Uncertainty-Aware Deep Learning System

DEDUCE is programmed with training sets of selected category images

CATEGORY: CAR

CATEGORY: TRUCK

DEDUCE assigns categories and confidence scores to new images it is requested to classify

CATEGORY:
CAR

CONFIDENCE SCORE:
90%

CATEGORY:
TRUCK

CONFIDENCE SCORE:
92%

UNCERTAIN

DEDUCE is unique among AI because it provides a principled way of obtaining uncertainty in predictions

MIT Lincoln Laboratory has developed DEDUCE (DEep Decision UnCertainty)—a novel method for training deep neural networks (DNN) to estimate how confident they are when making predictions. DEDUCE captures uncertainty inside and outside of training examples, enabling the DNN to detect with high confidence anomalous data and adversarial AI attacks (attempts to modify input data to break the AI). Furthermore, when the AI has insufficient evidence for a prediction and is likely to make an error, DEDUCE outputs high predictive uncertainty, allowing users to detect and potentially avoid errors. DEDUCE takes an important step in increasing user trust that DNNs can produce actionable insights.

KEY FEATURES

- Accurately estimates uncertainty in the predictions while maintaining high prediction accuracy

- Is computationally efficient

- Does not require a change of the network's architecture

- Does not rely on complex data augmentation, prior knowledge of anomalous data, or heuristics

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Deep neural networks (DNNs) are powerful artificial intelligence (AI) techniques that learn complex patterns from data and are used to make predictions. Such techniques have proven successful for large-scale object recognition and machine translation. However, a DNN has a major flaw – the network cannot let you know when it has low confidence in its predictions. Users may lose trust in these techniques if high-cost errors are not easy to predict. Therefore, capturing uncertainty is the key for DNNs to gain users' trust and increase the effectiveness of AI systems.

DEDUCE outperforms state-of-the-art DNNs by minimizing information captured by the DNN for decision outcomes that it cannot explain in a nontrivial manner and by providing robust predictive uncertainty estimates.

Applications

AUTONOMOUS VEHICLES

In autonomous vehicle applications, AI is enabling cars to navigate through traffic and handle various interactions with pedestrians, traffic lights and signs, and other vehicles. Because uncertainty propagates through its prediction and decision layers, our DNN can increase safety by accounting for the model's ignorance about the world and thus indicating low confidence when the AI system encounters distant or unfamiliar objects, or is operated under different weather conditions. Uncertainty in predictions can alert users to erroneous behavior (e.g., crashes) and improve controllers. We successfully demonstrated DEDUCE on benchmark image classification tasks for which it detected errors, anomalous data, and adversarial examples with high confidence.

MEDICAL DIAGNOSTICS

AI is being used for medical diagnostics, including recognizing cancerous tissue (oncology), diagnosing diseases on the basis of lab results and imagery (pathology), and even discovering new drugs or determining disease epidemiology. As healthcare systems continue to evolve, AI is becoming more prevalent. However, an overconfident wrong diagnosis can lead to complications or bad clinical outcomes. Predictive uncertainty in the AI supporting diagnosis is key to trusted detection of challenging or rare patient cases. Our method has shown improved confidence levels on a heart condition diagnosis that was based on electrocardiogram signals and on shock prediction for trauma patients that was based on vital signs and injury patterns. Our DNN detected challenging and ambiguous cases by ranking them on uncertainty.

PATENT PENDING
20210103814

More Information

T. Tsiligkaridis, "Information Aware Max-Norm Dirichlet Networks for Predictive Uncertainty Estimation," *Neural Networks*, vol. 135, 2021, pp. 105–114.

INTERESTED IN WORKING WITH MIT LINCOLN LABORATORY?

Scan the QR code to learn more
www.ll.mit.edu/partner-us

| Contact the Technology Transfer Office
tto@ll.mit.edu