IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 1 May 2025, accepted 18 May 2025, date of publication 21 May 2025, date of current version 2 June 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3572196

RESEARCH ARTICLE

Advancements in Landmine Detection: Deep Learning-Based Analysis With Thermal Drones

DANIEL HEUSCHMID, OLIVER WACKER, YANNICK ZIMMERMANN, PASCAL PENAVA, AND RICARDO BUETTNER, (Senior Member, IEEE)
Chair of Hybrid Intelligence, Helmut-Schmidt-University/University of the Federal Armed Forces Hamburg, 22043 Hamburg, Germany

Corresponding author: Ricardo Buettner (buettner@hsu-hh.de)

This work was supported by the Open-Access-Publication-Fund of the Helmut-Schmidt-University/University of the Federal Armed Forces Hamburg.

ABSTRACT The pervasive threat of landmines across conflict-affected regions necessitates advancements in detection technologies to enhance safety and efficiency in demining efforts. Furthermore, the development of a solution that can be effectively utilized in both resource-rich and resource-poor environments is essential, as many conflict-affected regions lack the financial means or technical infrastructure to deploy expensive and complex detection technologies. This paper introduces a deep learning-based approach utilizing uncrewed aerial vehicles equipped with thermal imaging cameras for landmine detection. Leveraging the MobileNetV3-Large architecture and building upon it, we propose a lightweight, yet powerful, machine learning model that can be seen as a safe, cost-efficient and fast landmine detection method. We evaluated our approach using a unique dataset comprising 2,700 thermographic images captured by a DJI Matrice 100 drone with a Zenmuse XT infrared camera. Through meticulous data pre-processing and augmentation strategies, we enhance the model's ability to generalize across different terrains and mine types, addressing one of the primary challenges in landmine detection. The evaluation of our model on a rigorously curated test set demonstrates promising results, achieving a training accuracy of 96.97 %, a validation accuracy of 97.19 %, and a test accuracy of 96.14 %. These metrics not only demonstrate the model's effectiveness in identifying landmines from thermal images but also highlight its potential to outperform other thermal approaches in real-world demining operations. This supports the application of uncrewed aerial vehicles and deep learning technologies in humanitarian and environmental challenges.

INDEX TERMS Convolutional neural network, deep learning, landmine detection, MobileNetV3, thermal imaging, uncrewed aerial vehicles.

I. INTRODUCTION

The clearance of landmines is a critical and complex challenge that has persisted for decades, with its relevance underscored by the ongoing conflicts and the resultant hazardous leftovers of wars [1], [2], [3]. Landmines, designed as explosive devices to be buried underground, aim to destroy or disable enemy targets, including soldiers, vehicles, and tanks [4], [5], [6]. Their deployment in conflict zones is intended to inflict damage through direct blasts, posing a significant threat not only during wars but also long after the conflicts have ended [5]. The indiscriminate nature of landmines, combined with their potential to cause harm to any individual who comes into contact with them, highlights the urgency of addressing this issue.

The impact of landmines is devastating, with a recorded 4,710 casualties in 2022 alone, nearly half of whom were children among the civilian victims [7], [8], [9], and a concerning 22 % increase to 5,757 casualties in 2023, with civilians comprising 84 % of all recorded casualties and children accounting for 37 % of civilian casualties where age was recorded, with a total of 580 casualties reported in Ukraine that year [10]. Antipersonnel mines continue to contaminate at least 58 countries and territories globally, with 33 signatories to the Mine Ban Treaty facing ongoing clearance responsibilities and 22 non-signatory states plus three

The associate editor coordinating the review of this manuscript and approving it for publication was Turgay Celik.

additional areas also affected. Improvised mines specifically pose challenges in no fewer than 25 countries that have joined the Treaty [10]. This data sheds light on the disproportionate effect of landmines on civilians, underscoring the necessity for enhanced clearance efforts and protective measures for these vulnerable populations [9]. In addition, this statistic not only illustrates the immediate dangers posed by landmines but also the long-term societal repercussions, particularly for the younger generation. Despite efforts to mitigate these dangers, the global challenge remains monumental. In 2019, approximately 164,000 anti-personnel mines were cleared worldwide, yet the complete eradication of landmines is anticipated to take decades due to the tens of millions that remain buried across various regions, assuming no additional mines are laid [11]. The persistence of landmines and unexploded ordnances as a severe problem is exacerbated by the increasing number of war zones and conflicts worldwide, like Afghanistan, Colombia or Ukraine, signaling a prolonged period of risk for the affected countries [7], [8], [10], [12]. The presence of these devices in diverse environments poses significant challenges for clearance efforts, which are influenced by factors such as operator experience, technological capabilities, and environmental conditions [13]. Despite their detection capabilities, existing GPR systems present significant practical limitations in real-world humanitarian demining operations. GPR require specialized experts for complex data interpretation, who are rarely available in conflict-affected regions [14], [15], [16], and the technology remains prohibitively expensive to acquire and maintain, making implementation particularly challenging in resource-constrained areas where landmine contamination is often most severe—a strong contrast considering that clearing a single mine costs 300-1.000 USD, while production costs are 3-30 USD [17].

Technological advancements in Artificial Intelligence, particularly the development and application of uncrewed aerial vehicles (UAVs) and deep learning (DL) techniques, have significantly improved landmine detection and clearance. UAVs, now widely used in military and civilian applications for aerial imaging and sensing, coupled with DL algorithms, are revolutionizing the mapping and analysis of the Earth's surface, including landmine detection [3], [5], [18], [19]. DL enhances the accuracy of thermal image recognition, crucial for landmine detection [20]. UAVs with thermal cameras capture heat emissions from objects on or beneath the surface, and DL algorithms process these images to identify patterns and anomalies indicating landmines [21], [22]. This combination allows for precise and efficient detection by analyzing large datasets and recognizing patterns beyond human capability [2]. Integrating these technologies improves landmine detection and localization, enhancing the safety and efficiency of demining operations [23].

Numerous UAV-based mine detection methodologies employ ground penetrating radar (GPR) technology, which utilizes electromagnetic waves to examine subsurface characteristics [24], [25], [26]. Despite GPR's efficacy in identifying mines, it presents notable drawbacks: (i) the requirement for proximity to potentially explosive devices elevates detonation risks [27]; (ii) elevated costs stemming from the UAV's additional payload and power demands; (iii) diminished operational speed, time and area coverage due to the substantial weight and measurement sensitivity [25]; and (iv) augmented complexity in UAV navigation and data interpretation [28]. Conversely, leveraging UAVs equipped with infrared cameras for thermographic imaging emerges as a rapid, cost-effective, and safe alternative that could offer significant advantages in conflict zones and economically disadvantaged regions [25], [27]. This work aims to bridge these gaps by integrating UAVs, infrared imaging, and advanced deep learning algorithms and to develop a lightweight yet powerful deep learning model that can be deployed on UAVs equipped with infrared cameras to improve landmine detection. This paper addresses the research gap in developing an efficient, cost-effective, and safe method for landmine detection across various environments, specifically targeting its suitability for financially constrained countries that suffer from high landmine contamination. As a result, the most important contributions of this study are the following:

- *Achievement of high accuracy and reliability in landmine detection (Training Accuracy: 96.97%, Validation Accuracy: 97.19%, Test Accuracy: 96.14%).*
- *Development of a lightweight MobielNetV3 approach for landmine detection, with comparable results to state-of-the-art deep learning architectures.*
- *Design of an approach using cost-effective UAVs with infrared cameras for rapid detection especially for economically disadvantaged countries.*

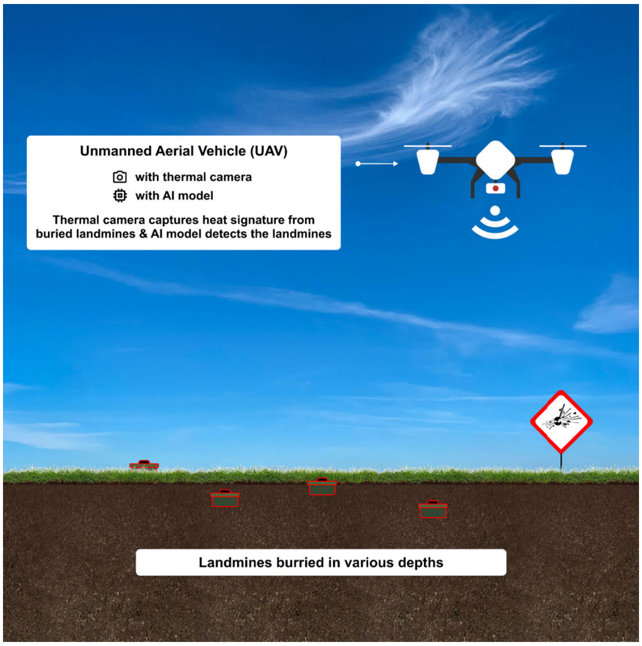


FIGURE 1. Exemplary field usage of UAV thermal imaging with AI model.

The structure of this document is outlined as follows: Chapter two provides an overview of the theoretical foundation concerning the detection of landmines using UAVs and the existing techniques for identifying buried landmines. The third chapter gives a comprehensive explanation of our research methodology, which is then followed by the fourth chapter, where we reveal our findings. Subsequently, we explore the significance of these results, acknowledge the constraints of our study, and suggest directions for forthcoming investigations. The document culminates with a summary in the conclusion, presented in the sixth chapter.

II. RESEARCH BACKGROUND

A. FUNDAMENTALS OF UAV-BASED LANDMINE DETECTION SYSTEMS

Transitioning from the introduction about the global landmine crisis, it's crucial to explore innovative approaches to tackle this issue. The advancement of drone technology is significantly impacting the field of landmine detection [5], [36], [37], [38], [39].

In the last few years, drones have seen a transformation, becoming more advanced and versatile. Their enhanced capabilities in flying longer distances, carrying more weight, and navigating autonomously have opened up new possibilities. The cost-effectiveness of these aerial devices has also increased their accessibility, allowing their application in various sectors. From monitoring crop health in precision agriculture to conducting detailed surveys of vast forest areas, drones are proving to be invaluable assets [5], [36], [37]. Their utility extends to detailed land mapping, providing essential data for development projects [38], and even in specialized tasks like testing electromagnetic compatibility and antenna measurements [39].

When it comes to detecting landmines, drones have introduced a game-changing approach. The traditional methods of finding and removing landmines are known for being dangerous and inefficient. However, drones offer a safer, nondestructive method to inspect and image regions that are suspected to contain landmines [5]. Employing drones in this critical task brings several significant advantages. They have the unique ability to reach and scan areas that are otherwise inaccessible or risky for human teams, ensuring comprehensive land coverage. This approach significantly lowers the chances of accidental detonations, making the detection process much safer for everyone involved [24]. Moreover, the ability of drones to cover large areas swiftly makes the detection operation not just safer, but also faster and more thorough [31].

Recent advancements in deep learning have further enhanced the capabilities of UAV-based landmine detection systems. Bajić and Potoč nik [34] demonstrated the use of YOLOv5 for UAV thermal imaging to detect unexploded ordnance, achieving over 90% recognition accuracy across eleven object classes. Similarly, Priya et al. [32] employed a Region Convolutional Neural Network (RCNN) approach to process thermal images for detecting buried objects and mines, reporting a 90% accuracy in target location prediction. Furthermore, Vivoli et al. [29] introduced a real-time landmine detection system utilizing YOLOv8 models for optical imaging, achieving high recall rates with minimal missed targets. Qiu et al. [35] proposed a joint fusion and detection framework using deep learning in UAV-borne multispectral sensing, enhancing the detection performance for occluded landmines by combining RGB and NIR image data.

