

COMPARATIVE ANALYSIS OF TRANSFORMER-BASED AND CONVENTIONAL CONVOLUTIONAL NEURAL NETWORK (CNN) MODELS FOR DEFECT DETECTION IN CAST PRODUCTS

***Igwe Lazarus Uduma and Oluwafemi Ayodeji Olugboji**

Department of Mechanical Engineering, Federal University of Technology, Minna, Niger State

*Corresponding authors' email: igwe.m1605340@st.futminna.edu.ng

ABSTRACT

Casting is a widely used manufacturing process which is frequently affected by surface and internal defects that compromise product quality and structural integrity. Conventional inspection methods, such as manual visual inspection, rely heavily on human expertise and are often slow, subjective, and prone to oversight. To address these limitations, this study develops a computer vision system for the detection of defects in cast products using Transformer-based and conventional convolutional neural network (CNN) models. The performance of both models was evaluated in terms of accuracy, precision, recall, specificity, sensitivity, and F1-score. The models were trained on a dataset of 7,348 grayscale images using the Google Colab platform. Experimental results show that the Transformer-based model outperformed the traditional CNN, achieving an accuracy of 98.4%, precision of 96.7%, recall of 94.9%, F1-score of 96.2%, specificity of 95.9%, and sensitivity of 97.8%. The proposed system enhances quality assurance, reduces manufacturing waste, and supports continuous process optimization, offering significant benefits for medium-sized foundries seeking improved efficiency and product reliability.

Keywords: Defect Detection, Cast, Transformer-Based Model, Convolutional Neural Network (CNN), Computer Vision

INTRODUCTION

The foundry industry plays a critical role in modern industrial society by producing cast products such as engine blocks, turbine blades, and structural components that support technological advancement and infrastructure development. Casting remains one of the most widely used manufacturing processes due to its ability to produce metal components with complex geometries, high dimensional accuracy, and desirable mechanical properties. The casting process involves pouring molten metal into a mould cavity, where it solidifies into the required shape (Craft Mach Engineered Solutions Inc., 2023). This manufacturing method is extensively applied in the automotive, aerospace, and heavy machinery industries because it enables the production of strong, durable, and intricately designed components. In the automotive sector, casting is commonly used to manufacture critical components such as engine blocks, cylinder heads, and transmission systems, which require high strength and dimensional precision (Butt et al., 2017). Similarly, the aerospace industry relies on casting to produce lightweight yet high-strength engine and structural components, while heavy machinery manufacturers use casting to fabricate large and durable parts for construction and mining equipment. Given the safety-critical nature of these components, maintaining strict quality standards in cast products is essential. Defects arising during the casting process can compromise structural integrity, reduce service life, and negatively affect product performance.

Despite its advantages, the casting process is susceptible to various surface and internal defects, including blow holes, short runs, shrinkage cavities, and uneven solidification. These defects often result from variations in process parameters such as temperature fluctuations, mould quality, and impurities in molten metal. Detecting and eliminating such defects at an early stage is therefore crucial to ensure compliance with quality and safety requirements. Traditionally, inspection in foundry workshops relies on manual visual inspection conducted by skilled personnel. However, manual inspection methods are time-consuming,

subjective, and prone to human error, particularly in high-volume production environments. Recent advances in computer vision and artificial intelligence have introduced new possibilities for automated defect detection in manufacturing systems. Computer vision techniques, when integrated with machine learning and deep learning algorithms, enable rapid and accurate analysis of cast product images, allowing defects to be detected with high reliability. Deep learning approaches, particularly convolutional neural networks (CNNs), have demonstrated strong performance in surface defect detection tasks within industrial settings (Pierdicca et al., 2020). More recently, Transformer-based architectures have gained increasing attention due to their ability to capture global contextual information and improve feature representation, making them suitable for complex visual inspection applications. The implementation of automated visual inspection systems offers several benefits, including reduced inspection time, improved consistency, lower production costs, and reduced material wastage. These advantages support the transition toward smart manufacturing environments where intelligent systems complement human expertise. The adoption of computer vision-based inspection aligns with the principles of Industry 5.0, which emphasize human-centric, resilient, and sustainable manufacturing systems (Noor-Arham et al., 2022).

