



JOURNAL OF SCIENCE, TECHNOLOGY AND EDUCATION (JSTE)



**A PUBLICATION OF THE
DEPARTMENT OF SCIENCE,
TECHNOLOGY AND MATHEMATICS
EDUCATION (STME),
NASARAWA STATE UNIVERSITY, KEFFI, NIGERIA**



VOLUME 10

ISSN: 2651-5539

A COMPARATIVE EVALUATION OF AN ENHANCED LIGHTWEIGHT IoT-BASED PIPELINE LEAK DETECTION MODEL AGAINST INCEPTIONV3, MOBILENETV2, AND RESNET-50 USING DEEP LEARNING AND THERMAL IMAGING

Ayuba, A.,^{1*} Gambo, F. L.,² Abdullahi, A. A.,¹ and Muhammad, A. U.¹

¹ Department of Computer Science, Federal University Dutse, Nigeria.

² Department of CyberSecurity, Federal University Dutse, Nigeria.

Corresponding author: ayubaabida@gmail.com

Citation: Ayuba, A., Gambo, F. L., Abdullahi, A. A., & Muhammad, A. U. 1. (2026). A comparative evaluation of an enhanced lightweight IoT-based pipeline leak detection model against InceptionV3, MobileNetV2, and ResNet-50 using deep learning and thermal imaging. *Journal of Science, Technology, and Education (JSTE)*; www.nsukjste.com/. 10(28), 362–374.

Abstract

Pipeline infrastructure is vital to the global energy sector, yet leakages remain a persistent threat with severe environmental and economic consequences. This study compares an enhanced lightweight deep learning model for IoT-based oil pipeline leak detection with three established architectures: InceptionV3, MobileNetV2, and ResNet-50. The proposed model integrates Convolutional Neural Networks (CNN), Knowledge Distillation (KD), and an Autoencoder (AE) to achieve high detection accuracy while maintaining a compact model size suitable for edge device deployment. A locally collected thermal image dataset of 1,506 images, augmented to 9,036 samples, was used for training and evaluation.

Performance was assessed at 10 and 20 training epochs using accuracy, precision, recall, and F1-score. The proposed CNN + KD + AE model consistently outperformed all baseline models at both epoch milestones, achieving an accuracy, precision, recall, and F1-score of 0.98 with a compact model size of 4.3 MB, compared to 8.49 MB for MobileNetV2 and 10.63 MB for ResNet-50. These results confirm the suitability of the proposed framework for real-time, resource-constrained pipeline monitoring applications.

Keywords: Pipeline leak detection, convolutional neural network, knowledge distillation, autoencoder, thermal imaging, IoT, edge computing.

Introduction

Pipeline infrastructure serves as a critical artery for transporting crude oil, natural gas, and water across vast distances. Its integrity is fundamental to national economic stability, public safety, and environmental sustainability. Pipeline leaks, whether caused

by corrosion, material fatigue, mechanical damage, or deliberate sabotage, pose severe risks including environmental degradation, resource loss, and potential loss of life (Maia et al., 2024). Conventional detection methods such as pressure drop analysis, acoustic emission sensing, and visual inspection are

often retrospective and limited in their ability to detect gradual or small-scale leaks in real time (Yuan et al., 2023; Singh & Patel, 2023). Recent advances in the Internet of Things (IoT) and deep learning have opened new possibilities for intelligent, automated pipeline monitoring. Thermal imaging has emerged as a particularly effective modality for detecting temperature anomalies associated with fluid leakage (Maia et al., 2024). When combined with CNN architectures, thermal images provide a rich basis for automatic leak classification. However, the computational burden of state-of-the-art models limits deployment on resource-constrained edge devices. Model compression strategies, including Knowledge Distillation (KD) and Autoencoder (AE)-based compression, have been proposed to address this challenge (Correa da Silva & Almeida, 2024; Farooq et al., 2023), yet comparative evaluations of lightweight hybrid frameworks against established pre-trained architectures in this domain remain limited.

Pre-trained models such as InceptionV3, MobileNetV2, and ResNet-50 are widely applied to image classification but have not been rigorously benchmarked against lightweight hybrid alternatives in thermal pipeline leak detection. This study addresses this gap by developing a unified CNN + KD + AE framework and systematically comparing its performance against all three

baselines at 10 and 20 training epochs. InceptionV3, MobileNetV2, and ResNet-50 were selected as baselines because they represent three distinct design philosophies: multi-scale parallel feature extraction, efficiency-optimised depth-wise separable convolutions, and residual skip-connection-based deep training, respectively, providing a thorough and diverse benchmark for evaluation.

Research Questions

This study is guided by the following research questions, which aim to address the challenges of improving accuracy, efficiency, and real-time performance in IoT-based pipeline leak detection systems:

- (i) How can a thermal image dataset for pipeline leak detection be effectively collected and organized?
- (ii) How can a lightweight CNN be integrated with Knowledge Distillation (KD) and an Autoencoder (AE) be developed for IoT edge devices?
- (iii) How does the proposed CNN + KD + AE model compare with InceptionV3, MobileNetV2, and ResNet-50 in terms of accuracy, precision, recall, F1-score, and model size?
- (iv) How does the performance of the proposed model vary at 10 and 20

training epochs in terms of consistency and generalization?

Aim and Objectives

The aim of this study is to develop and comparatively evaluate a lightweight deep learning model for IoT-based pipeline leak detection using thermal imaging, with emphasis on accuracy, computational efficiency, and edge deployment suitability. The objectives are to:

- (i) collect and organize a thermal image dataset;
- (ii) develop a lightweight CNN + KD + AE model suitable for IoT edge devices;
- (iii) comparatively evaluate the proposed model against InceptionV3, MobileNetV2, and ResNet-50 using accuracy, precision, recall, F1-score, and model size; and
- (iv) assess performance at 10 and 20 training epochs to evaluate consistency and generalization.

