2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

Roofus: Learning-based Robotic Moisture Mapping on Flat Rooftops with Ground Penetrating Radar

Kevin Lee[1,2], Wei-Heng Lin[2], Talha Javed[1*], Sruti Madhusudhan[1*], Bilal Sher[1*], Chen Feng[1,2]

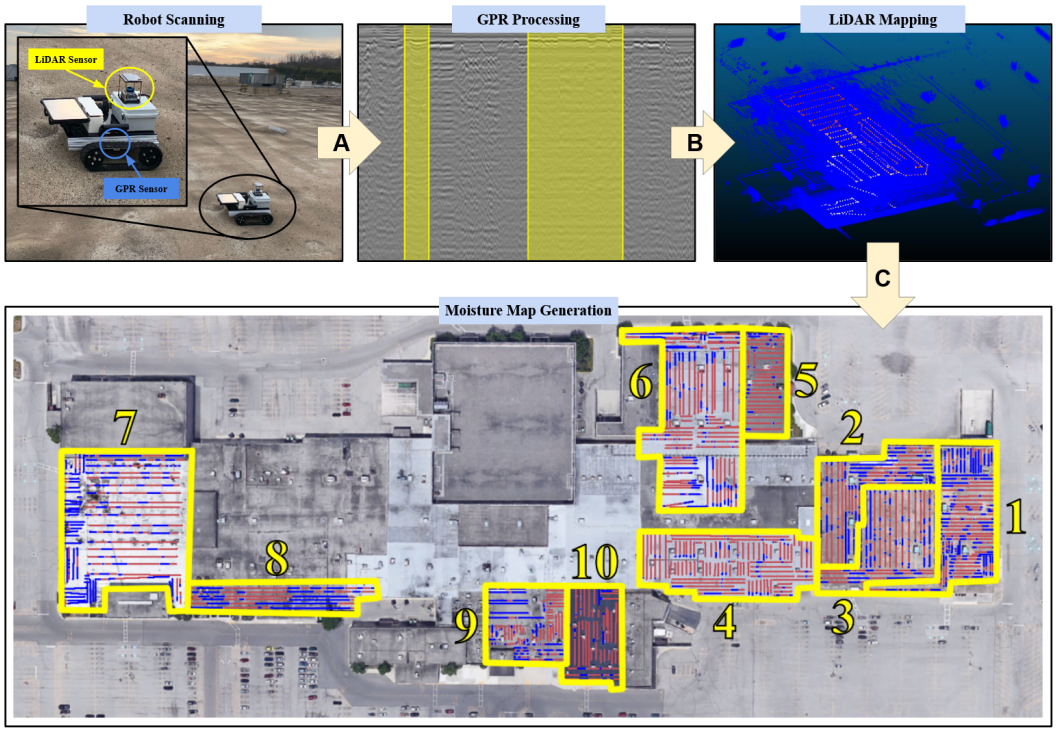

Fig. 1: **System overview of Roofus.** The system consists of a robotic platform for rooftop navigation and scanning, supported by a learning-based ground penetrating radar moisture detection model and LiDAR-based robot path mapping back-end. The dataset and reference report discussed in this paper are organized into 10 distinct rooftop sections, as illustrated.

***Abstract*— Robust moisture detection is crucial for building maintenance and cost reduction. Current methods are often limited by the type of roofing material or are cumbersome and expensive. Ground Penetrating Radar (GPR) has shown promise in recent works in moisture detection due to its effectiveness across a broader range of materials, its compactness and lightweight nature, and its ability to image the subsurface. We introduce Roofus, an integrated robotic moisture detection system for flat rooftops, designed to overcome traditional method limitations. It combines a remotely controlled robot with deep learning GPR data processing and automatic map generation. Real-world data is collected and manually annotated for supervised learning. We investigate a novel approach to interpreting GPR data via deep learning using Transformer-based classifiers. LiDAR inertial odometry is employed to integrate multiple individual GPR scans into a holistic moisture map over the rooftop. When evaluated against existing methods such as infrared thermal imaging, electrical capacitance surveys, and nuclear moisture gauges, our system shows promising viability for industry application.**

*indicate equal contributions.
[1]Building Diagnostic Robotics, Inc., Brooklyn, NY 11201, USA
[2]New York University, Brooklyn, NY 11201, USA {k.lee, whl318, cfeng}@nyu.edu

I. Introduction

Water damage has historically posed significant costs to the commercial and retail building industry [1], highlighting the critical need for effective moisture detection methods. Rooftop moisture detection is integral not only for maintaining the structural integrity of buildings but also for mitigating health hazards from mold growth, improving energy efficiency, and ensuring compliance with insurance requirements [2].

Traditional methods for moisture detection on rooftops, such as visual inspection, infrared thermal imaging (IRT), electrical capacitance (EC) surveys, and nuclear moisture gauges, are labor-intensive and have limitations that affect the accuracy and reliability of the results [3]. For instance, visual inspection and IRT cannot detect damage below the surface, while EC surveys are impractical on roofs with electrically conductive membranes. Nuclear moisture gauges require careful handling and disposal of radioactive components. Introducing a safe and widely effective system for moisture scanning can greatly reduce the costs and risks associated

with current methods

Recent studies suggest that Ground Penetrating Radar (GPR) may be a promising solution for effectively conducting rooftop moisture detection [3]. GPR is a non-destructive technology that uses electromagnetic (EM) waves to detect changes in the EM properties of the scanned medium. As the presence of water affects the EM properties of materials (for instance, wet versus dry concrete), GPR becomes a potential solution for moisture detection. Additionally, its ability to scan without physically damaging the surface further reinforces its suitability for rooftop moisture detection.