In many local foundry industries, including those operating in Afikpo, Ebonyi State, Nigeria, quality inspection is still predominantly performed using manual methods. This practice often results in inefficiencies, production delays, increased waste, and higher operational costs. There is therefore a pressing need for an efficient, and reliable computer vision-based inspection system tailored to the operational realities of small and medium-sized enterprises. Addressing this need can significantly enhance quality assurance processes, reduce material wastage, and improve product reliability in local foundry workshops. This study aims to develop a computer vision system for automated defect detection in cast products and to conduct a comparative performance analysis of deep learning models, specifically

conventional convolutional neural networks and Transformer-based models. The models are evaluated based on accuracy, processing efficiency, and industrial applicability. The proposed system is intended to support quality inspection engineers and improve quality control practices in small and medium-scale foundries within Afikpo, Ebonyi State.

MATERIALS AND METHODS

Data Collection

The dataset used in this research comprises 7,348 grayscale images of cast products collected from the foundry workshop at Akanu Ibiam Federal Polytechnic, Unwana. Image acquisition was performed under stable and controlled lighting conditions to ensure consistency and reduce illumination-induced variability. Figures 2.1, 2.2, and 2.3 show representative examples of defective and non-defective cast samples. Each image was captured as a top-view shot with a spatial resolution of 300×300 pixels. For the purpose of binary classification, the dataset was divided into two classes: defective and non-defective, as summarized in Table

2.1. The defective samples exhibit surface defects, visually characterized by open cavities, pits, and irregular surface conditions. Based on visual inspection, these defects correspond to surface porosity and blowhole-type defects, which are common gas-related anomalies in cast products. Non-defective samples consist of cast products with uniform surface texture and no visible cavities, pits, or surface discontinuities. All images were obtained from a single foundry environment involving similar casting materials, component types, and production conditions. While this controlled setup ensures consistency in image quality, it also implies that the dataset represents a focused inspection scenario rather than broad multi-site industrial variability. Class labels were assigned based on visual inspection of surface characteristics conducted during data collection at the foundry workshop. The final dataset distribution includes 3,758 defective samples and 2,875 non-defective samples, reflecting a slight class imbalance that was preserved to maintain realism with typical inspection outcomes in foundry operations.



Figure 1: Non-Defect Cast Sample



Figure 2: Defective Cast Sample 1



Figure 3: Defective Cast Sample 2

Table 1: Dataset Collection Summary

S/N	Class	Number of Images	Description	Label
1	Defective	3758	Cast images with visible defects	0
2	Non-Defective	2875	Cast images without visible defects	1

Data Preprocessing

During data preprocessing, data augmentation techniques were applied to the training dataset only to increase sample diversity and improve model robustness. Augmentation was employed to mitigate overfitting and enhance the models' ability to generalize to unseen data, particularly given the controlled nature of image acquisition. The augmentation pipeline consisted of a combination of geometric and photometric transformations. Random rotations were applied within a range of ± 15 degrees to simulate variations in specimen orientation during inspection. Horizontal and vertical flipping were applied with a probability of 0.5, accounting for changes in viewing angles commonly encountered in practical inspection scenarios. Image scaling was performed within a range of 0.9 to 1.1 of the original size to reflect minor variations in camera distance and focus. To further improve robustness, Gaussian noise with low variance was added randomly to a subset of training images to simulate sensor noise and surface irregularities. In addition, brightness adjustments were applied by varying pixel intensity within a range of $\pm 20\%$ to account for moderate illumination differences. All augmentation operations were applied on the training dataset, while the validation and test datasets were left unaltered to ensure unbiased performance evaluation and prevent data leakage. Collectively, these augmentation strategies increased the variability of the training data and exposed the models to a broader range of defect appearances and imaging conditions. This preprocessing approach supported stable model training and improved resilience to

minor distortions and environmental variations commonly encountered in foundry inspection environments.

Data Splitting

For the purpose of this study the data was split into 3 main subsets which includes the training set, test set and validation set. The training set was used to fit the model parameters and enable the learning of feature from the data, while the validation set served to fine-tune hyperparameters and prevent overfitting by providing feedback on model performance during training. Finally, the test set was reserved for final evaluation to assess the model's on unseen data. The data splitting followed ratio of 70 % for training, 10 % for validation, and 20 % for testing.

