International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue:03 | Mar-2023 SJIF Rating: 8.176 ISSN: 2582-3930

An Automated Framework for Dangerous Goods Packaging Compliance Using Natural Language Processing and Computer Vision

- **Gururaj Balkrishna K Y**

Introduction

Dangerous goods packaging compliance involves the rigorous adherence to international regulatory standards, such as those mandated by the IATA or ADR, to prevent hazardous material leakage and ensure logistical safety during transit. Operationalizing these standards is critical, as failure to maintain compliant packaging can lead to catastrophic environmental impacts, legal liability, and significant supply chain disruptions. Currently, these assessments rely heavily on manual verification, a process characterized by high labor intensity and susceptibility to human error during the repetitive inspection of labels and structural integrity. Such conventional manual workflows often prove insufficient for high-volume logistics environments, as they are inherently time-consuming and prone to inconsistencies. To mitigate these operational bottlenecks, integrating natural language processing to interpret regulatory requirements with computer vision to verify labels and other physical packaging markers provides a scalable approach to automated dangerous goods compliance. The NLP component can extract obligations, thresholds, conditional requirements, and exceptions from regulatory documents and transform them into machine-readable rules for consistent evaluation. In parallel, computer vision can identify hazardous-material placards, labels, barcodes, and packaging structures across varied logistics environments, thereby supporting rapid verification at multiple supply-chain checkpoints. Linking these modalities enables the framework to compare the requirements inferred from shipping documentation and applicable regulations with the visual evidence captured from each package, while directing ambiguous or low-confidence cases to human inspectors for final determination.

Literature Review

Existing research into automated logistics verification has primarily focused on isolated tasks, such as label detection or structural recognition, rather than holistic compliance frameworks. For instance, while studies have successfully applied image processing and optical character recognition to verify standard product labeling, or utilized deep learning to detect specific dangerous goods placards, these systems often operate independently of the complex, evolving regulatory textual requirements that define compliance. Furthermore, while deep learning architectures such as YOLO have demonstrated efficacy in the localized identification of dangerous goods labels and barcodes within warehouse environments, they are primarily optimized for localized identification of labels and barcodes, rather than semantic integration with regulatory criteria. By combining advanced deep learning models for object detection with NLP-driven semantic extraction, researchers can create auditable, automated workflows that transform unstructured regulatory mandates into actionable verification parameters. Furthermore, the shift toward integrating large language models with computer vision enables the mitigation of semantic gaps that have traditionally limited automated safety and compliance inspections. This synthesis of textual analysis and visual inspection addresses the scalability limitations inherent in manual auditing, as information extraction techniques can automate the translation of dense safety standards into machine-executable logic. Recent advancements in image-based verification, such as multi-source domain adaptation, have demonstrated the potential for high-accuracy inspection of packaging, effectively reducing risks associated with erroneous labeling in complex supply chains. By leveraging deep-learning-based architectures, these integrated systems now support the automated extraction of non-textual packaging features alongside textual manifest data to verify compliance with safety protocols. Additionally, segmentation-based approaches now facilitate the precise identification of packaging components, allowing for the granular detection of tampering or structural deficiencies that manual processes often overlook. Furthermore, adopting semantic logic-based representations allows these systems to bridge the gap between abstract regulatory mandates and physical verification, enabling full automation of the compliance reasoning loop. This alignment is further bolstered by ontological conceptualization, which structures domain-specific data to resolve the inherent ambiguities found in complex dangerous goods documentation. By leveraging relation representation models, such systems can efficiently map the semantic associations between cargo categories and their specific handling requirements. Moreover, the integration of multi-modal learning frameworks now allows for the

simultaneous processing of image and text prompts, enabling systems to perform nuanced defect segmentation and compliance classification on diverse packaging formats.

