THE JOURNAL OF HEALTH SCIENCE AND MEDICAL TECHNOLOGY
2026 Vol. 1, No. 2, ISSN(O): 3154-7869



Robust Disease Classification under Spurious Shortcut Shift Using Causal Shortcut Suppression

Oliver Bennett[1*], Emily Carter[1], James Holloway[1]
[1]Department of Medical Physics and Biomedical Engineering, University College London, Gower Street, London WC1E 6BT, United Kingdom
Corresponding Author*: Oliver Bennett Email: oliver.bennett@ucl.ac.uk

ARTICLE INFO

Keywords
Shortcut learning
causal invariance
medical image classification
distribution shift
chest radiography
robust deep learning
counterfactual feature masking

Published
02 December 2026

ABSTRACT

Shortcut learning remains a major obstacle in medical image classification because deep neural networks may rely on acquisition artifacts, scanner-specific patterns, patient-position markers, or background structures instead of disease-related morphology. This study investigates shortcut identification and removal for robust chest radiograph classification under distribution shift. We propose a Causal Shortcut Suppression Network (CSS-Net), which first estimates candidate shortcut variables through environment-wise risk variation and representation sensitivity analysis, and then reduces their influence by applying a causal invariance penalty and counterfactual feature masking during training. The model was evaluated on a multi-source chest X-ray dataset containing 112,480 images from 30,805 patients, including 14 disease labels and metadata related to scanner type, projection view, patient age, sex, and hospital source. To test robustness, we constructed three external validation settings: cross-hospital testing, projection-shift testing, and shortcut-amplified testing, where shortcut-related correlations were artificially increased from 0.35 to 0.80. Compared with a standard ResNet-50 classifier, CSS-Net improved mean AUC from 0.842 to 0.879 in internal validation and from 0.731 to 0.812 in cross-hospital validation. Under shortcut-amplified distribution shift, the false positive rate decreased by 18.6%, while sensitivity for disease-positive cases increased from 72.4% to 79.8%. Feature attribution analysis further showed that CSS-Net reduced attention on image borders and device markers by 41.3% and increased lesion-region activation by 27.6%. These results indicate that causally guided shortcut suppression can improve both predictive stability and clinical interpretability when image models are deployed across heterogeneous medical environments.

1. Introduction

Deep learning has been widely applied to chest radiograph interpretation and has achieved strong performance in the detection of pneumonia, pleural effusion, cardiomegaly, pneumothorax, pulmonary edema, atelectasis, and other thoracic abnormalities. Large-scale convolutional and transformer-based models can learn complex visual patterns from chest X-rays and support automated multi-label disease classification [1]. However, high performance on an internal test set does not necessarily translate into stable **performance** in other hospitals, imaging systems, or

Citation: Bennett, O., Carter, E., & Holloway, J. (2026). Robust Disease Classification under Spurious Shortcut Shift Using Causal Shortcut Suppression. *The Journal of Health Science and Medical Technology*, 1(2), 21-30.

3154-7869/© The Authors. Published by J&L Academic Group PLT. This is an open access article under the CC BY 4.0 license.
https://doi.org/10.64744/tjhsmt.2026.327.

patient populations. A major reason is shortcut learning, in which a model relies on features that are correlated with disease labels in the training data but are not causally related to the underlying pathology. Previous studies have shown that chest X-ray classifiers may exploit hospital-specific artifacts, acquisition characteristics, image-processing patterns, or equipment-related cues instead of clinically relevant disease features [2]. Demographic information may also be encoded in medical-image representations even when the visual basis is not apparent to human observers [3]. More generally, globally informed representation-selection methods have demonstrated that identifying important and redundant information from the overall feature context can improve the efficiency and selectivity of deep models [4]. This principle is relevant to shortcut suppression because unstable or noncausal representations should be identified in relation to the complete predictive context rather than treated as isolated image artifacts. Subsequent studies have further shown that demographic attributes may remain embedded in chest radiograph features and influence disease prediction across patient groups [5]. Shortcut signals in chest radiographs can arise from multiple sources. Some are visually obvious, such as portable-device markers, laterality labels, tubes, lines, pacemakers, image borders, or text annotations. Others are less apparent and may originate from scanner manufacturers, detector characteristics, post-processing algorithms, projection geometry, examination protocols, or institutional acquisition practices. Models can also exploit correlations between disease labels and patient age, sex, race, care setting, or hospital source when these variables are unevenly distributed across datasets [6,7]. Hospital identity is particularly important because clinical populations and imaging procedures often differ systematically between institutions. Combining datasets from several hospitals can therefore introduce artificial hospital–disease associations that are highly predictive during training but unstable after deployment [8,9]. A model trained under such conditions may appear highly accurate while actually learning a mixture of pathological signals and site-specific confounders. When the association between a shortcut and a disease changes in a new environment, classification performance may deteriorate even though the pathological appearance of the disease itself has not changed.

