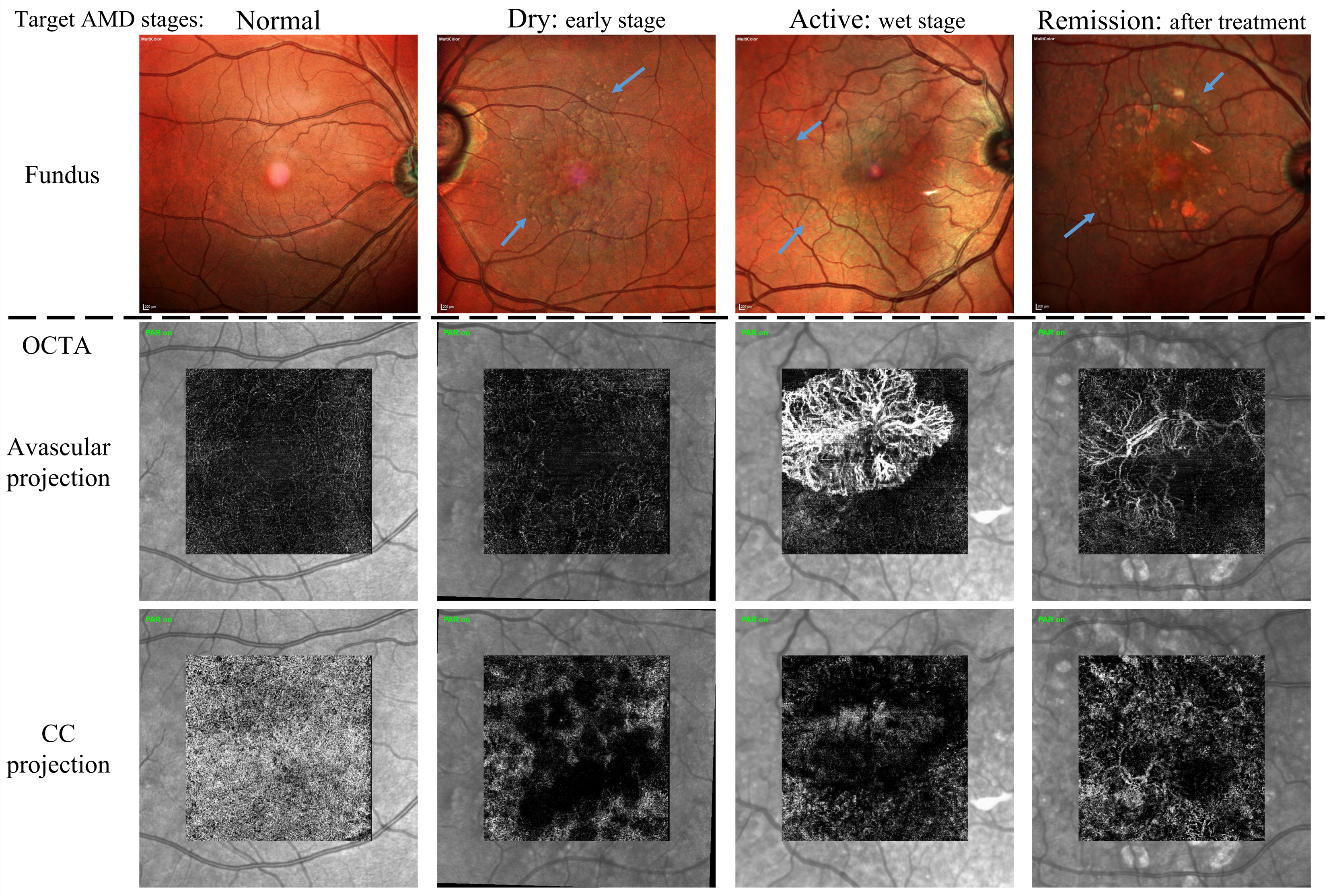
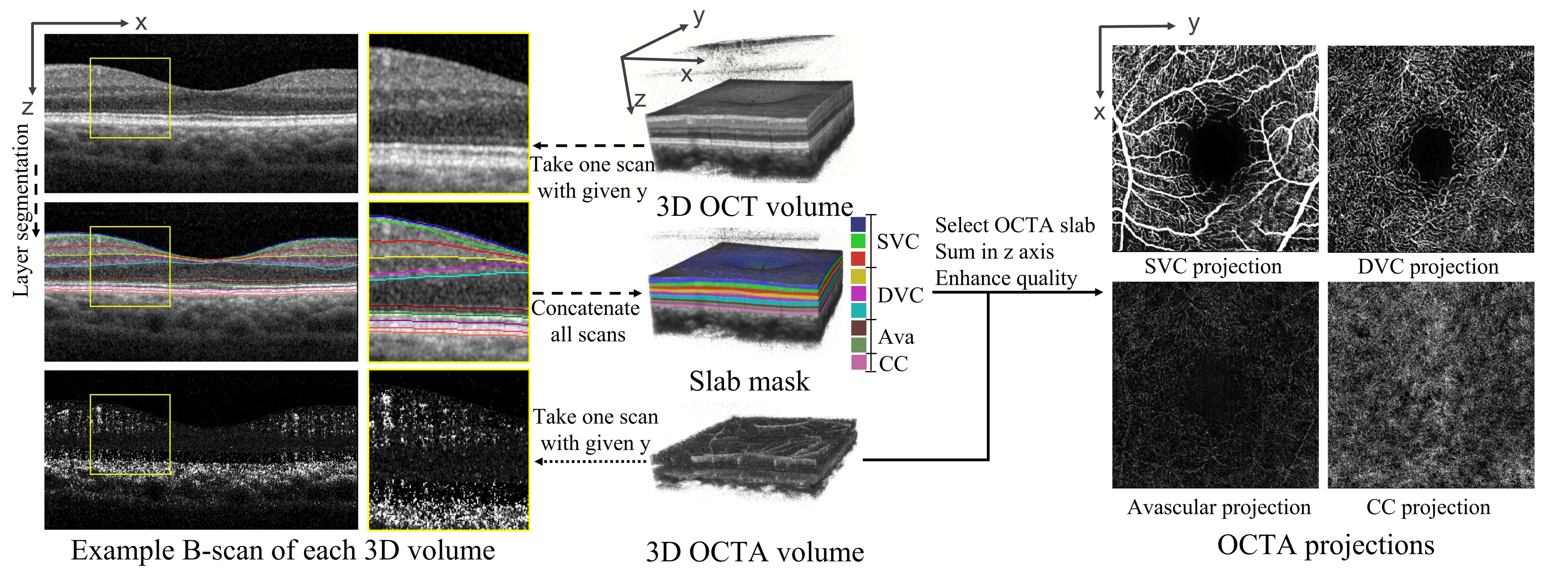
Robust AMD Stage Grading with Exclusively OCTA Modality Leveraging 3D Volume

