Deep Learning for Lung Cancer: A Comprehensive Systematic Review of Tasks, Modalities, and Architectures

Manoj Dhiman
Assistant Professor, Department of Computer Science & Informatics, Central University of Himachal Pradesh (CUHP), University in Dadroli, Himachal Pradesh

Dr. Pradeep Chouksey
Department of Computer Science & Informatics, Central University of Himachal Pradesh (CUHP), University in Dadroli, Himachal Pradesh

Abstract:

Background: Deep learning (DL) is now the dominant computational approach to lung-cancer imaging and pathology, but the field's true methodological breadth, validation practices, and translational readiness have never been mapped at scale. Narrow, task-specific reviews dominate the secondary literature. In a companion benchmarking exercise, no existing review covered more than about half of a comprehensive 78-dimension framework.

Methods: Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guideline, we searched seven databases (ScienceDirect, Scopus, Web of Science, SpringerLink, IEEE Xplore, ACM Digital Library, PubMed), retrieving 10,615 records. After de-duplication and title/abstract/keyword screening, 3,046 primary articles entered full-text assessment. Studies were included when lung cancer was a target, DL was the principal method, methodological novelty was present, and quantitative results were reported. Each included study was coded against 78 prespecified dimensions in eleven thematic groups. Risk of bias was assessed with QUADAS-2 and PROBAST.

Results: 1,223 studies met all criteria. Output grew steeply from 2017 to 2025, with a clear shift from convolutional networks (the principal family in 83% of studies) to transformers (15%) and, from 2024, foundation, generative, and state-space models (collectively 3.4%). Diagnostic CT was used in 68% of studies and nodule classification was the commonest task. Coverage was strikingly uneven: omics (7%), multimodal fusion (8%), pathology grading (2%), and Mamba/state-space models (0.3%) were rare. Reported performance was near-ceiling for detection and classification (median area under the receiver-operating-characteristic curve [AUC] approximately 0.96) but markedly lower and more variable for prognostic and biomarker tasks (median AUC approximately 0.82-0.85). External validation was present in only 18% of studies, code in 7%, and a random seed in fewer than 0.3% (three studies). Prospective validation (0.3%), regulator-cleared-tool reference (0.2%), and cost-effectiveness analysis (0.8%) were almost absent. Overall risk of bias was low in 15%, of some concern in 59%, high in 25%, and unclear in the remaining 1%. First authors were concentrated in China (40%) and India (20%).

Conclusions: DL for lung cancer is large, fast-growing, but methodologically narrow, anchored on convolutional analysis of a few public CT benchmarks. The dominant barriers to clinical translation are

weak external and prospective validation, near-absent regulatory engagement, and poor reproducibility. Routine reporting of code and random seeds is the cheapest field-wide remedy. We provide a reusable 78-dimension taxonomy, with the fully extracted dataset available on reasonable request to support future synthesis.

Keywords:
Deep learning, lung cancer, systematic review, convolutional neural networks, vision transformers.