Online supplement

Deep Learning Detection of Penumbral Tissue on Arterial Spin Labeling in Stroke

Table I. Demographics of Training Dataset

Sex		
	Female	89/167 (53.3%)
	Male	78/167 (46.7%)
Age, years		70.7 (14.7)
Clinical features		
	NIHSS score*	14 (5-19)
	Systolic blood pressure, mmHg	152.0 (28.3)
	Diastolic blood pressure, mm Hg	85.0 (17.6)
	Large Vessel Occlusion	105/137 (76.6%)
Field strength		
	1.5 T	93/167 (55.7%)
	3.0 T	74/167 (44.3%)
Scan Time Point		
	Baseline	121/167 (72.5%)
	Time point 2	38/167 (22.8%)
	Time point 3	8/167 (4.8%)
Scan time from the onset of symptoms, h	Whole group	11.7 (6.0-28.2)
	Baseline	8.2 (4.6-22.6)
	Time point 2	25.5 (13.0-33.0)
	Time point 3	34.2 (30.6-84.6)
Affected Territory†		
	Internal Carotid Artery	33/126 (26.2%)
	Middle Cerebral Artery	76/126 (60.3%)
	Posterior Cerebral Artery	11/126 (10.3%)
	Anterior Cerebral Artery	7/126 (5.6%)
	Basilar Artery	1/126 (0.7%)
	Common Carotid Artery	3/126 (2.4%)
	Vertebral Artery	4/126 (3.2%)
	Normal appearance	21/126 (16.7%)
	Multiple, bilateral	5/126 (4.0%)

Data are in n (%), mean (SD), or median (IQR). NIHSS = National Institutes of Health Stroke Scale. *Scores range from 0 to 42, with higher scores indicating more severe neurological deficit. †Affected territory (territories) designation according to the clinician's interpretation of MRA, if available, at baseline.

Table II. Demographics of Stanford Validation Dataset

Sex		
	Female	8/12 (66.7%)
	Male	4/12 (33.3%)
Age, years		68.3 (11.3)
Clinical features		
	NIHSS score*	14 (9 - 18)
	Systolic blood pressure, mmHg	132(30)
	Diastolic blood pressure, mm Hg	69 (18)
	Large Vessel Occlusion	12/12 (100%)
Field strength		
	1.5 T	1/12 (8.3%)
	3.0 T	11/12 (92.7%)
Scan Time Point		
	Baseline	12/12 (100%)
Scan time from the onset of symptoms, h	Whole group	5.8 (3.4-9.9)
Affected Territory†		
	Middle Cerebral Artery	11/12 (92.7%)
	Internal Carotid Artery	1/12 (8.3)

Data are in n (%), mean (SD), or median (IQR). NIHSS = National Institutes of Health Stroke Scale. *Scores range from 0 to 42, with higher scores indicating more severe neurological deficit. †Affected territory (territories) designation according to the clinician's interpretation of MRA, if available, at baseline.

Table III Confusion Matrix of treatment decision of DL model in the UCLA cohort

		DSC	
		Positive	Negative
Inference	Positive	17	1
	Negative	2	18

Of the total 38 subjects, 35 were correctly classified, leading to an accuracy of 0.92 (95% CI: [0.79, 0.98]).

Table IV. Summary of classification indices of the DL model and 6 ML algorithms.

	Accuracy	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	AUC	Cohen's Kappa
DL model	0.92	0.89	0.95	0.94	0.90	0.95	0.84
Linear Regression	0.84	0.68	1.00	1.00	0.76	0.947	0.68
Ridge Regression	0.79	0.58	1.00	1.00	0.70	0.915	0.58
Kernel Ridge Regression	0.79	0.58	1.00	1.00	0.70	0.931	0.58
Neural Network	0.66	0.37	0.95	0.88	0.60	0.94	0.31
SVM with RBF Kernel	0.82	0.63	1.00	1.00	0.73	0.949	0.63
Random Forests	0.74	0.53	0.95	0.91	0.67	0.924	0.47

When the cut-off determined by voxel-wise training was applied, our DL model achieved significantly higher accuracy for treatment eligibility, compared with ML algorithms. When the cutoff threshold was varied to generate a ROC curve, the DL model still yielded the highest AUC of 0.950, while the AUC of the ML algorithms ranged from 0.915 to 0.949. Cohen's Kappa coefficient also supported that the DL model has the most consistent output with Tmax label compared with ML algorithms.

Table V Confusion Matrix of treatment decision of DL model in the Stanford cohort.

		DSC	
		Positive	Negative
Inference	Positive	3	0
	Negative	1	8

Of the total 12 subjects, only 1 subject was misclassified, yielding an accuracy of 0.92 (95% CI: [0.62, 0.99]).

Table VI Mean CBF ± SD of infarct core, penumbral tissue and contralateral region at 1.5T and 3T

CBF (ml/100g/min)	ASL 1.5T	ASL 3T
Infarct CBF	11.8±13.4	9.9±9.7
Penumbral CBF	15.4±16.6	12.7±11.1
Contralateral CBF	38.9±26.9	34.5±15.4

Although no specific CBF threshold are required for the DL model to learn, the mean CBF values increase from the infarct core, to penumbral tissue **(**based on DL inference**)**, and to the contralateral region at both 1.5 and 3T.