Haochen Zhang, Anna Heinke, Carlo Galang, Daniel Deussen, Bo Wen, Dirk-Uwe Bartsch, William Freeman, Truong Nguyen, Cheolhong An

Computer Vision for Automated Medical Diagnosis
ICCV 2023 Workshop

Background





- **Left**: Comparison between fundus and OCTA w.r.t. AMD stages. All AMD stages exhibit drusens in fudus while OCTA displays pathological vessels in different retina layers directly. It is still an ongoing challenge to tell active stage from remission.
- **Top**: Interrelationships among OCT and OCTA raw volume, B-scans, and OCTA projection. OCT B-Scan layer segmentation influences OCTA projection generation.

Contribution

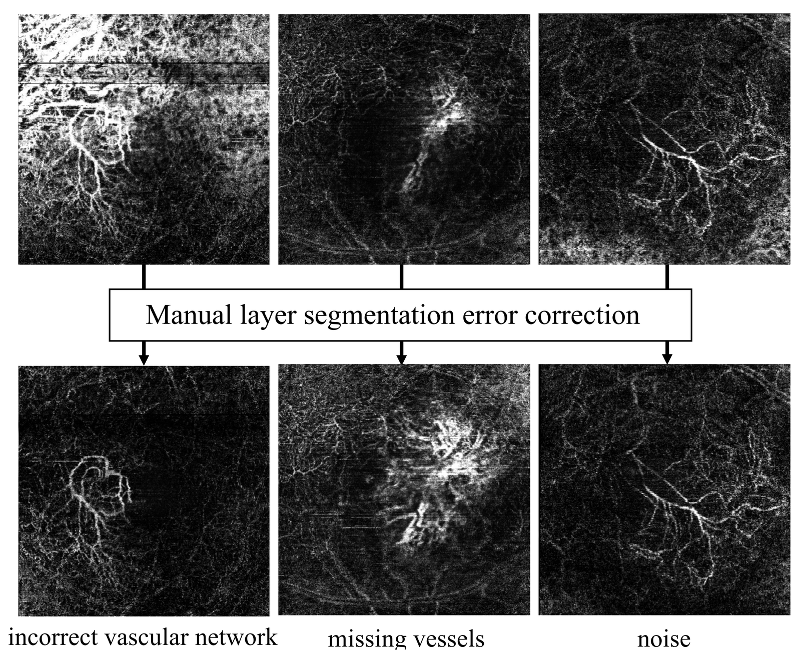
- ✓ We experimentally verified that the OCTA projections, which ophthalmologists usually use for diagnosis, are easily affected by layer segmentation errors. Those errors degrade the classification performance.
- ✓ We propose to use 3D raw OCTA volume to avoid the impacts of those errors. To achieve this, we modify a pretrained 2D network to perform volume classification. We also adopt an additional projection supervision to facilitate training of shallow feature extractor.
- ✓ Experimental results show that the proposed classifier can achieve the accuracy of more than 80%, regardless of the presence of layer segmentation errors. These results prove the effectiveness of our methods and suggest that OCTA is a promising modality to distinguish various stages of AMD disease.

Layer Segmentation Error

➢Error Population

sample type	Active	Remission	Dry	Normal	Total
# w/ seg. error	138	91	57	2	288
# w/o seg. error	52	39	90	61	242
error percentage	72.6%	70%	38.8%	3.2%	54.3%

➢Visual Inspection



➢Classification Accuracy

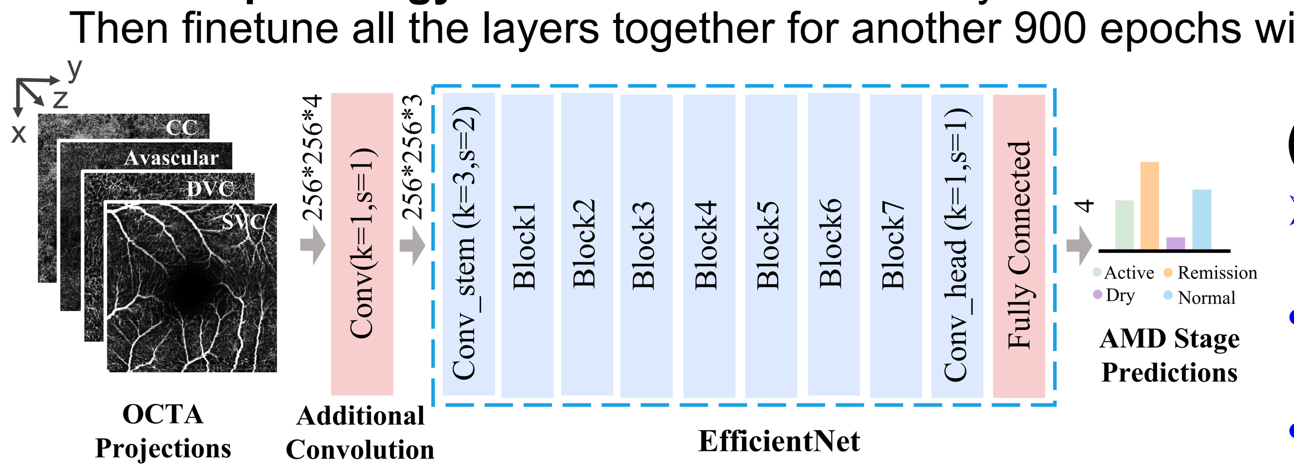
Train on	Test on	Accuracy
Clean set	Clean set	69.64%
Clean set	Mixed set	53.57%
Mixed set	Clean set	64.29%
Mixed set	Mixed set	57.14%

- R1,2: Hard to generalize from clean to mixed
- R1,2,4: Possible to learn the joint distribution
- R3,4 Samples with errors are hard to learn.