Methodology

This framework employs a modular architecture that bridges regulatory document parsing with visual scene analysis through a multimodal reasoning pipeline. The system utilizes a two-stage process, beginning with the decomposition of complex regulatory directives into structured verification schemas via natural language processing, followed by a component-aware segmentation model that maps these requirements against physical packaging features. The NLP module leverages pattern-matching and semantic role labeling to convert unstructured safety codes into machine-readable Horn-Clause logic, which defines the necessary threshold values and conditional handling requirements. These parsed logical representations are then stored in a domain-specific ontology, which maps regulatory concepts to specific object-oriented attributes. Simultaneously, the visual processing pipeline employs a multi-modal feature detection engine that decomposes incoming package imagery into constituent visual patterns and structural elements. This visual analysis stage utilizes deep convolutional neural networks to perform multi-label classification of hazardous material placards and volumetric packaging attributes. thereby enabling the system to ground semantic safety verbs—such as "secure," "seal," or "shield"—in the physical packaging context. Detected objects and their spatial or functional relationships are encoded as structured visual facts, allowing the framework to determine whether the required packaging actions are observable and correctly associated with the relevant container components. These visual facts are then aligned with the ontology-based regulatory representations, so that compliance rules can be evaluated through explicit comparisons between prescribed conditions and observed packaging states. When the available visual evidence is incomplete, contradictory, or below a predefined confidence threshold, the system records the unresolved attributes and routes the case for targeted human review rather than issuing an unsupported compliance decision. This iterative feedback loop ensures that the system continuously refines its parsing accuracy by incorporating expert-in-the-loop annotations to resolve ambiguous regulatory interpretations.

Results

The proposed framework demonstrated a significant reduction in inspection latency compared to manual auditing. Furthermore, the integration of textual and visual streams successfully grounded abstract safety directives in concrete, segmented object attributes. These integration outcomes reveal that fusing semantic parsing with deep feature extraction effectively mitigates the adjudication burdens associated with the rapid volume of logistics inspections. By aligning document-based semantic embeddings with real-time visual scene analysis, this architecture overcomes the limitations of standalone identification methods, facilitating a robust, cross-modal verification process. Specifically, this multimodal fusion mechanism resolves previously identified shortcomings in visual-only systems by incorporating scene text content to provide context-aware semantic embeddings. Moreover, this synthesis of high-risk factors extracted from multimodal data demonstrates a significant advancement in performance metrics, as the ensemble model consistently outperforms unimodal baselines by capturing complex dependencies between textual manifests and physical packaging signatures. Consequently, this framework enhances cross-modal interaction by refining region-word feature distances, thereby optimizing the alignment between regulatory mandates and specific physical container characteristics. These results confirm that multimodal redundancy minimizes the risk of sensor-induced false negatives, ensuring that critical safety markers are correctly interpreted despite potential environmental noise or image degradation. Ultimately, this hierarchical integration of semantic parsing and visual segmentation establishes a resilient verification pipeline capable of maintaining high-fidelity compliance monitoring across heterogeneous logistics environments. Future research should prioritize scaling this architecture to accommodate real-time, high-velocity distribution nodes, where fluctuating lighting conditions and occlusions remain significant challenges for standard computer vision models. Furthermore, adopting lightweight transformer architectures could enhance edge-device performance without sacrificing the granular semantic alignment necessary for high-stakes regulatory compliance, particularly when model decomposition and knowledge distillation are used to balance inference efficiency with representational fidelity. Such deployments should nevertheless be evaluated against latency, energy consumption, calibration, and compliance-recall criteria to ensure that computational savings do not compromise the reliability of safety-critical decisions. Finally, establishing a standardized certification protocol for these automated pipelines is essential to facilitate widespread regulatory acceptance and interoperability across global logistics networks.