This problem is especially challenging in multi-label chest radiograph classification because several shortcuts may coexist within the same image. A portable anteroposterior radiograph, for example, may simultaneously contain device-related features, hospital-specific image processing, a characteristic scanner pattern, and demographic correlations associated with disease prevalence. These factors can interact rather than act independently. A model may therefore distribute its predictive dependence across several weak shortcut signals, making the resulting bias difficult to detect through a single predefined confounder [10,11]. In addition, shortcut effects may vary across disease labels. Device presence may be strongly associated with cardiomegaly or pulmonary edema, whereas projection view or acquisition setting may have a greater influence on other findings. Evaluating shortcut learning therefore requires methods that can identify unstable representations across multiple environments and multiple labels rather than assuming that one shortcut mechanism dominates the entire dataset. Several approaches have been proposed to detect or reduce such behavior. Saliency maps, attention maps, gradient-based explanations, and related visualization techniques can indicate whether a classifier responds strongly to image borders, text annotations, devices, or regions outside the lungs [12,13]. Explanation-guided training has also been used to discourage models from relying on unwanted image regions [14,15]. These methods are useful for visual inspection, but spatial attention alone does not establish causal dependence. A model may attend to a region without using it as a decisive predictive feature, while distributed acquisition-related signals may influence predictions without producing a clearly localized saliency pattern. As

a result, shortcut detection based only on visual explanations may miss nonlocal or latent confounding information encoded throughout the representation space. Domain-generalization methods address the problem from another perspective by attempting to learn representations that remain stable across multiple training environments. Benchmark studies have shown that domain-generalization performance is highly sensitive to model selection, training conditions, and evaluation protocols [16,17]. Other methods improve out-of-distribution robustness by encouraging more consistent optimization behavior or reducing domain-specific representation differences [18]. These approaches can improve generalization when training domains provide sufficient variation, but they do not necessarily identify which features function as shortcuts. A model may achieve improved average performance across several domains while still retaining substantial dependence on demographic, acquisition, or institutional variables. Causal learning provides a more direct framework for distinguishing disease-related information from environment-specific associations. Causal intervention methods have been used to suppress domain-specific information in medical-image representations [19], while causal graphical models and instrumental-variable approaches have been introduced for chest X-ray classification [20]. By modeling relationships among disease labels, imaging variables, and potential confounders, these methods seek to reduce dependence on spurious correlations. Counterfactual analysis extends this idea by modifying selected features or latent representations and observing whether the predicted disease probability changes. Such perturbations can provide stronger evidence of functional dependence than conventional attention visualization because they directly test the sensitivity of a prediction to a controlled change [21,22]. Recent studies, however, also indicate that removing a known shortcut in one dataset does not necessarily guarantee improved performance in another population [23,24]. Shortcut mitigation must therefore address not only a predefined artifact but also the broader instability of learned representations across environments. Several limitations remain in current chest radiograph shortcut research. Many studies evaluate one confounder at a time, such as hospital source, projection view, demographic attributes, or medical devices. This design may underestimate the complexity of real-world imaging data, where multiple shortcuts can coexist and change simultaneously. Evaluation is also frequently based on random internal splits or a single external cohort. Such settings reveal whether a model generalizes to a specific dataset but provide limited information about how performance changes as shortcut correlations become progressively stronger or weaker. In addition, shortcut detection and shortcut suppression are often treated as separate procedures. A model may be analyzed after training to determine whether it uses undesirable features, while the identified instability is not directly incorporated into the optimization process. Conversely, some robust-training methods reduce domain dependence without explicitly determining which representations are unstable. This separation limits the ability to construct a unified framework in which shortcut evidence obtained during training directly guides representation suppression.