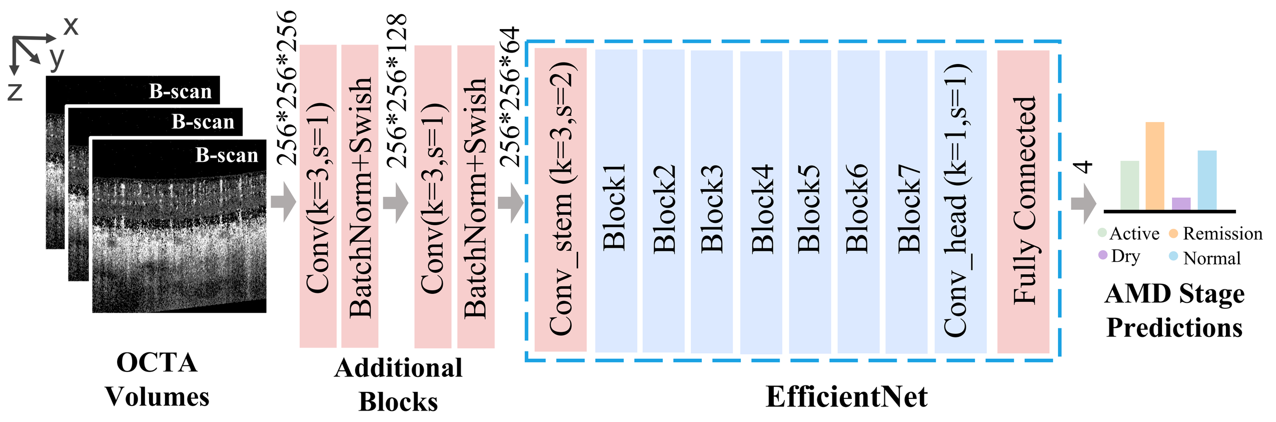
Methods

(a) 2D OCTA Projection Input

- ➢ **Classifier structure**: One additional convolution layer to convert input channels; Adjusted the output of the last FC to match the number of categories.
- ➢ **Warmup strategy**: first freeze all the blue layers and train only the red ones for 600 epochs; Then finetune all the layers together for another 900 epochs with a smaller learning rate.



