Uncertainty Estimation in Medical Image Deep Learning
Jasmine Sun$^1$ and Laura Brattain, PhD$^2$

$^1$Washington University in St. Louis, $^2$University of Central Florida

**Introduction**

**Goal:**
- Identify key anatomical landmarks in the neck for emergency airway access guidance
- Develop a real-time deep learning bounding box detector for neck ultrasound data
- Ensure reliable AI is a priority for rapid and safe procedure guidance
- Quantify AI model uncertainty to flag overconfident detections for human attention

**Motivation:**
- AI models often overestimate their confidence
- Uncertainty estimation can help identify erroneous detection
  - Minimizes risk of misdiagnosis/mistreatment
  - Helps build trust in medical AI tools

**Materials & Methods**

**Dataset:**
- 6 subjects, 3,060 images
  - Neck ultrasound images from various angles around the neck to simulate airway assessment
  - 11,142 annotations
  - Bounding boxes and class labels generated by airway specialists using YOLOMark, an opensource annotation tool

**4 classes**
- Model trained on thyroid cartilage (471 samples, 4.2%), strap muscle (6,465 samples, 58.0%), tracheal ring (1,694 samples, 15.2%), and thyroid gland (2,512 samples, 22.5%)

**Computing Resources:**
- Utilized Newton Visualization GPU Cluster at UCF ARCC

**Object Detector:**
- YOLOv9c (You Only Look Once) [2]
  - Chosen for its speed and accuracy despite being an efficient, lightweight model
  - Train and fine-tune using 6-fold cross validation

**Uncertainty Estimator:**
- Gaussian Mixture Model based Uncertainty Estimation [3]
  - Extracts epistemic uncertainty from an object detector
  - During training: model uses an Anchor loss term to learn a structured logit space
  - After training: defines a set of Gaussian Mixture Models for each known class
  - During deployment: obtain epistemic uncertainty measurement by computing log likelihood that a detected object belongs to a class

**Initial Model:**
- YOLOv9c
- Image size: 640x640
- Optimizer: AdamW
- Learning rate: 0.001
- Epochs: 200
- Early stopping
- Batch size: 32
- Dropout: 0.1
- Freeze: 10
- Object loss gain: 0.5
- Average mAP50: 0.788

**Finetuning hyperparameters:**
- Built off initial model
- Class loss gain: 1.0
- Object loss gain: 0.5
- Dual focal loss: 2.0
- Implemented data augmentation:
  - Scaling: 0.2
  - Shear: 0.2
  - Left-right flip: 0.2
  - Mosaic: 0.5
  - Mixup: 0.5

**Addressing class imbalance (final model):**
- Strap muscle makes up 58% of samples whereas thyroid cartilage makes up only 4.2% of samples
- Built off fine-tuned model
- Class weights used: [5.914, 0.431, 1.644, 1.109] for thyroid cartilage, strap muscle, tracheal ring, and thyroid gland, respectively
  - Calculated by: $w_i = \frac{N_i}{\sum_j N_j}$

**Experimental Results**

**Goal:**
- Improve classification, decrease overfitting, predict accurate bounding boxes, and improve AI model robustness from different ultrasound angles

**Uncertainty Estimation Results**

**Right:**
- Detections of tracheal ring w/ 0.30 confidence, 0.76 uncertainty and thyroid gland w/ 0.50 confidence, 1.16 uncertainty
- Flag for human attention

**Below:**
- Red lines indicate need for human attention (low confidence/high uncertainty)

**Conclusion**

- Developed real-time neck landmark detection AI models
- Finetuned YOLO parameters such as loss, data augmentation, and implemented class weight-base loss function
- Improved area under curve (AUC) and bounding box detections and classifications across classes
- Decreased amounts of missed detections and false positives
- Implemented GMM-based real-time uncertainty estimations to catch overconfident model predictions

**Future Work**

- Compare the GMM-based uncertainty estimator with other methods
- Compare YOLO performance with other object detectors
- Implement model in ultrasound device for real-time detection and uncertainty for use in clinical settings

**References**


**Acknowledgements**

- Dr. Mubarak Shah, Dr. Niels da Vitoria Lobo, and Melody Halbert at the UCF CRCV
- Dr. Greg Dion at the University of Cincinnati