This study proposes the Causal Shortcut Suppression Network (CSS-Net), a unified framework for identifying and reducing shortcut dependence in multi-label chest radiograph classification. CSS-Net evaluates model behavior across heterogeneous environments and uses environment-wise risk variation together with representation sensitivity analysis to identify candidate shortcut variables and unstable feature dimensions. Rather than assuming that one predefined confounder is responsible for distribution shift, the framework jointly considers scanner type, projection view, patient age, sex, hospital source, and their associated representation changes. A causal invariance penalty is then introduced to encourage disease-relevant representations to remain stable across environments, while counterfactual feature masking selectively perturbs

shortcut-sensitive features and penalizes predictions that change excessively in response to nonpathological information. This design links shortcut discovery and shortcut suppression within the same training framework. The proposed model is evaluated on 112,480 chest radiographs from 30,805 patients with 14 thoracic disease labels and metadata describing scanner type, projection view, age, sex, and hospital source. Three complementary evaluation settings are constructed to examine different forms of distribution shift: cross-hospital testing measures generalization across institutional environments, projection-shift testing evaluates robustness to changes in acquisition geometry, and shortcut-amplified testing deliberately strengthens shortcut–label associations to determine how rapidly model performance deteriorates as spurious correlations increase. In addition to conventional discrimination metrics, the evaluation examines subgroup stability, shortcut sensitivity, representation invariance, and changes in performance under counterfactual perturbation. The central objective is to determine whether direct identification and suppression of unstable shortcut-related representations can preserve disease classification performance when noncausal correlations vary across clinical environments. By integrating environment-level risk analysis, causal invariance, and counterfactual feature masking, CSS-Net aims to provide a more systematic approach to shortcut mitigation and to improve the reliability of chest radiograph models under realistic cross-hospital and cross-population distribution shifts.

2. Materials and Methods

2.1 Study Cohort and Image Sources

This study included 112,480 chest radiographs from 30,805 patients collected from several hospitals. Each image was linked to one or more of 14 thoracic disease labels. The available patient and imaging information included scanner type, projection view, age, sex, and hospital source. Only frontal chest radiographs with valid disease labels and complete key information were included. Both posteroanterior (PA) and anteroposterior (AP) images were retained because projection view was considered a possible shortcut factor. Images with damaged pixel data, duplicate records, missing patient IDs, or unclear projection information were excluded. Disease labels from different hospitals were converted to the same 14-label system. All images from the same patient were placed in only one dataset to avoid information overlap between training and testing.

2.2 Experimental Design and Control Experiments

The dataset was divided into training, validation, and internal test sets at the patient level. Disease prevalence and hospital source were considered during data splitting. A standard ResNet-50 model was used as the control and was trained with the same images, data splits, optimizer settings, and stopping rules as CSS-Net. Three distribution-shift tests were designed. For cross-hospital testing, images from selected hospitals were excluded from training and used only for testing. For projection-shift testing, the proportions of AP and PA images were changed between the training and test sets. For shortcut-amplified testing, the association between selected shortcut factors and disease labels was increased from 0.35 to 0.80 by controlled sampling. Ablation tests were also performed by removing the causal invariance loss or counterfactual feature masking from CSS-Net. These tests were used to measure the effect of each model component.

2.3 Image Processing, Measurement, and Quality Control

All chest radiographs were resized to 224 × 224 pixels and normalized using the mean and standard deviation of the training set. Pixel values below the 0.5th percentile or above the 99.5th percentile were clipped to reduce the effect of extreme values. Random horizontal flipping, small rotations, image shifts, and slight intensity changes were used only during training. No image

augmentation was applied to the validation or test sets. Shortcut sensitivity was measured by comparing model outputs before and after shortcut-related information was masked or changed. The tested factors included scanner-related patterns, projection information, demographic information, image borders, and device markers. Feature attribution maps were used to measure model activation in lesion-related regions and non-disease regions. Quality control included duplicate checking, patient-level data separation, label checking, metadata checking, and verification of the three distribution-shift test sets. All preprocessing settings were obtained from the training data and then kept unchanged for validation and testing.