Table 1 summarizes the following studies and gives an overview of related work. Multi-spectral and thermal infrared sensors integrated on UAVs significantly enhance wide-area landmine detection and mapping, as demonstrated by Baur et al. [22]. The use of an automated UAV-Survey system showcases a significant advancement, with high testing accuracy indicating a promising direction for future humanitarian demining initiatives [22]. Forero-Ramirez et al. [33] also discusses an innovative method for detecting antipersonnel landmines (APL) using UAVs equipped with thermal cameras. The technique involves digital image processing and pattern recognition applied to thermal images acquired from natural terrains. It achieved notable success rates in identifying areas with landmines, demonstrating UAVs' potential for safe and efficient landmine detection in challenging environments.

Also, the integration of GPR systems on drones elevates their effectiveness in this sensitive task. GPR systems are adept at creating detailed images of what lies among the surface, proving particularly useful in spotting mines that are made with minimal metal and are tricky to detect. This feature is a game-changer in regions where such mines pose a constant threat to communities [24]. Nonetheless, incorporating GPR systems on drones isn't without challenges. The main issue is the limited flight time of drones, which restricts the area that can be surveyed in one go [24].

Despite these challenges, introducing UAVs into the landmine detection process marks a significant stride forward. Compared to the traditional methods, the drone-based approach is not only quicker but also considerably safer, allowing for the survey of extensive areas without exposing the detection teams to the dangers of landmines [31]. The fusion of advanced machine learning with drone technology is setting new benchmarks in landmine detection, offering hopes for a future where the process is not only faster but also markedly safer and more accurate [31].

As we navigate through the challenges of landmine detection, it's evident that UAVs, equipped with advanced sensing technologies like GPR and machine learning algorithms, are supporting this transformative journey. These technologies promise not just to enhance the efficacy and safety of landmine detection operations but also to significantly cut down the number of casualties and injuries caused by these hidden threats. We aim to harness the immense potential of these technologies to forge a safer, landmine-free world.

TABLE 1. Concept matrix of related work.

Source	Year	UAV based Technology	Methodology	Target Detection	Performance
Fernandez et al. [24]	2018	GPR and SAR algorithm	Clutter removal technique; High-resolution underground imaging	Metallic landmines	Not specified
Vivoli et al. [29]	2019	Optical Imaging	YOLOv8 models for real-time detection	PFM-1 and PMA-2	98.4% accuracy
Baur et al. [22]	2020	Multi-spectral & thermal sensors for wide-area surveys	Automated technique with Faster R-CNN	Scatterable plastic landmines (PFM-1)	99.3% accuracy (partially withheld set) and 71.5% (completely withheld set)
Almutiry et al. [3]	2020	Tomographic Synthetic Aperture Radar (TSAR)	3D imaging	Nonmetallic landmine	Not specified
Alvey et al. [30]	2021	Simulation data (Unreal Engine & Microsoft's AirSim plugin)	NN model for automatic EHD	Not specified	Not specified
Yoo et al. [31]	2021	Magnetometer system and laser altimeter	Magnetometer on a pendulum to minimize noise; Detection of magnetic anomalies	Metal antipersonnel landmines (M16)	Not specified
Priya et al. [32]	2021	Thermal Imaging	RCNN for image processing	Buried objects and mines	90% accuracy in target location prediction
Forero et al. [33]	2022	Thermal camera	Multilayer Perceptron (MLP)	Antipersonnel landmines (APL)	97.1% (1 m from ground) and 88.8% (higher altitudes)
Bajic et al. [34]	2023	Thermal Imaging	YOLOv5 for detection	Unexploded ordnance	Over 90% recognition accuracy across eleven object classes
Qiu et al. [35]	2023	Multispectral sensing (RGB and NIR)	Joint fusion and detection via deep learning	Scatterable landmines	Not specified

B. ANALYSIS OF DEEP LEARNING METHODS IN LANDMINE DETECTION RESEARCH

Baur et al. [22] emphasized the power of high-resolution multi-spectral remote sensing combined with Faster R-CNN in the detection and mapping of scatterable APL. Their approach extends the capabilities of low-cost plug-and-play multi-spectral sensors to detect surface-laid antipersonnel mines across various spectral bands, showcasing the potential to decrease time and cost while increasing accuracy [22]. Although Forero-Ramirez et al. [33] primarily focuses on digital image processing and a multilayer perceptron (MLP) classifier. Its success underscores the growing importance of advanced computational methods in landmine detection. The findings suggest a promising avenue for incorporating deep learning in future research to further enhance detection capabilities, given its success in various image recognition tasks.

Alvey et al.'s [30] research ventured into the complex world of computational intelligence, focusing on the pivotal role of neural networks in automatic object detection. This study highlighted the intricacies involved in deciphering the inner workings of these networks. Alvey et al. [30] addressed this challenge by leveraging high-quality graphics simulations to generate a vast pool of accurately labelled data. This approach enabled a detailed assessment of the strengths and weaknesses of neural network models designed for explosive hazard detection via UAVs. The reliance on simulated scenarios may not fully encapsulate the unpredictability and varied nature of real-life environments, pointing towards a necessity for models that are capable of adapting and learning from genuine field data [30].

The utility of hyper-spectral data in improving the detection capabilities of UAVs, focusing on the accurate representation of explosive devices in spectral scenes for enhanced identification [40]. Bajic et al. [40] discusses the innovative modeling and simulation methods for the spectral data of explosive targets, highlighting the importance of precise data acquisition and processing for successful detection. They also underline the significance of advanced data analysis techniques, like Spectral Angle Mapping, in interpreting the hyper-spectral data collected by UAVs, illustrating the progress in analytical methods accompanying UAV technological advancements [40].

Meanwhile, Fernández et al. [24] introduced an innovative system employing GPR on UAVs to capture images from beneath the ground. This system, fortified with a synthetic aperture radar algorithm and a clutter removal technique, displayed proficiency in detecting a range of targets, both metallic and non-metallic. Validated in both controlled and real-world settings, the system demonstrated its capability to conduct safe inspections of challenging terrains and detect dangerous buried objects like landmines. While this system signifies a significant leap forward, its adaptability and performance in varying and possibly harsh field conditions remain areas ripe for exploration. This observation highlights the ongoing need for detection systems that are not only precise but also versatile and robust across different environments [24].

Yoo et al. [31] took a unique route by integrating a magnetometer system with a UAV to pinpoint metal antipersonnel landmines in Korea's demilitarized zone, an area largely untouched by development. The system's design featured an enhanced laser altimeter for stable flight over natural terrains peppered with obstacles like dust and bushes. A magnetometer, mounted on a pendulum, minimized the interference from magnetic noise and drone vibrations [31]. This setup demonstrated its utility by successfully identifying numerous magnetic anomalies linked to metallic objects, showcasing its potential to mitigate risks for detection personnel and speed up the landmine detection process. However, this system's focus on metal APL and its testing in a specific geographic locale underscore the need for more universal and versatile detection solutions that can operate effectively across various global regions and conditions [31].

These studies, while making substantial contributions to the field of landmine detection through various technological approaches, reveal several critical gaps in current research. While GPR-based systems demonstrate high accuracy, their complexity, cost, and operational limitations make them impractical for widespread deployment in resource-constrained environments. Additionally, existing thermal imaging approaches, though promising, have not been fully optimized for efficient real-time detection, particularly in economically disadvantaged regions where the landmine threat is often most severe. Current research summarised in Table 1 has primarily focused on high-performance but complex detection methods, leaving a significant gap in the development of efficient, cost-effective solutions that maintain reliable detection capabilities while being accessible to regions with limited resources. This gap is particularly critical given that many landmine-affected areas are in economically challenged regions that cannot afford or maintain sophisticated detection systems. Furthermore, there is limited research on the integration of lightweight deep learning models with thermal imaging for landmine detection, an approach that could potentially address these accessibility and efficiency challenges while maintaining detection reliability.

C. GROUND PENETRATING RADAR

GPR is a sophisticated, non-invasive technique utilized for subsurface examinations. It functions by emitting electromagnetic waves into the ground and measuring the echoes that bounce back from different subsurface structures. This method is particularly effective in detecting the contrasts in dielectric properties between different materials, which allows for the identification of both metallic and non-metallic objects, such as plastic mines [41], [42]. These plastic components, often used in modern mines, can be distinctly identified due to their lower dielectric constants compared to surrounding soil materials [43]. GPR is highly valued in mine detection for its ability to explore deeper than traditional metal detectors, particularly in non-conductive, dry soils where its penetration capabilities are maximized [44]. The technology's depth of reach and ability to discern different types of mines make it a vital tool in areas where undetected mines pose a serious threat. However, the application of GPR is not without challenges. The technology's effectiveness can be significantly hampered by soil composition and topographical irregularities, which can create noisy data or clutter. This clutter necessitates the use of complex signal processing techniques to clarify the data and achieve reliable results [45], [46]. The reflective nature and penetration depth of the radar signals can be adversely affected by moist or highly conductive soils, complicating the detection process [42], [45].

D. COMPARISON OF GROUND PENETRATING RADAR AND THERMAL IMAGING

When comparing GPR and thermal imaging for landmine detection, several critical factors highlight the advantages of thermal imaging over GPR, particularly in operational efficiency and practical deployment. GPR systems, although detailed in their subsurface scanning capabilities, are burdened with several disadvantages. The technology's sensitivity to environmental conditions, such as soil moisture, mineral content and (topographical) terrain characteristics, can drastically affect its performance, requiring significant calibration and adjustment on-site [47]. Moreover, GPR equipment in combination with an UAV is generally bulky and heavy as exemplary shown in Fig. 2, which limits the flight time when these systems are mounted for aerial surveys. This weight issue results in reduced operational efficiency, as frequent landings are necessary to recharge or change batteries, thereby limiting the area covered during each flight [48].

Additionally, the complexity of GPR data necessitates specialized experts for both processing and interpretation, which can hinder quick deployment and immediate action in mine detection operations [14], [15]. The scarcity of trained professionals is exacerbated in economically disadvantaged and high-risk regions, like Cambodia and Afghanistan, where the risks and living conditions deter potential experts from working [7], [10]. Furthermore, establishing training programs in these areas is fraught with difficulties, including limited access to educational resources, inadequate infrastructure, and ongoing security threats [16]. These factors collectively hinder the rapid and effective implementation of GPR technology for mine detection, necessitating the search for simpler and more accessible alternatives.

The financial aspect of deploying GPR is also a notable disadvantage. GPR systems are expensive to acquire and maintain, making them less accessible for use in underfunded or economically constrained regions [49]. This cost includes not only the direct expenses of locating and safely removing the mines but also the support and logistic costs necessary to ensure the safety of deminers and the effectiveness of the operation [49].

