

Supplementary Table S1. Discriminative performance of the multi-task deep-learning system in the classification of diabetic macular edema (DME) and center-involved DME (CI-DME) at the B-scan level across primary validation and external testing datasets

Classification task	OCT device	Dataset	AUROC (95% CI)	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Accuracy, % (95% CI)
Presence vs. absence of DME	Spectralis OCT	Primary	0.940 (0.934–0.945)	88.7 (86.9–90.3)	99.2 (99.0–99.5)	97.2 (96.8–97.6)
		External-4	0.958 (0.951–0.964)	92.9 (89.6–95.3)	98.7 (98.3 – 99.1)	98.2 (97.7–98.6)
		External-5	0.959 (0.953–0.964)	95.5 (94.5–96.4)	96.3 (95.6–96.9)	96.0 (95.5–96.6)
		External-6	0.899 (0.893–0.905)	80.4 (78.9–81.8)	99.4 (99.2–99.6)	93.6 (93.0–94.0)
	Triton OCT	Primary	0.937 (0.931–0.942)	88.5 (86.9–90.0)	98.8 (98.5–99.0)	96.8 (96.4–97.2)
		External-7	0.947 (0.941–0.953)	97.5 (97.0–98.0)	92.0 (90.4–100.0)	96.1 (95.6–96.1)
CI-DME vs. non-CI-DME	Spectralis OCT	Primary	0.940 (0.927–0.954)	90.3 (85.6–93.8)	98.1 (97.0–98.8)	96.1 (95.4–97.6)
		External-4	0.959 (0.932–0.978)	96.4 (89.9–99.3)	95.4 (92.0–97.7)	95.7 (92.9–97.6)
		External-5	0.956 (0.945–0.965)	93.8 (90.8–96.0)	97.4 (96.3–98.1)	96.6 (95.6–97.4)
		External-6	0.905 (0.892–0.916)	99.4 (98.7–99.8)	81.6 (79.4–83.6)	89.2 (87.9–90.4)
	Triton OCT	Primary	0.951 (0.938–0.961)	93.6 (91.9–95.0)	96.5 (94.1–98.1)	94.4 (93.0–95.5)
		External-7	0.952 (0.945–0.959)	99.5 (99.2–99.7)	90.1 (86.5–94.3)	99.0 (98.6–99.3)

AUROC: area under the receiver operating characteristic curve; CI-DME: center-involved diabetic macular edema, CI: confidence interval; OCT: optical coherence tomography.

Supplementary Table S2. Discriminative performance of the multi-task deeplearning system in the classification of non-DME retinal abnormalities at the B-scan level across primary validation and external testing datasets

Classification task	OCT device	Dataset	AUROC (95% CI)	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Accuracy, % (95% CI)
Presence vs. absence of non-DME retinal abnormalities	Spectralis OCT	Primary	0.975 (0.971–0.979)	95.5 (92.3–97.6)	90.7 (90.0–91.4)	90.9 (90.2–91.6)
		External-4	0.977 (0.972–0.982)	97.0 (93.5–98.9)	89.3 (88.2–90.3)	89.7 (88.7–90.7)
		External-5	0.988 (0.985–0.991)	95.7 (93.3–97.4)	94.8 (94.1–95.4)	94.9 (94.2–95.5)
		External-6	0.924 (0.919–0.929)	86.7 (85.4–88.0)	92.6 (92.0–93.2)	91.1 (90.5–91.6)
	Triton OCT	Primary	0.962 (0.958–0.966)	95.3 (92.8–97.1)	94.1 (93.5–94.6)	94.2 (93.6–94.7)
		External-7	0.948 (0.942–0.954)	86.8 (83.9–89.3)	91.9 (91.1–92.6)	91.3 (90.5–92.0)

AUROC: area under the receiver operating characteristic curve; CI-DME: center-involved diabetic macular edema, CI: confidence interval; OCT: optical coherence tomography.

Supplementary Table S3. Diagnostic performance of the deep-learning system in DME detection when only one scan from one eye was included in the primary dataset.

Classification task	OCT device	Dataset	AUROC (95% CI)	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Accuracy, % (95% CI)
Presence vs. absence of DME	Cirrus HD-OCT	Primary	0.915 (0.867–0.950)	82.9 (67.9–92.8)	100.0 (97.7–100.0)	96.5 (92.9–98.6)
	Spectralis OCT	Primary	0.938 (0.877–0.975)	88.9 (73.9–96.9)	98.7 (93.1–100.0)	95.6 (90.1–98.6)
	Triton OCT	Primary	0.989 (0.962–0.998)	98.5 (92.1–100.0)	99.2 (95.7–100.0)	99.0 (96.3–99.9)
CI-DME vs. non-CI-DME	Cirrus HD-OCT	Primary	0.966 (0.839–0.999)	93.1 (77.2–99.2)	100.0 (47.8–100.0)	94.1 (80.3–99.3)
	Spectralis OCT	Primary	0.952 (0.813–0.996)	90.5 (69.6–98.8)	100.0 (71.5–100.0)	93.8 (79.2–99.2)
	Triton OCT	Primary	0.974 (0.902–0.997)	100.0 (92.6–100.0)	94.7 (74.0–99.9)	98.5 (92.0–100.0)

AUROC: area under the receiver operating characteristic curve; CI: confidence interval; OCT: optical coherence tomography.

Supplementary Table S4. Diagnostic performance of the deep-learning system in non-DME retinal abnormalities detection performances when only one scan from one eye was included in the primary dataset.