In this paper, we introduce Roofus, a robust moisture detection system leveraging GPR that can be remotely deployed on building rooftops. Specifically, we design a remotely controllable robot to collect GPR scan data. This data is then analyzed by a deep learning back-end. LiDAR inertial odometry (LIO) is used in conjuction with the processed data to produce high-quality maps. The maps effectively visualize the specific locations on the roof where moisture is detected. The integration of these technologies allows for unprecedented increases in efficiency and cost-effectiveness for moisture detection and mapping.

Using Roofus, we gather and process real-world data, comparing our system's performance against a third-party moisture report obtained through traditional methods. We manually label the data, creating a dataset of 470 GPR B-scans (or 885K A-scans, also known as traces) from 10 distinct rooftops. Training a Transformer-based classifier on this dataset, we depart from the image-based approach [3–6], focusing on trace-level classification. Our approach is supported by previous studies [7], and we compare its performance against an image-based variant using a similar architecture. By combining model predictions with LIO results, we create accurate visualizations of moisture detection regions. Our system performs comparably to existing methods and demonstrates promise in detecting moisture on surfaces where traditional methods fail.

Our contributions are summarized as follows:

1) A novel investigation into determining the most effective deep learning method for processing GPR data.
2) The integration of LIO and GPR to generate detailed and easily interpretable moisture maps.
3) The design and development of a remotely operated robotic system capable of navigating and collecting data from building rooftops.

II. Related Work

A. Traditional Rooftop Moisture Detection Methods

Rooftop moisture detection is traditionally done in a destructive manner, in which inspectors take samples from a roof by cutting out sections of the roofing to examine its underlying layers for moisture. This process inevitably damages the roof and requires post-inspection repairs. Given the costly and time-consuming nature of the traditional destructive methods for rooftop moisture inspection, recent studies have been motivated to develop non-destructive testing (NDT) techniques. For instance, [8] has explored the use of IRT to detect energy leaks in electrical and mechanical components, whereas [9, 10] specifically use IR to detect moisture in building components based on changes in surface temperature. A similar technique is the use of EC sensors, a method that has been explored in literature for moisture evaluation in material such as wood [11] or soil [12]. The method has also been extensively used in industry for building moisture detection. However, there are limitations in both of these techniques, as they are highly subjective to environmental conditions and surface material that is being scanned. In IRT's case, the accuracy can be largely affected by sunlight and ambient temperature, whereas EC surveys cannot be done on electrically conductive materials. Furthermore, IR is unable to penetrate deeply to capture the presence of moisture directly. Another example of a NDT method is the Nuclear Moisture Gauges (NMG), which emit neutron or gamma rays that pass through the material being tested. However, given that NMG utilizes radioactive materials, this operation is much more restricted to use and hence less accessible compared to other methods. Ground Penetrating Radar (GPR) emerges as a competitive option, offering non-destructive imaging of the subsurface, independence from ambient temperature and sunlight, and ease of handling and safety, as summarized in Table I.

TABLE I: Moisture Detection Method Comparison

Method	Core Cuts	IRT	EC	NMG	GPR
Non-destructive	✗	✓	✓	✓	✓
Subsurface Imaging	✓	✗	✗	✗	✓
Temperature and Sunlight Agnostic	✓	✗	✓	✓	✓
Non-radioactive	✓	✓	✓	✗	✓

B. Deep Learning for Automated GPR Processing

Traditionally, analyzing GPR data has been a labor-intensive task, heavily reliant on experts in the field. This dependence arises from the complex nature of GPR signals, which require nuanced interpretation to accurately identify and classify subsurface features. To reduce the need of manual interpretation, deep learning has naturally become popular technique in automating the processing of GPR data [3–7, 13], owing to its inherent capability to learn and discern underlying features in the input data. Specifically, the use of vision-based models as Convolutional Neural Network (CNN) and its variations have demonstrated notable competence. For instance, [4, 5] treat B-scans as images and train a region-based CNN (R-CNN) model to identify the hyperbolic features of subsurface objects. Other works as [6] train a CNN to detect whether the input B-scan contains traces of buried landmines, and [3] employed sequence-to-sequence (seq2seq) architecture with ResNet and UNet++ as the backbone networks to perform segmentation on GPR images. While vision-based methods have shown their effectiveness in GPR analysis, we have found limitations in such methods throughout our experiments. Given the characteristics of GPR data, where data is structured temporally

and spatially, capturing features with kernels in a patch based manner might not be ideal. Therefore, we propose to employ the seq2seq architecture but use traces as input data directly, without converting the B-scans into images.