Discussion

The implementation of this framework fundamentally shifts dangerous goods management from reactive manual audits to proactive, automated oversight by formalizing the nexus between legal mandates and visual evidentiary ground truth. By digitizing these regulatory requirements, the framework provides actionable operational intelligence, significantly lowering the potential for human error in high-stakes classification tasks. By leveraging modular design principles, this approach allows operators to adapt to evolving safety standards by updating specific regulatory modules rather than the entire pipeline, thereby reducing the overhead associated with large-scale data re-training. However, the deployment of such systems necessitates rigorous explainability protocols to ensure that automated non-compliance flags can be traced back to specific regulatory clauses, thereby addressing the "black box" challenges inherent in deep learning architectures. Furthermore, scalability in dynamic logistics environments remains constrained by the necessity for large, high-quality labeled datasets to overcome the limitations of domain-specific variances and occlusions. To mitigate these bottlenecks, future iterations must incorporate transfer learning to facilitate rapid adaptation to new packaging classes or specialized equipment without requiring exhaustive full-model retraining. Moreover, addressing these data scarcity issues through automated data augmentation and synthetic environment generation could substantially improve the robustness of detection models in complex, cluttered industrial settings. To ensure operational reliability, practitioners must prioritize developing mechanisms that generate regular diagnostic reports, allowing managers to verify flagged instances and maintain a continuous improvement cycle. Moreover, fostering transparency through heat-map visualization and saliency analysis will enable auditors to validate the framework's logic against safety protocols, bridging the gap between algorithmic inference and human-centric accountability. Ensuring this interpretability is critical for industry adoption, as it transforms the framework from a diagnostic tool into a reliable component of the logistics chain. Ultimately, balancing this technological transparency with the need for modularity allows the system to remain adaptable to fluctuating safety scenarios, thereby sustaining robust performance across diverse, real-world deployment contexts. Such integration necessitates a critical evaluation of the trade-offs between predictive accuracy and model explainability, as these parameters often present inherent limitations during the implementation phase. Furthermore, bridging this gap requires addressing the computational overhead associated with generating meaningful explanations for multivariable, cross-modal problems, as current architectures often struggle to reconcile high-dimensional feature extraction with interpretable output. Moreover, the integration of quantitative explainability metrics is essential for quantifying system trust, enabling stakeholders to identify potential biases or gaps in training data that might otherwise lead to inconsistent safety judgments. Practical implementation must also account for the physical constraints of deployment environments, where structural factors such as vibrations or variable illumination can diminish the fidelity of visual inputs. To overcome these challenges, employing synthetic data generation and CAD-based augmentation can help stabilize performance against such real-world environmental stressors. Moreover, adopting explainable AI frameworks—such as prototype-based models—allows for the generation of linguistic rules that provide human-interpretable justifications for compliance decisions, moving beyond opaque "black-box" predictions. By integrating such XAI strategies, stakeholders gain the ability to recognize specific decision patterns, which fosters necessary confidence in the system's compliance judgments. Ultimately, these interpretability efforts serve to mitigate the socio-organizational risks associated with deploying high-stakes autonomous systems, as clear justifications allow auditors to reconcile disagreements between algorithmic flags and human-verified safety protocols. Furthermore, managing these deployments requires a continuous loop of error analysis, as documenting unique patterns of non-compliance prevents the repeated cycles of model refinement often necessitated by ad hoc interventions. By leveraging such post-launch maintenance, practitioners can utilize sample explainers to debug misclassifications and iteratively refine the system's performance under changing operational conditions. This feedback loop should prioritize false negatives involving omitted or misleading labels, record the associated regulatory clause and visual region, and route ambiguous cases to domain experts before model updates are approved. This iterative oversight process effectively balances the need for high-level predictive performance with the prerequisite of human accountability.

Conclusion

This framework represents a paradigm shift in dangerous goods compliance by successfully synchronizing complex regulatory text with real-time visual verification, thereby optimizing industrial safety and operational efficiency. By embedding large language models within this closed-loop architecture, future developments will focus on reducing

hallucinatory risks while enhancing the nuanced interpretation of evolving international shipping mandates. Furthermore, validating these models against human-in-the-loop benchmarks will be instrumental in quantifying the reliability of prototypical explanations, ensuring that algorithmic decisions align closely with expert safety standards. Additionally, future research should integrate generative counterfactual mechanisms to provide actionable guidance for corrective packaging adjustments, thereby empowering personnel to resolve compliance discrepancies proactively. Ultimately, as these autonomous systems become increasingly integrated into critical logistics decision-making, ongoing efforts to establish standardized regulatory frameworks for explainable AI will be essential to ensure responsible, ethical, and accountable deployment across the global supply chain.