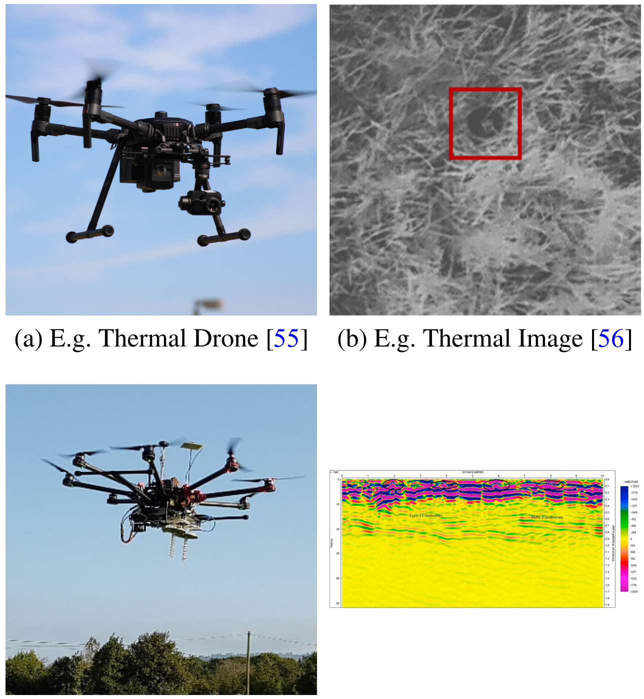
Since 1989, over 18 million explosive remnants of war and more than 730,000 anti-personnel mines have been removed in Afghanistan. Despite these efforts, the casualty rate remains high, with monthly casualties increasing from about 36 in 2012 to over 150 by 2017, according to Patrick Fruchet, Programme Manager of UNMAS, the United Nations Mine Action Service [50]. The current landmine contamination is categorized as massive, affecting more than 100 km^2 of land [7], [10]. The high costs associated with advanced demining technologies such as GPR are prohibitive for Afghanistan, which has a gross domestic product of approximately 14.27 billion USD [51]. This financial strain is exacerbated by the country's limited economic resources, making it challenging to allocate sufficient funds for comprehensive demining operations. Consequently, there is a pressing need for more cost-effective solutions to address the pervasive landmine threat in Afghanistan and other financially limited countries efficiently and affordably.

Field studies highlight the complementary strengths of both technologies. García-Fernández et al. [25] reported that GPR achieves superior detection rates of 87% for deeply buried mines at 15 cm depth compared to thermal imaging's significantly lower performance (31%) under such conditions. GPR maintains consistent performance regardless of weather conditions or time of day, offering reliable detection even during adverse environmental circumstances. Conversely, Bajic et al. [34] demonstrated thermal imaging's excellent capabilities for surface-laid mines with 94% detection accuracy during optimal dawn/dusk thermal transitions.

While the majority of scientific publications focus on high-performance but expensive and complex methods of landmine detection, we identified and closed the research gap by developing a deep learning model combined with thermal drones that emphasizes efficiency and cost-effectiveness. Thermal imaging presents several operational advantages, making it particularly suitable for rapid deployment in diverse environments [52]. Unlike GPR, thermal imaging relies on capturing infrared radiation released by buried items, providing a more straightforward and less complex method for mine detection [43] as shown in Fig. 2.

Additionally, this technology is less susceptible to soil disturbances and requires fewer technical adjustments and data processing, thus offering a more cost-effective and user-friendly solution [47]. Typically lighter than GPR systems, thermal imaging systems enable UAVs to operate for extended periods and cover larger areas without the need for frequent recharging [53]. This increased operational efficiency makes thermal imaging more suitable for extensive area surveillance and rapid response scenarios. Furthermore, thermal imaging detects temperature differences emitted by objects on or near the surface, indicating the presence of landmines without the need for direct soil contact or complex equipment adjustments [43].

The simplicity of data generated by thermal imaging facilitates quicker interpretation and requires less specialized training compared to GPR data (Fig. 2), thereby enhancing decision-making speed and operational responsiveness [54]. Moreover, the cost-effectiveness of thermal imaging is a significant advantage; thermal cameras generally entail lower initial investment and maintenance costs than GPR systems [34]. This affordability is crucial for broader deployment, particularly in regions with limited financial resources, thus addressing the urgent need for efficient and accessible demining solutions.



(a) E.g. Thermal Drone [55] (b) E.g. Thermal Image [56]

(c) E.g. GPR Drone [57] (d) E.g. GPR B-Scan [58]

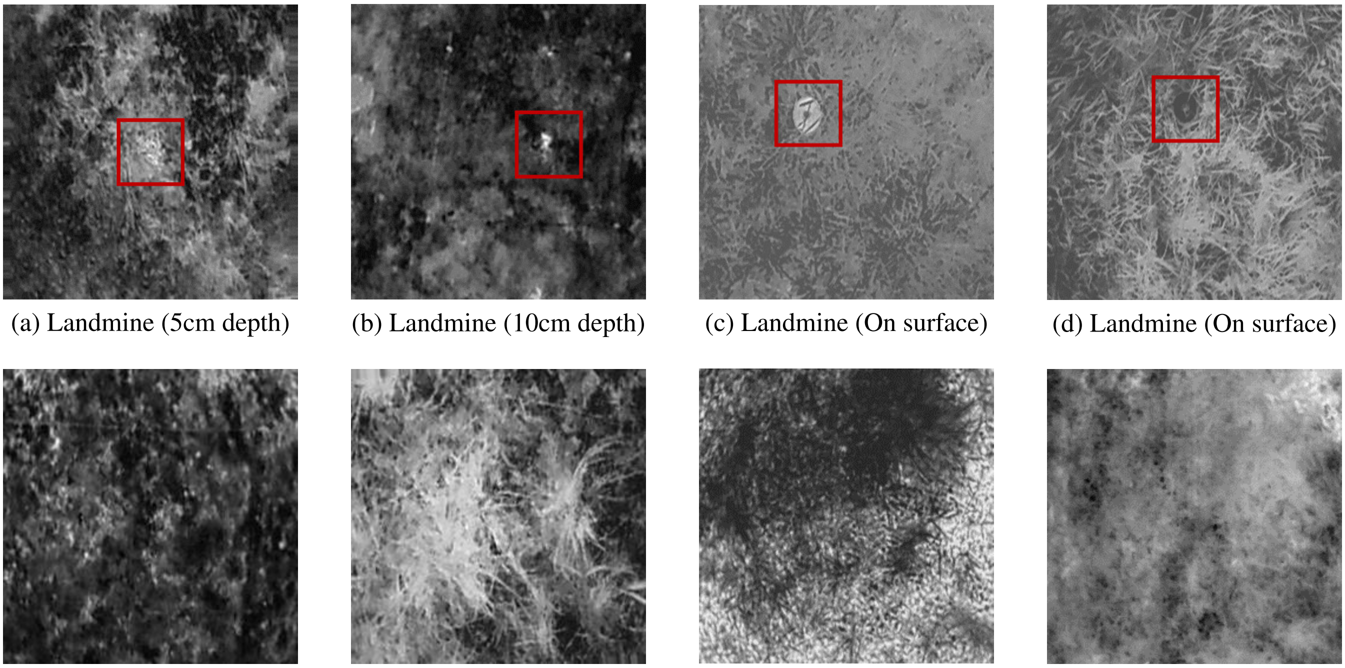
FIGURE 2. System and imaging examples.

III. METHODOLOGY

While GPR provides detailed subsurface information, the operational and logistical advantages of thermal imaging—such as lower cost, ease of use, and greater adaptability to diverse environmental conditions—make it a preferred choice for rapid and effective landmine detection in a variety of settings [43], [47]. The advantages of thermal imaging underscore the importance of further research and development in this field. Addressing the technological and financial gaps can optimize thermal imaging, providing a more effective and accessible solution for landmine detection, especially in economically disadvantaged and high-risk regions.

A. MACHINE LEARNING APPROACH

Leveraging advancements in drone technology and artificial intelligence for mine detection, this study introduces an approach utilizing TensorFlow's MobileNetV3-Large architecture for object detection. MobileNetV3 is a convolutional neural network (CNN) designed primarily for use in mobile and edge devices where computational resources are limited [59]. The MobileNetV3-Large architecture is optimized specifically for mobile phone CPUs, providing excellent accuracy with minimal latency. Our novelty lies in



(a) Landmine (5cm depth) (b) Landmine (10cm depth) (c) Landmine (On surface) (d) Landmine (On surface)

(e) No Landmine Ex. 1 (f) No Landmine Ex. 2 (g) No Landmine Ex. 3 (h) No Landmine Ex. 4

FIGURE 3. **Example images for the 2 dataset classes. Red bouding boxes are added by us.**

applying MobileNetV3-Large architecture to thermal-based landmine detection in humanitarian contexts. It enables on-device processing that enhances privacy, preserves battery life through reduced power consumption, and delivers strong performance across classification, detection, and segmentation tasks, all while maintaining computational efficiency for resource-constrained mobile environments [60]. CNNs are engineered for DL algorithms capable of identifying, recognizing, and classifying objects, as well as detecting patterns in data [61]. A standard CNN comprises convolutional layers for feature extraction, pooling layers for dimensionality reduction, and is finalized by a flattening layer connecting to hidden dense layers for classification [62]. These layers employ filters, which are learnable weight matrices and also known as kernels, to systematically scan and extract spatial hierarchies and patterns from input images [63]. The selection of a lightweight architecture for the machine learning model, deployed in UAV-based landmine detection, was motivated by considerations of safety and operational efficiency. The MobileNetV3 architecture was chosen because it outperforms heavyweight models like VGG16 and InceptionV3 in edge computing and applications with limited computational resources [64]. With this machine-learning architecture, scientists achieved new state-of-the-art results for mobile classification, detection and segmentation in the context of mobile networks [60]. Known for its rapid processing capabilities, this model is ideal for analyzing the massive amounts of data that drones gather. It can swiftly sift through the data, pinpointing potential landmines with remarkable precision. MobileNet use a simple and efficient architecture that utilizes depth-wise separable convolutions to create powerful yet lightweight deep neural networks [60]. MobileNetV3 has two new variants in comparison to its predecessor: MobileNetV3-Large and MobileNetV3-Small [60]. These models address high and low-resource use cases respectively.

B. MODEL FOUNDATION AND DESIGN

To ensure the model can recognise the input images and to get the best possible results, the data must be resized and reshaped to 224 × 224 pixels, a configuration recommended by Howard et al. [60]. The input shape is crucial as it determines the dimensionality of the input data used by the network. The MobileNetV3-Large model is applied without its top layer, enabling the integration of a custom top model designed specifically for this application. This choice reflects a strategic preference for the larger version of MobileNet, aligning with the deployment of advanced UAVs in mine detection operations, where the model's enhanced capabilities are essential for effective detection. Especially in the context of landmines, it is crucial that the model achieves outstanding results, while still being applicable to UAVs to avoid devastating consequences.

Although the MobileNetV3-Large model incurs higher energy and memory consumption, it is still considered a lightweight option. Its advantage lies in the increased number of parameters and layers, which significantly enhance accuracy in mine detection tasks, compared to its MobileNetV3-Small counterpart [60]. In addition to the parameters mentioned above, some other parameters are not mentioned because of the use of default values in the

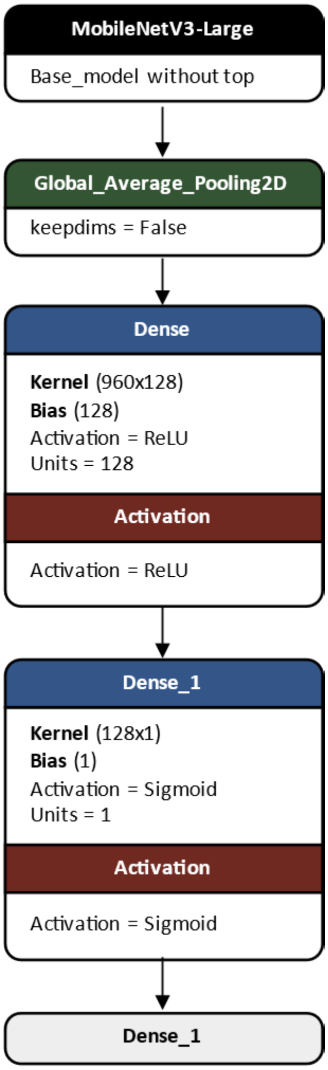


FIGURE 4. **Model architecture.**

base models or because they become irrelevant due to the removal of the default top layer. Pre-trained weights from the ImageNet dataset are loaded to leverage existing feature patterns, enhancing the speed and accuracy of the model in recognizing and classifying new landmine images [65]. To avoid overfitting, a dropout rate of 30 % is chosen, which proved to be optimal after several iterative adjustments. The base model's weights are kept non-trainable, a strategic choice in the transfer learning approach that maintains the integrity of pre-learned features during the training phase.

