Attention2AngioGAN: Synthesizing Fluorescein Angiography from Retinal Fundus Images using Generative Adversarial Networks

University of Nevada, Reno

Sharif Amit Kamran[1], Khondker Fariha Hossain[2], Alireza Tavakkoli[1], and Stewart Lee Zuckerbrod[3]

[1]University of Nevada, Reno, NV, USA

[2]Deakin University, Melbourne, Australia

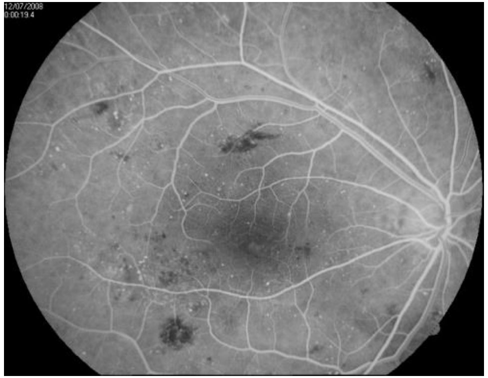
[2]Houston Eye associates, Houston, TX, USA

Presented By: Sharif Amit Kamran

Session Title: Medical Image Analysis I

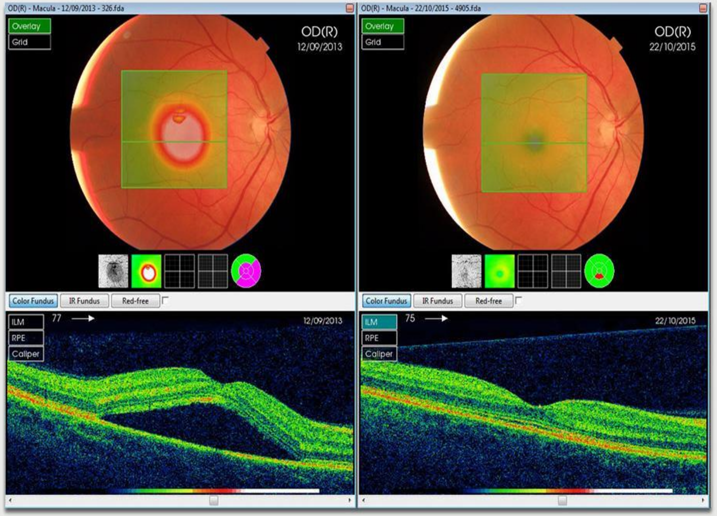
What is Fundus Fluorescein Angiography ?

- Retinal Degeneration
- Intravenous Fluorescent dye
- Affordable and wide-spread
- Complications: Nausea, Vomiting, Anaphylactic Shock, Death



Current Solution: SD-OCT

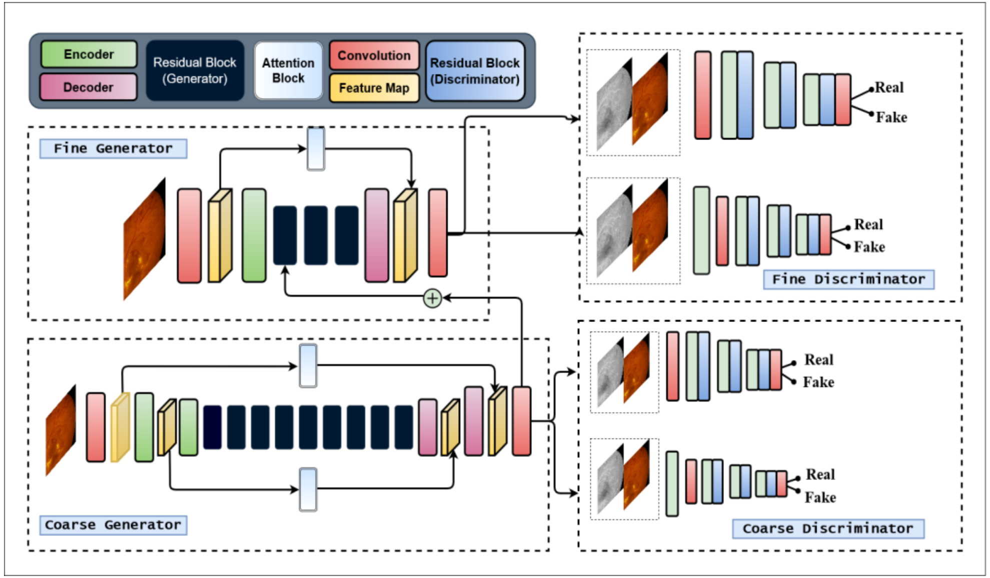
- A tool used for viewing the morphology of the retinal layers
- Differential diagnosis conducted by an expert
- Expensive
- Not available in developing countries