Model Training

The training of the models was carried out in the Google Collaboratory environment because of its free access to GPU resources and its compatibility with TensorFlow and Keras frameworks. This platform provided the required computing power to handle the dataset, to build, train, and evaluate the models efficiently. For this study, two models were selected, namely Convolutional Neural Network (CNN) and Transformer. They were chosen because of their strengths in image classification tasks. CNN was selected for its ability to extract spatial features such as edges, textures and local patterns, which are important in identifying casting defects. The Transformer was selected because of its self-attention mechanism that captures long-range dependencies and global

context in images. This makes it possible to detect subtle or scattered defects that may not be obvious through local features alone. The CNN architecture used in this study was made up of several convolutional layers with ReLU activation, followed by pool as shown in table 2.2 below. To prevent overfitting, dropout and L2 regularization were applied, and early stopping was employed based on validation

performance. Model checkpointing was used to retain the best-performing weights during training. Performance was monitored using accuracy, precision, recall, and F1-score on the validation set. Each epoch required approximately 2.5 minutes, and full training completed in approximately 1 hour and 47 minutes. Both models exhibited stable convergence, with no evidence of training instability.

Table 2: Model Training Specifications

S/N	Hyperparameter	Traditional CNN	Transformers
1.	Batch-Size	64	64
2.	Loss Function/Optimizer	Adam Optimizer	Adam Optimizer
3.	Epoch	50	50
4.	Activation function	Sigmoid function (output layer)	Sigmoid function (output layer)
5.	Network architecture	Inception Net	Attention based
6.	Learning rate	0.0001	0.0001
7.	Numbers of layers	32	47
8.	Number of hidden nodes	128	64
9	Input size	300×300×1	300×300×1
9	Pooling layer	Max pooling	Max pooling
9	Patch size	-	16 × 16
9	Dropout	0.5	-

Model Evaluation

The following metrics were used to evaluate the performance of the transformer based model and the traditional CNN models:

Accuracy

This shows the proportion of correct cast class predictions relative to the total number of input cast samples.

$$\text{Accuracy} = \frac{\text{Number of Correct casting class predictions}}{\text{Total number of cast predictions made}} \quad (1)$$

Recall

Recall, also known as sensitivity or true positive rate measures the model's ability to capture and correctly classify all positive instances in a dataset. It is calculated by dividing the number of correctly identified positive samples by the total number of actual positive samples that should have been recognized

$$\text{Sensitivity} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (2)$$

Precision

Precision quantifies how many of the positive predictions made by the model are actually correct. It is calculated by dividing the number of correctly identified positive samples

by the total number of positive samples that the classifier predicted.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (3)$$

F1-Score

F1-score is a metric used to evaluate the overall accuracy of the cast classification test. It represents the harmonic mean between recall and precision, providing insight into both the precision and recall of the classification. The F1-score is particularly useful when you want to balance the trade-off between precision and recall. If a model has high precision but low recall, it implies that it achieves high accuracy but may miss a significant number of relevant instances. Conversely, a high recall but low precision indicates that it captures many relevant instances but may also include a large number of false positives. It can be calculated using the following formula:

$$\text{Accuracy} = 2 * \frac{1}{\frac{1}{\text{precision}} + \frac{1}{\text{sensitivity}}} \quad (4)$$

RESULTS AND DISCUSSION

The performance of the model was evaluated on the test dataset. Table 4.2 below shows a summary of the results :

Table 3: Comparison Between Traditional CNN and Transformers

Evaluated Metrics	Traditional CNN	Transformer Model
Accuracy	95.6387	98.4194
Precision	0.9432	0.9667
F1-score	0.9523	0.9617
Recall	0.9317	0.9491
Specificity	0.9338	0.9594

In terms of accuracy, the transformer model outperformed the traditional CNN, achieving an accuracy rate of 98.42%, compared to 95.64%, resulting in a performance improvement of about 2.5%. The precision metric, which measures the accuracy of positive casting defect predictions, also favored the transformer model, with a precision of 0.9667, surpassing the traditional CNN's precision of 0.9432 and resulting in an improvement of approximately 2.2%. Similarly, the F1-score, a harmonic mean of precision and recall, demonstrated the

transformer model's superiority with a score of 0.9617, while the traditional CNN achieved a slightly lower F1-score of 0.9523, indicating an improvement of about 1%. Additionally, the recall metric, which assesses the model's capability to capture true positive defect classifications, also favored the transformer model with a recall of 0.9491, outperforming the traditional CNN's recall of 0.9317, leading to an improvement of about 1.7%.

