Enhancing AI-assisted Stroke Emergency Triage with Adaptive Uncertainty Estimation

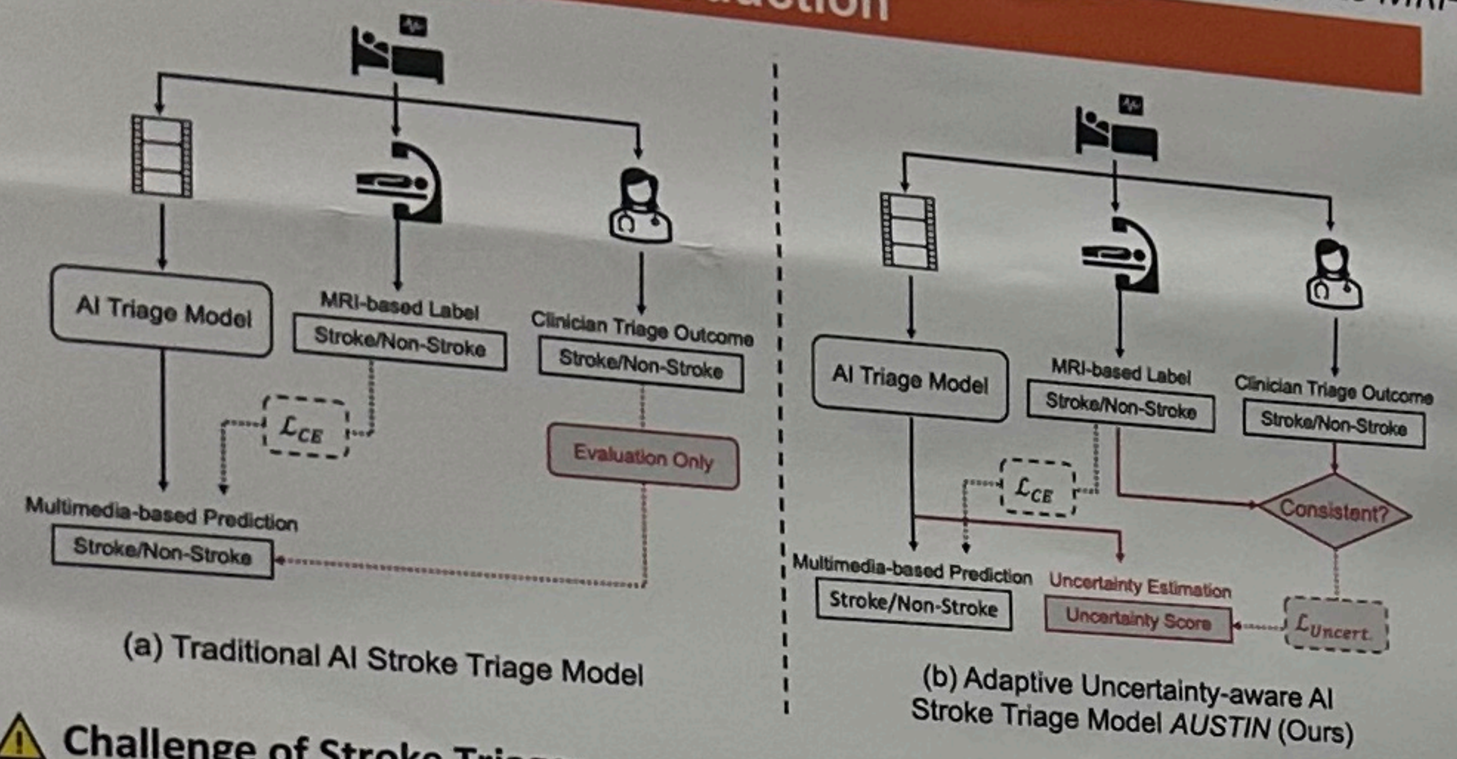
Shuhua Yang[1*], Tongan Cai[1*], Haomiao Ni[2*], Wenchao Ma[1], Yuan Xue[3], Kelvin Wong[4], John Volpi[4], James Z. Wang[1], Sharon X. Huang[1], Stephen T.C. Wong[4]

1. The Pennsylvania State University 2. University of Memphis 3. Ohio State University 4. Houston Methodist Hospital

MICCAI2025

"AUSTIN: An adaptive, uncertainty-aware multimedia model that exploits MRI–triage disagreements to boost ER stroke triage accuracy and yield confidence scores."