Instead of treating the problem as a classification or segmentation task, several studies [14–16] have framed the problem as one of anomaly detection. These works all share a common premise that anomalous data is rare compared to normal data in their respective tasks. One of the approaches used to overcome this scarcity of anomalous data samples is the reconstructive method. This approach relies on the assumption that a model trained exclusively on normal data will poorly reconstruct anomalous regions in the input data during inference, due to its unfamiliarity with such anomalies. Based on this assumption, one can generate an anomaly score map by computing the discrepancy between the input data and its reconstruction. However, this approach also has its limitations, as the models could become too good at generalizing both normal and anomalous data, resulting in false negatives. This issue has also been noted in multiple unsupervised anomaly detection studies [17, 18].

C. GPR as Sequential Data with Transformers

When it comes to the seq2seq architecture, the Transformer [19] has proven itself to be a game-changer as a backbone module in Natural Language Processing (NLP) [20–22], with its application further extending into the field of Computer Vision [23–25]. Unlike CNN or RNN, which primarily focuses on local features through convolution or sequential processing, Transformers excels at capturing the global context. This capability stems from their self-attention mechanism, which allows the model to weigh and integrate information from all parts of the input sequence simultaneously, regardless of their positional distances. This distinction is particularly advantageous for GPR data analysis as GPR data is characterized by complex, spatially distributed subsurface structures, which require a nuanced understanding of the entire context to accurately identify and classify the underlying features. The global perspective offered by Transformers enables a more comprehensive analysis of GPR data, facilitating the detection of subtle, yet significant, anomalies and patterns that might be overlooked by models focusing solely on local dependencies. Therefore, our work will be leveraging these advantages of Transformer to achieve a more robust GPR interpretation.

D. LiDAR for Localization on Flat Rooftops

Localization poses significant challenges for robots, particularly in dynamic environments like rooftops. Existing methods, such as GPS with Real-Time Kinematic (RTK), offer accuracy but come with limitations, such as the need for a base station within a limited range and unobstructed sky view [26]. Combining GPS with an Inertial Measurement Unit (IMU) provides robust localization [26] but struggles in GPS-denied environments or due to IMU drift, requiring costly navigational-grade IMUs. Visual odometry, while capable of mapping the environment [27, 28], faces challenges on rooftops, including limited features, depth perception issues, and difficulties with reflective surfaces.

For rooftops, which often lack features and have obstructed GPS signals, LiDAR emerges as a more accurate and robust localization solution, despite being more expensive than cameras [29]. Combining LiDAR with an IMU improves localization accuracy, creating detailed maps of the environment and the robot's path. While LiDAR's performance can be impacted by rain, fog, or dust, these limitations can be mitigated by fusing LiDAR with a cost-effective IMU or GPS, and a camera, resulting in higher localization accuracies.

III. Robot Design and Sensors

Our robot, pictured in Figure 1, is designed with a focus on maximizing the efficiency and accuracy of rooftop inspections and data collection. It embodies a compact, yet robust design, making it adept at navigating various environments to systematically gather essential data for analyzing rooftop moisture conditions.

The design features a compact footprint measuring 0.46 m by 0.46 m by 0.41 m and a light weight of 9.07 kg, facilitating portability and maneuverability across different rooftop terrains. Mobility is achieved through track wheels, enabling the robot to overcome obstacles up to 10.16 cm tall and perform efficiently on diverse terrains. For navigation and guidance, the robot is equipped with advanced hardware, including LiVOX Mid360 LiDAR, Insta360 Air Camera, and u-blox ZED-F9P GPS-RTK. These components, combined with algorithms that integrate LiDAR, visual, and IMU data, enable precise localization and systematic coverage of the inspection area.

At the heart of its design, the robot incorporates a Proceq GP8800 SFCW GPR sensor, rigidly mounted approximately 2.54 cm above the surface for optimal and consistent data acquisition. The integration ensures synchronized collection of GPR, LiDAR, IMU, and GPS data as the robot moves at an average walking speed of about 1.4 m/s, facilitating detailed analysis and comprehensive scanning of the rooftop. The robot is equipped with a 6S Lithium Polymer Battery, which enables an operational endurance of up to 2 hours on a single charge and is hot-swappable. This allows for extensive data collection sessions, minimizing interruptions for battery recharges and enhancing the thoroughness of rooftop inspections. Controlled remotely via an Arduino dashboard app, the robot navigates in predetermined straight segments approximately 1 m apart, ensuring systematic coverage.

IV. Software and Dataset

A. Data Collection and Dataset Creation

Data collection. Data was collected using Roofus across a multi-level, multi-section retail mall complex, encompassing over 46.5K square meters of roofing divided into ten distinct sections. These sections, depicted in Figure 1, showcased a variety of roofing materials and compositions, offering a broad spectrum for analysis. Despite the absence of original floor plans and modification records, our team managed