The development of the model top (Fig. 4) involves the addition of three layers, chosen for their compatibility with binary classification problems. The architecture is augmented with a new top layer, incorporating a GlobalAveragePooling2D layer to reduce parameters and complexity, thereby mitigating overfitting by concentrating on global image features. This is followed by a fully connected (dense) layer of 128 neurons, employing a rectified linear unit (ReLU) activation function, which enables the network to recognise complex patterns in the data. The purpose of this layer is to enable the network to learn complicated patterns and connections in the data. Finally, a single-neuron dense layer with a sigmoid activation function is implemented for the output layer, facilitating binary classification through logistic regression, which assigns input values to probabilities between 0 and 1, allowing for continuous adaptation of model weights during training [66]. Fig. 5 illustrates a block diagram of our proposed framework.

C. EVALUATION PROCEDURE

To mitigate the risk of data leakage, the landmine dataset is first randomly split into 80% training, 10% validation, and 10% testing using the Python library split-folders with a fixed seed of 42, ensuring reproducibility. This split is performed before applying the augmentation process. Utilizing the 10% of images, which have not been previously screened by the model for its evaluation, ensures a rigorous assessment of the model's performance. By conducting data augmentation subsequent to the initial data split, any overlap of augmented data between the data splits is prevented, thereby preserving the integrity of the evaluation process. To comprehensively evaluate the model's performance, the accuracy, balanced accuracy, precision, recall, and a confusion matrix are analyzed. The confusion matrix, in particular, offers detailed insights into true positives (correctly identified landmines), true negatives (correctly identified areas without landmines), as well as false positives (areas incorrectly classified as containing landmines, i.e., type I errors) and false negatives (landmines incorrectly classified as safe areas, i.e., type II errors), providing a nuanced understanding of the model's efficacy and potential areas for improvement.

It is important to note that while our test set comprises 233 images (approximately 10% of the total dataset), this sample size follows standard machine learning evaluation protocols and provides statistically valid assessment of model performance. The specialized nature of landmine detection using thermal imaging inherently limits dataset sizes due to the controlled conditions required for data collection, safety protocols, and specialized equipment needed, as documented by Tenorio-Tamayo et al. [56].

To create a basis for comparison with non-lightweight models, we performed the exact same procedure for training and evaluation with the same data splits for three state-of-the-art deep learning models ResNet50 [67], Inception [68], [69] and VGG19 [70]. This allows us to compare the evaluation of the results and thus assess whether the light-weight approach is comparable with the heavy-weight approaches, which are not optimized for use in drones. For the fine-tuning of the alternative models, we used different amounts of layers due to the different depths of the models. For ResNet we unfroze the last 20 layers, for Inception the last 50 and for VGG19 the last 8 layers.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (1)$$

$$BalancedAccuracy = \frac{TPR + TNR}{2} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$Cohen's\ Kappa = \frac{P_0 - P_e}{1 - P_e} \tag{5}$$

TP = True Positives
TN = True Negatives
FP = False Positives
FN = False Negatives
TPR = True Positive Rate
TNR = True Negative Rate
P_0 = Observed Agreement
P_e = Expected Agreement

D. DATA AUGMENTATION AND PRE-PROCESSING

The augmentation and pre-processing phase is vital for preparing the data before training. The image dataset is meticulously categorized into two distinct classes: images depicting landmines and those without. This rigorous selection process culminates in a refined dataset comprising 1,993 high-quality images, deemed suitable for subsequent augmentation. Furthermore, the methodology employed facilitates the equalization of previously existing imbalances within the dataset. The pre-processing and augmentation of this curated dataset are efficiently handled by TensorFlow's Keras ImageDataGenerator class [71], which allows for real-time data augmentation by performing selected transformations on the images.

During the pre-processing phase, the images undergo a series of precisely calibrated transformations to improve the model's ability to generalize across various conditions and to enhance the quantity of data. Firstly, the images are randomly rotated within the range of $[-10, 10]$ degrees. This is done to simulate diverse viewing angles, which would give the model rotational invariance and enable it to recognize objects reliably regardless of their orientation. Moreover, in order to replicate the variance in object size as perceived by the observer, a random zoom of up to 10 % is applied to the images. This zoom factor was carefully selected to ensure the preservation of essential object features while simultaneously introducing a degree of scale variance into the dataset. To address the empty spaces resulting from the aforementioned zoom and rotation operations, the 'nearest' fill mode is used. This method successfully fills in missing areas by assigning the value of the nearest pixels, resulting in a smooth transition in the modified regions. Initially, images are resized to a uniform format of 224×224 pixels with three color channels, ensuring consistency in input data. In addition the ImageDataGenerator significantly expands the dataset size by generating altered versions of the original images. By applying transformations including rotation, scaling, and flipping, the dataset size is increased tenfold, creating nine alternative versions for each original image. This expansion enriches the model's training data, enhancing its ability to generalize from varied representations of landmines. It is important to note that no horizontal or vertical shifts are made, preserving the original framing of the subjects in the images. Another improvement to the dataset is made by adjusting the image brightness. Brightness levels are randomly varied between 70 and 100 %, introducing a diverse range of lighting conditions into the dataset. This strategic variation not only introduces an element of lighting variance but also significantly enhances the model's adaptability to fluctuations in illumination, which is a critical factor for real-world applications. To further augment the dataset and simulate different viewing perspectives, both horizontal and vertical flipping are employed. This step was crucial in enabling the recognition of objects by the model irrespective of their orientation in the image frame.

Another important step is the normalization of pixel values, achieved by rescaling the standard range of [0, 255] to [0, 1] [72]. This normalization, essential for processing by neural networks. It transforms the data into a range that is more suitable for processing by the neural network, ensuring that the gradient descent steps during the model's training are in an optimal scale, which in turn streamlines the computational demand. Such uniformity facilitates stable and efficient model training, ultimately enhancing performance and predictive accuracy. With these settings, a total number of 1,993 images is obtained, distributed over three splits at the end of the augmentation.

Given the imbalance in the used dataset, where the number of images with landmines significantly outnumbers those without landmines, asymmetrical augmentation is applied, and the class weights of [0.9765] for the non-mine class and [1.0246] for the mine class are additionally computed. This technique helps to mitigate the remaining imbalance by assigning a higher penalty to misclassifications of the underrepresented class, thereby ensuring that the presented model pays more attention to detecting landmines accurately.

A binary cross-entropy loss function is utilized, which is suitable for binary classification problems such as the presented [73], where the task is to differentiate between images containing landmines and those that do not. The model is configured using the Adam optimizer, as this is well suited for binary CNNs in terms of performance [74], and the learning rate is set to 0.0001, a relatively low value that allows for gradual and more stable convergence during training, avoiding jumps in accuracy.

The model is trained for 20 epochs with 25 steps per epoch. A dropout rate of 30% is chosen to prevent overfitting, a common problem in machine learning where models perform well on training data but are unable to generalize sufficiently, thereby encouraging the model towards greater generalization during training for improved performance on unseen data [75], [76]. Since too small batch sizes can lead to high gradient variance, making learning unstable, and too

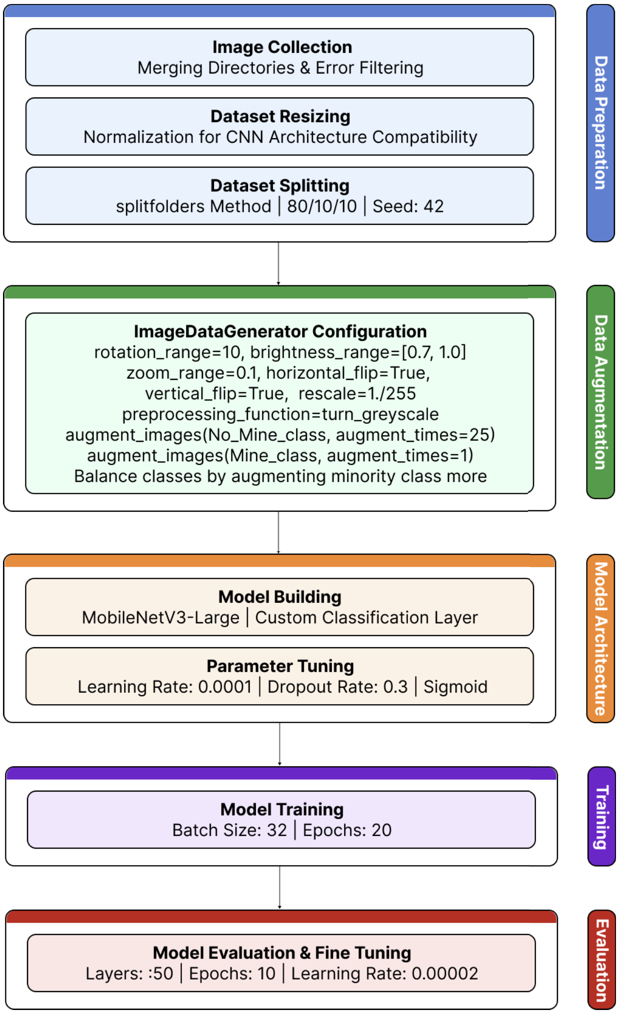


FIGURE 5. Block diagram.

large batch sizes can reduce variance but increase the risk of not finding the best solution [77], a batch size that aims to balance efficiency and stability in the model's learning process is chosen [78]. Therefore, after iterative adjustments, a batch size of 64 is selected. After training the model, we fine-tuned it for another ten epochs. To do this, we divided the learning rate by five and unfroze the last 50 layers of the model in order to adjust them minimally with a reduced learning rate.