The specificity metrics, which measure the model's ability to correctly identify defective and non-defective instances, respectively, also favored the transformer model. It achieved a specificity of 0.9594, outperforming the traditional CNN's specificity of 0.9338 with an improvement of about 2.0% in specificity. Collectively, these performance metrics highlight the enhanced accuracy, sensitivity, and precision of the

transformer model, showing its efficacy in the casting defect classification task compared to the traditional CNN.

The confusion matrix in Figure 4.2 shows that the model performed very well in distinguishing between defective and non-defective samples, as shown by the high values of True Positives and True Negatives.

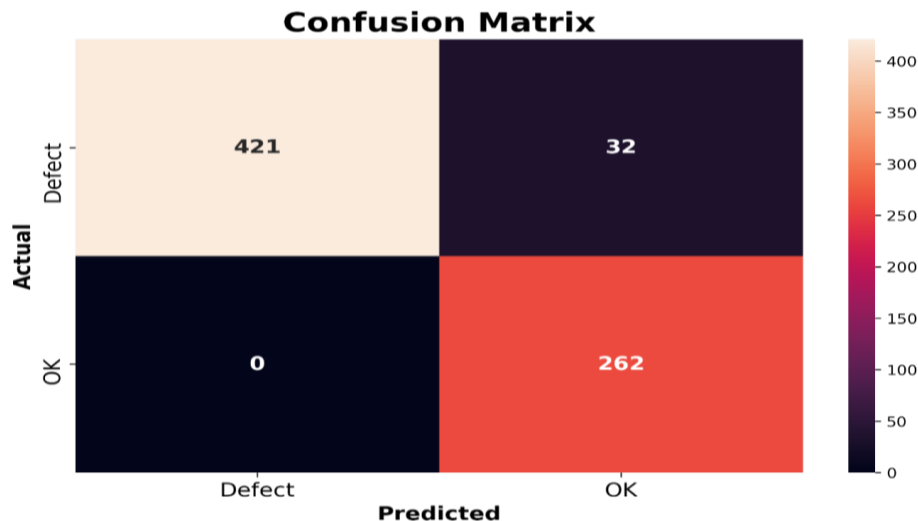


Figure 4: System Confusion Matrix Plot

The confusion matrix shown in the results section provides insights into the performance of the model by presenting a summary of prediction outcomes across the two classes.

True Positives (TP): The top-left cell with a value of 421 indicates the number of defective samples correctly classified as defective.

False Positives (FP): The top-right cell with a value of 32 represents the non-defective samples incorrectly classified as defective. This is also known as a Type I error.

False Negatives (FN): The bottom-left cell has a value of 0, indicating that there were no defective samples misclassified as non-defective. This result implies that the model perfectly avoided Type II errors.

True Negatives (TN): The bottom-right cell, with a value of 262, shows the number of non-defective samples correctly classified as non-defective.

The absence of False Negatives (0) is particularly noteworthy, as it means the model did not miss any defective samples, which is crucial in quality control where missing defective items could lead to significant losses. However, there is a moderate number of False Positives (32), which indicates that some non-defective items were mistakenly classified as defective, possibly leading to unnecessary re-inspection or rejection of quality products. This result is highly favorable for an automated inspection system, as it minimizes the risk of defective items passing through quality control while maintaining a high accuracy rate.

Discussion

The results of this study indicate that transformer-based models can provide improved performance for automated casting defect detection when compared with traditional convolutional neural networks under the evaluated experimental conditions. The transformer model achieved an overall classification accuracy of 98.4%, demonstrating its potential effectiveness for visual inspection tasks in controlled foundry environments. While this level of performance is encouraging, it should be interpreted in the

context of the dataset characteristics and evaluation setup. Compared to the traditional CNN model, the transformer-based approach consistently achieved higher values across all evaluated metrics, including accuracy, precision, recall, sensitivity, and specificity. Although the observed improvements are relatively modest in magnitude, their consistency suggests that the transformer architecture offers a more balanced classification behavior. This can be attributed to the self-attention mechanism employed by transformers, which enables the model to capture global contextual information across the cast surface. Such capability is particularly relevant for surface-related defects, such as porosity and blowholes, which may appear as spatially dispersed or irregular patterns that are not easily captured using purely local feature extraction. The absence of false negative predictions observed for the transformer model in the evaluated test split indicates that, under the controlled conditions of this experiment, defective samples were not misclassified as non-defective. From a quality control perspective, this is a desirable outcome, as false negatives pose a higher operational risk than false positives. However, this result should not be interpreted as a guarantee of error-free performance, as deep learning models are sensitive to data variation, initialization, and operating conditions. Further evaluation using multiple data splits or larger and more diverse datasets would be necessary to confirm the robustness of this behavior in real-world deployment scenarios. Overall, the results suggest that the observed performance gains are primarily driven by architectural differences rather than dataset-specific artifacts.