Literature Review

Deep learning has become foundational in pipeline leak detection. Kim et al. (2025) proposed a CNN-LSTM-genetic algorithm hybrid achieving 99.69% accuracy on acoustic emission data, though the approach demands substantial compute and is unsuitable for thermal imaging. Yuan et al. (2023) introduced an attention-based LSTM for pressure-based leak localization with an

AUC of 0.99, but without thermal data or edge optimization. Rao et al. (2023) and Singh and Patel (2023) applied YOLOv5 to thermographic leak detection with high precision, while Xiong et al. (2022) achieved 95.33% accuracy classifying thermal anomalies from underground gas micro-leaks. Ekong et al. (2023) combined IoT sensors with CNN for 97% detection accuracy, though edge deployment was not explored. These studies demonstrate strong performance but reveal a consistent gap: high-accuracy architectures are rarely optimized for edge deployment, and lightweight alternatives are seldom benchmarked systematically against established baselines within the thermal pipeline domain.

InceptionV3 (Szegedy et al., 2016) applies parallel convolution filters of varying sizes (1×1 , 3×3 , 5×5) to capture multi-scale spatial features, supplemented by filter factorization and 1×1 bottleneck convolutions for efficiency. With approximately 23.9 million parameters, its strength is rich multi-scale feature extraction, though its high memory footprint constrains edge deployment. MobileNetV2 (Sandler et al., 2018) uses inverted residual blocks with linear bottlenecks and depth wise separable convolutions, achieving competitive accuracy with approximately 3.4 million parameters (8.49 MB). Correa da Silva and Almeida (2024) demonstrated its viability for

real-time thermal classification on Raspberry Pi, though it can underperform on complex datasets. ResNet-50 (He et al., 2016) introduced skip connections to address the vanishing gradient problem, enabling reliable training of deep 50-layer networks with approximately 25.6 million parameters (10.63 MB custom variant). Its principal limitation is high model size and inference cost for edge hardware.

Knowledge distillation (Hinton et al., 2015) trains a compact student to replicate a larger teacher's soft probability outputs, preserving predictive capability at reduced cost. Farooq et al. (2023) demonstrated 4 ms inference on Raspberry Pi using a distilled thermal model. An autoencoder complements distillation by compressing inputs through a bottleneck encoder-decoder, reducing dimensionality and filtering noise, with Dang et al. (2023) confirming high-fidelity image compression via convolutional autoencoders. IoT communication protocols such as LoRaWAN (Kheir & Shaaban, 2024) and NB-IoT (Shi et al., 2022) support pipeline monitoring but impose bandwidth constraints that make image compression essential. Three key gaps motivate this study: high-accuracy models rarely address edge feasibility; InceptionV3, MobileNetV2, and ResNet-50 have not been simultaneously evaluated in a single thermal

pipeline framework; and combined KD and AE compression for thermal pipeline monitoring remains underexplored (Dang et al., 2023; Sun et al., 2023).

Methodology

The dataset consists of 1,506 thermal images collected locally using a FLIR C3 thermal camera under both leak and non-leak pipeline conditions. Images were captured at 320×240 pixels in JPEG format, distributed across 506 leak and 1,000 non-leak samples. Each image was augmented into six variations using rotation, zooming, horizontal flipping, width shift, height shift, and rescaling, expanding the dataset to 9,036 images. Pixel intensity values were normalised to [0, 1] by dividing by 255.

The proposed system integrates a CNN for spatial feature extraction, a KD framework for transferring representations from a large teacher model to a compact student, and an AE for image compression and noise reduction. Thermal images are first compressed by the AE before being classified by the CNN student model. This pipeline enables accurate real-time inference on resource-constrained edge devices such as the Raspberry Pi 4 Model B. Figure 1 shows the system schematic, and Figure 2 shows the training framework.

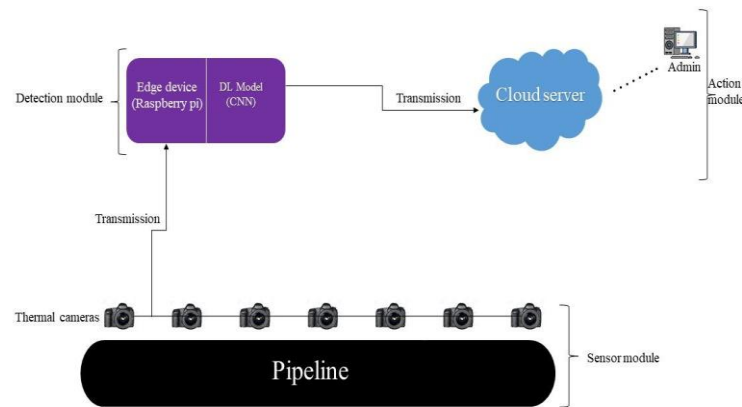


Figure 1: Schematic diagram of the proposed CNN + KD + AE system.

