Clinically Reliable Deep Learning Model for Differentiation of Glioblastoma from Solitary Brain Metastasis : Empirical Estimation of Uncertainty with Deep Ensembles

PRESENTER: **Sujeong Eom, Seng Chan You**

INTRO

- Differentiating glioblastoma (GBM) from solitary brain metastasis (SBM) through magnetic resonance image (MRI) is important but difficult due to their **similar MRI features**.
- Deep learning–based model may be a supportive tool but there remains an issue of reliability as **overconfident** predictions are often made.

OBJECTIVE

- We aimed to address both **uncertainty** and **interpretability** in deep learning model to provide reliable prediction.

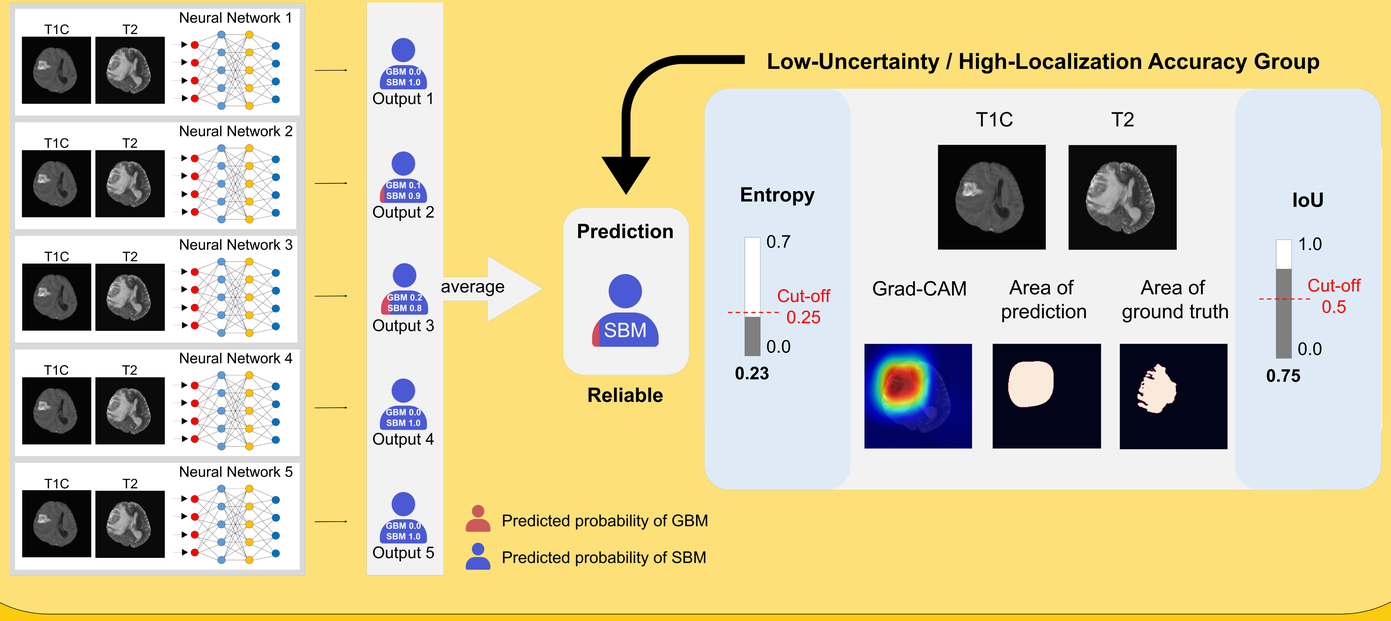
METHODS

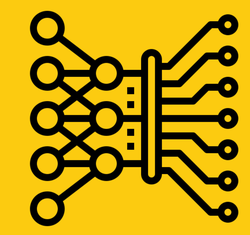
- MRIs from 469 (300 GBM, 169 SBM) and 143 (101 GBM, 42 SBM) patients are used for training and external validation. Out-of-distribution (OOD) dataset from unknown class, meningioma, is used.
- **Deep ensembles**, a simple method for **high-quality predictive uncertainty** with five parallel pre-trained neural networks is trained on MRIs. A single net is also trained for comparison.
- Diagnostic performance is evaluated with area under the curve (AUC). Uncertainty is measured by **entropy** value. Interpretability addressed as localization accuracy is calculated by **Intersection over Union (IoU)** of prediction area (Gradient-weighted Activation Mapping [Grad-CAM]) and ground truth area (segmentation mask).

Sujeong Eom[1,2], Seng Chan You[1,2], Yae Won Park[3], Seungwoo Kim[4], Sungbin Lim[4], Sung Soo Ahn[3], Seung-Ku Lee[3]

[1]Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, Korea
[2]Institute for Innovation in Digital Healthcare, Yonsei University, Seoul, Korea
[3]Department of Radiology and Research Institute of Radiological Science and Center for Clinical Imaging Data Science, Yonsei University College of Medicine, Seoul, Korea
[4]Artificial Intelligence Graduate School, UNIST, Ulsan, Korea

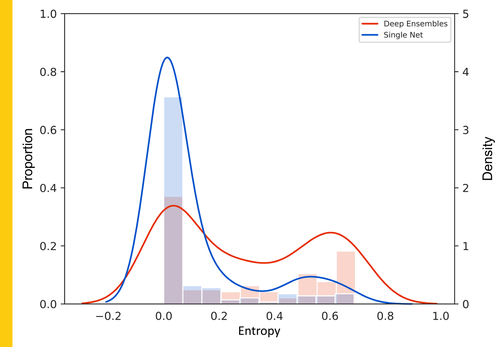
Reliable predictions of deep learning model by the estimation of **uncertainty** and **interpretability**



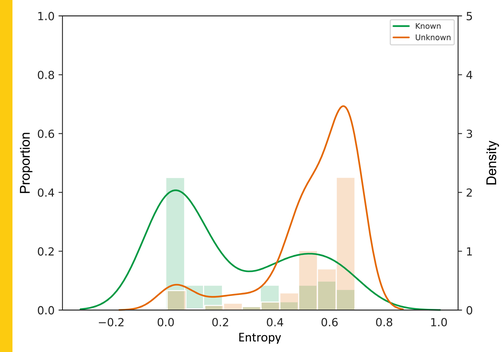


Take a picture to **download** the **GitHub Code**

Deep ensembles vs Single net uncertainty on external validation



Deep ensembles uncertainty on Known vs Unknown data



RESULTS

- Diagnostic performance of deep ensembles

Group	AUC	95% CI
External validation (n=143)	0.828	0.762-0.895
Low-uncertainty subgroup (n=71)	0.905	0.833-0.978
Low-uncertainty & High-localization accuracy subgroup (n=35)	**0.976**	0.928-1.000

*Uncertainty cut-off : 0.25
Localization accuracy cut-off : 0.5

- Proportion of data in uncertainty subgroup

Group	Known	Unknown
Low-uncertainty	62.0	9.3
High-uncertainty	38.0	**90.7**

*Uncertainty cut-off : 0.25

CONCLUSION

Our findings indicate empirical assessment of uncertainty and interpretability provides evidence for prediction's robustness for deep learning trained with relatively small dataset. For further study, we are planning on corporation with OHDSI medical imaging working group (WG), to apply the study's empirical assessment of reliability on the data with WG medical imaging extension.

YONSEI UNIVERSITY
UNIST ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY
OHDSI