2.4 Data Processing and Analytical Formulas

Prediction scores were processed using the same workflow for all training environments. The classification risk was calculated separately for each environment. The mean risk was then used to measure the variation in model performance across different imaging conditions. The environment invariance loss was calculated as:

$$L_{inv}=\frac{1}{E}\sum_{e=1}^{E}(R_e-\bar{R})^2$$

where E is the number of training environments, R_e is the classification risk in environment e, and $\bar{R}$ is the mean risk across all training environments. A lower L_{inv} indicates more stable prediction across different imaging environments.

Counterfactual feature masking was used to measure changes in disease prediction after shortcut-related information was removed. The difference between the original prediction and the prediction after masking was calculated as:

$$L_{cf}=\frac{1}{N}\sum_{i=1}^{N}(p_i-p_i^{cf})^2$$

where N is the number of training samples, p_i is the prediction probability from the original image, and p_i^{cf} is the prediction probability after shortcut-related features were masked. The classification loss, environment invariance loss, and counterfactual consistency loss were jointly used for model training. Only complete image and metadata records were analyzed. Statistical significance was set at $p < 0.05$.

2.5 Model Training and Statistical Analysis

All models were initialized with ImageNet-pretrained ResNet-50 weights. Training used the AdamW optimizer with an initial learning rate of 1×10^{-4} and a batch size of 64. Each model was trained for up to 50 epochs. Early stopping was applied according to validation AUC, and the learning rate was reduced when validation performance no longer improved. Model performance was measured using AUC, sensitivity, specificity, false positive rate, and F1 score. Mean AUC across the 14 disease labels was used as the main evaluation measure. Results were reported for the internal test, cross-hospital test, projection-shift test, and shortcut-amplified test. Bootstrap sampling with 1,000 repeats was used to calculate 95% confidence intervals. Paired bootstrap tests were used to compare AUC values between CSS-Net and ResNet-50, while McNemar's test was used to compare paired classification results. A two-sided P value below 0.05 was considered statistically significant.

3. Results and Discussion

3.1 Internal Classification Performance

CSS-Net achieved a mean AUC of 0.879 on the internal test set, compared with 0.842 for ResNet-50. The increase of 0.037 suggests that shortcut suppression did not reduce the model's ability to detect disease-related features. This is relevant because removing shortcut information may also remove useful features when the two are strongly related. Yuan et al. (2026) reported that demographic information can remain in chest X-ray features without being directly used for disease prediction. Their results suggest that shortcut control should reduce harmful dependence rather than remove all information related to patient or imaging factors [25,26]. The higher AUC of CSS-Net suggests that the causal penalty and feature masking retained useful disease information while limiting unstable image cues. The result also suggests that controlling shortcut effects at the disease-risk level can improve classification without reducing performance on the original data distribution.

3.2 Cross-Hospital Generalization

The performance difference between CSS-Net and ResNet-50 became larger during cross-hospital testing. The mean AUC of ResNet-50 decreased from 0.842 to 0.731, while that of CSS-Net decreased from 0.879 to 0.812. The corresponding performance gaps were 0.111 and 0.067. CSS-Net therefore reduced the cross-hospital gap by 0.044. Similar results were reported by Iqbal et al. (2025), who found that chest X-ray model AUC decreased from 0.85 on MIMIC-CXR to 0.73 on an external CheXpert dataset. Their study linked much of this decrease to acquisition-related shortcut information, as illustrated in Fig. 1. Bao et al. (2026) also found that hospital source could become associated with disease prevalence and reduce performance in some groups. The present findings follow the same pattern. However, the smaller decrease obtained by CSS-Net suggests that reducing environment-specific dependence can improve transfer between hospitals. The remaining difference between internal and external AUC may be related to changes in patient groups, disease prevalence, image quality, and imaging procedures [27,28].

