International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 11, Issue 2, 2023

Enhancing Diagnostic Accuracy In Medical Imaging Through Computer Vision Methods

S Ramesh[1], Dr. Dushyant Sharma[2], Dr. Anurag Pandey[3]

[1]Research Scholar, Department of Computer Science and Engineering, CCS University, Meerut

[2]Research Supervisor, Department of Computer Science and Engineering, CCS University, Meerut

[3]Research Co-Supervisor, Department of Computer Science and Engineering, CCS University, Meerut

Abstract

The integration of computer vision methods into medical imaging has marked a paradigm shift in diagnostic medicine. This paper investigates how deep learning-based computer vision techniques principally convolutional neural networks (CNNs) enhance diagnostic accuracy across key imaging modalities, including chest X-ray (CXR), computed tomography (CT), magnetic resonance imaging (MRI), and retinal fundus photography. Objectives include evaluating quantitative performance gains offered by CNN architectures and benchmarking algorithmic accuracy against radiologist performance. A systematic secondary review methodology was adopted, aggregating peer-reviewed studies reporting sensitivity, specificity, and area under the receiver operating characteristic curve (AUC). Results drawn from published data up to 2022 indicate that deep learning models achieve pooled sensitivity of 0.95, specificity of 0.97, and AUC of 0.97 across ophthalmological imaging. DenseNet-121 achieved 98.7% accuracy for diabetic retinopathy grading. These findings confirm that computer vision-based diagnostic systems represent a clinically viable complement to traditional radiology, significantly improving early detection rates while reducing diagnostic error.

Keywords: *Computer vision, Convolutional neural network, Medical image analysis, Diagnostic accuracy, Transfer learning*

1. Introduction

Medical imaging constitutes the backbone of modern clinical diagnosis, enabling clinicians to visualize internal anatomical structures and detect pathological abnormalities non-invasively. However, the global volume of medical images has expanded at an extraordinary rate, creating a substantial burden on radiological workforces and increasing the risk of diagnostic errors attributable to reader fatigue and inter-observer variability (Zhou et al., 2021). Computer vision, particularly deep learning operationalized through convolutional neural networks (CNNs), has emerged as a transformative solution to these systemic challenges. The history of AI in radiology traces to computer-aided detection systems in the 1990s, which relied on handcrafted features and classical machine learning. These early systems were constrained by limited generalizability and required extensive domain-specific feature engineering (Chan et al., 2020). The 2012 ImageNet challenge catalyzed a broader shift toward deep learning, as CNN-based models dramatically outperformed traditional methods on large-scale visual recognition tasks. By 2015–2016, transfer learning techniques enabled the medical imaging community to leverage pre-trained models from non-medical datasets and fine-tune them for clinical applications, substantially accelerating deployment timelines (Liu et al., 2021). Landmark studies on diabetic retinopathy detection soon demonstrated specialist-level performance, establishing the feasibility of algorithmic diagnostic systems in ophthalmology (Liu et al., 2019).

CNNs differ fundamentally from earlier approaches in their capacity to learn hierarchical, task-relevant representations directly from raw pixel data without manual feature extraction (Bhatt et al., 2021). Architectures including VGG, ResNet, InceptionNet, and DenseNet have demonstrated remarkable versatility across imaging modalities chest X-rays, CT scans, MRI sequences, and retinal fundus photographs. Transfer learning has further accelerated progress, allowing models pre-trained on large general datasets to be fine-tuned for specific clinical purposes even when annotated medical data is limited. The clinical stakes of accurate medical image interpretation are profound. Missed pulmonary nodules on CT scans, undetected diabetic retinopathy, or misclassified brain tumors on MRI can result in delayed treatment with life-altering consequences for patients. A growing body of evidence, including large-scale meta-analyses spanning over 500 peer-reviewed studies, confirms that deep learning models frequently match or exceed human expert performance on well-defined diagnostic tasks (Soffer et al., 2021). In contexts such as India, where specialist radiologist density remains critically low relative to population size, AI-assisted diagnostic tools hold unique promise for expanding diagnostic access and improving healthcare equity.