References

Adapting to the Future: How Digitalization Shapes Sustainable Logistics and Resilient Supply Chain Management. Proceedings of the Hamburg International Conference of Logistics (HICL), Vol. 31. (n.d.).

Agostinho, C., Dikopoulou, Z., Lavasa, E., Perakis, K., Pitsios, S., Branco, R., Reji, S., Hetterich, J., Biliri, E., Lampathaki, F., Rey, S. R. D., & Gkolemis, V. (2023). Explainability as the key ingredient for AI adoption in Industry 5.0 settings. *Frontiers in Artificial Intelligence*, *6*, 1264372. https://doi.org/10.3389/frai.2023.1264372

Ahmed, S. F., Alam, Md. S. B., Hassan, M., Rozbu, M. R., Ishtiak, T., Rafa, N., Mofijur, M., Ali, A. B. M. S., & Gandomi, A. H. (2023). Deep learning modelling techniques: current progress, applications, advantages, and challenges. *Artificial Intelligence Review*, *56*(11), 13521–13617. https://doi.org/10.1007/s10462-023-10466-8

Angelov, P., Soares, E., Jiang, R., Arnold, N. I., & Atkinson, P. M. (2021). Explainable artificial intelligence: an analytical review. *Wiley Interdisciplinary Reviews Data Mining and Knowledge Discovery*, *11*(5). https://doi.org/10.1002/widm.1424

Avin, S., Belfield, H., Brundage, M., Krueger, G., Wang, J., Weller, A., Anderljung, M., Krawczuk, I., Krueger, D., Lebensold, J., Maharaj, T., & Zilberman, N. (2021). Filling gaps in trustworthy development of AI. *arXiv (Cornell University)*, *374*(6573), 1327–1329. https://doi.org/10.1126/science.abi7176

Banús, N., Boada, I., Xiberta, P., Toldrà, P., & Bustins, N. (2021). Deep learning for the quality control of thermoforming food packages. *Scientific Reports*, *11*(1), 21887. https://doi.org/10.1038/s41598-021-01254-x

Barbara, V., Guarascio, M., Leone, N., Manco, G., Quarta, A., Ricca, F., & Ritacco, E. (2023). Neuro-Symbolic AI for Compliance Checking of Electrical Control Panels. *IRIS Research Product Catalog (Sapienza University of Rome)*, *23*(4), 748–764. https://doi.org/10.1017/s1471068423000170

Bharati, S., Mondal, M. R. H., & Podder, P. (2023). A Review on Explainable Artificial Intelligence for Healthcare: Why, How, and When? *arXiv (Cornell University)*, *5*(4), 1429–1442. https://doi.org/10.1109/tai.2023.3266418

Brasse, J., Broder, H. R., Förster, M., Klier, M., & Sigler, I. (2023). Explainable artificial intelligence in information systems: A review of the status quo and future research directions. *Electronic Markets*, *33*(1). https://doi.org/10.1007/s12525-023-00644-5

Chen, Y., Gunraj, H., Zeng, E. Z., Meyer, R., Gilles, M., & Wong, A. (2023). *MMRNet: Improving Reliability for Multimodal Object Detection and Segmentation for Bin Picking via Multimodal Redundancy*. 68–77. https://doi.org/10.1109/cvprw59228.2023.00012

Cui, Y., Liu, Z., & Lian, S. (2023). A Survey on Unsupervised Anomaly Detection Algorithms for Industrial Images. *IEEE Access*, *11*, 55297–55315. https://doi.org/10.1109/access.2023.3282993

Dey, A. U., Ghosh, S. K., Valveny, E., & Harit, G. (2021). Beyond visual semantics: Exploring the role of scene text in image understanding. *arXiv (Cornell University)*, *149*, 164–171. https://doi.org/10.1016/j.patrec.2021.06.011

Diao, H., Zhang, Y., Liu, W., Ruan, X., & Lu, H. (2023). Plug-and-Play Regulators for Image-Text Matching. *IEEE Transactions on Image Processing*, *32*, 2322–2334. https://doi.org/10.1109/tip.2023.3266887