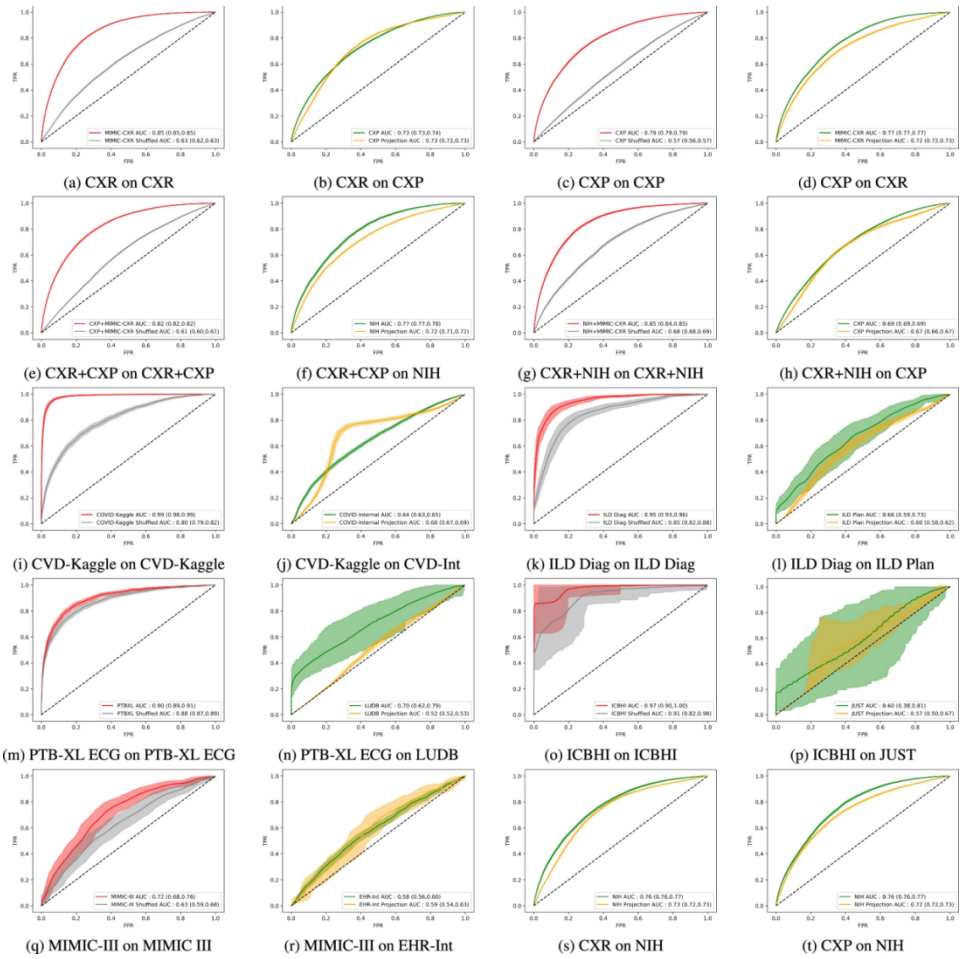


Fig. 1. Cross-hospital classification performance under shortcut shift.

3.3 Performance under Amplified Shortcut Shift

When the shortcut-label correlation was increased from 0.35 to 0.80, CSS-Net reduced the false positive rate by 18.6% and increased sensitivity from 72.4% to 79.8%. The 7.4 percentage-point increase in sensitivity means that more disease-positive cases were correctly identified when shortcut signals became stronger. This test is stricter than a standard random split because the relation between shortcut features and disease labels was changed directly [29,30]. Khaled et al. (2025) used counterfactual chest X-rays to test a similar problem. They found that changing one pathological feature could alter predictions for other unrelated diseases, as illustrated in Fig. 2. Their results showed that chest X-ray models may use off-target findings as shortcuts. They also reported better internal and external performance after counterfactual training. CSS-Net follows a related idea but uses feature masking rather than generated pathology changes. The lower false positive rate and higher sensitivity suggest that the model was less affected by shortcut signals when these signals became strongly associated with the target labels.

