

Results:

| 2D Input | Setting | | Error-free | | Error-prone | |
|-------------------------|--------------|--------------|--------------|---------------|-------------|---------------|
| | MM | PT | Accuracy | AUC | Accuracy | AUC |
| Thakoor et. al. [30] | X | X | 55.36 | 0.8159 | 57 | 0.8176 |
| | ✓ | X | 62.5 | 0.8512 | 66 | 0.8428 |
| ours(2D) | X | X | 73.21 | 0.8565 | 62 | 0.8065 |
| | X | ✓ | 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 | X | X | 75 | 0.9489 | 69 | 0.8841 |
| Med.Net34 [7] | ✓ | X | 73.21 | 0.9238 | 73 | 0.9009 |
| ours(3D) | ✓ | X | 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.