E. DATASET

The dataset employed in this study was developed by Tenorio-Tamayo et al. [56], comprising 2,700 monochrome thermographic images captured in a terrain with low vegetation percentage and clay soil composition (35.6% sand, 35.4% clay, and 29% silt with a moderately acidic pH of 5.73). These images were collected using a Zenmuse XT infrared camera (7.5-13.5 μ m wavelength range and a thermal sensitivity of less than 50 milliKelvins) mounted on a DJI Matrice 100 drone, following established protocols for optimal thermal contrast detection. The dataset encompasses significant variability across multiple dimensions to ensure model robustness. Nine distinct zones within a 10×10 m area contained simulated antipersonnel mines of type Legbreaker, with systematic variation in burial depths: three zones with landmines at 1 cm depth, two zones at 5 cm depth, one zone at 10 cm depth, two zones with surface-placed landmines, and one control zone without landmines. The simulated mines were constructed with PVC cylinder housing, equipped with a 5 ml syringe for detonation, and loaded with 20 g of nails and 400 g of anthracite coal. The anthracite coal was specifically selected for its thermal characteristics closely resembling those of TNT in terms of specific heat and thermal conductivity, while providing safety during experimentation. To account for different operational scenarios, images were systematically captured at altitudes ranging from 1 to 10 meters (at 1-meter intervals), with additional temporal variability introduced by acquiring thermograms across three different days. For enhanced dataset diversity, 11 additional thermograms were captured at 1-meter height in each zone (one image per second), accounting for drone position variations due to wind and thermal camera noise. All images were acquired between 5 and 6 pm on mostly sunny, cloudless days with ambient temperatures between 26 and 30°C and wind speeds of 3-5 m/s. These diverse thermal conditions ensured heterogeneous temperature contrasts across the dataset, a critical factor for thermal imagery where temperature gradients directly determine detection efficacy. The varying wind speeds introduced slight blurriness in some images, further enhancing model robustness by training it to recognize mines under different environmental disturbances. The raw thermographic images have not been modified, calibrated, or subjected to any pre-processing, preserving the authenticity of the thermal signatures. Equation (6) shows the mathematical model with which the noise-to-signal ratios for the thermal radiations can be calculated for each recording day for the thermal images. To do this, the thermal signals (S) of the target and the background, which are specified in the data set for each day, must be subtracted and set in relation to the standard deviation of the camera noise.

$$Noise - to - Signal - Ratio = \frac{S_{target} - S_{background}}{\sigma_{noise}} \tag{6}$$

In Figure 3, a selection of representative images is showcased, categorized into two classifications: "Mine" (a-d) and "No Mine" (e-h). These examples illustrate the terrain characteristics and detection challenges. The red bounding boxes indicating landmine locations were manually added by the research team after analysis solely for the purpose of visualization but do not represent an automated localization process. In summary, these annotations were created by the authors of this paper and do not derive from the dataset. For consistency in presentation, these images have been standardized to a 1:1 format.

F. DATA FILTER ALGORITHM

To ensure the reliability and functionality of the proposed binary image detection model, a meticulous pre-processing of the dataset was inevitable. This integral phase involved removing images due to mislabeling, defective image files and a few other factors, which are briefly described below. The methodological procedure is documented in Fig. 6. The methodological approach begins with adjustments based on the file properties of the data. Then, a process inspired by André Berchtold's [79] test-retest method is utilized to maintain high data quality. This method ensures the consistency of the exclusion criteria through independent evaluations (here by the three authors), where an image is excluded only if at least two deem it non-representative, achieving a reliability of 95.62%. Further, Schuler et al.'s [80] workflow for test-retest reliability in metric conjoint experiments is incorporated. This workflow involves repeated assessments to verify the consistency of the exclusion criteria over time, ensuring rigorous data handling procedures. First, all TIFF files are removed as they are not readable. Furthermore, some images, mainly in the class with existing mines, have blue markings (Fig. 9). Since these markings could cause the model to focus on them during training and thus impair its robustness, images including this particularly hex color code are filtered out. Through detailed analysis, it was discovered that images taken from altitudes below one meter are excessively sensitive, frequently producing visuals that appear overly bright or, in some instances, blurred (Fig. 7). Conversely, images taken from heights above five meters predominantly reveal the quadrant demarcations of the survey area and frequently depict mines from adjacent zones (Fig. 8). This complexity in the aerial imagery not only challenges the delineation of surveyed quadrants but also paves the way for misinterpretations in data classification. Consequently, the presence of images labeled as 'No Mine' that visually contain mines indicate a misclassification issue within the dataset. This mislabeling presents a significant challenge, as it could lead to the model incorrectly learning the association between image features and the presence or absence of mines, thereby reducing the model's detection accuracy [81]. To ensure the sorting process remained objective, valid, and aligned with the goal of maximizing the model's predictive performance, the test-retest method inspired by André Berchtold's is applied [79]. Across all images concerned, the authors achieved an overall reliability of 95.62 %. Furthermore, images with considerable blurriness are encountered, a phenomenon observed irrespective of the altitude from which they are taken. These images are deemed non-representative and are removed, as well as data that exhibit strong deviations in contrast settings, especially within a single flight's dataset.

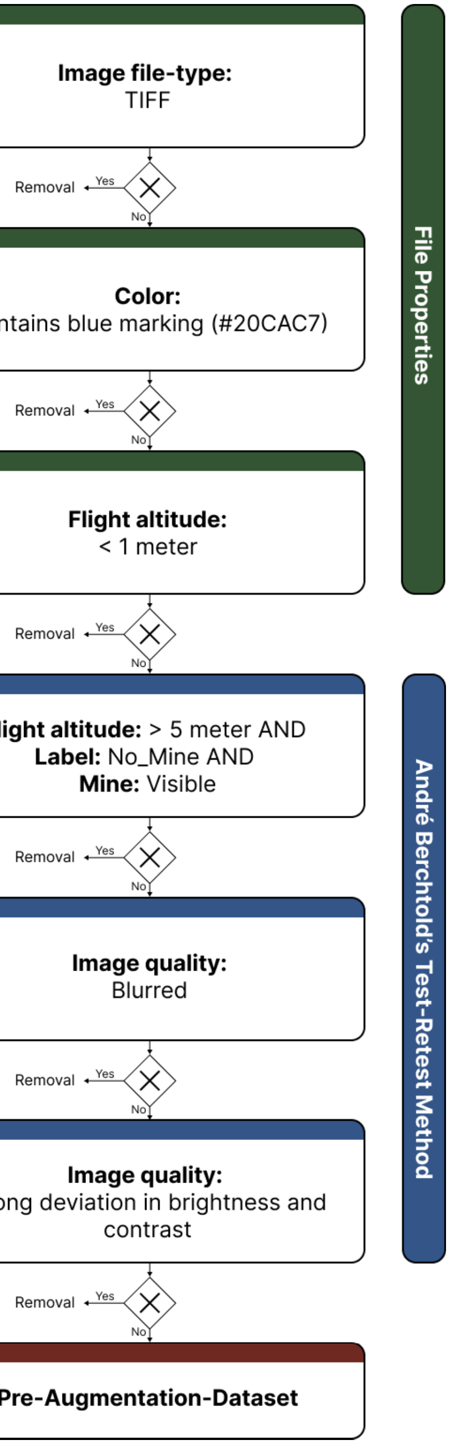


FIGURE 6. Adjustment algorithm.

IV. RESULTS

The model results, achieving a training accuracy of 96.97%, a validation accuracy of 97.19% (Fig. 10), and a test accuracy of 96.14%, demonstrate a high level of efficacy in landmine detection using thermal imaging from UAVs. The low training and validation losses of 0.1028 and 0.0894 (Fig. 11), respectively, further corroborate the model's ability to generalize well to unseen data, reducing the risk of overfitting. The accuracy of 96.14% and the balanced accuracy of 95.95% (Tab. 2), which accounts for any class

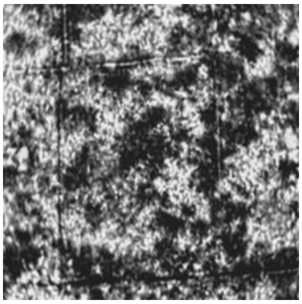


FIGURE 7. Contrast discrepancies.

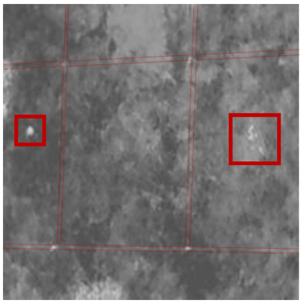


FIGURE 8. Overlapping zones.

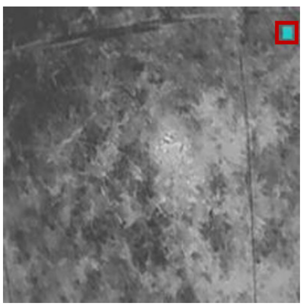


FIGURE 9. Marked images.

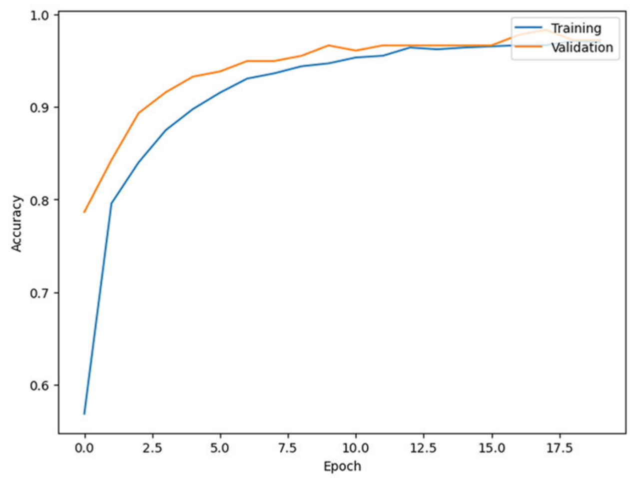


FIGURE 10. Model accuracy.

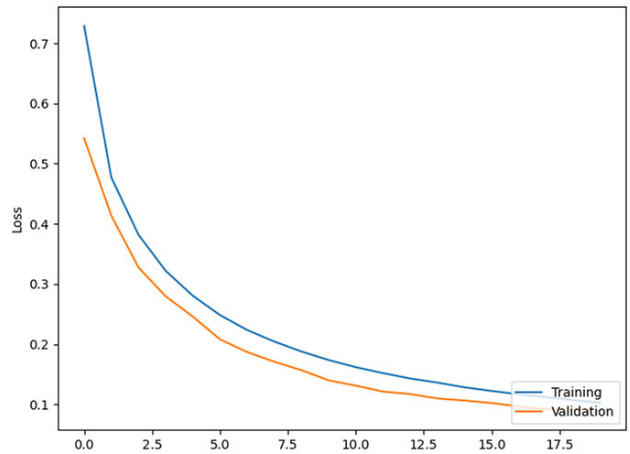


FIGURE 11. Model loss.

imbalance by averaging the recall obtained on each class, ensure reliable performance in both identifying landmines and avoiding false positives. This metric indicates that the model is equally proficient in correctly identifying landmines and non-landmines, minimizing the risk of false negatives and false positives. In a detailed evaluation involving 233 images, our model demonstrated high efficacy in landmine detection, as evidenced by the confusion matrix shown in Figure 12. The matrix reveals that the model correctly identified 98 true positives (mines) and 126 true negatives (non-mines), while only misclassifying 6 actual mines as no-mines (false negatives) and incorrectly classifying 3 non-mines as mines (false positives). This performance translates to a recall value of 94.23% (Tab. 2), indicating that the model accurately detects the vast majority of mines, which is crucial for ensuring safety in demining operations.

The Cohen's Kappa score of 0.9216 (Tab. 2), which approaches the perfect agreement mark of 1.0, reflects a strong concordance between the model's predictions and the actual classifications, beyond what would be expected by chance. This high Kappa score signifies that the model's performance is consistent and reliable, even in varying and challenging conditions encountered in real-world demining operations. Such a high level of agreement underscores the model's robustness and its capability to generalize well across different terrains and environmental scenarios, which is critical for ensuring the safety and effectiveness of landmine detection efforts in diverse operational contexts.