This paper synthesizes quantitative performance data from published benchmarks to assess how computer vision methods enhance diagnostic accuracy in medical imaging, to compare algorithmic versus radiologist performance across modalities, and to identify pathways toward responsible clinical deployment of these technologies in resource-constrained healthcare environments.

2. Literature Review

The application of computer vision to medical imaging has generated a rich body of literature over the past decade. Huang et al. (2021) conducted a comprehensive examination of CNN algorithms for object detection and semantic segmentation in oncology imaging, demonstrating that DeepLab v3+-based systems achieved internal validation accuracy of 95.5% for upper gastrointestinal cancer detection in endoscopic images. Their work underscored the centrality of semantic segmentation for precise tumor margin delineation a task critical for surgical planning and radiation therapy targeting. In the domain of lesion-level object detection, He et al. (2020) advanced the Mask R-CNN framework, which introduced pixel-level instance segmentation to complement bounding box detection, enabling more granular pathological localization in CT and MRI volumes. This framework was extended by Dogan et al. (2021), who applied Mask R-CNN in conjunction with 3D U-Net to achieve high-precision pancreatic segmentation in CT imaging, demonstrating that pipeline integration of complementary deep learning architectures yields superior boundary delineation compared to single-model approaches. For brain tumor segmentation specifically, Guan et al. (2022) proposed the 3D AGSE-VNet framework an attention-guided squeeze-and-excitation volumetric network achieving competitive Dice similarity coefficients on standard MRI benchmark datasets.

Regarding regulatory and adoption landscapes, Benjamens et al. (2020) documented over 64 FDA-approved AI-based medical devices and algorithms, with the majority concentrated in radiology and cardiology. This signals that clinical validation of deep learning systems had reached sufficient maturity for regulatory acceptance, even as generalizability and dataset bias remained active research concerns. In chest radiology, Yoo et al.

(2021) conducted a multi-reader study using the National Lung Screening Trial (NLST) dataset, demonstrating that AI assistance increased radiology residents' sensitivity for visible lung cancer detection and simultaneously reduced unnecessary CT recommendations by specialist radiologists for cancer-negative cases, thus optimizing both detection yield and resource utilization. Wang (2022) further documented that DenseNet-201 achieved 97.0% accuracy, 96.2% sensitivity, and an AUC of 0.968 on CT datasets for pulmonary nodule classification, reinforcing the dominance of dense skip-connection architectures for this task.

For diabetic retinopathy (DR), Farag et al. (2022) implemented DenseNet combined with a convolutional block attention module, reporting classification metrics above 98% on the APTOS 2019 dataset. Maqsood et al. (2021) proposed a 3D CNN for retinal hemorrhage localization achieving 97.71% average accuracy across 1,509 fundus images from multiple databases, confirming robustness across dataset heterogeneity. Jang et al. (2021) demonstrated deep learning-based gastric carcinoma subclassification from whole-slide histopathology images, reflecting the cross-organ scalability of CNN architectures and reinforcing the potential of these methods beyond canonical imaging domains.

3. Objectives

1. To evaluate the quantitative diagnostic performance of CNN-based computer vision models measured by accuracy, sensitivity, specificity, and AUC across medical imaging modalities including CXR, CT, MRI, and fundus photography using data from studies published up to 2022.
2. To compare the diagnostic accuracy of deep learning algorithms against radiologist benchmarks and assess the clinical applicability of computer vision methods for enhancing early disease detection in resource-limited healthcare environments such as India.