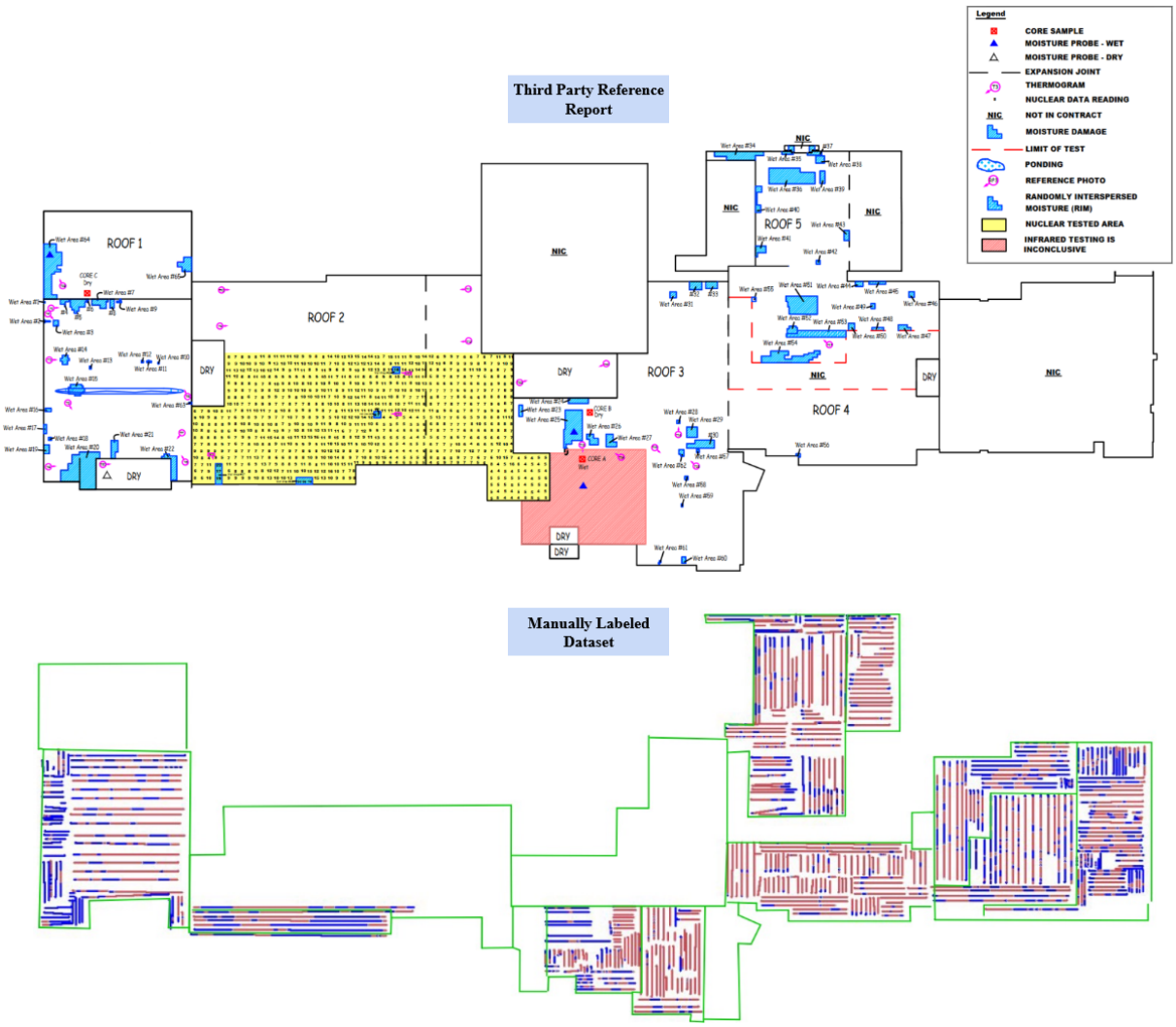

Fig. 2: **Comparison between third-party Reference report and our manual labels.** Blue in the manual labels indicates positive moisture labels (moisture is present) and red indicates negative labels. Roof numbers should be disregarded, as they are internal identifications used by the third party. Note the label correspondences between the reference and our labels, particularly in Sections 4, 6, and 10. This image is best viewed in color.

to identify the majority of roofing types through on-site observations.

Sections 1 through 3 featured modified bituminous (MB) membrane roofs enhanced with a liquid membrane to extend their lifespan, where visible signs of moisture-related degradation, such as bubbling, were noted. Section 4 also had an MB roof but lacked the liquid membrane treatment found in the preceding sections. Section 5 was characterized by a single-ply Ethylene Propylene Diene Monomer (EPDM) rubber membrane roof topped with gravel, a setup posing significant challenges to conventional moisture detection methods like infrared and capacitance sensing, thereby underscoring the potential of GPR-based techniques.

Section 6 mirrored Section 4 but included additional obstacles like piping and ductwork. Section 7, presumed to have Thermoplastic Polyolefin (TPO) over MB roofing, showed mild bubbling indicative of layer separation, complicating infrared imaging. Section 8 was similar to Section 7 in composition, though the difference is not visible in 1 due to an outdated satellite image. The makeup of Section 9 remained uncertain, though it likely featured a liquid membrane akin to Sections 1 through 3. Lastly, Section 10 was equipped with an EPDM rubber membrane roof.

The robotic scans were systematically planned and executed to maximize coverage within each section. Initially positioned at a corner, the robot's path was strategically plotted to navigate around obstacles such as HVAC units and ductwork. As the robot currently lacks full automation, it was remotely operated by a team member throughout the data collection phase. Upon completing each section, the data was promptly uploaded to a cloud-based repository for subsequent processing.