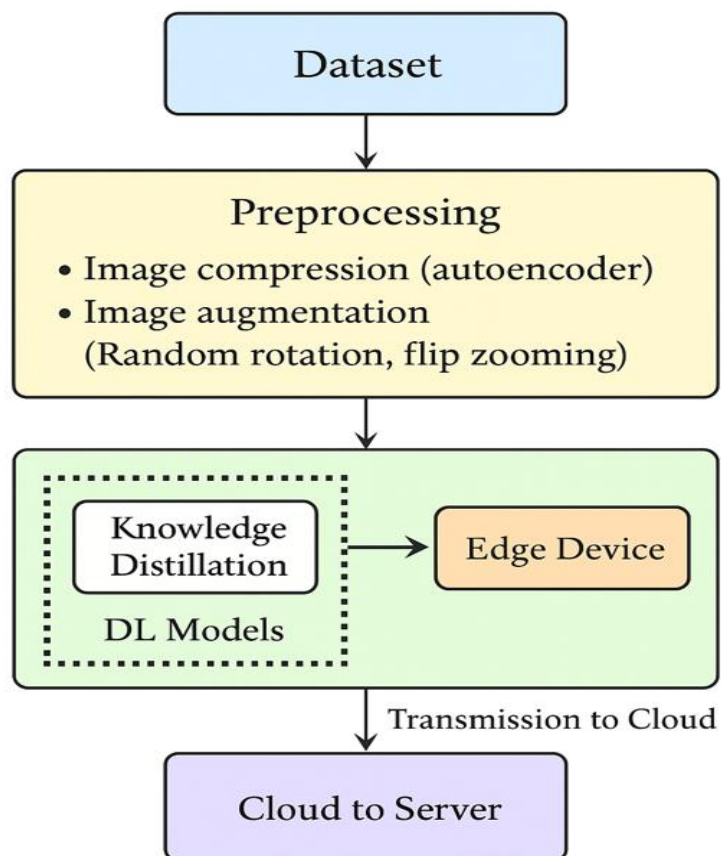


Figure 2: Proposed system framework showing the training and deployment workflow

Let $x \in R^{(H \times W \times C)}$ be an input thermal image. The autoencoder encodes x into $z = E_{\phi}(x)$ and reconstructs $\hat{x} = D_{\psi}(z)$, minimizing $L_{AE} = \|x - \hat{x}\|^2$. The CNN classifier f_{θ} produces logits $s = f_{\theta}(z)$, with cross-entropy loss $L_{CE} = -\sum y_i \log(\text{softmax}(s_i))$. Knowledge distillation

transfers the teacher's soft outputs via $L_{KD} = T^2 \cdot KL(\sigma(t/T) \parallel \sigma(s/T))$, where t and s are teacher and student logits and T is the temperature. The total loss is $L_{total} = \lambda_{AE} \cdot L_{AE} + \lambda_{CE} \cdot L_{CE} + \lambda_{KD} \cdot L_{KD}$. Figure 3 shows the training flowchart.

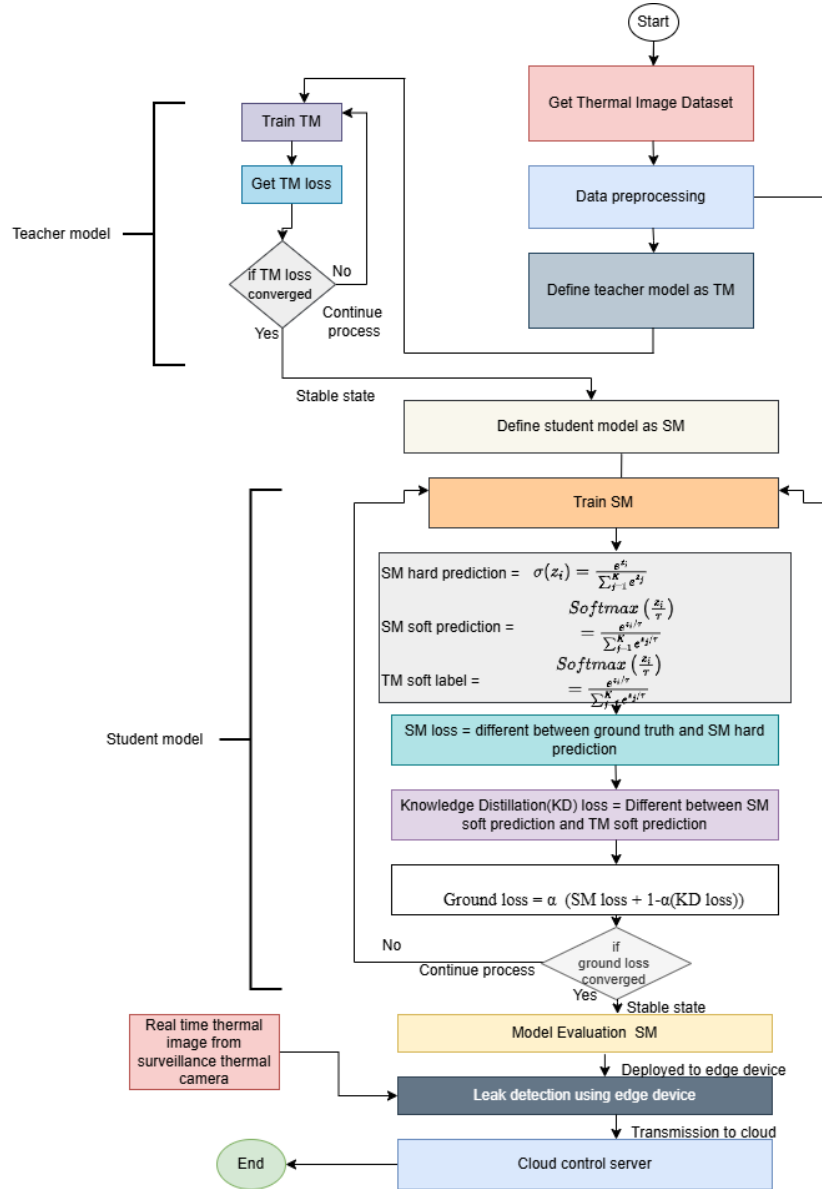


Figure 3: System flowchart illustrating the teacher-student training process.