4. Methodology

This study adopts a systematic secondary review design, aggregating quantitative performance data from peer-reviewed studies published between 2018 and 2022. The review followed Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) principles to ensure transparency and reproducibility. Databases searched included PubMed, Scopus, IEEE Xplore, and Google Scholar, using search terms such as "convolutional neural network," "medical image classification," "diagnostic accuracy," "sensitivity," "specificity," "AUC," "chest X-ray," "CT scan," "MRI," "diabetic retinopathy," and "deep learning." Peer-reviewed articles reporting quantitative diagnostic accuracy metrics accuracy, sensitivity, specificity, AUC, F1-score for deep learning models applied to medical imaging were included. Studies using standard benchmark datasets such as LIDC-IDRI, LUNA16, APTOS 2019, Kaggle DR, Messidor-2, DIARETDB, DRIVE, and NLST were prioritized to ensure comparability and reproducibility across tabulated findings.

A total of 503 studies were included in the foundational meta-analysis upon which pooled statistical data are drawn, supplemented by performance data from fourteen additional primary benchmark studies covering lung, retinal, and histopathological imaging tasks. The primary analytical tools referenced across included studies are CNN architectures ResNet-50, VGG-19, InceptionV3, DenseNet-121/169/201, MobileNet-V2, and EfficientNet-B0 implemented in TensorFlow and PyTorch frameworks. Performance metrics were extracted and tabulated directly from published results. Statistical pooling was conducted

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 11, Issue 2, 2023

via random-effects meta-analysis models. Data augmentation, transfer learning, and hyperparameter tuning were dominant methodological features across included primary studies. All data presented pertain exclusively to validation or test sets reported in peer-reviewed sources with confidence intervals where available.

5. Results

Table 1: Pooled Deep Learning Diagnostic Accuracy Across Medical Imaging Specialties

Specialty	Studies (n)	Patient Cohorts (n)	Pooled Sensitivity (95% CI)	Pooled Specificity (95% CI)	Pooled AUC (95% CI)
Ophthalmology	82	94	0.95 (0.91–0.97)	0.97 (0.93–0.99)	0.97 (0.96–0.99)
Breast Disease	82	100	0.87 (0.80–0.94)	0.89 (0.85–0.93)	0.90 (0.87–0.94)
Respiratory Disease	115	142	0.87 (0.82–0.92)	0.88 (0.84–0.92)	0.91 (0.88–0.95)
Overall (all specialties)	503	—	0.95 (0.91–0.97)	0.97 (0.93–0.99)	0.97 (0.96–0.99)

Source: Soffer et al. (2021)

As demonstrated in Table 1, deep learning models across 503 peer-reviewed studies show consistently high pooled diagnostic performance. Ophthalmology achieved the highest pooled AUC of 0.97 (95% CI: 0.96–0.99), driven by diabetic retinopathy detection on retinal fundus images. Respiratory disease studies encompassing 115 trials across 142 patient cohorts returned a pooled AUC of 0.91, indicating strong but slightly lower performance, attributable to greater imaging heterogeneity across CT and CXR modalities. Breast imaging AUC ranged between 0.868 and 0.909 across mammography and MRI modalities.

Table 2: Deep Learning Performance on Chest X-Ray for Lung Pathology Detection

Pathology	Modality	AUC (95% CI)	Sensitivity	Specificity
Abnormal CXR	CXR	0.917 (0.869–0.966)	0.873	0.894
Pneumothorax	CXR	0.910 (0.863–0.957)	0.718	0.918
Lung Cancer/Mass	CXR	0.864 (0.827–0.901)	0.873	0.820
Lung Nodules	CT Scan	0.887 (0.847–0.928)	0.860	0.775

Source: Soffer et al. (2021)

Table 2 presents stratified diagnostic performance for chest pathology detection derived from pooled meta-analytic data. Pneumothorax detection on CXR yielded the highest specificity (0.918), indicating a reliable tool with low false-positive burden. Lung nodule detection via CT returned an AUC of 0.887, though with reduced specificity (0.775), reflecting the inherent challenge of differentiating benign from malignant nodules within volumetric CT data. Abnormal CXR detection reached an AUC of 0.917 with sensitivity of 0.873, confirming DL's utility as a triage tool in chest radiology workflows.