Classification task	OCT device	Dataset	AUROC (95% CI)	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Accuracy, % (95% CI)
Presence vs. absence of non- DME retinal abnormalities	Cirrus HD-OCT	Primary	0.938 (0.895–0.967)	93.3 (81.7–98.6)	88.9 (82.8–93.4)	89.9 (84.8–93.3)
	Spectralis OCT	Primary	0.993 (0.955–0.999)	91.7 (61.5–99.8)	100.0 (96.5–100.0)	99.1 (95.3–100.0)
	Triton OCT	Primary	0.978 (0.946–0.994)	100.0 (82.4–100.0)	92.6 (87.6–96.0)	93.3 (88.8–96.4)

AUROC: area under the receiver operating characteristic curve; CI: confidence interval; OCT: optical coherence tomography.

Supplementary Table S5. Characteristics of subjects with non-DME retinal abnormalities in the primary dataset.

<u>Dataset</u>	<u>OCT device</u>	<u>Total sample n=14,246</u>	<u>Sample for DME classification module</u>		
			No DME	Non-CI-DME	CI-DME
Primary	Cirrus HD-OCT	No. of OCT volumes	253	29	101
		No. of eyes	69	14	28
		No. of subjects	41	10	21
		Gender (male, %)	16 (39.0)	5 (50.0)	9 (42.9)
		Age (years, SD)	70.4 (11.6)	69.7 (8.9)	63.9 (8.2)
	Spectralis OCT	No. of OCT volumes	189	60	104
		No. of OCT B-scans	3,452	1,025	223
		No. of eyes	96	29	37
		No. of subjects	60	20	23
		Gender (male, %)	32 (53.3)	10 (50.0)	14 (60.9)
		Age (years, SD)	62.8 (10.3)	63.9 (10.5)	69.6 (9.8)
	Triton OCT	No. of OCT volumes	665	162	342
		No. of OCT B-scans	6,661	648	1854
		No. of eyes	239	48	140
		No. of subjects	104	25	98
		Gender (male, %)	66 (63.5)	12 (48.0)	46 (46.9)
		Age (years, SD)	58.9 (13.0)	64.9 (12.6)	63.0 (11.4)

CI-DME: center-involved diabetic macular edema; OCT: optical coherence tomography; SD: standard deviation.

Supplementary Table S6. Characteristics of subjects with DME in the primary dataset.

<u>Dataset</u>	<u>OCT device</u>	<u>Total sample</u> n=14,454	<u>Sample for non-DME retinal abnormalities</u> <u>classification module</u>	
			Absence	Presence
Primary	Cirrus HD-OCT	No. of OCT volumes	1,078	130
		No. of eyes	242	42
		No. of subjects	148	31
		Gender (male, %)	89 (60.1)	14 (45.2)
		Age (years, SD)	62.6 (9.8)	63.8 (9.6)
	Spectralis OCT	No. of OCT volumes	367	164
		No. of OCT B-scans	4,194	1,248
		No. of eyes	111	66
		No. of subjects	67	43
		Gender (male, %)	43 (64.2)	24 (55.8)
		Age (years, SD)	62.0 (10.6)	64.4 (11.0)
	Triton OCT	No. of OCT volumes	707	504
		No. of OCT B-scans	5,302	2,502
		No. of eyes	197	188
		No. of subjects	100	123
		Gender (male, %)	50 (50.0)	58 (47.2)
		Age (years, SD)	62.9 (11.6)	65.1 (10.1)

CI-DME: center-involved diabetic macular edema; OCT: optical coherence tomography; SD: standard deviation.

Supplementary Table S7. Diagnostic performance of the deep-learning system in DME detection among eyes with non-DME retinal abnormalities across primary dataset.

Classification task	OCT device	Dataset	AUROC (95% CI)	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Accuracy, % (95% CI)
Presence vs. absence of DME	Cirrus HD-OCT	Primary	0.964 (0.912–0.990)	92.9 (66.1–99.8)	100.0 (96.4–100)	99.1 (95.3–100.0)
	Spectralis OCT	Primary	0.846 (0.664–0.952)	69.2 (38.6–90.9)	100 (79.4 –100.0)	86.2 (68.3–96.1)
	Triton OCT	Primary	0.935 (0.851–0.980)	96.8 (83.3–99.9)	90.2 (76.9–97.3)	93.1 (84.5–97.7)

AUROC: area under the receiver operating characteristic curve; CI: confidence interval; OCT: optical coherence tomography.

Supplementary Table S8. Diagnostic performance of the deep-learning system in the detection of non-DME retinal abnormalities among eyes with DME across primary dataset.

Classification task	OCT device	Dataset	AUROC (95% CI)	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Accuracy, % (95% CI)
Presence vs. absence of non-DME retinal abnormalities	Cirrus HD-OCT	Primary	0.927 (0.883–0.958)	92.9 (66.1–99.8)	92.5 (87.9–95.7)	92.5 (88.1–95.7)
	Spectralis OCT	Primary	0.917 (0.869–0.952)	87.5 (61.7–98.4)	96.0 (91.9–98.4)	95.3 (91.2–97.8)
	Triton OCT	Primary	0.947 (0.922–0.967)	97.6 (87.1–99.9)	91.9 (88.7–94.4)	92.5 (89.5–94.8)

AUROC: area under the receiver operating characteristic curve; CI: confidence interval; OCT: optical coherence tomography.