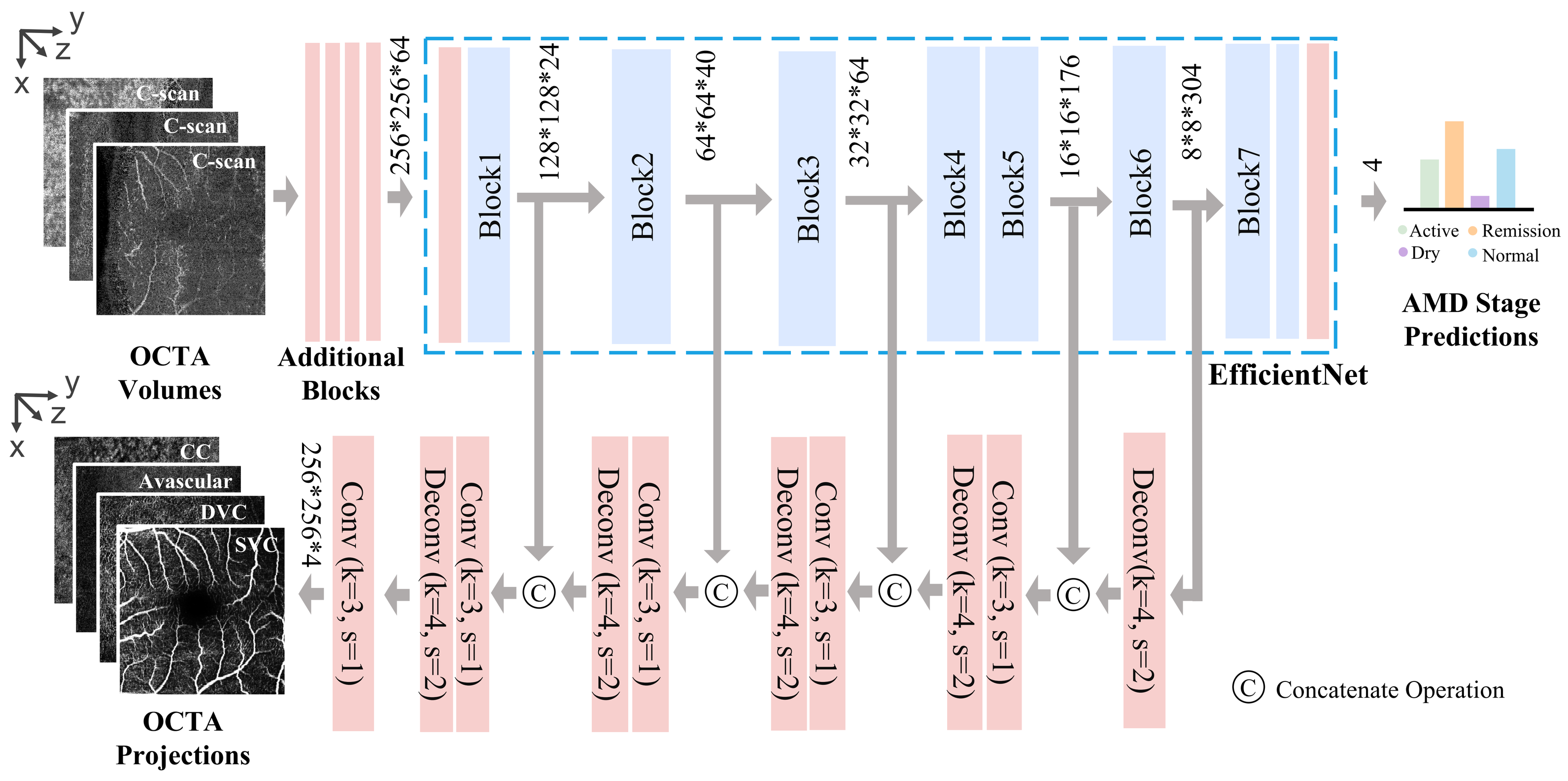
(a) 2D OCTA projection input



(b) 3D OCTA volume input

(b,c) 3D OCTA volume Input

- ➢ **2D convolutional backbone for volume classification**:
  - Take one dimension of 3D as channel dimension (y-axis > z-axis);
  - Use additional convolutional blocks to reduce channels.
- ➢ **Projection supervision**:
  - Add another Unet decoder branch onto the EfficientNet backbone to generate OCTA projections from 3D OCTA volume;
  - Preserve the information necessary for displaying vessel patterns and aiding in AMD grading.



(c) 3D OCTA volume input w/ projection supervision

Experimental Results

Data Preparation:

- The dataset consists of 697 raw OCTA volumes with projections: active 182, remission 187, dry 188 and normal 140.
- Error-free subset only has samples with no layer segmentation errors; Error-prone subset contains numerous samples with errors.

Results:

2D Input	Setting		Error-free		Error-prone	
	MM	PT	Accuracy	AUC	Accuracy	AUC
Thakoor et. al. [30]	✗	✗	55.36	0.8159	57	0.8176
	✓	✗	62.5	0.8512	66	0.8428
ours(2D)	✗	✗	73.21	0.8565	62	0.8065
	✗	✓	**80.36**	**0.9264**	**72**	**0.8697**
Human	-	-	58.92	-	60	-