Table 3: CNN Architecture Comparison for Diabetic Retinopathy (DR) Detection

Model	Dataset	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
InceptionV3	APTOS 2019	96.18	95.40	97.20	0.971
ResNet-50 + RF	Messidor-2	96.00	94.10	97.30	0.963
VGG-19 (VGGNet-s)	Kaggle DR	95.68	93.20	96.80	0.952
DenseNet-121	APTOS 2019	98.70	97.80	99.10	0.986
MobileNet-V2	APTOS 2019	93.09	91.50	94.80	0.941

Source: Lahmar & Idri (2022); Yaqoob et al. (2021)

Table 3 documents CNN model performance for DR detection across standard ophthalmological benchmark datasets. DenseNet-121 demonstrated the highest accuracy at 98.70% with an AUC of 0.986, followed by InceptionV3 at 96.18%. ResNet-50 combined with Random Forest classification on Messidor-2 achieved 96.00% accuracy, validating the hybrid feature extraction-classification approach. MobileNet-V2, while returning lower accuracy (93.09%), offers computational efficiency critical for deployment on mobile diagnostic devices in low-resource Indian healthcare settings, illustrating the accuracy-efficiency tradeoff that governs architecture selection in practice.

Table 4: AI vs. Radiologist Performance: Lung Nodule Detection on CT Scan

Metric	AI Models (Range)	Radiologists (Range)	Difference
Sensitivity (%)	86.0–98.1	68.0–76.0	AI +10 to +22%
Specificity (%)	77.5–87.0	87.0–91.7	AI −4 to 0%
Accuracy (%)	64.96–92.46	73.31–85.57	Variable
AUC (Range)	0.864–0.937	Not routinely reported	AI Superior

Source: Wang (2022); Park et al. (2020)

Table 4 presents a direct comparative analysis of AI versus radiologist performance for lung nodule detection on CT scans. AI models achieved markedly superior sensitivity (86.0–98.1% versus 68–76% for radiologists), meaning that deep learning algorithms are substantially less likely to miss malignant nodules. However, lower AI specificity (77.5–87.0%) compared to radiologists (87–91.7%) reveals a higher false-positive tendency. These findings support AI deployment as a high-sensitivity first-line screener complemented by radiologist confirmation rather than as an autonomous replacement diagnostic system.

Table 5: Deep Learning Performance for Lung Cancer Detection on CT Scan

Model	Dataset	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
CNN (PSO-optimized)	LIDC-IDRI (17,870 samples)	96.54	87.79	98.22	0.931
DenseNet-201	LUNA16/CT	97.00	96.20	97.50	0.968
DL Algorithm (CXR)	NLST Dataset	—	100.00	97.00	0.990
VGG19 + CNN	Histological CT	—	97.50	98.30	0.982

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 11, Issue 2, 2023

Source: Wang (2022); Park et al. (2020); Yoo et al. (2021)

Table 5 documents deep learning performance specifically for lung cancer detection. DenseNet-201 achieved the highest balanced performance with 97.0% accuracy and AUC of 0.968 on CT datasets. The DL algorithm evaluated on the NLST chest radiograph dataset by Park et al. (2020) achieved perfect sensitivity (100%) for clearly visible lung cancers with 97% specificity and AUC of 0.990. The PSO-optimized CNN on the 17,870-sample LIDC-IDRI dataset achieved 96.54% accuracy and 0.931 AUC, confirming architectural optimization value on heterogeneous, real-world nodule imaging data.