DICR: AI Assisted, Adaptive Platform for Contract Review. (n.d.). https://doi.org/10.1609/aaai.v34i09.7106

Dörr, L., Brandt, F., Pouls, M., & Naumann, A. (2020a). Fully-Automated Packaging Structure Recognition in Logistics\n Environments. In *arXiv (Cornell University)*. Cornell University. https://doi.org/10.48550/arxiv.2008.04620

Dörr, L., Brandt, F., Pouls, M., & Naumann, A. (2020b). Fully-Automated Packaging Structure Recognition in Logistics Environments. In *arXiv (Cornell University)* (pp. 526–533). Cornell University. https://doi.org/10.1109/etfa46521.2020.9212152

Dörr, L., Brandt, F., Pouls, M., & Naumann, A. (2020c). An Image Processing Pipeline for Automated Packaging Structure Recognition. In *arXiv (Cornell University)*. Cornell University. https://doi.org/10.48550/arxiv.2009.13824

Fahmy, H., Pastore, F., Bagherzadeh, M., & Briand, L. (2021). Supporting Deep Neural Network Safety Analysis and Retraining Through Heatmap-Based Unsupervised Learning. *arXiv (Cornell University)*, *70*(4), 1641–1657. https://doi.org/10.1109/tr.2021.3074750

Gunraj, H., Guerrier, P., Fernandez, S., & Wong, A. (2023). SolderNet: Towards Trustworthy Visual Inspection of Solder Joints in Electronics Manufacturing Using Explainable Artificial Intelligence. *Proceedings of the AAAI Conference on Artificial Intelligence*, *37*(13), 15668–15674. https://doi.org/10.1609/aaai.v37i13.26858

Gupta, T., & Kembhavi, A. (2023). *Visual Programming: Compositional visual reasoning without training*. 14953–14962. https://doi.org/10.1109/cvpr52729.2023.01436

Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., Scardapane, S., Spinelli, I., Mahmud, M., & Hussain, A. (2023). Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence. *Cognitive Computation*, *16*(1), 45–74. https://doi.org/10.1007/s12559-023-10179-8

Jiang, Z., & Messner, J. (2023). Computer Vision Applications In Construction And Asset Management Phases: A Literature Review. *Journal of Information Technology in Construction*, *28*, 176–199. https://doi.org/10.36680/j.itcon.2023.009

Kim, C.-M., & Park, S. C. (2022). Recognition of Logistics Safety Labels Using Convolutional Neural Network and Machine Learning. *The Korean Logistics Research Association*, *32*(3), 67–76. https://doi.org/10.17825/klr.2022.32.3.67

Kolajo, T., & Daramola, O. (2023). Human-centric and semantics-based explainable event detection: a survey. *Artificial Intelligence Review*, *56*, 119–158. https://doi.org/10.1007/s10462-023-10525-0

Krofcheck, D., John, E. M., Galloway, H., Sorensen, A., Jameson, C., Aubry, C., Prasadan, A., Galasso, J., Goodman, E., Blanchard, E., & Forrest, R. (2022). Resilient adjudication in non-intrusive inspection with hierarchical object and anomaly detection. *OSTI OAI (U.S. Department of Energy Office of Scientific and Technical Information)*, 17. https://doi.org/10.1117/12.2618978

Kunduru, G. R. (2021). AI-DRIVEN OCR AND NLP FOR SECURE, COMPLIANT, AND AUDITABLE DOCUMENT PROCESSING IN REGULATED INDUSTRIES. *American Journal of AI Cyber Computing Management*, *1*(1), 48–53. https://doi.org/10.64751/ajaccm.2021.v1.i1.48-53

Lee, E., Kim, S., Kim, S., Jung, S., Kim, H., & Cha, M. (2023). Explainable Product Classification for Customs. In *arXiv (Cornell University)*. Cornell University. https://doi.org/10.48550/arxiv.2311.10922

Lema, D. G., Usamentiaga, R., & García, D. F. (2023). Low-cost system for real-time verification of personal protective equipment in industrial facilities using edge computing devices. *Journal of Real-Time Image Processing*, *20*(6). https://doi.org/10.1007/s11554-023-01368-7

