFNA
3. 27.06.1996
4a. 25.08.2016
4b. 24.05.2024
Blood Group B+
SL
12.
A. H. K. Jagath Chandrasiri
7. Holder's Signature
4c. Commissioner General of Motor Traffic



5. B3178653 4d. 961790891 V

1,2. PANCHALINGAM

THINESHAN

8. NO 2A CHETTY STREET NALLUR

JAFFNA

SL

3. 27.06.1996

4a. 25.08.2016

4b. 24.05.2024

[redacted]

Blood Group B+

12.


Blood Group O+
4b. 28.11.2022
4a. 31.01.2015
3. 18.07.1996
KOPAY SOUTH KOPAY JAFFNA
8. VEERAPATHIRAR TEMPLE VIEW
1,2. SATHGURUNATHA SHARMA HARIHARAN
5. B2768170
4d. 962002862 V

Methodology - Image Chunking

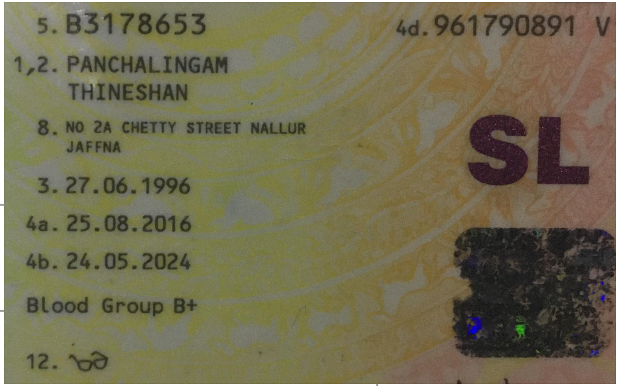
Chunking Accuracy for an image=
# of chunks after using chunking algorithm / # of chunks we need in that image

- We got 97.21% chunking accuracy with 47 images.

Chunking method	OCR Accuracy	Entity Annotation Accuracy
Original Image + Cropping + Deskewing + Not resizing + Chunking	0.7462	0.7933
Original Image + Cropping + Not resizing + Chunking	0.764	0.7896

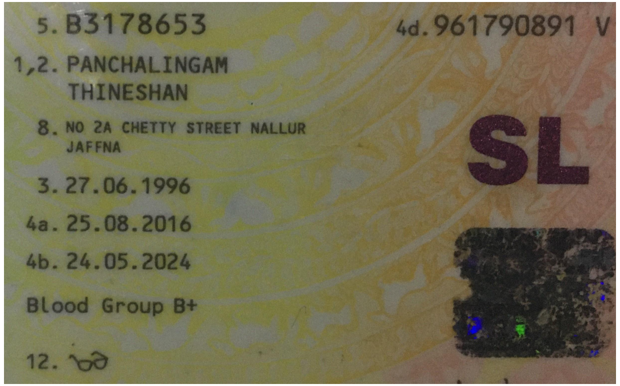
Methodology - Entity Extraction or Driving License

- We used a regex based approach to entity annotation



<table>
<tr><th>Entity</th><th>Regex Pattern</th></tr>
<tr><td>Licence Number</td><td>r'([B]\d{5,7}[^0-9][^\SD]{0,1})'
Ex: B2761530D, B4059906, B670543, B57231</td></tr>
<tr><td>Administrative Number
(NIC number)</td><td>r'([1-9]{1}[0-9]{8}[ ][v|V]|[1-2]{1}[0-9]{11})'
Ex: 961790891 V, 199617900681</td></tr>
<tr><td>Name</td><td>r'(?:1,2[\.|\,])(.*?)(?:8)'</td></tr>
</table>

<table>
<tr><td>Permanent Place of Residence</td><td>r'(?:8\.)(.*?)(?:3\.)'</td></tr>
<tr><td>Blood Group</td><td>r'(A\+|B\+|AB\+|O\+|A\-|B\-|AB\-|O\-|0\+|0\-)'</td></tr>
<tr><td>Restriction in code form (ex: Specs)</td><td>r'(?:12\.)'</td></tr>
</table>



Methodology - Entity Annotation for Dates

Dates: r'(\d{2}[.]\d{2}[.]\d{4})'