Figure 2 depicts the third-party moisture report, which primarily relied on EC surveys and IRT, supplemented by core cuts for verification. Select areas were additionally assessed using nuclear moisture gauges. We refer to this

report as the "reference" report. Our scanning overlaps with the third-party's in Sections 4, 6, 7, 8, and 10. Sections 1, 2, 3, and 5 were deemed out of scope for the third-party, either because they were not requested by their client or, particularly in the case of Section 5, were too challenging to scan. Despite this, the presence of standing water was confirmed in these sections. Section 9 was deemed inconclusive due to conflicting results from IRT and core cuts, although the presence of moisture was likely.

Manual Annotation. While not as intuitive as object detection, signs of moisture damage and ingress are clearly identifiable in GPR scans. In order to develop and train a robust, supervised classifier for moisture detection, high quality labels are necessary. Leveraging our experience and knowledge of moisture readings in GPR scans, we manually annotated the GPR scans collected from the robot to create a manually labeled dataset. To ensure the integrity of the manual labels, annotators were randomly assigned scans without knowledge of which section the scan originated from. Furthermore, annotators were prohibited from using the third-party report as reference. The overlapped comparison between the third-party reference report and our manual labels shown in Figure 4 demonstrates the similarity of our manual annotations to the third-party labels, thus reinforcing our confidence in the validity of the manually labeled dataset.

In total the dataset is comprised of 470 B-scans, or 885K traces. 369 of the B-scans contain moisture labels. There are 302K positive traces (traces that indicate the presence of moisture), which makes up approximately 34% of the total dataset.

B. Supervised Learning for GPR Processing

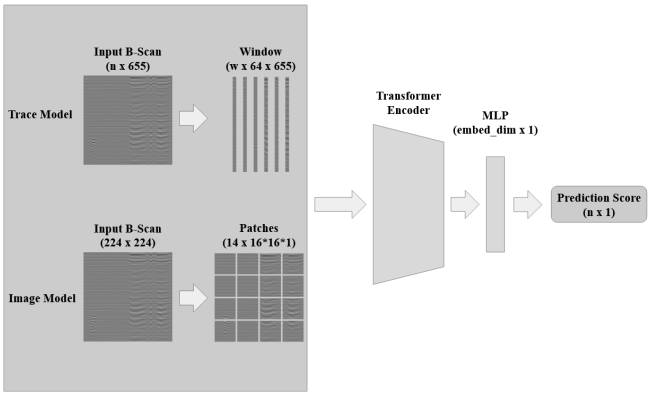

Fig. 3: Transformer-based GPR processing model architecture.

The manually annotated dataset provides labels for training, allowing us to adopt a supervised learning approach for the GPR processing back-end.

We explore GPR processing using deep learning with the Transformer architecture. The model's objective is to predict whether moisture is detected at the trace level. Our experiments compare the traditional image-based GPR data processing with deep learning, referred to as the Image model, against a novel approach that uses sequences of traces as input, designated as the Trace model, as depicted in Figure 3. Both models consist of a Transformer encoder and a multi-layer perceptron (MLP) binary classification head.

The Image model processes a GPR B-scan (n x 655, where n is the number of traces) by converting it to a grayscale image (224 x 224 px), which is then patched into size 16 x 16 patches for input into the Transformer encoder. In contrast, the Trace model divides the B-scan into w non-overlapping windows of 64 traces (w x 64 x 655), where $w = \text{ceil}(n \div 64)$. Prior to windowing, the B-scan is padded with its mean trace until $n \mod 64 = 0$ is satisfied. Both models output a predictive score between 0 and 1 for each column or trace in the input (224 labels for image-based, n labels for trace sequence), indicating the likelihood of moisture presence. The Image model's outputs are interpolated to the original length of the B-scan to produce results corresponding to each trace.

Although the input formats differ, the general structures of both models are analogous. They are trained using encoder layers with self-attention heads. We experiment with "base," "large," and "huge" variants of each model, containing approximately 85.6K, 302M, and 630M trainable parameters, respectively. These variants differ in the number of encoder layers, self-attention heads, and embedding sizes. Both models are trained over 50 epochs, with the Image model using a learning rate of 1e-4 and the Trace model using a learning rate of 1e-5. We employ the AdamW optimizer with decay = 0.05 and betas = (0.9, 0.95), along with a cosine decay learning rate scheduler. Training is conducted on a machine with an Nvidia RTX 8000 GPU. To address class imbalance and prioritize harder samples, we use Focal Loss [30] over Binary Cross Entropy.

C. LIO for Moisture Map Generation

To achieve an accurate moisture map from the scan, it is crucial to accurately track the robot's path in the environment. We utilize LIO for this purpose. The Livox MID-360 LiDAR is chosen for its lightweight, compact design, and relatively lower cost compared to other LiDARs on the market. Despite low-featured rooftops, the non-repetitive scanning pattern of the LiDAR enables it to generate a denser point cloud for each frame at a lower frequency. Moreover, the LiDAR is equipped with a built-in IMU, simplifying integration into the system.