Lettner, M., Ramoser, H., & Linauer, M. (2012). Detection of Dangerous Goods Placards for Road and Tunnel Safety. *19th ITS World CongressERTICO - ITS EuropeEuropean CommissionITS AmericaITS Asia-Pacific*. https://trid.trb.org/view.aspx?id=1262685

Li, Y., Han, W., Han, Z., Jiang, N., Wang, W., & Huang, J. (2022). Computer Vision-Based Hazard Identification of Construction Site Using Visual Relationship Detection and Ontology. *Buildings*, *12*(6), 857. https://doi.org/10.3390/buildings12060857

Liu, S., Hui, T., Huang, S., Wei, Y., Li, B., & Li, G. (2021). Cross-Modal Progressive Comprehension for Referring Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(9), 1. https://doi.org/10.1109/tpami.2021.3079993

Liu, T., Li, B., Du, X., Jiang, B., Xiao, J., Jin, L., & Zhao, Z. (2023a). Component-aware anomaly detection framework for adjustable and logical industrial visual inspection. In *arXiv (Cornell University)*. Cornell University. https://doi.org/10.48550/arxiv.2305.08509

Liu, T., Li, B., Du, X., Jiang, B., Xiao, J., Jin, L., & Zhao, Z. (2023b). Component-aware anomaly detection framework for adjustable and logical industrial visual inspection. *Advanced Engineering Informatics*, *58*, 102161. https://doi.org/10.1016/j.aei.2023.102161

Liu, X., Jing, X., Zhu, Q., Du, W., & Wang, X. (2023). Automatic Construction Hazard Identification Integrating On-Site Scene Graphs with Information Extraction in Outfield Test. *Buildings*, *13*(2), 377. https://doi.org/10.3390/buildings13020377

Liu, X., Meehan, J., Tong, W., Wu, L., Xu, X., & Xu, J. (2020). DLI-IT: a deep learning approach to drug label identification through image and text embedding. *BMC Medical Informatics and Decision Making*, *20*(1). https://doi.org/10.1186/s12911-020-1078-3

Martino, F. D., & Delmastro, F. (2022). Explainable AI for clinical and remote health applications: a survey on tabular and time series data. *Artificial Intelligence Review*, *56*(6), 5261–5315. https://doi.org/10.1007/s10462-022-10304-3

Moenck, K., Wendt, A., Prünte, P., Koch, J., Sahrhage, A., Gierecker, J., Schmedemann, O., Kähler, F., Holst, D., Gomse, M., Schüppstuhl, T., & Schoepflin, D. (2023). Industrial Segment Anything -- a Case Study in Aircraft Manufacturing, Intralogistics, Maintenance, Repair, and Overhaul. In *arXiv (Cornell University)*. Cornell University. https://doi.org/10.48550/arxiv.2307.12674

Monteiro, G. R., Camelo, L., Aquino, G., Fernandes, R. de A., Gomes, R. C. S., Printes, A. L., Torné, I. G., Silva, H. L. N. da, Oliveira, J. P. de, & Figueiredo, C. M. S. (2023). A Comprehensive Framework for Industrial Sticker Information Recognition Using Advanced OCR and Object Detection Techniques. *Applied Sciences*, *13*(12), 7320. https://doi.org/10.3390/app13127320

N, N. K. H., Kiran, Kotiyal, A., Jagadeesh, B., Shanthamallappa, M., & S, G. P. M. (2023). *Real-Time Automated Detection of Errors in the Product Labels Using Image Processing and OCR*. 1–6. https://doi.org/10.1109/aikiie60097.2023.10390320

Naumann, A., Hertlein, F., Doerr, L., Thoma, S., & Furmans, K. (2023). Literature Review: Computer Vision Applications in Transportation Logistics and Warehousing. In *arXiv (Cornell University)*. Cornell University. https://doi.org/10.48550/arxiv.2304.06009