All models were trained using the Adam optimizer with a learning rate of 0.001, a batch size of 32, and binary cross-entropy

loss, with a 70/15/15 train/validation/test split. The teacher CNN has 52,711,489 parameters (201.08 MB); the student has

126,273 parameters (493.25 KB). The proposed model achieves a deployable size of 4.3 MB. Performance was evaluated at 10 and 20 epochs using: Accuracy = $(TP + TN)/(TP + TN + FP + FN)$; Precision =

$TP/(TP + FP)$; Recall = $TP/(TP + FN)$; F1-Score = $2 \times (Precision \times Recall)/(Precision + Recall)$. Model size (MB) was additionally reported as a proxy for edge deployment feasibility.

Results

4.1 Comparative Performance Evaluation at 10 Epochs

Table 1 presents the comparative performance of all models at 10 epochs. The proposed CNN + KD + AE model achieved the highest accuracy, precision, recall, and F1-score of 0.98 while maintaining the

smallest model size at 4.3 MB. MobileNetV2 performed competitively at 0.97 accuracy. InceptionV3 and ResNet-50 recorded 0.88 and 0.87, respectively, indicating slower convergence at this early stage.

Table 1: Comparative Performance Evaluation of All Models at 10 Epochs

S/N	Model	Accuracy	Precision	Recall	F1-Score	Model Size
1	MobileNetV2	0.97	0.98	0.97	0.97	8.49 MB
2	InceptionV3	0.88	0.89	0.88	0.87	8 KB
3	Residual Network (ResNet-50)	0.87	0.86	0.88	0.87	10.63 MB
4	Proposed (CNN + KD + AE)	0.98	0.98	0.98	0.98	4.3 MB

Figures 4 to 7 present the training and validation accuracy and loss curves for all models at 10 epochs. The proposed model demonstrated rapid convergence with minimal gap between training and validation

curves. InceptionV3 and ResNet-50 showed greater fluctuation in validation loss. MobileNetV2 showed stable but slightly lower validation accuracy than the proposed model.

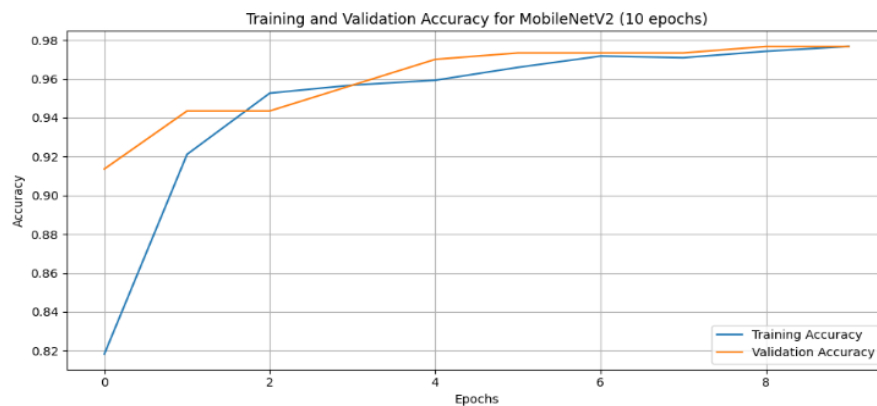


Figure 4: Training and validation curves — MobileNetV2 (10 epochs).

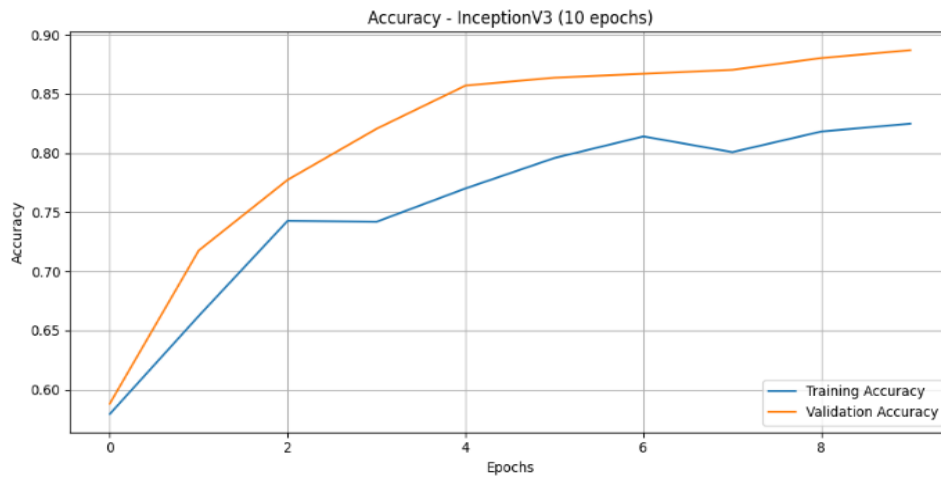


Figure 5: Training and validation curves — InceptionV3 (10 epochs).