Optical Coherence Tomography

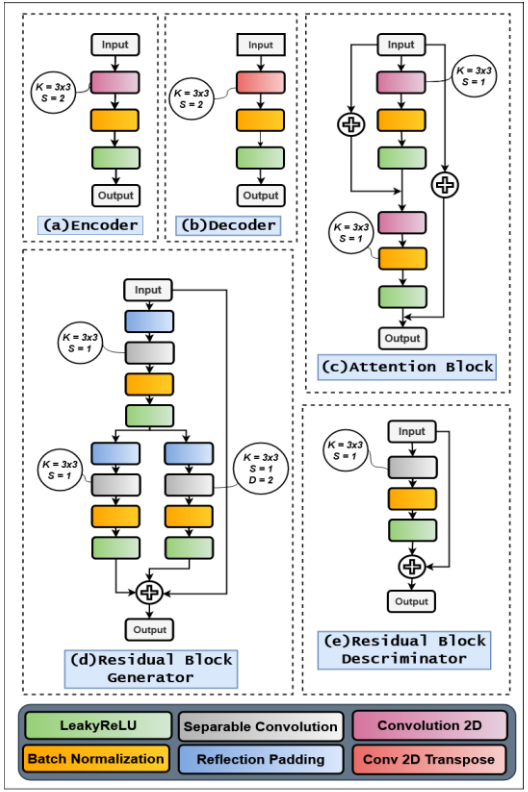
Proposed Architecture

- Conditional Generative Architecture
- Two Generators, G_{Coarse} G_{Fine}
- Four Discriminators $D1_{Coarse}$ $D1_{Fine}$ $D2_{Coarse}$ $D2_{Fine}$
- Multi-scale inputs & outputs (Spatial Dimension)
- Perceptual loss for G_{Fine} , G_{Coarse}
- Feature-matching loss for $D1_{Coarse}$ $D1_{Fine}$ $D2_{Coarse}$ $D2_{Fine}$
- Attention Block for retaining Manifold features



Proposed Blocks

- **Encoder:** Convolution followed by BatchNorm, Leaky ReLU
- **Decoder:** Separable Convolution followed by BatchNorm, Leaky ReLU
- **Discriminator Res-block :** Separable Convolution
- **Generator Res-Block:** Two branches, i) Convolution with Dilation ii) Separable Convolution
- **Attention Block :** Convolution followed by BatchNorm, Leaky ReLU x2



Perceptual and Feature Matching Loss

Here,

L_{perc} = Perceptual loss

L_{fm} = Feature-matching loss

F_{vgg} = VGG encoder

X = Real fundus

Y = Real Angio

G(x) = Fake Angio

N = No. of features from discriminators

M = No. of features from encoders

$$\mathcal{L}_{perc}(G) = \mathbb{E}_{x,y} \sum_{i=1}^{k} \frac{1}{M} \| F_{vgg}^{i}(y) - F_{vgg}^{i}(G(x)) \|$$

$$\mathcal{L}_{fm}(G, D_n) = \mathbb{E}_{x,y} \sum_{i=1}^{k} \frac{1}{N} \| D_n^{i}(x, y) - D_n^{i}(x, G(x)) \|$$