CONCLUSION

This study investigated the application of computer vision based deep learning models for automated defect detection in cast manufacturing, with particular emphasis on comparing transformer-based architectures and conventional convolutional neural networks. The experimental evaluation, conducted on a dataset of 7,348 grayscale images collected

under controlled conditions, indicates that the transformer-based model achieved consistently higher performance than the traditional CNN across multiple classification metrics, including accuracy, precision, recall, sensitivity, and specificity. The transformer model attained an overall accuracy of 98.4%, suggesting its potential suitability for visual inspection tasks in controlled foundry environments. While the observed performance improvements are relatively modest in magnitude, their consistency across evaluation metrics indicates that transformer architectures can provide a more balanced classification behavior for surface-related casting defects, such as porosity and blowholes. These findings support the potential of data-driven inspection systems to complement manual inspection practices by improving consistency and reducing subjectivity in quality assessment, particularly in medium-scale foundry operations. From a practical standpoint, the results suggest that automated inspection systems based on deep learning may contribute to improved quality control processes and reduced material waste when deployed under appropriate operating conditions. However, the findings of this study are limited to the evaluated dataset and experimental setup, and further validation using larger and more diverse datasets would be required to establish robustness and generalizability in real-world industrial scenarios. Future work may focus on enhancing system applicability by integrating automated material handling mechanisms, such as conveyor-based inspection, incorporating controlled LED illumination to stabilize imaging conditions, and expanding locally sourced datasets through collaboration with small and medium-scale foundries. Increasing dataset diversity and operational variability would further support the evaluation of model robustness and facilitate broader adoption of intelligent inspection systems in industrial manufacturing.