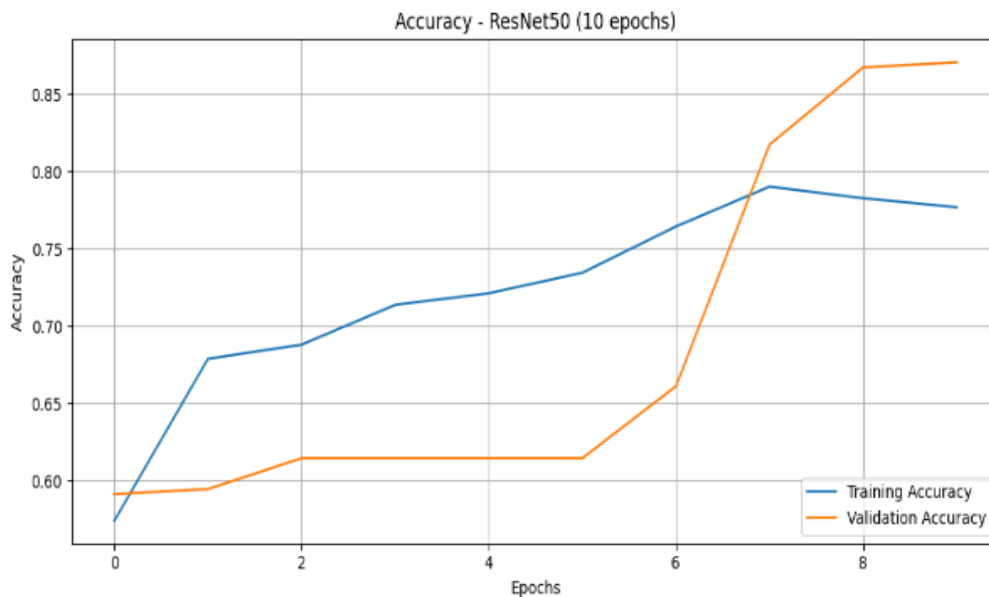


Figure 6: Training and validation curves — ResNet-50 (10 epochs).

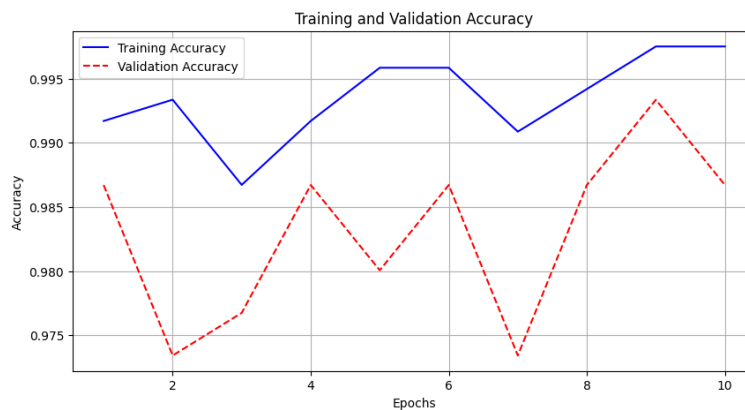


Figure 7: Training and validation curves-Proposed Model (10 epochs).

Figure 8 presents the confusion matrix grid at 10 epochs (test set: 185 non-leak, 116 leak).

The proposed model recorded only 2 misclassifications per class. MobileNetV2

misclassified 1 non-leak and 6 leak samples. InceptionV3 misclassified 28 non-leak and 6 leak samples, while ResNet-50 misclassified 32 non-leak and 7 leak samples. The

proposed model's near-perfect class separation confirms the effectiveness of the CNN + KD + AE framework.

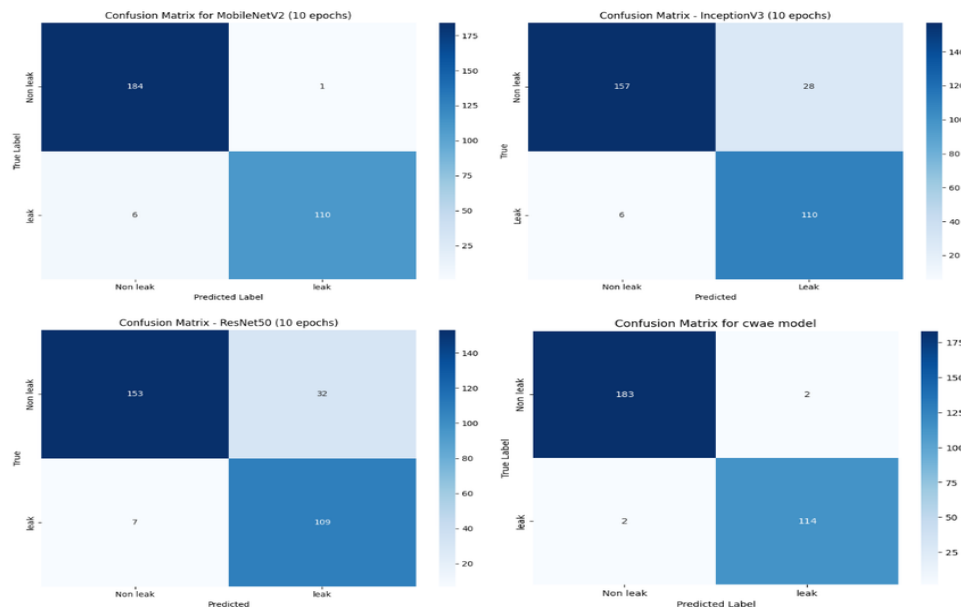


Figure 8: Confusion matrices at 10 epochs (top-left: MobileNetV2; top-right: InceptionV3; bottom-left: ResNet-50; bottom-right: Proposed Model).

4.2 Comparative Performance Evaluation at 20 Epochs

Table 2 presents the comparative performance at 20 epochs. The proposed model again achieved 0.98 across all metrics at 4.3 MB. MobileNetV2 improved to 0.92,

InceptionV3 reached 0.92, and ResNet-50 recorded 0.90. The proposed model not only converges faster but sustains superior performance as training progresses.