DATES<= get all dates from a license using regex
current_date <= get current date
check all dates in DATES array is valid (1<=days<=31 and 1<=month<=12)
If length of DATES == 3
    Date of Birth = DATES array 1st element
    Date of Issue = DATES array 2nd element
    Date of Expiry = DATES array 3rd element
Else
    For date in DATES:
        (1) If time difference between date and current_date >=18 years then date=Date of Birth
        (2) If date>current_date then date = Date of Expiry
        (3) If date<current_date and 0<=time difference in years<=8 then date = Date of Issue

Methodology - Date of Birth, Gender Validation Using NIC Number

- We validate our ocr predicted Date of Birth using actual text NIC number. Date of Birth is correctly predicted in each license image

Methodology - Experimental Evaluation for Driving License

Evaluation Matrix

OCR accuracy = number of correctly predicted words / number of words in actual text

Entity Annotation Accuracy = number of entities correctly predicted in JSON output of OCR text / number of entities in actual JSON text file

Evaluating Google OCR output with original license image

Accuracy	Percentage (%)
OCR_Accuracy	**83.83%**
Entity Annotation Accuracy	**85.79%**

Entities	Accuracy (after annotation)
Licence Number	0.97
Administrative Number (NIC number)	0.92
Name (Surname with other names)	0.79
Permanent Place of Residence	0.36
Date of Birth	0.95
Date of Issue of the License	0.92
Date of Expiry of the License	0.93
Blood Group	0.92
Restriction in code form (Specs)	0.97

Evaluating with different combinations of pre-processing

Pre-processing methods	OCR Accuracy	Entity Annotation Accuracy
None (Original Image)	0.8383	0.8579
Original Image + Deskewing + Cropping + Resizing	0.6038	0.6059
Original Image + Deskewing + Cropping + Not resizing	0.805	0.8004
Original Image + Cropping + Deskewing + Not resizing + Chunking	0.7462	0.7933
Original Image + Cropping + Not resizing + Chunking	0.764	0.7896

Completed Tasks

Month	Task Breakdown	Status
January	• Literature review on information extraction from ID card and Driving License systems. • Literature review on preprocessing techniques.	Completed
February	• Working on a project proposal.	Completed
March	• Finished the project proposal and worked with proposal methodology.	Completed
April	• Data annotation. • Converting image to text using existing OCR services. • Trying preprocessing approaches.	Completed
May	• Train a model with YOLOv3 to categorize NIC. • Evaluating OCR accuracy with different approaches.	Completed
June	• Try different best practices to implement cropping pipeline without machine learning approach	Completed
July	• Entity extraction using a regex based approach	Completed

Future Tasks

- Implementing a machine learning model for cropping the images.
- Implementing a GAN based model to remove glare/reflection from the image.
- Implementing a model which can extract name and address from driving license/NIC OCR output.
- Correcting the spelling errors from Google OCR output
- Improve chunking algorithm to use NIC chunking.
- Implementing a postprocessing and Entity Recognition to NIC images

References

[1]L. Yuan, C. Ruan, H. Hu and D. Chen, "Image Inpainting Based on Patch-GANs," in IEEE Access, vol. 7, pp. 46411-46421, 2019, doi: 10.1109/ACCESS.2019.2909553.

[2]D. Pathak, P. Krähenbühl, J. Donahue, T. Darrell and A. A. Efros, "Context Encoders: Feature Learning by Inpainting," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 2536-2544, doi: 10.1109/CVPR.2016.278.

[3]Chi, Zhixiang & Wu, Xiaolin & Shu, Xiao & Gu, Jinjin. (2018). Single Image Reflection Removal Using Deep Encoder-Decoder Network.

[4] C. Doersch, S. Singh, A. Gupta, J. Sivic, and A. Efros. What makes paris look like paris? ACM Transactions on Graphics, 2012. 6

[5]O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei. Imagenet large scale visual recognition challenge. IJCV, 2015. 2, 6

[6] P. Agrawal, J. Carreira, and J. Malik. Learning to see by moving. ICCV, 2015. 2, 7, 8

[7] X. Wang and A. Gupta. Unsupervised learning of visual representations using videos. ICCV, 2015. 2, 7, 8

[8] C. Doersch, A. Gupta, and A. A. Efros. Unsupervised visual representation learning by context prediction. ICCV, 2015. 2, 3, 7, 8

[9] Z. Liu, P. Luo, X. Wang, and X. Tang, ‘‘Deep learning face attributes in the wild,’’ in Proc. IEEE Int. Conf. Comput. Vis. (CVPR), Dec. 2015, pp. 3730–3738.