Introduction



(a) Traditional AI Stroke Triage Model

(b) Adaptive Uncertainty-aware AI Stroke Triage Model *AUSTIN* (Ours)

Challenge of Stroke Triage:

- **High-stake, subtle signs:** Stroke patients risk disability/mortality; symptoms can be subtle.
- **Limited access to MRI in ER:** MRI is gold standard but often unavailable in ERs.
- **Capacity gap:** NIHSS needs expertise; neurologist shortage can lead to misdiagnoses.

Motivation for our new model AUSTIN:

- **Utilize triage labels:** Treat clinician triage as a complementary, informative signal.
- **Quantify uncertainty:** Calibrated uncertainty σ to flag ambiguous patient cases.
- **Learn from mismatches:** Use MRI–triage disagreements to boost robustness.

Method: AUSTIN

Data & Setting

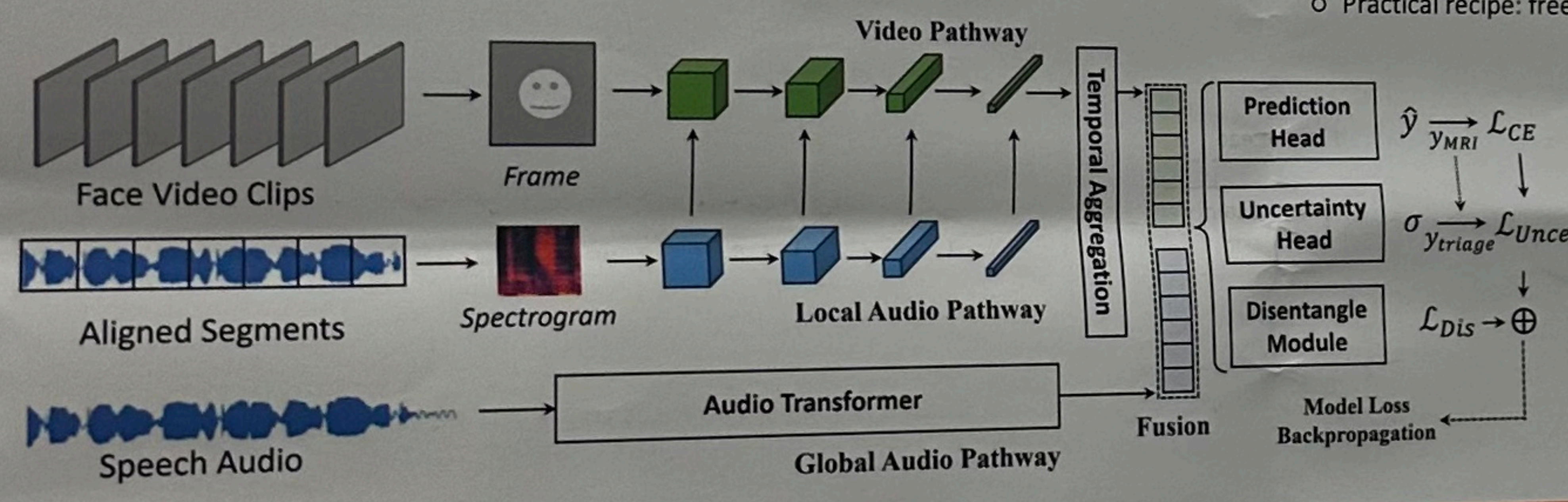
- **Cohort:** 249 ER patients (171+, 78-)
  - Diverse demographics
- **Temporal holdout:** 170/36/43 train/val/test.
- **Labels:**
  - MRI Label: binary groundtruth
  - Triage Label: ER clinician's impression label
- **Protocol:** "Cookie Theft" description video per NIHSS; paired audio recorded on mobile.