V. DISCUSSION

A. MODEL PERFORMANCE AND COMPARATIVE FRAMEWORK ANALYSIS

In our study, we employed TensorFlow's MobileNetV3-Large architecture combined with UAV-based thermal imaging for landmine detection, aiming for real-time detection capabilities. Our approach, focusing on thermal imagery,

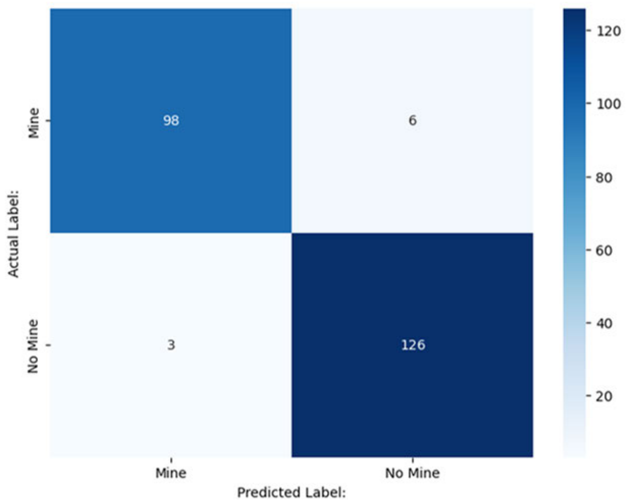


FIGURE 12. Confusion matrix.

contrasts with Baur et al. [22], who utilized UAV-based multi-spectral and thermal infrared sensors alongside a Faster R-CNN model. While Baur et al. [22] reported high accuracy in detecting surface-laid mines, our methodology highlights the efficiency and rapid deployment potential of lightweight models.The juxtaposition reveals a shared goal of enhancing landmine detection through advanced technologies but emphasizes different aspects: Baur et al. [22] focus on the precision of multi-spectral data, while our research emphasizes the agility and real-time application of machine learning models. Both studies underscore the evolving landscape of UAV-assisted demining, suggesting a future direction towards integrating diverse imaging techniques and machine learning advancements to tackle the complexities of landmine detection.

Our implementation of MobileNetV3-Large demonstrates several operational advantages over traditional GPR-based approaches discussed by Fernández et al. [24]. The lightweight nature of our model addresses key limitations identified in previous studies, particularly regarding deployment efficiency and resource utilization. The balanced accuracy of 95.95 %, achieved across test images containing landmines at various burial depths (surface to 10cm), indicates consistent performance across different deployment scenarios. This is particularly significant for practical demining operations where landmines may be buried at varying depths. This versatility is significant when compared to the limitations of GPR systems noted by García-Fernández et al. [25], especially regarding operational speed and equipment complexity.

Our approach achieves 96.14% accuracy, comparable to benchmarks by Vivoli et al. [29] (98.4%) and Baur et al. [22] (99.3%). While some methods report higher accuracy, they use more complex architectures with greater computational demands. Our methodology addresses a novel problem domain not previously examined in this manner. This unique problem-methodology combination delivers higher accuracy than most studies while maintaining consistent performance between validation and test data. This efficiency-accuracy balance makes our approach suitable for resource-constrained applications like UAV systems.

The study by Forero-Ramirez et al. [33] developed an automatic detection system for antipersonnel landmines using aerial thermographic images processed by a MLP classifier. This system achieved high detection success rates, demonstrating average success percentages of 97.1 % for images captured at one meter above the ground and 88.8 % at higher altitudes. The methodology highlights the efficacy of MLP classifiers in pattern recognition from thermal images acquired via UAVs, underscoring the potential for non-invasive, accurate detection of landmines in natural terrains. The comparison reveals complementary approaches: Forero-Ramirez et al.'s [33] work underscores the precision and reliability of MLP classifiers in handling thermal imagery for landmine detection, while our research advocates for the agility and operational advantages of deploying advanced lightweight models in UAV systems. Both studies contribute valuable insights into the integration of machine learning and UAV technology for landmine detection, yet each explores different aspects of this integration, from classifier selection to model efficiency and deployment strategies.

Our classification-focused approach is purposely designed for humanitarian demining's most critical need: reliable identification of potential danger zones across large areas with minimal resources. At 96.14% accuracy, our system effectively bridges the gap between theoretical capabilities and practical field deployment, operating on cost-effective hardware while extending drone operation time. Our thermal approach also detects non-metallic mines often missed by traditional methods, delivering an immediately deployable solution where accessibility and reliability create tangible humanitarian impact. Additionally, our method is simpler compared to accurate but resource-intensive approaches like GPR. The good results of the approach can be attributed to the combination of MobileNetV3-Large's efficient feature extraction capabilities with its depth-wise separable convolutions [60]. This architecture provides an optimal balance between computational efficiency and feature detection performance, allowing the model to identify subtle thermal signatures while remaining deployable on resource-constrained UAV platforms unlike more complex GPR-based alternatives while still maintaining high accuracy, making it more accessible for humanitarian endeavors with constrained resources.

The results of our study underscore the significance and robustness of our machine learning model in the detection of landmines through thermographic imaging from drones. With a remarkable validation accuracy of 97.19 % (Fig. 10), our model demonstrates high proficiency in distinguishing between the presence and absence of mines. The high validation accuracy demonstrates that the model

TABLE 2. Evaluation metrics.

Evaluation Metric	Value
Accuracy	96.14 %
Balanced Accuracy	95.95 %
Precision	97.03 %
Recall	94.23 %
Cohens Kappa	0.9216

has successfully learned the distinctive thermal signatures of buried landmines, a crucial capability for reliable detection. Moreover, this performance level aligns favorably with other state-of-the-art detection systems, such as those reported by Bajic et al. [34], who achieved over 90 % recognition accuracy in their study of unexploded ordnance detection. The relatively low validation loss of 0.0894 (Fig. 11) further attests to the model's limited error potential, indicating a reliable performance in mine detection tasks. This low loss value is particularly meaningful as it quantifies the model's confidence in its predictions. A validation loss below 0.1 suggests that the model not only makes correct classifications but does so with high confidence, minimizing uncertainty in the detection process. This is essential in the context of landmine detection, where confident and accurate predictions are crucial for operational safety.

The confusion matrix (Fig. 12) provides a comprehensive overview of the model's performance, categorizing the outcomes into true positives, true negatives, false positives, and false negatives. This categorization enables a nuanced understanding of the model's characteristics and optimization potential. In analyzing the model's error distribution patterns, we observe an instructive relationship between false positives (i.e., error I) and false negatives (i.e., error II) in safety-critical landmine detection operations. While our model achieves exceptional performance metrics across standard evaluation criteria, the distribution between false positives (3) and false negatives (6) presents an opportunity for further optimization in future research. False negatives, where the model fails to identify actual landmines, are particularly relevant in safety-critical applications, while false positives, where the model incorrectly identifies a non-mine area, primarily affect operational efficiency through additional verification requirements.

The demonstrated recall value of 94.23% underscores the model's strong detection capabilities, while also highlighting potential avenues for further enhancement in future iterations. This error distribution pattern suggests opportunities for architectural refinement, particularly regarding the model's sensitivity to subtle thermal signatures indicative of buried explosives. Future research directions could explore advanced neural architecture modifications specifically designed to optimize detection performance based on established principles in high-stakes detection systems where detection reliability takes precedence over resource efficiency. To advance these capabilities, systematic approaches through expanded synthetic data augmentation, multi-modal sensing capabilities, and advanced training strategies could further enhance the model's already strong performance. This development pathway aligns with best practices in safety-critical systems, where continuous optimization of detection reliability remains a key focus, particularly for humanitarian demining operations where operational safety is paramount.

TABLE 3. Evaluation metrics - alternative architectures.

Evaluation Metric	ResNet	VGG19	Inception
Accuracy	97.85 %	99.57 %	91.84 %
Balanced Accuracy	97.68 %	99.51 %	92.54 %
Cohens Kappa	0.9564	0.9913	0.8375

In order to be able to compare the results achieved with our light-weight approach with state-of-the-art non-light-weight models, the results of three such models using the same methodology are listed in table 3. For all three metrics, the table shows that we achieve comparable and competitive results with our approach. While the light-weight model performs approx. 2 % worse compared to ResNet and approx. 3 % worse compared to VGG19, the model even performs approx. 3 % better compared to the Inception architecture. Even though every percentage is highly relevant for a landmine detection application, this comparison shows the potential of a light-weight approach using remote-controlled drones on which these models can be used in a mobile manner without putting people in danger.

B. PRACTICAL IMPLICATIONS

The translation of our research findings into real-world applications reveals several significant operational advantages that enhance the practical utility of our proposed system in field deployments.

The practical implementation of our system benefits from the lightweight nature of the MobileNetV3-Large architecture, addressing computational resource limitations often encountered in field operations. This advantage is particularly relevant when considering the challenges of deploying complex detection systems in remote or conflict-affected areas, as noted by Barnawi et al. [5]. The system's scalability is enhanced by several key characteristics that contribute to its practical applicability. Firstly, the minimal training requirements for operators make it accessible to a broader range of personnel, reducing the need for extensive specialized training programs. Additionally, the system demonstrates reduced infrastructure needs compared to GPR systems, making it more suitable for deployment in resource-constrained environments. The faster deployment capabilities further enhance its practical utility, allowing for rapid response in critical situations. Moreover, the lower maintenance requirements contribute to long-term operational sustainability, reducing both technical complexity and associated costs of system upkeep.

The operational efficiency is further enhanced by the system's demonstrated performance in thermal imagery analysis. The ability to conduct non-invasive detection through UAV-based thermal imaging significantly reduces the risk to personnel compared to traditional ground-based methods. This advantage becomes particularly relevant in areas with difficult terrain or in post-conflict zones where physical access may be limited or dangerous. The system's capability to process aerial thermographic images effectively at various heights, as demonstrated by our results, provides operational flexibility in different deployment scenarios. The integration of UAV technology with our lightweight model addresses several practical challenges identified in traditional demining approaches. The combination of aerial deployment capabilities with real-time processing potential reduces exposure time for personnel and enables rapid assessment of larger areas compared to conventional methods. This integration particularly benefits time-critical operations and preliminary surveys of suspected hazardous areas, allowing for more efficient allocation of ground-based resources and personnel.

These practical advantages collectively position our system as a viable solution for large-scale deployment in humanitarian demining operations, particularly in resource-limited and challenging operational environments where traditional detection methods face significant constraints. The system's demonstrated reliability, coupled with its operational efficiency and reduced resource requirements, provides a sustainable approach to addressing the global challenge of landmine detection and clearance.

The presented metrics collectively suggest that UAVs equipped with thermal cameras and advanced deep learning models like MobileNetV3-Large can provide a reliable, efficient, and scalable solution for landmine detection. This approach offers significant advantages over traditional methods, including reduced operational risks, cost efficiency, and the ability to cover large areas quickly, making it particularly suitable for use in conflict zones and economically disadvantaged regions. The success of this method highlights its potential for real-world applications, offering a safer and more effective means of addressing the pervasive threat of landmines.

In addition to the advantages shown, the lightweight approach also has disadvantages in terms of practical application. As shown in the table 3, the complex models currently perform slightly better than the lightweight model in two cases. Howard et al. [60] also showed this trade-off between model complexity and efficiency. Even if the complex models perform better, the higher processing times and computational resource requirements of non-lightweight models would make the mobile and remote use of the method significantly more difficult, which is also shown by the fact that even modest increases in model complexity can lead to disproportionate growth in computational demands, with larger models requiring specialized hardware acceleration for practical deployment [60].

VI. CONCLUSION

A. SIGNIFICANCE OF THE RESULTS

The continuing presence of unexploded landmines poses an ongoing threat to the safety of people and even vehicles in various regions of the world [7], [10]. Our research, as detailed in this publication, reveals the unfortunate reality that landmines continue to claim lives, often without warning, particularly in territories burdened by war or those recovering from recent conflicts. The indiscriminate nature of these explosives makes them a relentless threat, requiring urgent and effective clearance methods to protect innocent lives.