Adversarial Loss

Here,

$L_{adv}(G)$ = Generator loss

$L_{adv}(D)$ = Discriminator loss

$L_{adv}(G,D)$ = Adversarial loss

ƛ = Weight multipliers

X = Real fundus

Y = Real Angio

G(x) = Fake Angio

$$\mathcal{L}_{adv}(D) = -\mathbb{E}_{x,y}\left[\ \min(0, -1 + D(x,y))\right] \ - \mathbb{E}_{x}\left[\ \min(0, -1 - D(x, G(x)))\right]$$

$$\mathcal{L}_{adv}(G) = -\mathbb{E}_{x,y}\left[(D(G(x), y))\right]$$

$$\mathcal{L}_{adv}(G, D) = \mathcal{L}_{adv}(D) + \lambda_{adv}(\mathcal{L}_{adv}(G))$$

$$\min_{G_f, G_c}\Big(\max_{D_f, D_c}(\mathcal{L}_{adv}(G_f, G_c, D_f, D_c)) + \lambda_{rec}\left[\ \mathcal{L}_{rec}(G_f, G_c)\right] + \lambda_{fm}\left[\ \mathcal{L}_{fm}(G_f, G_c, D_f, D_c)\right]$$

$$+ \lambda_{perc}\left[\ \mathcal{L}_{perc}(G_f, G_c)\right]\ \Big)$$

Hyperparameters

Batch size, **b** = 4

Epochs, **e** = 100

Learning rate, $\boldsymbol{\alpha}$= $2e^{-4}$, β_1= 0.5, β_2= 0.999

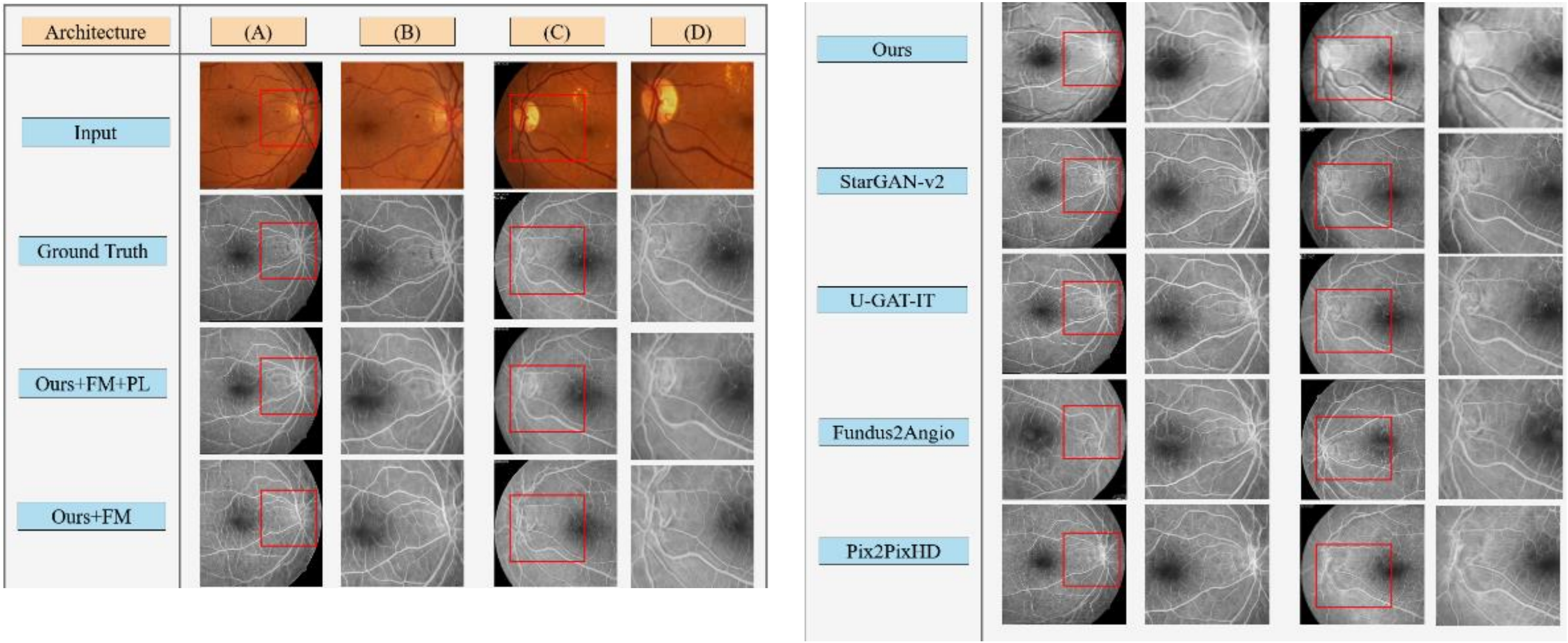
Weight multiplier, λ_{fm}= 1 , λ_{rec}= 10, λ_{perc}= 10, λ_{adv}= 10