[10] Y. Netzer, T. Wang, A. Coates, A. Bissacco, B. Wu, and A. Y. Ng, ‘‘Reading digits in natural images with unsupervised feature learning,’’ in Proc. NIPS Workshop Deep Learn. Unsupervised Feature Learn., vol. 2011, no. 2, 2011, p. 5.

[11] J. Krause, M. Stark, J. Deng, and L. Fei-Fei, ‘‘3D object representations for fine-grained categorization,’’ in Proc. IEEE Int. Conf. Comput. Vis. Workshops, Jun. 2013, pp. 554–561.

[12] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba, ‘‘Places: A 10 million image database for scene recognition,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 6, pp. 1452–1464, Jun. 2018

[13] P. Saunier, "GaN for Next Generation Electronics," 2014 IEEE Compound Semiconductor Integrated Circuit Symposium (CSICS), La Jolla, CA, 2014, pp. 1-4, doi: 10.1109/CSICS.2014.6978558.

[14] K. Gai, Z. Shi, and C. Zhang. Blind separation of superimposed moving images using image statistics. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 34(1):19–32, 2012

[15]Y. Li and M. S. Brown. Exploiting reflection change for automatic reflection removal. In IEEE International Conference on Computer Vision (ICCV), pages 2432–2439, 2013. 1, 2

[16]J. Kim, S. Song and S. Yu, "Denoising auto-encoder based image enhancement for high resolution sonar image," 2017 IEEE Underwater Technology (UT), Busan, 2017, pp. 1-5, doi: 10.1109/UT.2017.7890316.

[17] Medium. 2020. How To Train Yolov3 On Google Colab To Detect Custom Objects (E.G: Gun Detection).

[18] GitHub. 2020. Developer0hye/Yolo_Label.

[19] Li X, Hu G, Teng X, Xie G. Building structured personal health records from photographs of printed medical records. AMIA ... Annual Symposium proceedings. AMIA Symposium. 2015;2015:833-842

[20] Nguyen Thanh Cong, Nguyen Dinh Tuan, Tran Quoc Long, "Information extraction from ID card via computer vision techniques", 2018.

[21] W. Xia, E. C. S. Chen, S. E. Pautler and T. M. Peters, "A Global Optimization Method for Specular Highlight Removal From a Single Image," in IEEE Access, vol. 7, pp. 125976-125990, 2019.

22] Meslouhi, O.E. & Kardouchi, Mustapha & Allali, Hakim & Gadi, Taoufiq & Ait Benkaddour, Yassir. (2011). Automatic detection and inpainting of specular reflections for colposcopic images. Open Computer Science. 1. 341-354. 10.2478/s13537-011-0020-2.

[23] Yang, Qingxiong & Tang, Jinhui & Ahuja, Narendra. (2015). Efficient and Robust Specular Highlight Removal. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 37. 1304-1311. 10.1109/TPAMI.2014.2360402.

[24] A. C. S. Souza, M. C. F. Macedo, V. P. Nascimento and B. S. Oliveira, "Real-Time High-Quality Specular Highlight Removal Using Efficient Pixel Clustering," *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, Parana, 2018, pp. 56-63. doi: 10.1109/SIBGRAPI.2018.00014

[25] Abiko, Ryo & Ikehara, Masaaki. (2019). Single Image Reflection Removal based on GAN with Gradient Constraint. IEEE Access. PP. 1-1. 10.1109/ACCESS.2019.2947266.

[26]Setiawan, Agung & Mengko, TIr & Santoso, Oerip & Suksmono, Andriyan. (2013). Color retinal image enhancement using CLAHE. 1-3. 10.1109/ICTSS.2013.6588092.

[27]Putnam, Christopher. (2017). OVERVIEW OF GLARE TYPES AND THEIR RELATIONSHIP WITH MACULAR PIGMENT OPTICAL DENSITY.. International Journal of Advanced Research. 5. 1131-1140. 10.21474/IJAR01/5165.

[28]"opencv/opencv",*GitHub*,2020.[Online].Available:https://github.com/opencv/opencv/tree/master/samples/dnn/face_detector. [Accessed: 13- Aug- 2020].

THANK YOU