Table 2: Comparative Performance Evaluation of All Models at 20 Epochs

S/N	Model	Accuracy	Precision	Recall	F1-Score	Model Size
1	MobileNetV2	0.92	0.93	0.92	0.92	8.49 MB
2	InceptionV3	0.92	0.93	0.93	0.93	8 KB
3	Residual Network (ResNet-50)	0.90	0.91	0.89	0.90	10.63 MB
4	Proposed (CNN + KD + AE)	0.98	0.98	0.98	0.98	4.3 MB

Figures 9 to 12 show accuracy and loss curves at 20 epochs. The proposed model's curves remained tightly aligned between training and validation, confirming robust

generalization. MobileNetV2 and InceptionV3 showed gradual improvement, while ResNet-50 continued to exhibit greater variability.

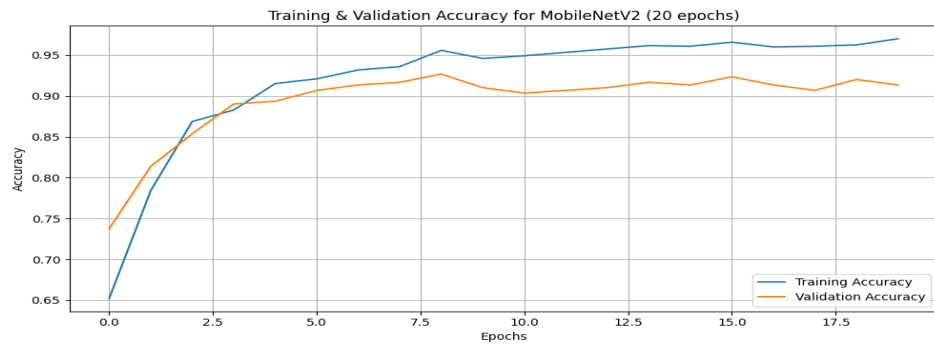


Figure 9: Training and validation curves — MobileNetV2 (20 epochs).

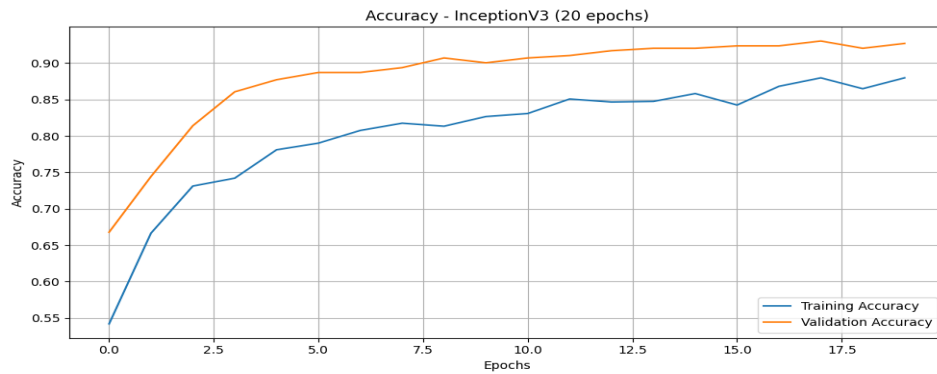


Figure 10: Training and validation curves — InceptionV3 (20 epochs).

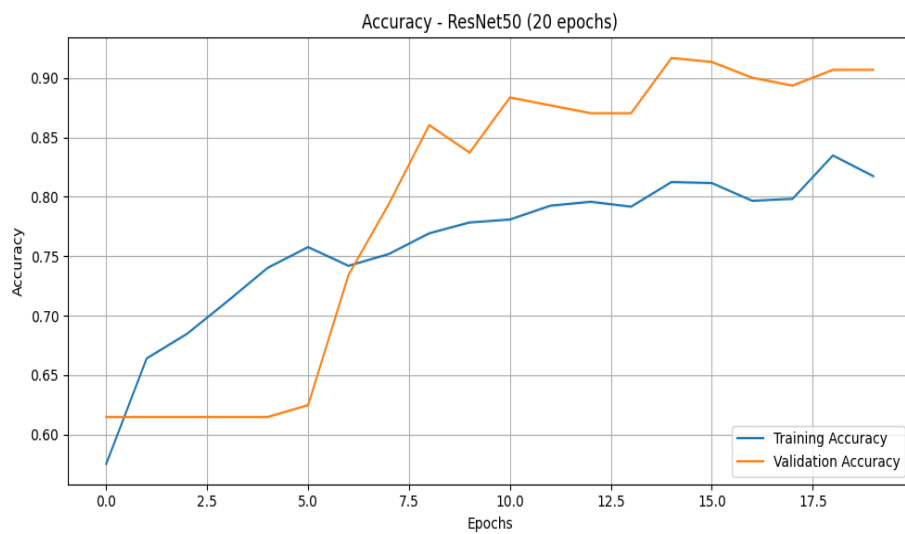


Figure 11: Training and validation curves — ResNet-50 (20 epochs).