Table 6: Transfer Learning Model Performance Across Medical Imaging Tasks (2020–2022)

Transfer Learning Model	Task	Dataset	Accuracy (%)	AUC
Inception-ResNet-V2	DR Detection	EyePACS (>90,000 images)	~96.00	0.986
DenseNet-169	DR Grading	APTOS 2019	100.00	1.000
VGG-16	Skin Cancer Detection	ISIC/Dermoscopy	94.50	0.940
InceptionV3	Gastric Carcinoma	Whole-Slide Histopathology	92.30	0.921
ResNet-50 + RF	Diabetic Macular Edema	Messidor-2	96.00	0.963

Source: Lahmar & Idri (2022); Fraiwan & Faouri (2022); Jang et al. (2021); Yaqoob et al. (2021)

Table 6 consolidates transfer learning performance across diverse imaging tasks from 2020 to 2022. The DenseNet-169 model achieved perfect accuracy of 100% for DR grading on APTOS 2019 through ensemble integration with DenseNet-201, enabled by data augmentation and global average pooling. Inception-ResNet-V2, trained on over 90,000 EyePACS images, returned AUC of 0.986 with sensitivity of 0.958. VGG-16 applied to dermoscopy-based skin cancer detection achieved 94.5% accuracy, while InceptionV3 for gastric carcinoma subclassification from whole-slide images returned 92.30% accuracy, demonstrating broad cross-domain architectural transferability.

6. Discussion

The results synthesized in this paper reinforce a central finding emerging from the global medical imaging literature: deep learning-based computer vision methods substantially enhance diagnostic accuracy across multiple imaging modalities, with performance frequently matching and in specific domains exceeding that of trained radiologists. The pooled meta-analytic evidence reported by Soffer et al. (2021), encompassing 503 peer-reviewed studies, establishes a clear statistical foundation pooled sensitivity of 0.95 and specificity of 0.97 across ophthalmological imaging represents performance levels historically unattained by conventional diagnostic technologies. The superiority of AI models in sensitivity over radiologists for lung nodule detection (86.0–98.1% versus 68–76%) is a clinically consequential finding, as lung cancer outcomes are profoundly stage-dependent. Early detection facilitated by high-sensitivity screening translates directly into improved survival rates (Wang, 2022). Simultaneously, lower AI specificity relative to radiologists (77.5–87.0% versus 87–91.7%) suggests a tendency toward false positives which, if unaddressed, risks overdiagnosis, unnecessary patient anxiety, and procedural harm. This finding precisely aligns with the recommendations of Yoo et al. (2021), whose NLST multi-reader study identified AI as most effective when deployed as a second reader improving resident sensitivity for visible lung cancers while reducing unnecessary CT recommendations by specialist radiologists, thus simultaneously enhancing detection and resource efficiency.

The diabetic retinopathy findings are equally instructive. DenseNet-121's 98.7% accuracy and 0.986 AUC on the APTOS dataset, as documented by Farag et al. (2022), reflect the maturity of CNN-based ophthalmic screening. In India, where ophthalmologist-to-patient ratios remain severely inadequate in rural and semi-urban regions, such models can serve as front-line screening tools to prioritize specialist referrals, potentially preventing thousands of preventable blindness cases annually. The 3D CNN proposed by Maqsood et al. (2021) for retinal hemorrhage localization, achieving 97.71% accuracy across 1,509 multi-database fundus images, further confirms the robustness of deep learning methods across dataset heterogeneity a property essential for real-world deployment across diverse clinical settings. Transfer learning emerges consistently as the dominant methodological approach enabling high performance particularly in data-scarce medical imaging domains (Chan et al., 2020). The Inception-ResNet-V2 model's performance at AUC 0.986 on over 90,000 EyePACS images highlights how population-scale training produces generalizable diagnostic representations (Lahmar & Idri, 2022). The cross-organ scalability demonstrated by InceptionV3 for gastric carcinoma classification from whole-slide histopathology images (Jang et al., 2021) and VGG-16 for skin cancer detection from dermoscopy data (Fraiwan & Faouri, 2022) further signals that foundational representational learning in deep CNNs transfers meaningfully even across pathologically distinct domains.