For LIO, we select LIO-SAM [31] for its high-accuracy mapping capabilities and Fast-LIO [32] for its adaptability to dynamic environmental conditions. These algorithms are used for post-processing the LiDAR scans, generating the path and map of each section. Timestamps are employed to mark the start and end of each GPR B-scan. These timestamps are then synchronized with the path produced by LIO, enabling the precise overlay of the B-scan on the path and resulting in more accurate positioning of GPR data.

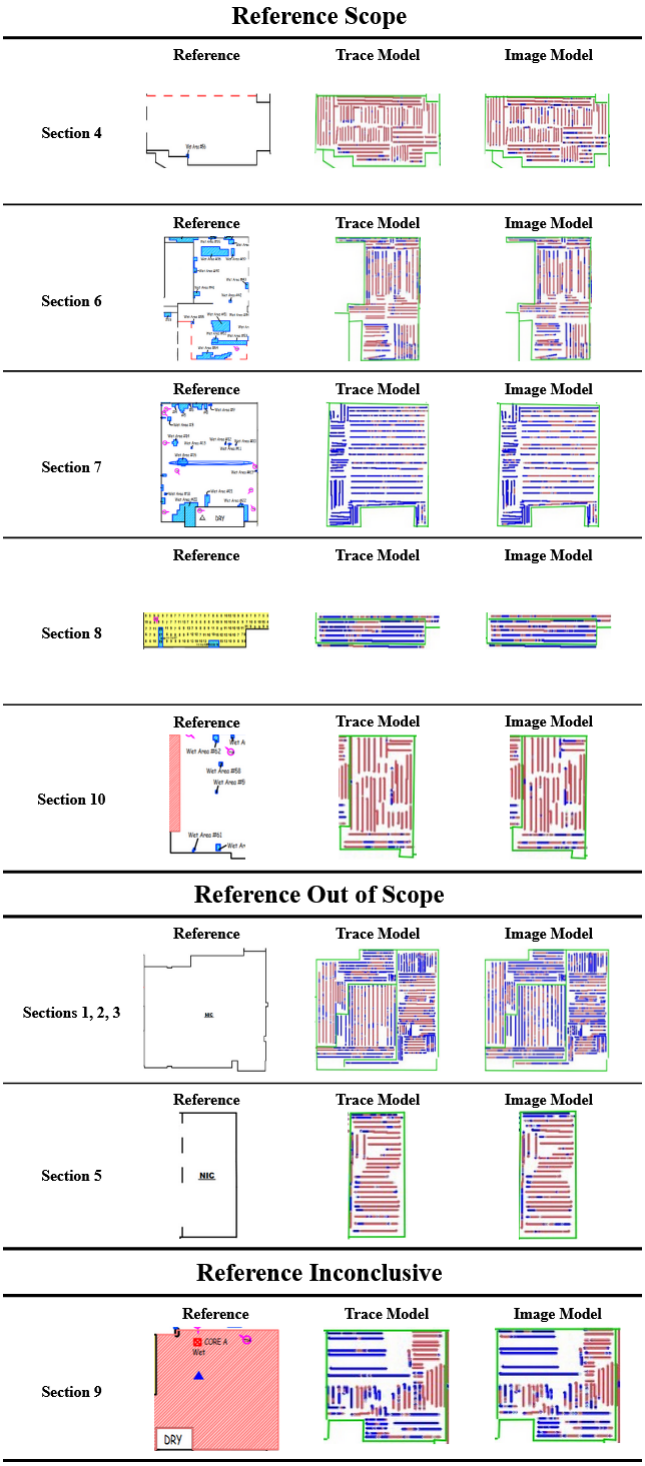

Fig. 4: **Visual report comparison with reference.** Blue indicates positive labels whereas red indicates negative labels. Recall Sections 1, 2, 3, and 5 were out of scope in the reference report but had visible standing water. Section 5 in particular poses significant challenges to traditional detection methods due to the added layer of gravel. Section 9 was determined inconclusive; however, core cuts confirmed the presence of moisture.

V. Experiments

We assess the performance of moisture detection by comparing the generated moisture map with the reference report. The manually labeled dataset is split by B-scan into training, validation, and testing subsets (64%/16%/20%, respectively). The split is stratified to maintain a consistent ratio of B-scans without moisture labels to B-scans with at least one moisture label in each subset. Before training, the B-scans are standardized using the average trace and standard deviation calculated from the training subset. This standardization not only improves training quality and efficiency but also enhances the interpretability of GPR scans for human inspection.

We use the Area Under the Receiver Operator Characteristic (AUROC) as a quantitative measure of model performance. AUROC considers the trade-off between True Positive Rate (TPR) and False Positive Rate (FPR) across different thresholds, with values ranging from 0.0 to 1.0. A higher AUROC value indicates better model performance, with an ideal model achieving a value close to 1.0. TPR and FPR are chosen as appropriate metrics because our goal is to accurately identify regions with moisture damage while minimizing false alarms.

Additionally, we conduct a qualitative assessment of our models by visually comparing the generated moisture maps to the reference report. We consider the model to perform well if its results align with the report, showing significant overlap between the predicted labels and the labeled regions in the report. These qualitative assessments are conducted using the seed and variant that demonstrate the best quantitative performance.