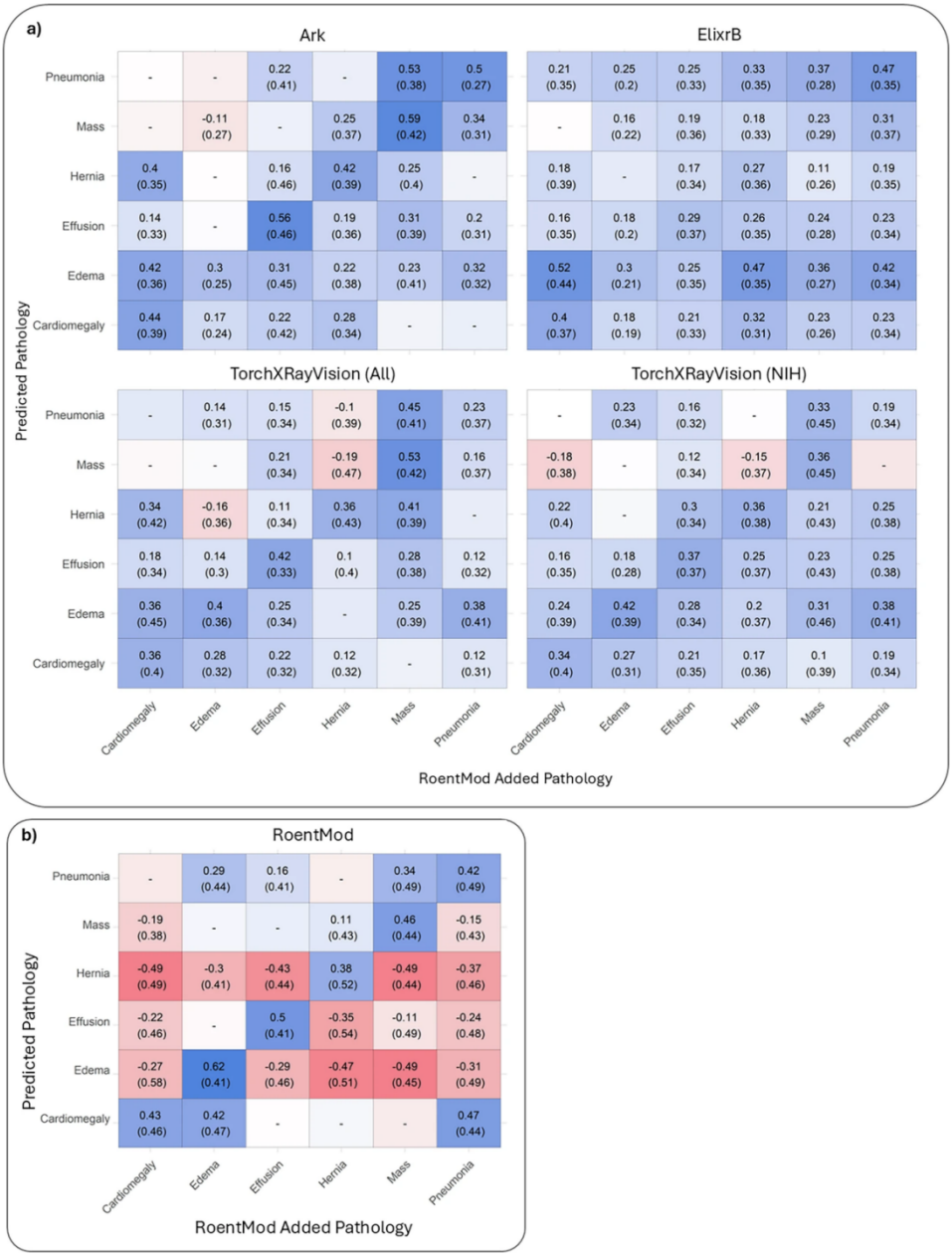


Fig. 2. Prediction changes after masking shortcut-related image features.

