

**Table S9. Summary statistics for the grading prediction performance (LOCS- III based grading) of the deep learning system on retro-illumination images (PSC).**

	AUC	Sensitivity	Specificity	Accuracy
<b>Grade 0 vs Others</b>	0.9391 (0.9220 – 0.9563)	81.41% (72.58%, 90.25%)	86.78% (83.53%, 90.03%)	84.67% (80.81%, 88.54%)
<b>Grade 1 vs Others</b>	0.7441 (0.6925 – 0.7959)	72.94% (61.58%, 84.30%)	93.33% (90.03%)	90.83% (87.69%, 93.96%)
<b>Grade 2 vs Others</b>	0.7716 (0.7186 – 0.8247)	63.36% (55.47%, 71.25%)	96.39% (93.05%, 99.73%)	90.05% (86.50%, 93.60%)
<b>Grade 3 vs Others</b>	0.8532 (0.7885 – 0.9179)	63.77% (54.34%, 73.21%)	94.18% (90.52%, 97.84%)	86.95% (83.13%, 90.77%)
<b>Grade 4 vs Others</b>	0.9663 (0.9556 – 0.9772)	71.58% (57.82%, 85.34%)	91.74% (87.72%, 95.75%)	89.20% (84.82%, 93.58%)
<b>Grade 5 vs Others</b>	0.9848 (0.9781 – 0.9914)	100.0% (100.0%, 100.0%)	100.0% (100.0%, 100.0%)	100.0% (100.0%, 100.0%)
<b>Overall</b>	0.9174 (0.9055 – 0.9295)	71.49% (66.58%, 76.41%)	91.64% (90.36%, 92.92%)	87.89% (85.58%, 90.20%)
<b>Overall (ResNet18)</b>	0.8952 (0.8801 – 0.9103)	67.87% (63.24%, 72.50%)	91.06% (89.93%, 92.20%)	86.37% (84.44%, 88.31%)
<b>Overall (WideResNet50-2)</b>	0.9087 (0.8970 – 0.9203)	70.88% (65.77%, 75.99%)	91.21% (89.98%, 92.44%)	87.16% (84.93%, 89.40%)
<b>Overall (ResNext50)</b>	0.9141 (0.9034 – 0.9248)	68.55% (63.23%, 73.86%)	91.10% (90.01%, 92.19%)	87.61% (85.53%, 89.68%)

Asymptotic two-sided 95% CIs, adjusted for clustering by patients, were calculated and presented as proportions (sensitivity, specificity, accuracy) and the AUC.

Abbreviations: LOCS-III = Lens Opacities Classification System III; PSC = posterior subcapsular opacity; CI = confidence interval; AUC = area under the receiver operating characteristic curve