Dataset

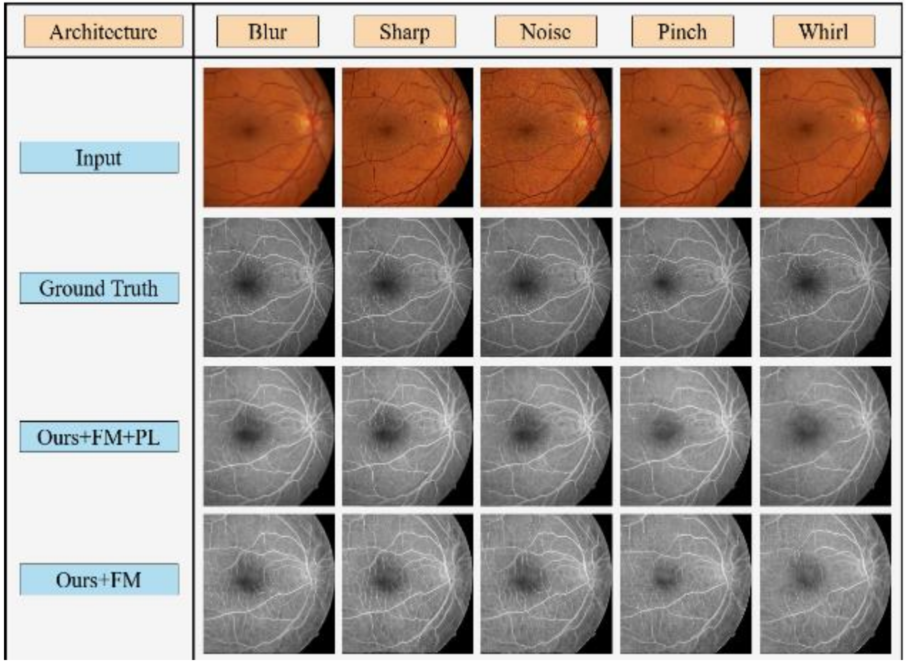
- Data: **Hajeb et al** [1]
- 17 Normal and Abnormal Fundus-Angio pairs
- Fundus resolution : 576x720x3
- Angio resolution : 576x720x1
- 50 random crops : Total image 50x17 = 850

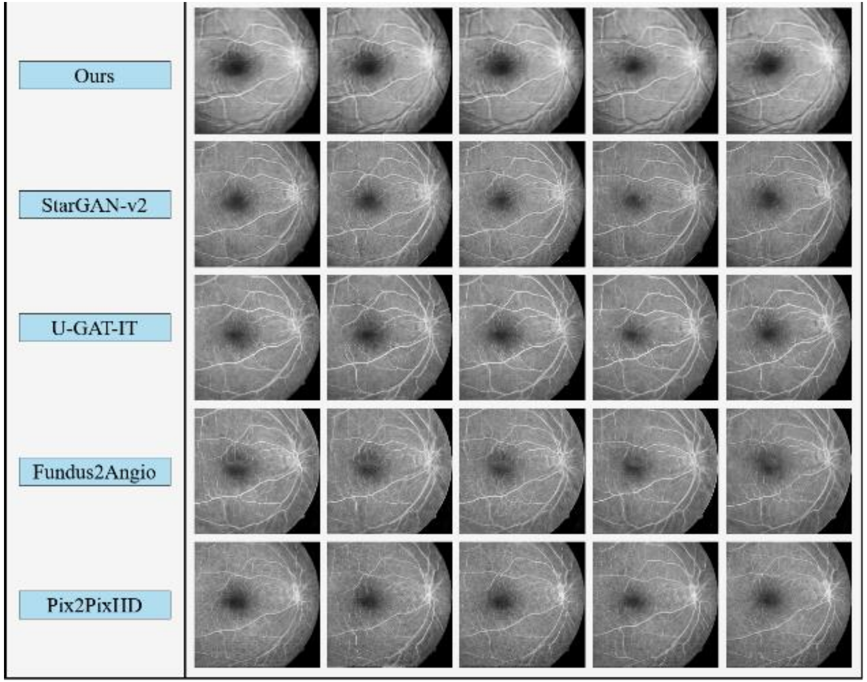
[1] SH. Hajeb, H. Rabbani, MR. Akhlaghi, “Diabetic Retinopathy Grading by Digital Curvelet Transform", Computational and Mathematical Methods in Medicine, vol. 2012, Article ID 761901, 11 pages, 2012.1607-1614, July 2012.