Model Framework

- **Multimedia encoder:**
  - Vido frame pathway: face frames → ResNet-50 (face-pretrained).
  - Local audio pathway: aligned spectrogram slices → ResNet-18, fused into frames.
  - Global audio pathway: full-utterance embedding → One-Peace transformer.
- **Temporal aggregation** with **S4** state-space layer; late fusion to final feature **h**.
- **Two heads:** *prediction* (stroke vs non-stroke) and *uncertainty* (σ).
- **Identity disentangling** via *adversarial discriminator* to reduce spurious cues.
- **Overall Training Objectives:**

$$\mathcal{L}_{\mathrm{Dis}} = \sum_{i,j} \|\delta_{ij} - \mathrm{D}(h_i, h_j)\|_2 , \qquad \mathcal{L}_{\mathrm{adv,E}} = -\sum_{i,j} \|0.5 - \mathrm{D}(h_i, h_j)\|_2 , \quad (2)$$

  - Penalize by learned uncertainty σ and MRI–triage agreement.
  - Practical recipe: freeze ResNet/One-Peace backbones; train S4, heads, and D.



Adaptive Uncertainty-Aware Loss

$$\mathcal{L}_{\mathrm{Uncert.}} = \frac{1}{2\sigma^2}\mathcal{L}_{\mathrm{CE}} + w \log(\sigma + \epsilon) , \quad w = \exp(-\alpha|y_{\mathrm{MRI}} - y_{\mathrm{triage}}|), \quad (1)$$

- **Labels Agree (w≈1):** Model is confident, σ is pushed lower.
- **Labels Disagree (w≈0): Model is uncertain, σ remains high.**
- The loss dynamically adapts to **clinician-MRI inconsistency.**
- Keep MRI as supervision but **learn to be uncertain** when MRI↔triage labels disagree.
  - σ both **regularizes training** (avoid overfitting to ambiguous cases) and **guides workflow** (route to imaging/expert).

Experiments & Results

Setup

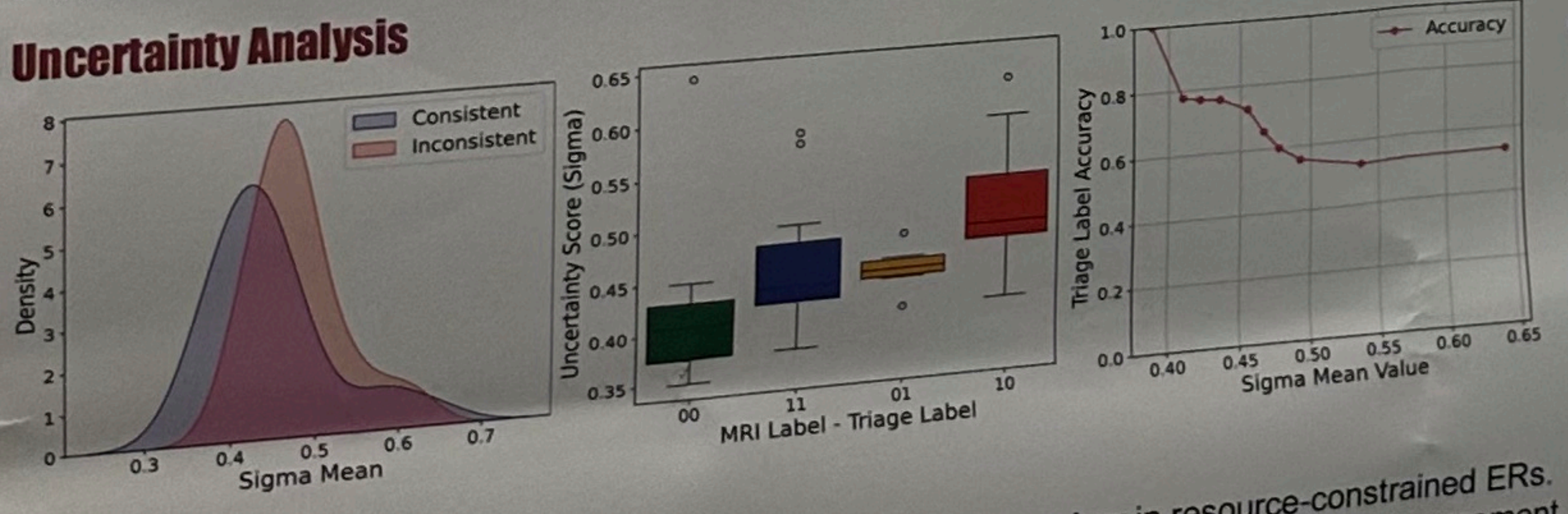
- **Preprocess:** PFLD face operations; N=7 clips, L=64 frames/clip; mel (M=128).
- **Backbones:** ResNet-50 (face video, FairFace-pretrained), ResNet-18 (local audio, ESC-50-pretrained), One-Peace (global audio).
- **Optimization:** batch size 32, lr 1e-4, dropout 0.2, 100 epochs on V100 (~6h).
- **Metrics:** Accuracy, Specificity, Sensitivity, **AUC**