Nauta, M., & Seifert, C. (2023). The Co-12 Recipe for Evaluating Interpretable Part-Prototype Image Classifiers. In *arXiv (Cornell University)*. Cornell University. https://doi.org/10.48550/arxiv.2307.14517

Nguyen, H.-S., Huynh, C.-D., & Bui, N.-Q. (2023). Digital transformation for shipping container terminals using automated container code recognition. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, *21*(3), 535. https://doi.org/10.12928/telkomnika.v21i3.24137

Pegu, B., Singh, M., Kant, K., Singh, K., & Bhowmik, T. (2020). *MoDest: Multi-module Design Validation for Documents*. 332–340. https://doi.org/10.1145/3430984.3431001

Petković, D. (2023). It is Not "Accuracy vs. Explainability"—We Need Both for Trustworthy AI Systems. *IEEE Transactions on Technology and Society*, *4*(1), 46–53. https://doi.org/10.1109/tts.2023.3239921

Rao, V. N., Zhen, X., Hovsepian, K., & Shen, M. (2021). *A First Look: Towards Explainable TextVQA Models via Visual and Textual Explanations*. 19–29. https://doi.org/10.18653/v1/2021.maiworkshop-1.4

Reshadat, V., Akçay, A., Zervanou, K., Zhang, Y., & Jong, E. de. (2023). SCRE: special cargo relation extraction using representation learning. *Neural Computing and Applications*, *35*(25), 18783–18801. https://doi.org/10.1007/s00521-023-08704-9

Rožanec, J. M., Novalija, I., Zajec, P., Kenda, K., Ghinani, H. T., Suh, S., Bian, S., Veliou, E., Papamartzivanos, D., Giannetsos, T., Menesidou, S. A., Alonso, R., Cauli, N., Meloni, A., Recupero, D. R., Kyriazis, D., Sofianidis, G., Theodoropoulos, S., Fortuna, B., … Soldatos, J. (2022). Human-centric artificial intelligence architecture for industry 5.0 applications. *International Journal of Production Research*, *61*(20), 6847–6872. https://doi.org/10.1080/00207543.2022.2138611

Ruediger-Flore, P., Glatt, M., Hussong, M., & Aurich, J. C. (2023). CAD-based data augmentation and transfer learning empowers part classification in manufacturing. *The International Journal of Advanced Manufacturing Technology*, *125*, 5605–5618. https://doi.org/10.1007/s00170-023-10973-6

Sampaio, I. G. B., Viterbo, J., & Guérin, J. (2023). Improving robustness of industrial object detection by automatic generation of synthetic images from CAD models. *Computational Intelligence*, *39*(3), 415–432. https://doi.org/10.1111/coin.12572

Sewada, R., Jangid, A., Kumar, P., & Mishra, N. (2023). Explainable Artificial Intelligence (XAI). *Journal of Nonlinear Analysis and Optimization*, *13*(1), 41–47. https://doi.org/10.36893/jnao.2022.v13i02.041-047

Shahin, M., Chen, F. F., Hosseinzadeh, A., Koodiani, H. K., Bouzary, H., & Shahin, A. (2023). Enhanced safety implementation in 5S + 1 via object detection algorithms. *The International Journal of Advanced Manufacturing Technology*, *125*, 3701–3721. https://doi.org/10.1007/s00170-023-10970-9

Shajalal, M., Boden, A., & Stevens, G. (2022). Explainable product backorder prediction exploiting CNN: Introducing explainable models in businesses. *Electronic Markets*, *32*(4), 2107–2122. https://doi.org/10.1007/s12525-022-00599-z

Singh, S. P., Burgess, G. J., & Singh, J. (2003). Package performance testing of dangerous goods in high-altitude shipments. *Packaging Technology and Science*, *16*(3), 119–130. https://doi.org/10.1002/pts.619

Tang, S., & Golparvar-Fard, M. (2017). *Joint Reasoning of Visual and Text Data for Safety Hazard Recognition*. 450–457. https://doi.org/10.1061/9780784480847.056

Thota, M., Kollias, S., Swainson, M., & Leontidis, G. (2020). Multi-source domain adaptation for quality control in retail food packaging. *Computers in Industry*, *123*, 103293. https://doi.org/10.1016/j.compind.2020.103293