REFERENCES

- Adi, K., Pujiyanto, S., Nurhayati, O. D., & Pamungkas, A. (2017). Beef quality identification using thresholding method and decision tree classification based on Android smartphone. *Journal of Food Quality*, 2017, 1–10. <https://doi.org/10.1155/2017/1674718>
- Adi, K., Pujiyanto, S., Nurhayati, O. D., & Pamungkas, A. (2017). Beef marbling identification using color analysis and decision tree classification. *Advanced Science Letters*, 23(7), 6618–6622. <https://doi.org/10.1166/asl.2017.9697>
- Akanmu, A. A., Anumba, C. J., & Ogunseju, O. O. (2021). Towards next-generation cyber-physical systems and digital twins for construction. *Journal of Information Technology in Construction*, 26, 505–525. <https://doi.org/10.36680/j.itcon.2021.027>
- Alaa, K. A., Kotb, Y., & Kotb, A. (2024). Automated detection of defects on metal surfaces using vision transformers. *arXiv preprint arXiv:2410.04440*. <https://doi.org/10.48550/arXiv.2410.04440>
- Alaiz-Rodríguez, R., & Parnell, A. C. (2020). A machine learning approach for lamb meat quality assessment using FTIR spectra. *IEEE Access*, 8, 52385–52394. <https://doi.org/10.1109/ACCESS.2020.2974623>
- Avian, C., Leu, J. S., Prakosa, S. W., & Faisal, M. (2022). An improved classification of pork adulteration in beef based on electronic nose using modified deep extreme learning with principal component analysis. *Food Analytical Methods*, 15(11), 3020–3031. <https://doi.org/10.1007/s12161-022-02361-9>
- Baduge, S. K., Thilakarathna, S., Perera, J. S., Arashpour, M., Sharafi, P., Teodosio, B., Shringi, A., & Mendis, P. (2022). Artificial intelligence and smart vision for building and construction 4.0: Machine and deep learning methods and applications. *Automation in Construction*, 141, 104440. <https://doi.org/10.1016/j.autcon.2022.104440>
- Bellamy, R. K. E., Dey, K., Hind, M., Hoffman, S. C., Houde, S., Kannan, K., Lohia, P., Martino, J., Mehta, S., Mojsilovic, A., Nagar, S., Ramamurthy, K. N., Richards, J., Saha, D., Sattigeri, P., Singh, M., Varshney, K. R., & Zhang, Y. (2019). AI fairness 360: An extensible toolkit for detecting and mitigating algorithmic bias. *IBM Journal of Research and Development*, 63(4/5), 4:1–4:15. <https://doi.org/10.1147/JRD.2019.2942287>
- Bila, S., Fitrianto, A., & Sartono, B. (2021). Image classification of beef and pork using convolutional neural network in Keras framework. *International Journal of Science, Engineering and Information Technology*, 5(2), 245–248. <https://doi.org/10.21107/ijseit.v5i02.9864>
- Butt, S. I., Asgher, U., Mushtaq, U., Ahmed, R., Zhang, F., Ayaz, Y., Jamil, M., & Amjad, M. K. (2017). Intelligent machine vision-based modeling and positioning system in sand casting process. *Advances in Materials Science and Engineering*, 2017, 1–11. <https://doi.org/10.1155/2017/3192672>
- Cáceres, C., & Selling, B. (1996). Casting defects and the tensile properties of an AlSiMg alloy. *Materials Science and Engineering A*, 220(1–2), 109–116. [https://doi.org/10.1016/S0921-5093\(96\)10433-0](https://doi.org/10.1016/S0921-5093(96)10433-0)
- CraftMach Engineered Solutions Inc. (2023). Why metal casting parts are preferred for high-performance applications across all major industries. CraftMach. <https://craftmach.com/why-metal-casting-parts-are-preferred-for-high-performance-applications-across-all-major-industries/>
- Delgado, J. M. D., Oyedele, L., Demian, P., & Beach, T. (2020). A research agenda for augmented and virtual reality in architecture, engineering and construction. *Advanced Engineering Informatics*, 45, 101122. <https://doi.org/10.1016/j.aei.2020.101122>
- Di Leo, G., Liguori, C., Pietrosanto, A., & Sommella, P. (2017). A vision system for the online quality monitoring of industrial manufacturing. *Optics and Lasers in Engineering*, 89, 162–168. <https://doi.org/10.1016/j.optlaseng.2016.05.007>
- Ferguson, M., Ak, R., Lee, Y. T., & Law, K. H. (2017). Automatic localization of casting defects with convolutional neural networks. In *Proceedings of the IEEE International Conference on Big Data* (pp. 1726–1735). IEEE. <https://doi.org/10.1109/BigData.2017.8258115>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., & Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern*

Recognition, 77, 354–377.
<https://doi.org/10.1016/j.patcog.2017.10.013>

Gupta, R., Anand, V., Gupta, S., & Koundal, D. (2023). Deep learning model for defect analysis in industry using casting images. *Expert Systems with Applications*, 232, 120758. <https://doi.org/10.1016/j.eswa.2023.120758>

Ibrahim, H., & Tapamo, J.-R. O. (2024). A survey of vision-based methods for surface defect detection and classification in steel products. *Informatics*, 11(2), 25. <https://doi.org/10.3390/informatics11020025>

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. <https://doi.org/10.1145/3065386>

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>

Noor-A-Rahim, M., Firyaguna, F., John, J., Khyam, M. O., Pesch, D., Armstrong, E., Claussen, H., & Poor, H. V. (2022). Toward industry 5.0: Intelligent reflecting surface in smart manufacturing. *IEEE Communications Magazine*, 60(10), 72–78. <https://doi.org/10.1109/MCOM.001.2200016>

Pierdicca, R., Paolanti, M., Felicetti, A., Piccinini, F., & Zingaretti, P. (2020). Automatic faults detection of photovoltaic farms: SolAIr, a deep learning-based system for thermal images. *Energies*, 13(24), 6496. <https://doi.org/10.3390/en13246496>

Tang, Z., Tian, E., Wang, Y., Wang, L., & Yang, T. (2021). Nondestructive defect detection in castings using spatial attention bilinear convolutional neural network. *IEEE Transactions on Industrial Informatics*, 17(1), 82–89. <https://doi.org/10.1109/TII.2020.2985159>

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (pp. 5998–6008).

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: An overview and application in radiology. *Insights into Imaging*, 9(4), 611–629. <https://doi.org/10.1007/s13244-018-0639-9>

Zhou, L., Zhang, L., & Konz, N. (2023). Computer vision techniques in manufacturing. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 53(1), 105–117. <https://doi.org/10.1109/TSMC.2022.3166397>