3.4 Feature Attribution and Clinical Relevance

Feature attribution analysis found that CSS-Net reduced activation on image borders and device-marker regions by 41.3%. At the same time, activation in lesion-related regions increased by 27.6%. These changes were consistent with the better results in cross-hospital and shortcut-amplified testing. Rosenblatt et al. (2026) reported that chest X-ray models could use text, image edges, acquisition patterns, and other non-disease information for prediction. Gu (2026) further found that acquisition parameters can contain demographic information and influence model errors across patient groups. Liang et al. (2026) reported that shortcut-reduction methods that work well in one dataset may not give the same results in another population. The present study therefore used both external testing and controlled shortcut shift instead of relying only on attribution maps. The lower activation in border and marker regions was accompanied by higher cross-hospital AUC and better sensitivity under shortcut shift. This agreement between image attribution and classification results suggests that CSS-Net relied more on disease-related regions and less on unstable image features.

4. Conclusion

This study developed CSS-Net to reduce shortcut learning in chest radiograph classification under distribution shift. The model combines environment-based shortcut detection, an invariance loss, and counterfactual feature masking. CSS-Net achieved a mean AUC of 0.879 in internal testing and 0.812 in cross-hospital testing, compared with 0.842 and 0.731 for ResNet-50. When the shortcut-label correlation became stronger, the false positive rate decreased by 18.6%, while sensitivity increased from 72.4% to 79.8%. Feature attribution results also showed less attention to image borders and device markers and more attention to lesion-related regions. These results suggest that CSS-Net can reduce the effect of unstable image cues while keeping useful disease information. The main contribution of this study is the use of shortcut detection and shortcut suppression in the same training process. This method may improve the use of chest X-ray models across different hospitals, scanners, and imaging settings. However, the study focused only on chest radiographs and used a limited set of recorded metadata. Other shortcut factors may still affect model predictions. Future studies should test CSS-Net on larger multi-center datasets, other medical imaging tasks, and prospective clinical data.

References

[1]. Qi, C., & Qiao, X. (2026). Building and Operating a Large Scale Multi-Agent System: A Case Study from Industry. Available at SSRN 6795198.

[2]. Bundza, P. (2026). MCADS: Out-of-Distribution Generalization of Multi-Label Chest X-Ray Classification Models Across Seven Architectures (Doctoral dissertation, Vilniaus Gedimino technikos universitetas.).

[3]. Xu, T., Zhu, W., & Zhang, J. (2026). A Study on the Application of Alternative Data in Credit Assessment for the Unbanked Population.

[4]. Yao, G., Zheng, J., Wang, Z., Zhang, W., Han, R., Zhao, C., ... & Liu, R. (2026, March). V-pruner: A fast and globally-informed token pruning framework for vision transformer. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 40, No. 40, pp. 34396-34404).

[5]. Maiter, A., Taylor, J., Metherall, P., Hyde, G., Anderson, E., Hesketh, E., ... & Johns, C. S. (2026). Independent head-to-head comparison of commercial artificial intelligence devices for lung cancer detection on chest radiographs. Radiology, 319(2), e252205.

[6]. Du, Y., Liu, Q., Dong, Y., & Chen, Z. (2026). Cross-Domain Transfer and Few-Shot Recognition of Defect Images in Complex Industrial Settings.

[7]. Xiong, W., Guo, Y., & Zeng, Y. (2026). A Study on the Energy Efficiency of MEP Systems and Coordinated Dispatch Mechanisms for Multi-energy Systems in High-density Urban Buildings. Available at SSRN 7120518.

[8]. Hong, Z., Liang, S., Xu, T., & Chen, H. (2026). A Study on Failover Verification and Recovery Objective Prediction for Cross-Region Cloud Services.

[9]. Gharibi, Z. (2025). An Ensemble Learning Approach for Predicting Hospital Stay in Transplant Patients. Healthcare Analytics, 100444.

[10]. Li, J., Ma, R., & Gu, Y. (2026). From Audit Requirements to Computable Software Architecture: A Rule-Engine and Evidence-Indexing Method for Trusted Digital Infrastructure.

[11]. Huang, J., Yin, J., Xu, T., & Yang, J. (2026). Quality Assessment and User Adoption Prediction for Enterprise-Level Financial Reporting Data Products. Available at SSRN 7179118.

[12]. Zhang, Z., Wang, J., Li, Z., Wang, Y., & Zheng, J. (2025). Anncoder: A mti-agent-based code generation and optimization model. Symmetry, 17(7), 1087.