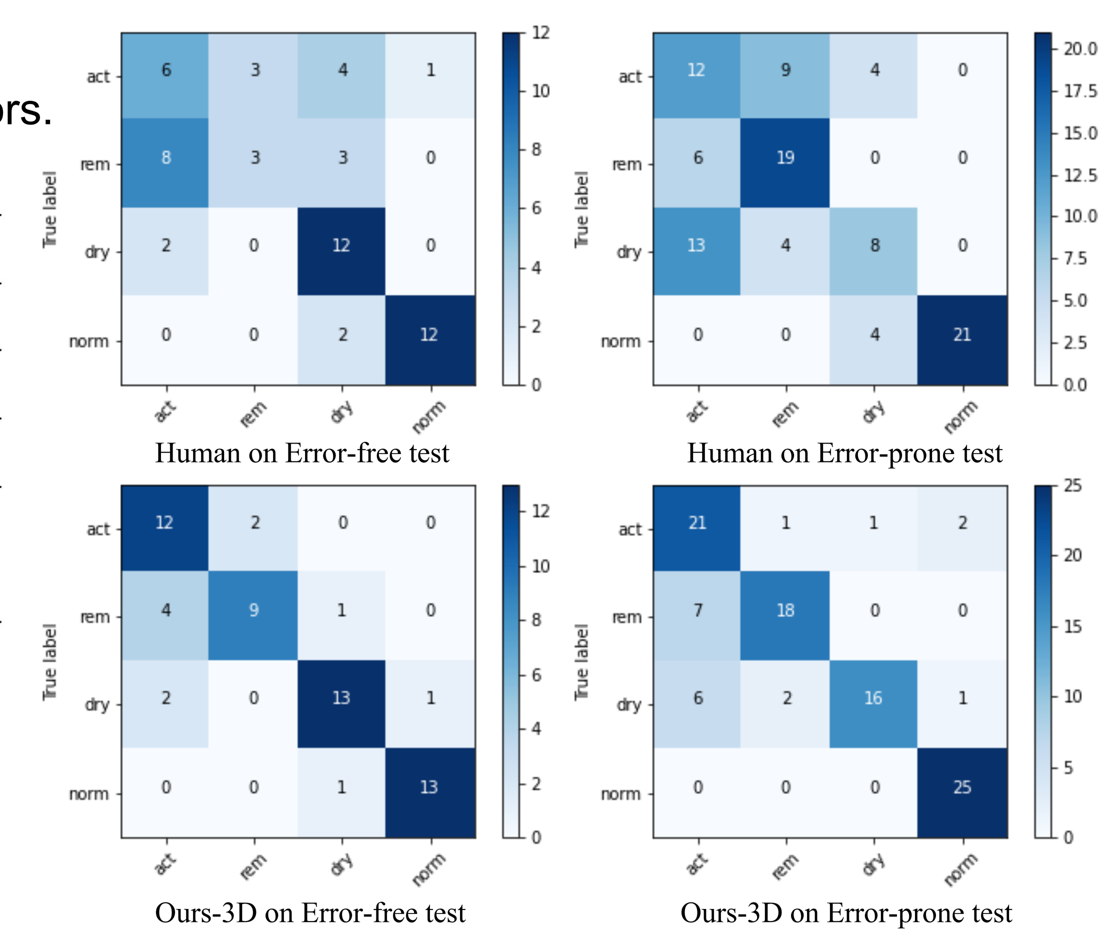
3D Input	Setting		Error-free		Error-prone	
	PT	PS	Accuracy	AUC	Accuracy	AUC
Effic.Net 3D	✗	✗	75	0.9489	69	0.8841
Med.Net34 [7]	✓	✗	73.21	0.9238	73	0.9009
ours(3D)	✓	✗	82.14	**0.9524**	74	0.9055
	✓	✓	**83.93**	0.9298	**80**	**0.912**

* MM: Multimodal information (including OCT B-scan, OCT and OCTA projections), PT: Pretraining. PS: Projection Supervision.

Discussion:

- [30] vs Ours(2D): EfficientNet backbone and Imagenet pretrained model helps.
- Human vs Ours(2D): Proving the potential of OCTA as a diagnostic modality for AMD.
- Ours(2D) vs Ours(3D), Error-free vs Error-prone: Directly analyzing 3D raw data benefits
- 3D Conv vs Ours(3D): Well-designed 2D CNN is better than 3D when training data is limited.
- Ours(3D): Our proposed projection supervision is helpful.

Confusion Matrices:



Confusion Matrices:

- Human struggled to distinguish remission from active;
- Layer segmentation error degrades accuracy on Dry;
- Ours-3D performs well and resist layer segmentation error better.