Our study introduces a method using drones equipped with thermal imaging to detect landmines, emphasizing safety and cost-efficiency. This approach not only enhances safety but also speeds up the process of clearing landmines, demonstrating the potential of technology to tackle critical issues effectively and with reduced risk. By employing advanced algorithms, we've shown that drones can accurately identify landmines, offering a significant improvement over traditional detection methods. This method stands out for its operational speed and expansive area coverage capabilities, making it an invaluable tool for preliminary mine detection. Its straightforward application and cost-effective nature render this approach accessible to economically disadvantaged nations, offering them a vital support in their demining efforts. While GPR remains the predominant technology for mine detection [5], our findings suggest that thermal infrared imaging serves as a robust alternative, especially under constraints of resources and capabilities. This underscores the potential of our model to significantly impact humanitarian demining operations, contributing to safer environments in regions plagued by the remnants of conflict.

However, it is crucial to acknowledge that in the context of mine clearance, where the stakes involve human lives, the aim should be to achieve an accuracy and a recall as close to 100 % as possible, with zero tolerance for false negatives to ensure no mine goes undetected. Despite this, the lightweight nature of our model, combined with its integration into UAVs equipped with thermal imaging technology, presents a significant advancement for the initial reconnaissance of minefields over extensive areas. This capability is especially critical for quick and efficient action in war zones, where the model's deployment could save lives by accelerating the mine detection process and reducing the risks for ground personnel involved in demining operations.

B. LIMITATIONS

Our research on explosive material detection through machine learning faced several constraints. While we utilized simulated Legbreaker-type landmines to ensure safety, these replicas cannot fully represent the detection characteristics of actual landmines. This gap between simulation and reality emphasizes the need for more comprehensive datasets incorporating diverse environmental conditions. The dataset suffered from class contamination, with mine depictions appearing in images labeled as mine-free. These inconsistencies impacted the model's learning process and limited its classification accuracy. Additionally, our approach is restricted to mines with metal components detectable in thermographic imaging. Various landmine types, particularly non-metallic ones, remain undetectable through our method, limiting its geographical applicability in regions where such mines are prevalent.

C. FUTURE WORK

In future work, we plan to enhance our lightweight landmine detection model through dataset augmentation encompassing diverse environmental conditions. This includes creating comprehensive data across various soil types, vegetation, and weather conditions typical of conflict zones, along with different landmine dummy types. A critical focus lies on addressing the system's false negative detections which, despite a recall of 94.23%, resulted in six undetected landmines during testing. Given the life-threatening implications in humanitarian demining, we propose three approaches: synthetic data augmentation targeting challenging scenarios, implementation of sophisticated validation protocols, and testing video stream analysis instead of individual images. The latter has shown promising results in related fields [82]. While this might increase false positives from the current three cases, such a trade-off is acceptable given the safety-critical application. Additionally, we aim to integrate real-time detection capabilities using YOLOv8 [83], enabling live landmine identification with bounding boxes during operations. In general, in future work we will carry out a reliable comparison with a standardized dataset that compares different approaches for landmine detection with thermal cameras such as ours, and those of Bajic et al. [34] or Forero-Ramirez et al. [33].

ACKNOWLEDGMENT

(Daniel Heuschmid, Oliver Wacker, and Yannick Zimmermann are co-first authors.)

REFERENCES

[1] S. Belkin, S. Yagur-Kroll, Y. Kabessa, V. Korouma, T. Septon, Y. Anati, C. Zohar-Perez, Z. Rabinovitz, A. Nussinovitch, and A. J. Agranat, "Remote detection of buried landmines using a bacterial sensor," *Nature Biotechnol.*, vol. 35, no. 4, pp. 308–310, Apr. 2017.
[2] J. S. Silva, I. F. L. Guerra, J. Bioucas-Dias, and T. Gasche, "Landmine detection using multispectral images," *IEEE Sensors J.*, vol. 19, no. 20, pp. 9341–9351, Oct. 2019.
[3] M. Almutiry, "UAV tomographic synthetic aperture radar for landmine detection," *Eng., Technol. Appl. Sci. Res.*, vol. 10, no. 4, pp. 5933–5939, Aug. 2020.
[4] E. S. Endesha, H. Khalil Murad, and Y. O. Gdura, "A hovercraft prototype for landmine detection," in *Proc. IEEE 1st Int. Maghreb Meeting Conf. Sci. Techn. Autom. Control Comput. Eng. MI-STA*, May 2021, pp. 136–141.
[5] A. Barnawi, I. Budhiraja, K. Kumar, N. Kumar, B. Alzahrani, A. Almansour, and A. Noor, "A comprehensive review on landmine detection using deep learning techniques in 5G environment: Open issues and challenges," *Neural Comput. Appl.*, vol. 34, no. 24, pp. 21657–21676, Dec. 2022.
[6] L.-S. Yoo, J.-H. Lee, S.-H. Ko, S.-K. Jung, S.-H. Lee, and Y.-K. Lee, "A drone fitted with a magnetometer detects landmines," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 12, pp. 2035–2039, Dec. 2020.
[7] Int. Campaign to Ban Landmines. (2023). *Landmine Monitor 2023*. [Online]. Available: https://backend.icblcmc.org/assets/reports/Landmine-Monitors/LMM2023/Downloads/Landmine-Monitor-2023_web.pdf
[8] Int. Campaign to Ban Landmines. (2020). *Landmine Monitor 2020*. [Online]. Available: https://backend.icblcmc.org/assets/reports/Landmine-Monitors/LMM2020/Translations/LM2020.pdf
[9] R. Keeley, "Improvised explosive devices (IED): A humanitarian mine action perspective," *J. Conventional Weapons Destruction*, vol. 21, no. 1, p. 3, Apr. 2017.
[10] Monitoring Res. Committee. (2024). *Landmine Monitor 2024*. [Online]. Available: https://backend.icblcmc.org/assets/reports/Landmine-Monitors/LMM2024/Downloads/Landmine-Monitor-2024-Final-Web.pdf
[11] N. Cumming-Bruce, A. Frost, and L. Pinches. (2020). *Clearing the Mines 2020*. [Online]. Available: https://www.mineactionreview.org/assets/downloads/907_NPA_Clearing_the_Mines_2020_WEB.pdf
[12] I. Makki, R. Younes, C. Francis, T. Bianchi, and M. Zucchetti, "A survey of landmine detection using hyperspectral imaging," *ISPRS J. Photogramm. Remote Sens.*, vol. 124, pp. 40–53, Feb. 2017.
[13] J. Baur, G. Steinberg, A. Nikulin, K. Chiu, and T. D. Smet, "How to implement drones and machine learning to reduce time, costs, and dangers associated with landmine detection," *J. Conventional Weapons Destruction*, vol. 25, no. 1, p. 29, Sep. 2021.
[14] M. Manataki, A. Vafidis, and A. Sarris, "GPR data interpretation approaches in archaeological prospection," *Appl. Sci.*, vol. 11, no. 16, p. 7531, Aug. 2021.
[15] M. Quinta-Ferreira, "Ground penetration radar in geotechnics. Advantages and limitations," *IOP Conf. Ser., Earth Environ. Sci.*, vol. 221, Mar. 2019, Art. no. 012019.
[16] S. Mallapaty, "Afghanistan's terrified scientists predict huge research losses," *Nature*, vol. 597, no. 7874, pp. 15–16, Sep. 2021.
[17] M. K. Habib, "Humanitarian demining: The problem, difficulties, priorities, demining technology and the challenge for robotics," in *Humanitarian Demining. I-Tech Education and Publishing*, London, U.K., Feb. 2008, doi: 10.5772/5407.
[18] L. P. Osco, J. M. Junior, A. P. M. Ramos, L. A. D. C. Jorge, S. N. Fatholahi, J. D. A. Silva, E. T. Matsubara, H. Pistori, W. N. Gonçalves, and J. Li, "Deep learning in UAV imagery," *Int. J. Appl. Earth Observ. Geoinformation*, vol. 102, pp. 1–38, Sep. 2021.
[19] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, Jun. 2016.
[20] X. Yang and Y. Qiao, "Infrared long-distance target detection based on deep learning," in *2nd Int. Conf. Softw. Eng. and Artif. Intell. (SEAI)*, 2022, pp. 1–5.
[21] G. Zheng, X. Wu, Y. Hu, and X. Liu, "Object detection for low-resolution infrared image in land battlefield based on deep learning," in *Proc. Chin. Control Conf. (CCC)*, Jul. 2019, pp. 8649–8652.
[22] J. Baur, G. Steinberg, A. Nikulin, K. Chiu, and T. S. de Smet, "Applying deep learning to automate UAV-based detection of scatterable landmines," *Remote Sens.*, vol. 12, no. 5, p. 859, Mar. 2020.
[23] S. Wu, C. Xiao, L. Wang, Y. Wang, J. Yang, and W. An, "RepISD-net: Learning efficient infrared small-target detection network via structural re-parameterization," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5622712.
[24] M. G. Fernández, Y. Á. López, A. A. Arboleya, B. G. Valdés, Y. R. Vaqueiro, F. Las-Heras Andrés, and A. P. García, "Synthetic aperture radar imaging system for landmine detection using a ground penetrating radar on board a unmanned aerial vehicle," *IEEE Access*, vol. 6, pp. 45100–45112, 2018.
[25] M. García-Fernández, G. Álvarez-Narciandi, Y. Álvarez López, and F. Las-Heras, "Array-based ground penetrating synthetic aperture radar on board an unmanned aerial vehicle for enhanced buried threats detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5104218.
[26] X. Shi, D. Cheng, Z. Song, and C. Wang, "A real-time method for landmine detection using vehicle array GPR," in *Proc. 17th Int. Conf. Ground Penetrating Radar (GPR)*, Jun. 2018, pp. 1–4.
[27] M. García-Fernández, G. Álvarez-Narciandi, A. Arboleya, C. V. Antuña, F. L. Andrés, and Y. Álvarez López, "Development of an airborne-based GPR system for landmine and IED detection: Antenna analysis and intercomparison," *IEEE Access*, vol. 9, pp. 127382–127396, 2021.
[28] X. Xie, H. Qin, C. Yu, and L. Liu, "An automatic recognition algorithm for GPR images of RC structure voids," *J. Appl. Geophysics*, vol. 99, pp. 125–134, Dec. 2013.

[29] E. Vivoli, M. Bertini, and L. Capineri, "Deep learning-based real-time detection of surface landmines using optical imaging," *Remote Sens.*, vol. 16, no. 4, p. 677, Feb. 2024.

[30] B. J. Alvey, D. T. Anderson, C. Yang, A. Buck, J. M. Keller, K. E. Yasuda, and H. A. Ryan, "Characterization of deep learning-based aerial explosive hazard detection using simulated data," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Dec. 2021, pp. 1–8.

[31] L.-S. Yoo, J.-H. Lee, Y.-K. Lee, S.-K. Jung, and Y. Choi, "Application of a drone magnetometer system to military mine detection in the demilitarized zone," *Sensors*, vol. 21, no. 9, p. 3175, May 2021.

[32] C. N. N. Priya, S. D. Ashok, B. Maji, and K. S. Kumaran, "Deep learning based thermal image processing approach for detection of buried objects and mines," *Eng. J.*, vol. 25, no. 3, pp. 61–67, Mar. 2021.