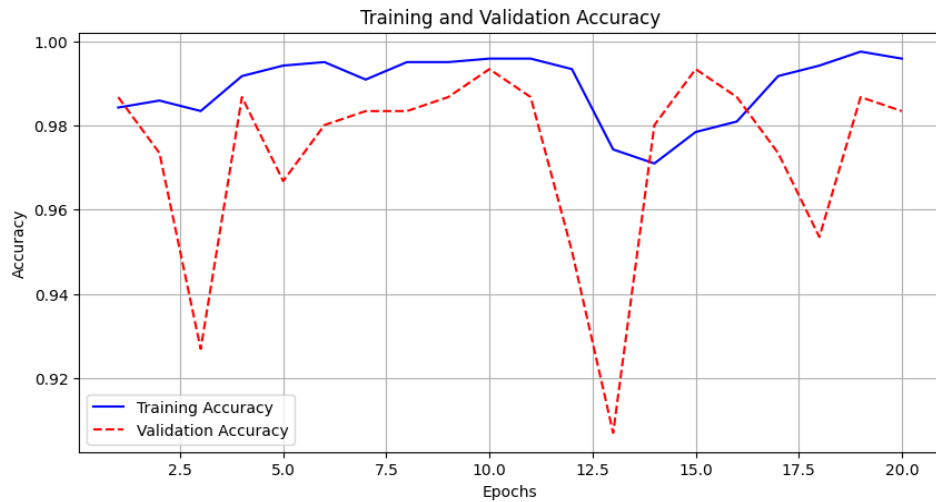


Figure 12: Training and validation curves — Proposed Model (20 epochs).

Figure 13 presents the confusion matrix grid at 20 epochs. The proposed model misclassified only 4 non-leak and 1 leak sample—the fewest across all models. MobileNetV2 misclassified 0 non-leak but 23 leak samples, reflecting a strong bias toward the non-leak class that is particularly

dangerous in a safety-critical monitoring context. InceptionV3 misclassified 16 non-leak and 6 leak samples, while ResNet-50 misclassified 11 non-leak and 17 leak samples. The proposed model's balanced error distribution demonstrates superior discrimination between both classes.

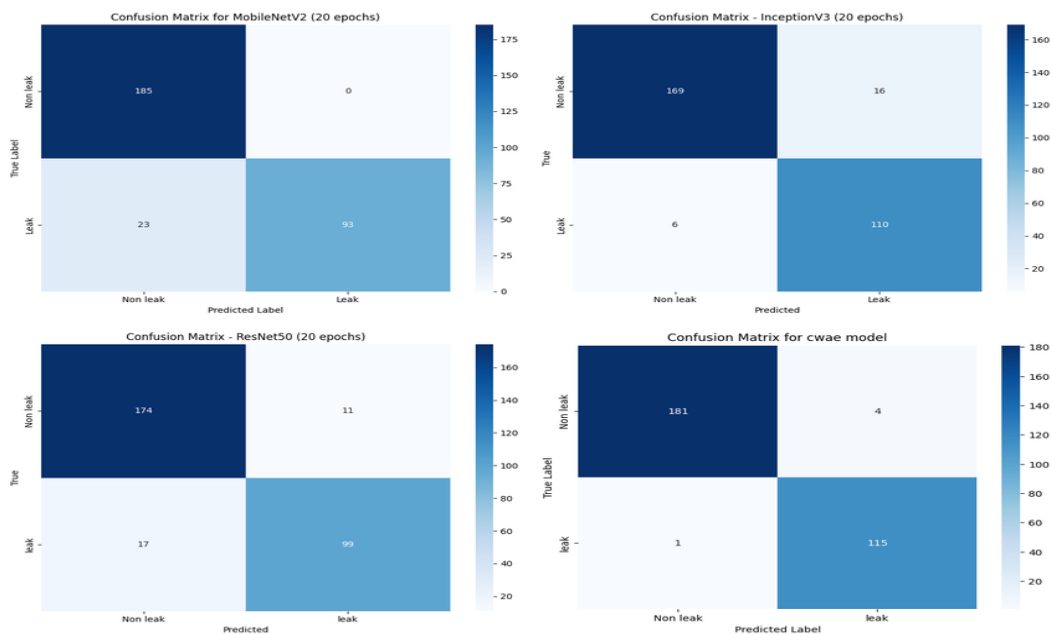


Figure 13: Confusion matrices at 20 epochs (top-left: MobileNetV2; top-right: InceptionV3; bottom-left: ResNet-50; bottom-right: Proposed Model).

Discussion

Across both epoch milestones, the proposed CNN + KD + AE model consistently outperformed InceptionV3, MobileNetV2, and ResNet-50 on all reported metrics while maintaining the smallest model footprint at 4.3 MB. The autoencoder reduces input noise and dimensionality, enabling the classifier to generalize better to unseen thermal variations. Knowledge distillation allows the student to benefit from the teacher's soft probability outputs, learning richer inter-class relationships than hard labels alone provide. ResNet-50 showed the most inconsistent performance, constrained by its larger parameter count, requiring more data to converge stably on this domain-specific dataset. MobileNetV2 was the strongest baseline, but its tendency to misclassify leak samples as non-leak at 20 epochs represents a meaningful safety limitation. The proposed model's low combined misclassification rate and balanced error distribution make it well-suited for deployment in safety-critical pipeline monitoring.

Conclusion

This study presented a comparative evaluation of a lightweight CNN + KD + AE model against InceptionV3 (Szegedy et al., 2016), MobileNetV2 (Sandler et al., 2018), and ResNet-50 (He et al., 2016) for IoT-based oil pipeline leak detection using thermal imagery. The proposed model

achieved consistent accuracy, precision, recall, and F1-score of 0.98 at both 10- and 20-epoch milestones, with a model size of 4.3 MB, outperforming all baselines across all metrics. Confusion matrix analysis confirmed the proposed model produced the fewest and most balanced misclassifications at both training durations. The findings confirm that a lightweight hybrid deep learning architecture combining CNNs, Knowledge Distillation, and Autoencoder-based compression can outperform established pre-trained baselines in thermal pipeline leak detection while requiring substantially fewer computational resources, making it directly suitable for deployment on resource-constrained IoT edge devices without a dependency on cloud infrastructure.