Main Experiment

Model	Accuracy	Specificity	Sensitivity	AUC
Clinician Triage Performance	0.5349	0.5385	0.5333	-
DeepStroke	0.6977	0.6154	0.7333	0.6564
Proposed Encoder w/o $\mathcal{L}_{\mathrm{Uncert.}}$	0.6047	0.6923	0.7000	0.6658
+ $\mathcal{L}_{\mathrm{Uncert.}}$ w/ Fix $w=1$	0.6976	0.6154	0.7333	0.7128
+ Adaptive w (AUSTIN)	**0.7442**	**0.7692**	**0.7333**	**0.7897**

- Multimedia encoder and uncertainty formulation (fixed-weight) improves over DeepStroke;
- **Adaptive** weighting yields large AUC boost.

Uncertainty Analysis



- **Higher σ** on cases where MRI & triage **disagree** (distribution shift in KDE).
- Combination (MRI, triage)=(1,0) shows **highest median σ** → MRI-confirmed stroke with subtle/no outward signs.
- As σ-threshold increases, **triage-MRI consistency** drops → σ tracks case difficulty.
- Clinical use: σ can gate **further imaging or expert review.**

Ablations

Vision	Global Audio	Local Audio	Accuracy	Specificity	Sensitivity	AUC
ResNet50	One-Peace	ResNet18	0.6047	**0.6923**	0.7000	**0.6658**
ResNet50	AST	ResNet18	**0.6977**	0.6154	**0.7217**	0.6564
ResNet50	×	ResNet18	0.5584	0.4846	0.6333	0.5821
×	One-Peace	×	0.5814	0.4615	0.6333	0.5897
FaceXFormer	One-Peace	×	0.6279	0.6154	0.6333	0.6051
MARLIN	One-Peace	×	0.6279	0.5846	0.6667	0.6129

Clinical Implications

- **Triage safety:** High-σ cases are chosen for **MRI / expert review** in resource-constrained ERs.
- **Interpretability:** σ acts as an **intuitive confidence signal**, aligned to clinician–MRI agreement.
- **Generalization:** The adaptive uncertainty paradigm can extend to other tasks where multimodal inputs and diagnostic labels disagree.

Conclusion

- AUSTIN integrates **adaptive uncertainty** with multimedia encoders for ER stroke triage.
- It delivers **SoTA** performance and **actionable σ** for workflow decisions.
- Its added interpretability improves reliability in resource-constrained ERs.

Future Work

- **Prospective, multi-site validation** and larger cohorts; fairness & subgroup analyses.
- **Calibrated uncertainty** (temperature scaling, conformal) and per-segment σ.
- **Active/triage routing:** learn policies for MRI image vs. observe vs. discharge.

References

Cai, T., Ni, H., Yu, M., Huang, X., Wong, K., Volpi, J., Wang, J.Z., Wong, S.T.: Deepstroke: An efficient stroke screening framework for emergency rooms with multimodal adversarial deep learning. Medical Image Analysis 80, 102522 (2022)

Cai, T., Wong, K., Wang, J.Z., Huang, S.X, Yu, X., Volpi, J.J., Wong, S.T.: M³ stroke: Multi-modal mobile AI for emergency triage of mild to moderate acute strokes. In: IEEE EMBS International Conference on Biomedical and Health Informatics (BHI). pp. 1–8 (2024). https://doi.org/10.1109/BHI62660.2024. 10913652

This project was supported by the National Institute of Neurological Disorders and Stroke, National Institutes of Health, USA, under award number R01NS140292, the T.T. and W.F. Chao Foundation, and the John S. Dunn Foundation.