단어 수(공백 미포함): 2396/2400

Title: Low-False-Positive Cerebral Aneurysm Detection on TOF-MRA Using Refined Prototype-Contrastive Auxiliary Learning: External Multi-Scanner Validation

Authors:

Burnyoung Kim, Jun-Ho Kim, Sung-Min Gho

Purpose

To evaluate whether targeted refinements to a prior prototype-contrastive auxiliary loss improve cerebral aneurysm detection sensitivity at low false-positive (FP) rates (≤1 FP/scan) on TOF-MRA in an external multi-scanner validation cohort.

Methods and Materials

A 3D RetinaUNet detector was trained on 1,879 TOF-MRA studies (2,329 lesions; Philips 77%, 3.0T 80%) from four institutions. The baseline used the prior prototype-contrastive RetinaUNet, and the proposed model added three refinements: an ε-regularized contrastive denominator, 3×3×3 ROI feature averaging, and a margin-based prototype separation loss. Postprocessing parameters were tuned on internal cross-validation and fixed before hold-out and external testing.

Both models underwent five-fold ensemble inference on an internal hold-out cohort (282 cases, 220 lesions) and an independent external cohort (334 cases, 438 expert-annotated lesions): Siemens (n=158), GE (n=82), Philips (n=79), Toshiba/Canon (n=15); field strength was available for 332 of 334 cases: 1.5T, n=193; 3.0T, n=139. Lesion-level FROC was computed at FP/scan ∈ {1/8, 1/4, 1/2, 1}. Mean FROC sensitivity (mFROC) was reported with 95% bootstrap CIs from 1,000 resamples, and paired-bootstrap testing was used to compare models.

Results

External mFROC was higher for the proposed model than for the baseline model: 0.720 (95% CI, 0.639–0.796) versus 0.668 (95% CI, 0.590–0.742), with a paired-bootstrap difference of +0.050 (95% CI, +0.024 to +0.079; $p<0.001$). The largest gain was observed at the lowest FP operating point. At 0.125 FP/scan, sensitivity increased from 0.493 (95% CI, 0.415–0.573) to 0.639 (95% CI, 0.559–0.720), with a paired difference of +0.144 (95% CI, +0.090 to +0.201; $p<0.001$).

Vendor-stratified gains on the external cohort were heterogeneous, ranging from +0.003 in Philips to +0.066 in GE. Field-strength analyses showed numerically larger improvement at 3.0T (+0.074) than at 1.5T (+0.038). On the internal hold-out cohort, the mFROC difference was smaller and not statistically significant (mFROC 0.905 vs. 0.881, p=0.146).

Conclusion

Three targeted refinements to a prior prototype-contrastive auxiliary loss improved lesion-level FROC for cerebral aneurysm detection on an independent external TOF-MRA cohort, with the largest gain at the lowest false-positive operating point. These findings support potential utility in high-confidence aneurysm detection workflows.

Clinical Relevance/Application

Improved sensitivity at low false-positive rates may support more reliable TOF-MRA aneurysm detection workflows. Vendor- and protocol-specific calibration remains necessary before broad clinical deployment.

Figures & Tables

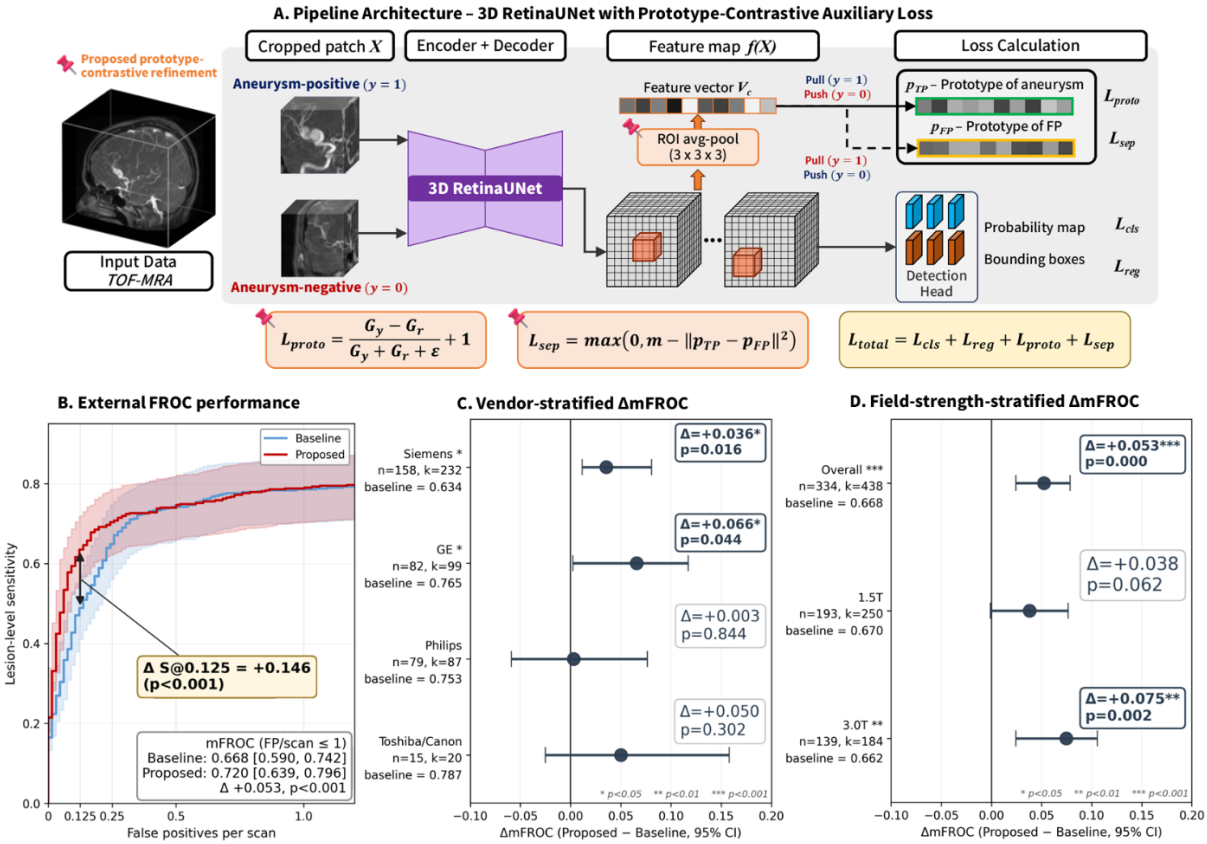

Figure 1. Pipeline architecture and external validation performance. (A) Schematic of the 3D RetinaUNet detector with a prototype-contrastive auxiliary loss module and three proposed refinements: ε-regularization, 3×3×3 ROI feature averaging, and explicit prototype separation loss (L_{sep}). (B) Lesion-level FROC curves on the independent external TOF-MRA cohort with 95% bootstrap confidence intervals. The largest improvement was observed at 0.125 FP/scan. (C) Vendor-stratified external performance shown as mFROC differences between the proposed and baseline models.

(D) Field-strength-stratified external performance shown as mFROC differences. Field-strength analysis excluded 2 cases with missing field-strength information. Error bars indicate 95% bootstrap confidence intervals. Subgroup analyses were exploratory.