[13]. Pushpalatha, N., Swathi, P., Chowdary, C. A., Gudipalli, T. L., & Godavarthi, B. (2026, April). Explainable Image Classification with Saliency Maps and Class Activation Mapping (CAM). In 2026 13th International Conference on Computing for Sustainable Global Development (INDIACom) (pp. 1-7). IEEE.

[14]. Bao, Y., & Wang, H. (2026). Orion: A High-Throughput, Fault-Tolerant Event Routing Architecture for Hyperscale Data Streams. Fault-Tolerant Event Routing Architecture for Hyperscale Data Streams (January 01, 2026).

[15]. Yang, Q., & Du, Y. (2026). Predicting Rollback Risks and Controlling Gradual Rollout Traffic for Online Ranking Feature Deployments.

[16]. Zhao, Z., & Welsch, R. E. (2026). Point-in-Time Financial RAG with Frozen LLMs and Market-Feedback Adaptive Retrieval. arXiv preprint arXiv:2605.31201.

[17]. Iqbal, S., Ali, D., & Tian, Q. (2025). Systematic Review of Domain Adaptation and Generalization in Image Classification and Face Recognition.

[18]. Yuan, Y., Huang, J., Yin, J., & Xu, T. (2026). Confidence Calibration and Semantic Error Analysis for Ambiguous Coreference Resolution in User-generated Social Media Text. Available at SSRN 7393238.

[19]. Gu, X. (2026). Identifying Causal Effects and Analyzing Heterogeneity of User Growth Interventions on Digital Platforms: Evidence from Large-Scale Behavioral Data. Available at SSRN 6809181.

[20]. Khaled, A., Sabir, M., Qureshi, R., Caruso, C. M., Guarrasi, V., Xiang, S., & Zhou, S. K. (2025). Leveraging MIMIC datasets for better digital health: a review on open problems, progress highlights, and future promises. arXiv preprint arXiv:2506.12808.

[21]. Liang, S., Hsu, K. H., & Liu, Z. (2026). Rolling Scheduling and Capacity Balancing for Server Assembly Under Engineering Change Disturbances.

[22]. Bao, Y., & Wang, H. (2026). Large-scale Metrics Migration: A Systematic Approach to Rebuilding Production Measurement Infrastructure. Available at SSRN 7185660.

[23]. Zhao, J., Fan, J., & Li, L. (2026). A Study on an Explainable Causal-Enhanced LLM Agent for Predicting the Forming Quality of Automotive Component Materials.

[24]. Rosenblatt, M., Foster, M. L., Adkinson, B. D., Tejavibulya, L., Khaitova, M., Ye, J., ... & Scheinost, D. (2026). External validation improves generalizability, replicability and reproducibility in predictive models for neuroimaging. Nature Methods, 1-11.

[25]. You, S. (2026). Verifiable Audit Mechanisms in AI Compliance Automation: Scalability. Available at SSRN 6547458.

[26]. Jiao, Y., Shi, T., Zhao, B., & Wang, A. (2026). The Impact of VR Sketching Simulations on the Assessment of Scale and Site Suitability in Public Art Spaces.

[27]. Suleman, M. U., Mursaleen, M., Khalil, U., Saboor, A., Bilal, M., Khan, S. A., ... & Tahir, M. (2025). Assessing the generalizability of artificial intelligence in radiology: a systematic review of performance across different clinical settings. Annals of Medicine and Surgery, 87(12), 8803.

[28]. Liang, S., Du, Y., Chen, W., & Liu, Z. (2026). Order Allocation and Emergency Transportation Decisions Amid Multi-Source Procurement Disruptions.

[29]. Masoudi, B., & Bazzi, S. (2026). Towards Trustworthy Diagnostics: An Interpretable Framework for Uncovering Shortcut Learning in Plant Disease Models. Journal of Soft Computing and Information Technology, e242549.

[30]. Yuan, Y., Xu, T., Yin, J., & Huang, J. (2026). Modeling Loading Anomalies and Identifying Root Causes in Media Consumption Workflows on Large Social Media Platforms. Available at SSRN 7393318.