Recommendations

The proposed model should be deployed on real IoT edge hardware in a live pipeline environment to validate real-time inference performance under field conditions. Future studies should collect larger, more diverse thermal image datasets that cover a range of pipeline materials, environmental conditions, and leak severities. Quantisation-aware training and weight pruning should be explored to further reduce inference time and energy consumption. Multi-sensor fusion combining thermal imaging with pressure or

acoustic data would improve detection reliability and enable leak localization. Extending the comparative framework to include EfficientNet-Lite and MobileNetV3

would provide a more comprehensive benchmark for edge-deployable pipeline monitoring systems.

References

- Correa da Silva, P. E., & Almeida, J. (2024). An edge computing-based solution for real-time leaf disease classification using thermal imaging. *IEEE Geoscience and Remote Sensing Letters*. <https://doi.org/10.1109/LGRS.2024.3456637>
- Dang, T. D., Nguyen, M. T., & Kim, J. (2023). Hyperspectral image compression using convolutional autoencoder with residual blocks. *Sensors*, 23(10), Article 4892. <https://doi.org/10.3390/s23104892>
- Ekong, A. P., James, G. G., & Ohaeri, I. (2023). Oil and gas pipeline leakage detection using IoT and deep learning algorithm. *Journal of Information Systems and Technology Management*, 6(1), 15–25.
- Farooq, M. A., Shariff, W., Khan, F., & Corcoran, P. (2023). Development, optimization, and deployment of thermal forward vision systems for advance vehicular applications on edge devices. *Proceedings of SPIE*, 12701, Article 127010M. <https://doi.org/10.1117/12.2679749>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778). IEEE. <https://doi.org/10.1109/CVPR.2016.90>
- Hinton, G., Vinyals, O., & Dean, J. (2015). Distilling the knowledge in a neural network. *arXiv*. <https://arxiv.org/abs/1503.02531>
- Howard, A., Sandler, M., Chu, G., Chen, L.-C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., Le, Q. V., & Adam, H. (2019). Searching for MobileNetV3. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 1314–1324). IEEE. <https://doi.org/10.1109/ICCV.2019.0140>
- Jadin, M. S., & Ghazali, M. F. (2014). Gas leak detection with thermal imaging. *International Journal of Advanced Computer Science and Applications*, 5(6), 123–129. <https://doi.org/10.14569/IJACSA.2014.050617>
- Kayan, H., Majib, Y., Alsaferi, W., Barhamgi, M., & Perera, C. (2022). AnoML-IoT: An end-to-end re-configurable multi-protocol anomaly detection pipeline for Internet of Things. *Internet of Things*, 19, Article 100481. <https://doi.org/10.1016/j.iot.2022.100481>
- Kheir, A. E., & Shaaban, E. (2024). Modelling the performance of a multi-hop LoRaWAN linear sensor network for energy-efficient pipeline monitoring systems. *Applied Sciences*, 14(20), Article 9391. <https://doi.org/10.3390/app14209391>

- Kim, J. M., Faisal, S., & Zahoor, A. (2025). Real-time pipeline leak detection: A hybrid deep learning approach using acoustic emission signals. *Applied Sciences*, 15(1), Article 185. <https://doi.org/10.3390/app15010185>
- Maia, D. M., Mendes, J. V. S., Silva, J. P. A. M., Bastos, R. F., Silva, M. S., Mirre, R. C., Neto, J. C. S., & Lepikson, H. A. (2024). IoT leak detection system for onshore oil pipeline based on thermography. *Sensors*, 24(21), Article 6960. <https://doi.org/10.3390/s24216960>
- Obaid, M. H., & Hamad, A. H. (2023). Deep learning approach for oil pipeline leakage detection using image-based edge detection techniques. *Journal of Electrical Systems and Automation*, 56(4), 416–425.
- Rao, S., Kumar, R., & Sharma, P. (2023). Pipeline leak detection using AI and thermal imaging. *International Journal of Advanced Computer Science and Applications*, 14(3), 112–119.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4510–4520). IEEE. <https://doi.org/10.1109/CVPR.2018.00474>
- Shi, Z., Wu, J., Zhang, Q., & Yao, J. (2022). ICT innovation to promote sustainable development goals: Implementation of smart water pipeline monitoring system based on narrowband Internet of Things. *Sustainability*, 16(22), Article 9683. <https://doi.org/10.3390/su16229683>
- Singh, A., & Patel, R. (2023). Leak detection in pipelines using thermal imaging. *Journal of Pipeline Engineering*, 22(2), 85–92.
- Sun, J., Zhang, L., & Wang, H. (2023). Leakage aperture identification of natural gas pipeline based on compressed acquisition and deep sparse autoencoder. *Measurement and Control*, 56(3), 123–130. <https://doi.org/10.1177/00202940221110>
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2818–2826). IEEE. <https://doi.org/10.1109/CVPR.2016.308>
- Xiong, K., Jiang, J., Pan, Y., Yang, Y., Chen, X., & Yu, Z. (2022). Deep learning approach for detection of underground natural gas micro-leakage using infrared thermal images. *Sensors*, 22(14), Article 5322. <https://doi.org/10.3390/s22145322>
- Yuan, Y., Li, X., & Wang, Z. (2023). Real-time pipeline leak detection using attention-based LSTM. *Engineering Applications of Artificial Intelligence*, 123, Article 106057. <https://doi.org/10.1016/j.engappai.2023.106057>
- Zhang, L., Xu, Y., & Tang, Y. (2023). Pipeline leak diagnosis using scalograms and deep learning. *Journal of Pipeline Science and Engineering*, 3(2), Article 100114. <https://doi.org/10.1016/j.jpse.2023.100114>