TABLE II: Model Performance

Models	Image			Trace		
	Base	Large	Huge	Base	Large	Huge
AUROC	0.7285	0.7289	0.7400	**0.8747**	0.8734	0.8726

AUROC Assessment. Table II presents the quantitative performance of each model tested across different sizes. The results, averaged over 5 different seeds, highlight the best performance achieved. The Trace model demonstrates an advantage over the Image model, with the best-performing variant achieving an average AUROC of 0.8747, compared to the Image model's best average AUROC of 0.7400 – a 16.6842% difference. This performance gap suggests the effectiveness of using sequences of traces over image-based approaches for supervised trace classification, diverging from the conventional approach of analyzing GPR B-scans as images. Interestingly, the Trace model exhibited minimal performance variation across the different model sizes tested. Despite our efforts to include a diverse range of roofing materials, this consistency may suggest that the current dataset does not fully capture the variability or representation needed to challenge the model. Expanding the dataset to

include an even broader array of roofing membranes and additional examples of existing materials could potentially lead to further performance improvements.

Visual Report Comparison. We present the moisture maps generated by the Trace model and Image model in Figure 4. The maps were generated using prediction thresholds of 0.319 for the Trace model and 0.259 for the Image model, determined through ROC analysis. Our observations reveal a strong agreement between the model predictions and the third-party report. The models accurately capture the sparsity of positive labels in Sections 4 and 10, as well as the higher density of positive labels in Sections 6 and 7. The results in Section 8 also generally mirror the locations of the hot-spots in the reference. These results demonstrate the effectiveness of our supervised models, suggesting they are at least as effective as traditional methods for moisture detection.

Moreover, the models' ability to generate plausible predictions for sections that were either out of scope or inconclusive in the ground truth (Particularly Sections 5 and 9) showcases the potential of our solution to address the limitations of traditional moisture detection methods.

VI. Conclusions

In this paper, we have introduced Roofus, an innovative robotic system designed for automated moisture detection on building rooftops, coupled with the generation of high-quality reports. Our approach has demonstrated the effectiveness of processing GPR data as a sequence of traces, rather than as images, for deep learning, resulting in a performance improvement of up to 16.68%. The generated experimental report showcases the system's ability to perform as well as current state-of-the-art methods, such as EC surveys and IRT, even in scenarios where these methods have failed. Additionally, we have showcased the application of LIO algorithms to produce detailed moisture report maps.

Despite these promising results, there are several limitations and avenues for future work to consider. The computational demands of LIO algorithms pose challenges related to battery life and size constraints. In the realm of deep learning, the reliance on manual annotation for training data remains a bottleneck, especially considering the vast amounts of GPR data that can be collected. Furthermore, expert analysis is still required to create these manual labels.

For future work, we propose several directions. Regarding LIO, achieving full autonomy similar to an autonomous robot vacuum could enhance the system's efficiency and usability. For deep learning, exploring self-supervised or unsupervised processing methods could mitigate the challenges associated with obtaining high-quality annotated data. Additionally, implementing real-time GPR processing on the robot itself would be a significant advancement, enhancing the system's ability to generate reports far more efficiently.

Acknowledgments

This work is supported by NSF grant #2322242 and in part through the NYU IT High Performance Computing resources, services, and staff expertise.

Disclosure

Bilal Sher, Sruti Madhusudhan, Talha Javed, and Chen Feng are founders of Building Diagnostic Robotics, Inc., a startup company that uses AI and robotics for building inspections.

References

[1] H. W. Busching, R. G. Mathey, W. J. Rossiter, Jr., and W. C. Cullen, "Effects of moisture in built-up roofing: a state-of-the-art literature survey," 1977. 1

[2] *Moisture Control Guidance for Building Design, Construction and Maintenance*. U.S. Environmental Protection Agency (EPA), 2013. 1

[3] B. A. Sher and C. Feng, "Deepgpr: Learning to identify moisture defects in building envelope assemblies from ground penetrating radar," in *Proceedings of the 40th International Symposium on Automation and Robotics in Construction*, B. García de Soto, V. Gonzalez-Moret, and I. Brilakis, Eds. Chennai, India: International Association for Automation and Robotics in Construction (IAARC), July 2023, pp. 561–568. 1, 2

[4] F. Hou, W. Lei, S. Li, and J. Xi, "Deep learning-based subsurface target detection from gpr scans," *IEEE Sensors Journal*, vol. 21, no. 6, pp. 8161–8171, 2021. 2

[5] R. Jaufer, A. Ihamouten, D. Guilbert, S. Todkar, T. Yaram, and X. Derobert, "Deep learning based automatic hyperbola detection on gpr data for buried utility pipes mapping," in *2021 11th International Workshop on Advanced Ground Penetrating Radar (IWAGPR)*, 2021, pp. 1–6. 2

[6] S. Lameri, F. Lombardi, P. Bestagini, M. Lualdi, and S. Tubaro, "Landmine detection from gpr data using convolutional neural networks," in *2017 25th European Signal Processing Conference (EUSIPCO)*, 2017, pp. 508–512. 2

[7] Z. Tong, J. Gao, and D. Yuan, "Advances of deep learning applications in ground-penetrating radar: A survey," *Construction and Building Materials*, vol. 258, p. 120371, 2020. 2

