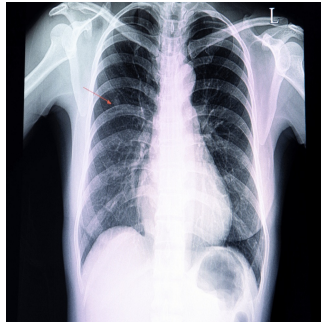

Pulmonary Nodule Risk Estimation With Deep Learning and Uncertainty Analysis



Lung cancer remains a significant cause of mortality, and early detection is essential to improving patient outcomes. The advent of lung cancer screening trials like the National Lung Screening Trial (NLST) and the Danish Lung Cancer Screening Trial (DLCST) has highlighted the potential of low-dose computed tomography (CT) scans in detecting lung nodules at earlier stages, which could significantly reduce mortality rates. However, the increasing demand for CT screenings has outpaced the availability of radiologists, leading to diagnosis delays. Deep Learning (DL) models have emerged as a solution to this challenge by providing automated assistance in interpreting medical images, specifically for pulmonary nodule malignancy risk estimation. While DL models have shown promise, their lack of uncertainty estimation can hinder their effectiveness in clinical practice, where precise and reliable predictions are crucial. A recent article in *European Radiology* explores the integration of uncertainty estimation into a previously developed DL model for pulmonary nodule malignancy risk estimation, assessing its performance across various nodule characteristics and evaluating its potential for real-world clinical application.

The Role of Uncertainty in Deep Learning

One of the critical limitations of current DL models is their inability to communicate uncertainty in their predictions. This is particularly important in medical applications, where decisions about patient care often hinge on accurate and reliable assessments. A DL model that can estimate its own uncertainty provides clinicians with an additional layer of information, helping them identify cases where further human evaluation is necessary. In this study, uncertainty estimation was integrated into a pre-existing DL model, which had already demonstrated high performance in estimating the malignancy risk of pulmonary nodules.

The researchers applied this uncertainty estimation to two datasets: the development set from the DLCST, which consists of lung cancer screening data, and a clinical dataset from a tertiary academic centre. The uncertainty was calculated using an entropy-based method, which measures the randomness in the model's predictions. High entropy indicates uncertainty, while low entropy suggests the model is confident in its classification. Two thresholds were set based on the 90th and 95th percentiles of the uncertainty distribution, dividing the nodules into certain and uncertain groups.

Performance of the Model in Uncertain Cases

The study found a significant drop in the model's performance when analysing uncertain cases. When the model was applied to the certain cases (those with low entropy), it achieved an area under the receiver operating characteristic curve (AUC) of 0.93 in the DLCST dataset, indicating excellent performance in identifying malignant nodules. However, the model's AUC dropped to 0.62 in the uncertain group, highlighting the challenges in accurately classifying these cases. This pattern was also observed when the model was validated on the external clinical dataset, where the AUC dropped from 0.90 for certain cases to 0.62 for uncertain cases.

Notably, the uncertain cases included larger benign nodules as well as part-solid and non-solid nodules, which are more challenging to classify compared to solid nodules. This suggests that nodule characteristics, such as size and composition, contribute significantly to the model's uncertainty. These findings underscore the importance of integrating uncertainty estimation into DL models, as it allows for a clearer understanding of when the model is likely to make errors, providing clinicians with a valuable tool to identify cases that require additional human review.

Implications for Clinical Practice

Integrating uncertainty estimation into DL models has profound implications for clinical practice. In a real-world setting, clinicians are more likely to trust the model's predictions when they are accompanied by a certainty score. In cases where the model is uncertain, clinicians can be alerted to the need for further review, thereby reducing the risk of misdiagnosis. This is particularly important in settings where radiologists are in short supply, as it allows the DL model to handle routine cases while flagging more complex cases for human evaluation.

The study also highlights the importance of validating DL models on external datasets. The significant difference in uncertainty levels between the screening data from the DLCST and the clinical dataset underscores the need to assess models on diverse datasets that better reflect the

variability encountered in clinical practice. This approach ensures the model is robust enough to handle the wide range of nodule types and sizes it will encounter in a real-world setting.

Furthermore, the study suggests that uncertainty estimation could be instrumental in optimising clinical workflows. By referring uncertain cases to human experts, the model can help prioritise cases that require immediate attention, potentially improving patient outcomes by accelerating the diagnostic process. However, the authors caution that the optimal uncertainty threshold may vary depending on the clinical setting and should be determined through multidisciplinary discussions involving both clinicians and data scientists.

Conclusion

The integration of uncertainty estimation into DL models represents a significant step forward in improving the safety and reliability of automated pulmonary nodule malignancy risk estimation. By identifying cases where the model is uncertain, this approach provides clinicians with a vital tool for making more informed decisions, reducing the risk of errors, and enhancing the overall trustworthiness of DL models in medical applications.

The findings of this study suggest that DL models with uncertainty estimation can perform well in some instances but struggle with more complex cases, such as larger benign nodules and subsolid nodules. This highlights the importance of continuously refining DL algorithms and expanding training datasets to improve performance across a broader range of nodule types. As the use of CT screenings continues to rise and lung cancer screening programmes become more widespread, the need for reliable and efficient DL models will only grow. Uncertainty estimation will play a crucial role in ensuring that these models can be safely and effectively integrated into clinical practice, allowing them to assist in reducing the workload of radiologists while maintaining high standards of patient care.

Further research should focus on expanding the application of uncertainty estimation across different imaging modalities and cancer types and exploring its impact on human readers through clinical studies. By combining the strengths of DL models and human expertise, we can move closer to a future where AI-driven diagnostics become a seamless and trusted part of medical practice.

Source Credit: [European Radiology](#)

Image Credit: [iStock](#)

Published on : Mon, 16 Sep 2024