Tziafas, G., & Kasaei, H. (2022). Enhancing Interpretability and Interactivity in Robot Manipulation: A Neurosymbolic Approach. In *arXiv (Cornell University)*. Cornell University. https://doi.org/10.48550/arxiv.2210.00858

Verifying Attention Robustness of Deep Neural Networks Against Semantic Perturbations. (n.d.). https://doi.org/10.1007/978-3-031-33170-1_3

Wang, X., & El-Gohary, N. (2023). Deep Learning–Based Named Entity Recognition and Resolution of Referential Ambiguities for Enhanced Information Extraction from Construction Safety Regulations. *Journal of Computing in Civil Engineering*, *37*(5). https://doi.org/10.1061/(asce)cp.1943-5487.0001064

Wu, J., Xue, X., & Zhang, J. (2023). Invariant Signature, Logic Reasoning, and Semantic Natural Language Processing (NLP)-Based Automated Building Code Compliance Checking (I-SNACC) Framework. *Journal of Information Technology in Construction*, *28*, 1–18. https://doi.org/10.36680/j.itcon.2023.001

Xu, G., Hao, Z., Luo, Y., Hu, H., An, J., & Mao, S. (2023). DeViT: Decomposing Vision Transformers for Collaborative Inference in Edge Devices. *IEEE Transactions on Mobile Computing*, *23*(5), 5917–5932. https://doi.org/10.1109/tmc.2023.3315138

Yang, J., Chen, K., Ding, K., Na, C., & Wang, M. (2022). Auto Insurance Fraud Detection with Multimodal Learning. *Data Intelligence*, *5*(2), 388–412. https://doi.org/10.1162/dint_a_00191

Yang, J., Fu, G., Zhu, W., Cao, Y., Cao, Y., & Yang, M. Y. (2020). A deep learning-based surface defect inspection system using multi-scale and channel-compressed features. *IEEE Transactions on Instrumentation and Measurement*, 1. https://doi.org/10.1109/tim.2020.2986875

Yao, B., Yang, X., Lin, L., Lee, M. W., & Zhu, S.-C. (2010). I2T: Image Parsing to Text Description. *Proceedings of the IEEE*, *98*(8), 1485–1508. https://doi.org/10.1109/jproc.2010.2050411

Zhang, J., & El-Gohary, N. (2013). Semantic NLP-Based Information Extraction from Construction Regulatory Documents for Automated Compliance Checking. *Journal of Computing in Civil Engineering*, *30*(2). https://doi.org/10.1061/(asce)cp.1943-5487.0000346

Zhang, J., & El-Gohary, N. (2015). Automated Information Transformation for Automated Regulatory Compliance Checking in Construction. *Journal of Computing in Civil Engineering*, *29*(4). https://doi.org/10.1061/(asce)cp.1943-5487.0000427

Zhang, J., & El-Gohary, N. (2016). Integrating semantic NLP and logic reasoning into a unified system for fully-automated code checking. *Automation in Construction*, *73*, 45–57. https://doi.org/10.1016/j.autcon.2016.08.027

Zhang, Q., Wen, Y., Han, D., Zhang, F., Zhou, C., & Xiao, C. (2019). *Ontology-based Construction and Application of Knowledge Graph of Maritime Dangerous Goods*. *21*, 303–308. https://doi.org/10.1109/ictis.2019.8883824

Zhang, R., Bahrami, Z., Feng, K., & Liu, Z. (2023). A Visual and Textual Information Fusion-Based Zero-Shot Framework for Hazardous Material Placard Detection and Recognition. *IEEE Transactions on Artificial Intelligence*, *5*(4), 1755–1768. https://doi.org/10.1109/tai.2023.3298588

Zoghlami, F., Kurrek, P., Jocas, M., Masala, G. L., & Salehi, V. (2020). Design of a Deep Post Gripping Perception Framework for Industrial Robots. *Journal of Computing and Information Science in Engineering*, *21*(2). https://doi.org/10.1115/1.4048204