[8] C. Balaras and A. Argiriou, "Infrared thermography for building diagnostics," *Energy and Buildings*, vol. 34, no. 2, pp. 171–183, 2002, tOBUS - a European method and software for office building refurbishment. 2

[9] E. Barreira, R. Almeida, and J. Delgado, "Infrared thermography for assessing moisture related phenomena in building components," *Construction and Building Materials*, vol. 110, pp. 251–269, 2016. 2

[10] I. Garrido, M. Solla, S. Lagüela, and N. Fernández, "Irt and gpr techniques for moisture detection and characterisation in buildings," *Sensors*, vol. 20, no. 22, 2020. 2

[11] S. K. Korkua and S. Sakphrom, "Low-cost capacitive sensor for detecting palm-wood moisture content in real-time," *Heliyon*, vol. 6, no. 8, p. e04555, 2020. 2

[12] X. Deng, H. Gu, L. Yang, H. Lyu, Y. Cheng, L. Pan, Z. Fu, L. Cui, and L. Zhang, "A method of electrical conductivity compensation in a low-cost soil moisture

sensing measurement based on capacitance," *Measurement*, vol. 150, p. 107052, 2020. 2

[13] J. K. Alvarez and S. Kodagoda, "Application of deep learning image-to-image transformation networks to gpr radargrams for sub-surface imaging in infrastructure monitoring," in *2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2018, pp. 611–616. 2

[14] F. Picetti, G. Testa, F. Lombardi, P. Bestagini, M. Lualdi, and S. Tubaro, "Convolutional autoencoder for landmine detection on gpr scans," in *2018 41st International Conference on Telecommunications and Signal Processing (TSP)*, 2018, pp. 1–4. 3

[15] G. V. Pillai, A. Verma, and D. Sen, "Transformer based self-context aware prediction for few-shot anomaly detection in videos," in *2022 IEEE International Conference on Image Processing (ICIP)*, 2022, pp. 3485–3489.

[16] Z. He, D. He, M. Xiao, A. Lou, and G. Lai, "Convolutional transformer-inspired autoencoder for hyperspectral anomaly detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1–5, 2023. 3

[17] V. Zavrtanik, M. Kristan, and D. Skocaj, "Draem - a discriminatively trained reconstruction embedding for surface anomaly detection," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2021, pp. 8330–8339. 3

[18] T. Defard, A. Setkov, A. Loesch, and R. Audigier, "Padim: A patch distribution modeling framework for anomaly detection and localization," in *Pattern Recognition. ICPR International Workshops and Challenges: Virtual Event, January 10–15, 2021, Proceedings, Part IV*. Berlin, Heidelberg: Springer-Verlag, 2021, p. 475–489. 3

[19] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, ser. NIPS'17. Red Hook, NY, USA: Curran Associates Inc., 2017, p. 6000–6010. 3

[20] A. Radford, K. Narasimhan, T. Salimans, I. Sutskever *et al.*, "Improving language understanding by generative pre-training," 2018. 3

[21] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever *et al.*, "Language models are unsupervised multitask learners," *OpenAI blog*, vol. 1, no. 8, p. 9, 2019.

[22] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei, "Language models are few-shot learners," in *Proceedings of the 34th International Conference on Neural Information Processing Systems*, ser. NIPS'20. Red Hook, NY, USA: Curran Associates Inc., 2020. 3

[23] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16x16 words: Transformers for image recognition at scale," in *International Conference on Learning Representations*, 2021. 3

[24] J.-B. Alayrac, J. Donahue, P. Luc, A. Miech, I. Barr, Y. Hasson, K. Lenc, A. Mensch, K. Millican, M. Reynolds *et al.*, "Flamingo: a visual language model for few-shot learning," *Advances in Neural Information Processing Systems*, vol. 35, pp. 23 716–23 736, 2022.

[25] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark *et al.*, "Learning transferable visual models from natural language supervision," in *International conference on machine learning*. PMLR, 2021, pp. 8748–8763. 3

[26] F. Aghili and A. Salerno, "Driftless 3-d attitude determination and positioning of mobile robots by integration of imu with two rtk gpss," *IEEE/ASME Transactions on Mechatronics*, vol. 18, no. 1, pp. 21–31, 2013. 3

[27] G.-S. Cai, H.-Y. Lin, and S.-F. Kao, "Mobile robot localization using gps, imu and visual odometry," in *2019 International Automatic Control Conference (CACS)*, 2019, pp. 1–6. 3

[28] J. Cobos, L. Pacheco, X. Cufí, and D. Caballero, "Integrating visual odometry and dead-reckoning for robot localization and obstacle detection," in *2010 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR)*, vol. 1, 2010, pp. 1–6. 3

[29] Z. Su, X. Zhou, T. Cheng, H. Zhang, B. Xu, and W. Chen, "Global localization of a mobile robot using lidar and visual features," in *2017 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2017, pp. 2377–2383. 3

[30] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 318–327, 2020. 5

[31] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and D. Rus, "Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE Press, 2020, p. 5135–5142. 5

[32] W. Xu and F. Zhang, "Fast-lio: A fast, robust lidar-inertial odometry package by tightly-coupled iterated kalman filter," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3317–3324, 2021. 5