From an architectural standpoint, DenseNet variants consistently outperformed ResNet and VGG counterparts in accuracy and AUC. This aligns with DenseNet's dense skip connections, which facilitate feature reuse and gradient propagation directly addressing the vanishing gradient problem in deep networks (Bhatt et al., 2021). The segmentation advances documented by Huang et al. (2021), He et al. (2020), and Dogan et al. (2021) further expand computer vision's clinical utility beyond classification toward precise lesion delineation for treatment planning, indicating that the field is progressing from detection toward comprehensive diagnostic characterization. Regulatory progress, as evidenced by the 64-plus FDA-approved AI medical algorithms documented by Benjamens et al. (2020), signals growing institutional confidence in these technologies. However, Liu et al. (2019) and Zhou et al. (2021) consistently caution that heterogeneity in methodology, dataset composition, and outcome reporting risks overestimation of real-world performance underscoring the urgent need for AI-specific reporting guidelines akin to the STARD standards for diagnostic accuracy research. Addressing these limitations through prospective, multicenter clinical trials will be essential before widespread autonomous deployment is warranted.

7. Conclusion

This paper demonstrates that computer vision methods, particularly CNN-based deep learning architectures, meaningfully enhance diagnostic accuracy across chest radiology, oncology, ophthalmology, and histopathology. Pooled evidence confirms that deep learning achieves AUC of 0.97 in ophthalmological imaging, with individual models such as DenseNet-121 reaching 98.7% accuracy for diabetic retinopathy grading. AI consistently outperforms radiologists in sensitivity for lung nodule detection, while specificity gaps indicate a complementary rather than replacement deployment paradigm is most clinically appropriate. Transfer learning and dense architectural designs are primary contributors to high performance across modalities. Future research should prioritize prospective multicenter validation, AI-specific reporting standards, and deployment frameworks tailored to resource-constrained healthcare systems, ensuring equitable translation of these diagnostic advances across India and comparable healthcare environments.

References

1 Benjamens, S., Dhunnoo, P., & Meskó, B. (2020). The state of artificial intelligence-based FDA-approved medical devices and algorithms: An online database. *npj Digital Medicine*, *3*, 118. https://doi.org/10.1038/s41746-020-00324-0

2 Bhatt, D., Patel, C., Ghetia, H., & Mazzeo, P. L. (2021). Convolutional neural networks in medical image understanding: A survey. *Evolutionary Intelligence*, *15*(1), 1–22. https://doi.org/10.1007/s12065-020-00540-3

3 Chan, H.-P., Samala, R. K., Hadjiiski, L. M., & Zhou, C. (2020). Deep learning in medical image analysis. *Advances in Experimental Medicine and Biology*, *1213*, 3–21. https://doi.org/10.1007/978-3-030-33128-3_1

4 Dogan, R. O., Dogan, H., Bayrak, C., & Kayikcioglu, T. (2021). A two-phase approach using Mask R-CNN and 3D U-Net for high-accuracy automatic segmentation of pancreas in CT imaging. *Computer Methods and Programs in Biomedicine*, *207*, 106141. https://doi.org/10.1016/j.cmpb.2021.106141

5 Farag, M. M., Fouad, M., & Abdel-Hamid, A. T. (2022). Automatic severity classification of diabetic retinopathy based on DenseNet and convolutional block attention module. *IEEE Access*, *10*, 38299–38308. https://doi.org/10.1109/ACCESS.2022.3165193

6 Fraiwan, M., & Faouri, E. (2022). On the automatic detection and classification of skin cancer using deep transfer learning. *Sensors*, *22*(13), 4963. https://doi.org/10.3390/s22134963

7 Guan, X., Yang, G., Yang, J., Xu, X., Jiang, W., & Lai, X. (2022). 3D AGSE-VNet: An automatic brain tumor MRI data segmentation framework. *BMC Medical Imaging*, *22*(1), 6. https://doi.org/10.1186/s12880-021-00728-8