Qualitative Evaluation - 1



Qualitative Evaluation - 2





Quantitative Evaluation - 1

Table 4.1: Test results for different architectures for Retinal Angiogram generation

Fréchet Inception Distance (FID)						
Architecture	Orig.	Noise	Blur	Sharp	Whirl	Pinch
A2AGAN + PL[1]+ FM[2]	24.6	21.6 (3.0↓)	30.0 (5.4↑)	25.6 (1.0↑)	40.0 (15.4↑)	24.9 (0.3↑)
A2AGAN + FM[2]	**20.7**	**20.8** (0.1↑)	**23.5** (2.8↑)	**24.9** (4.2↑)	**27.8** (7.1↑)	**19.5** (1.2↓)
A2AGAN	47.5	43.1 (4.4↓)	49.8 (2.3↑)	50.5 (3.5↑)	61.9 (14.5↑)	46.7 (0.8↓)
StarGAN-v2 [37]	27.7	35.1 (7.4↑)	32.6 (4.9↑)	27.4 (0.3↓)	32.7 (5.0↑)	26.7 (1.0↓)
U-GAT-IT [36]	24.5	26.0 (1.5↑)	30.4 (5.9↑)	26.8 (2.3↑)	33.0 (9.5↑)	29.1 (4.6↑)
Fundus2Angio	30.3	41.5 (11.2↑)	32.3 (2.0↑)	34.3 (4.0↑)	38.2 (7.9↑)	33.1 (2.8↑)
Pix2PixHD [35]	42.8	53.0 (10.2↑)	43.7 (1.1↑)	47.5 (4.7↑)	45.9 (3.1↑)	39.2 (3.6↓)
Kernel Inception Distance (KID)						
Architecture	Orig.	Noise	Blur	Sharp	Whirl	Pinch
A2AGAN + PL[1]+ FM[2]	**0.00087**	**0.05045**	**0.00235**	**0.05162**	0.05390	**0.04575**
A2AGAN + FM[2]	0.00392	0.05390	0.00505	0.05301	0.05657	0.05341
A2AGAN	0.00595	0.05237	0.00617	0.05298	0.05613	0.05419
StarGAN-v2 [37]	0.00118	0.05274	**0.00235**	0.05331	0.05539	0.05271
U-GAT-IT [36]	0.00131	0.05610	0.00278	0.05533	0.05815	0.05719
Fundus2Angio	0.00184	0.05328	0.00272	0.05267	**0.05278**	0.04985
Pix2PixHD [35]	0.00258	0.05613	0.00254	0.05788	0.06029	0.05838

[1] PL = Perceptual Loss; FM = Feature-Matching Loss

[2] FID: Lower is better; KID: Lower is better

Quantitative Evaluation - 2

TABLE II
RESULTS OF QUALITATIVE WITH UNDISCLOSED PORTION OF FAKE/REAL EXPERIMENT

		Results		Average		
Architectures		Correct	Incorrect	Missed[1]	Found[1]	Precision[2]
Ours + FM + PL	Fake Real	10% 80%	90% 20%	55%	45%	**47.1%**
Ours + FM	Fake Real	12% 82%	88% 18%	53%	47%	**48.2%**

[1] Missed higher is better; Found lower is better
[2] Precision Lower is better

Future Direction

- Synthesizing other modalities of opthalmological images
- Cross-domain information fusion
- Retinal vessel segmentation using GAN

Thank you