8 He, K., Gkioxari, G., Dollar, P., & Girshick, R. (2020). Mask R-CNN. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *42*(2), 386–397. https://doi.org/10.1109/TPAMI.2018.2844175

9 Huang, S., Yang, J., Fong, S., & Zhao, Q. (2021). Artificial convolutional neural network in object detection and semantic segmentation for medical imaging analysis. *Frontiers in Oncology*, *11*, 638182. https://doi.org/10.3389/fonc.2021.638182

10 Jang, H.-J., Song, I.-H., & Lee, S.-H. (2021). Deep learning for automatic subclassification of gastric carcinoma using whole-slide histopathology images. *Cancers*, *13*(15), 3811. https://doi.org/10.3390/cancers13153811

11 Lahmar, C., & Idri, A. (2022). On the value of deep learning for diagnosing diabetic retinopathy. *Health and Technology*, *12*, 89–105. https://doi.org/10.1007/s12553-021-00606-x

12 Liu, X., Faes, L., Kale, A. U., Wagner, S. K., Fu, D. J., Bruynseels, A., Mahendiran, T., Moraes, G., Shamdas, M., Kern, C., Ledsam, J. R., Schmid, M. K., Balaskas, K., Topol, E. J., Bhatt, D. L., Denniston, A. K., & Keane, P. A. (2019). A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: A systematic review and meta-analysis. *The Lancet Digital Health*, *1*(6), e271–

e297. https://doi.org/10.1016/S2589-7500(19)30123-2

13 Liu, X., Song, L., Liu, S., & Zhang, Y. (2021). Advances in deep learning-based medical image analysis. *Health Data Science*, *2021*, 8786793. https://doi.org/10.34133/2021/8786793

14 Maqsood, S., Damaševičius, R., & Maskeliūnas, R. (2021). Hemorrhage detection based on 3D CNN deep learning framework and feature fusion for evaluating retinal abnormality in diabetic patients. *Sensors*, *21*(11), 3865. https://doi.org/10.3390/s21113865

15 Park, S., Lee, S. M., Kim, N., Seo, J. B., Yun, J., & Jung, K. H. (2020). Performance of a deep learning algorithm compared with radiologic interpretation for lung cancer detection on chest radiographs in a health screening population. *Radiology*, *297*(3), 687–696. https://doi.org/10.1148/radiol.2020201240

16 Soffer, S., Ben-Cohen, A., Shimon, O., Amitai, M. M., Greenspan, H., & Klang, E. (2021). Diagnostic accuracy of deep learning in medical imaging: A systematic review and meta-analysis. *npj Digital Medicine*, *4*, 65. https://doi.org/10.1038/s41746-021-00438-z

17 Wang, L. (2022). Deep learning techniques to diagnose lung cancer. *Cancers*, *14*(22), 5569. https://doi.org/10.3390/cancers14225569

18 Yaqoob, M. K., Ali, S. F., Bilal, M., Hanif, M. S., & Al-Saggaf, U. M. (2021). ResNet based deep features and random forest classifier for diabetic retinopathy detection. *Sensors*, *21*(11), 3883. https://doi.org/10.3390/s21113883

19 Yoo, H., Lee, S. H., Arru, C. D., Doda Khera, R., Singh, R., Siebert, S., Kim, D., Lee, Y., Park, J. H., Eom, H. J., Digumarthy, S. R., & Kalra, M. K. (2021). AI-based improvement in lung cancer detection on chest radiographs: Results of a multi-reader study in NLST dataset. *European Radiology*, *31*(12), 9664–9674. https://doi.org/10.1007/s00330-021-08074-7

20 Zhou, S. K., Greenspan, H., Davatzikos, C., Duncan, J. S., van Ginneken, B., Madabhushi, A., Prince, J. L., Rueckert, D., & Summers, R. M. (2021). A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises. *Proceedings of the IEEE*, *109*(5), 820–838. https://doi.org/10.1109/JPROC.2021.3054390