[33] J. C. Forero-Ramírez, B. García, H. A. Tenorio-Tamayo, A. D. R. Girón, H. L. Correa, S. E. Nope-Rodríguez, A. B. López, and J. T. Buitrago-Molina, "Detection of 'legbreake' antipersonnel landmines by analysis of aerial thermographic images of the soil," *Infr. Phys. Technol.*, vol. 125, Aug. 2022, Art. no. 104307.

[34] M. Bajić and B. Potoč nik, "UAV thermal imaging for unexploded ordnance detection by using deep learning," *Remote Sens.*, vol. 15, no. 4, p. 967, Feb. 2023.

[35] Z. Qiu, H. Guo, J. Hu, H. Jiang, and C. Luo, "Joint fusion and detection via deep learning in UAV-borne multispectral sensing of scatterable landmine," *Sensors*, vol. 23, no. 12, p. 5693, Jun. 2023.

[36] A. Bhardwaj, L. Sam, Akanksha, F. J. Martín-Torres, and R. Kumar, "UAVs as remote sensing platform in glaciology: Present applications and future prospects," *Remote Sens. Environ.*, vol. 175, pp. 196–204, Mar. 2016.

[37] A. M. Cunliffe, R. E. Brazier, and K. Anderson, "Ultra-fine grain landscape-scale quantification of dryland vegetation structure with drone-acquired structure-from-motion photogrammetry," *Remote Sens. Environ.*, vol. 183, pp. 129–143, Sep. 2016.

[38] M.-L. Ding, C.-B. Ding, L. Tang, X.-M. Wang, J.-M. Qu, and R. Wu, "A W-Band 3-D integrated mini-SAR system with high imaging resolution on UAV platform," *IEEE Access*, vol. 8, pp. 113601–113609, 2020.

[39] M. García-Fernández, Y. Álvarez López, A. Arboleya, B. González-Valdés, Y. Rodríguez-Vaqueiro, M. E. De Cos Gómez, and F. Las-Heras Andrés, "Antenna diagnostics (Don't short) and characterization using unmanned aerial vehicles," *IEEE Access*, vol. 5, pp. 23563–23575, 2017.

[40] M. Bajić and M. Bajić, "Modeling and simulation of very high spatial resolution UXOs and landmines in a hyperspectral scene for UAV survey," *Remote Sens.*, vol. 13, no. 5, p. 837, Feb. 2021.

[41] A. M. Zoubir, I. J. Chant, C. L. Brown, B. Barkat, and C. Abeynayake, "Signal processing techniques for landmine detection using impulse ground penetrating radar," *IEEE Sensors J.*, vol. 2, no. 1, pp. 41–51, Feb. 2002.

[42] M. Sato, "Principles of mine detection by ground-penetrating radar," in *Anti-personnel Landmine Detection for Humanitarian Demining*. Springer, 2009, pp. 19–26.

[43] S. Lameri, F. Lombardi, P. Bestagini, M. Lualdi, and S. Tubaro, "Landmine detection from GPR data using convolutional neural networks," in *Proc. 25th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2017, pp. 508–512.

[44] I. Giannakis, A. Giannopoulos, and C. Warren, "A realistic FDTD numerical modeling framework of ground penetrating radar for landmine detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 1, pp. 37–51, Jan. 2016.

[45] J. Lee, H. Lee, S. Ko, D. Ji, and J. Hyeon, "Modeling and implementation of a joint airborne ground penetrating radar and magnetometer system for landmine detection," *Remote Sens.*, vol. 15, no. 15, p. 3813, Jul. 2023.

[46] F. Lombardi, H. D. Griffiths, M. Lualdi, and A. Balleri, "Characterization of the internal structure of landmines using ground-penetrating radar," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 2, pp. 266–270, Feb. 2021.

[47] O. A. Pryshchenko, V. Plakhtii, O. M. Dumin, G. P. Pochanin, V. P. Ruban, L. Capineri, and F. Crawford, "Implementation of an artificial intelligence approach to GPR systems for landmine detection," *Remote Sens.*, vol. 14, no. 17, p. 4421, Sep. 2022.

[48] M. Garcia-Fernandez, Y. Alvarez-Lopez, and F. Las Heras, "Autonomous airborne 3D SAR imaging system for subsurface sensing: UWB-GPR on board a UAV for landmine and IED detection," *Remote Sens.*, vol. 11, no. 20, p. 2357, Oct. 2019.

[49] M. Solla, V. Pérez-Gracia, and S. Fontul, "A review of GPR application on transport infrastructures: Troubleshooting and best practices," *Remote Sens.*, vol. 13, no. 4, p. 672, Feb. 2021.

[50] P. Fruchet. (2019). *Statement By UNMAS Programme Manager on Casualty Rates*. [Online]. Available: https://news.un.org/en/story/2019/02/1032141

[51] World Bank. (2024). *Afghanistan | Data*. Accessed: Jun. 18, 2024. [Online]. Available: https://data.worldbank.org/country/afghanistan

[52] A. Nikulin, T. S. d. Smet, J. Baur, W. Frazer, and J. Abramowitz, "Detection and identification of remnant PFM-1 'Butterfly Mines' with a UAV-based thermal-imaging protocol," *Remote Sens.*, vol. 10, no. 11, p. 1672, Oct. 2018.

[53] T. X. B. Nguyen, K. Rosser, and J. Chahl, "A review of modern thermal imaging sensor technology and applications for autonomous aerial navigation," *J. Imag.*, vol. 7, no. 10, p. 217, Oct. 2021.

[54] C. Biscarini, I. Catapano, N. Cavalagli, G. Ludeno, F. A. Pepe, and F. Ubertini, "UAV photogrammetry, infrared thermography and GPR for enhancing structural and material degradation evaluation of the Roman masonry bridge of ponte lucano in Italy," *NDT E Int.*, vol. 115, Oct. 2020, Art. no. 102287.

[55] ZLEA. (2024). *Dji Matrice 210 Fa34rt4mx9 At Embry-riddle Aeronautical University*. [Online]. Available: https://commons.wikimedia.org/wiki/File:DJI_M210_Matrice_210_(FA34RT4MX9)_(2-2-2024).jpg

[56] H. A. Tenorio-Tamayo, J. C. Forero-Ramírez, B. García, H. Loaiza-Correa, A. D. Restrepo-Girón, S. E. Nope-Rodríguez, A. Barandica-López, and J. T. Buitrago-Molina, "Dataset of thermographic images for the detection of buried landmines," *Data Brief*, vol. 49, Aug. 2023, Art. no. 109443.

[57] C. Noviello, G. Gennarelli, G. Esposito, G. Ludeno, G. Fasano, L. Capozzoli, F. Soldovieri, and I. Catapano, "An overview on down-looking UAV-based GPR systems," *Remote Sens.*, vol. 14, no. 14, p. 3245, Jul. 2022.

[58] S. M. Ebrahim, N. I. Medhat, K. K. Mansour, and A. Gaber, "Examination of soil effect upon GPR detectability of landmine with different orientations," *NRIAG J. Astron. Geophysics*, vol. 7, no. 1, pp. 90–98, Jun. 2018.

[59] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.

[60] A. Howard, M. Sandler, B. Chen, W. Wang, L.-C. Chen, M. Tan, G. Chu, V. Vasudevan, Y. Zhu, R. Pang, H. Adam, and Q. Le, "Searching for MobileNetV3," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 1314–1324.

[61] A. Ajit, K. Acharya, and A. Samanta, "A review of convolutional neural networks," in *Proc. Int. Conf. Emerg. Trends Inf. Technol. Eng. (ic-ETITE)*, Feb. 2020, pp. 1–5.

[62] J. Sachdev, S. Shekhar, and S. Indu, "Melanoma screening using deep neural networks," in *Proc. 3rd Int. Conf. Converg. Technol. (I2CT)*, Apr. 2018, pp. 1–5.

[63] B. Kim, J. Ponce, and B. Ham, "Deformable kernel networks for joint image filtering," *Int. J. Comput. Vis.*, vol. 129, no. 2, pp. 579–600, Feb. 2021.

[64] E. Kristiani, C.-T. Yang, and C.-Y. Huang, "ISEC: An optimized deep learning model for image classification on edge computing," *IEEE Access*, vol. 8, pp. 27267–27276, 2020.

[65] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

[66] S. Sharma, S. Sharma, and A. Athaiya, "Activation functions in neural networks," *Int. J. Eng. Appl. Sci. Technol.*, vol. 4, no. 12, pp. 310–316, May 2020.

[67] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[68] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.

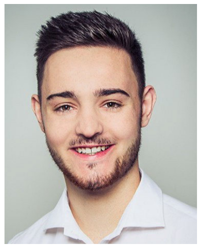
[69] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.

[70] Simonyan, "Very deep convolutional networks for large-scale image recognition," in *Proc. 3rd Int. Conf. Learn. Represent.*, Jan. 2015, pp. 1–21.

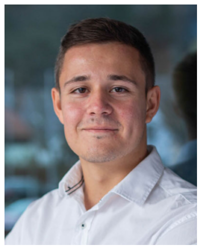
[71] TensorFlow. (2023). *ImageDataGenerator: TensorFlow API Documentation*. Accessed: Jun. 16, 2023. [Online]. Available: https://www.tensorflow.org/api_docs/python/tf/keras/preprocessing/image/ImageDataGenerator
[72] J. Mandal and D. Das, "Steganography using adaptive pixel value differencing (APVD) of gray images through exclusion of Overflow/Underflow," *Comput. Sci. Inf. Technol.*, vol. 2, pp. 1–10, Mar. 2012.
[73] U. Ruby and V. Yendapalli, "Binary cross entropy with deep learning technique for image classification," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 4, pp. 5393–5397, Aug. 2020.
[74] M. López-Sánchez, J. Hernandez, B. Hernández-Ocaña, and O. Chávez-Bosquez, "Comparative study of optimizers in the training of a convolutional neural network in a binary recognition model," *Res. Comput. Sci.*, vol. 150, pp. 73–82, Mar. 2022.
[75] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, Jan. 2014.
[76] K. Cobbe, O. Klimov, C. Hesse, T. Kim, and J. Schulman, "Quantifying generalization in reinforcement learning," in *Proc. Int. Conf. Mach. Learn.*, Dec. 2018, pp. 1–8.
[77] Z. Ma, Y. Xu, H. Xu, Z. Meng, L. Huang, and Y. Xue, "Adaptive batch size for federated learning in resource-constrained edge computing," *IEEE Trans. Mobile Comput.*, vol. 22, no. 1, pp. 37–53, Jan. 2023.
[78] N. B. Aldin and S. S. A. B. Aldin, "Accuracy comparison of different batch size for a supervised machine learning task with image classification," in *Proc. 9th Int. Conf. Electr. Electron. Eng. (ICEEE)*, Mar. 2022, pp. 316–319.
[79] A. Berchtold, "Test–retest: Agreement or reliability?" *Methodol. Innov.*, vol. 9, pp. 1–7, Jan. 2016.
[80] J. Schüler, B. S. Anderson, C. Y. Murnieks, M. Baum, and A. Küsshauer, "Test-retest reliability in metric conjoint experiments: A new workflow to evaluate confidence in model results," *Entrepreneurship Theory Pract.*, vol. 48, no. 2, pp. 742–757, Mar. 2024.
[81] T. X. Tran and R. S. Aygun, "WisdomNet: Trustable machine learning toward error-free classification," *Neural Comput. Appl.*, vol. 33, no. 7, pp. 2719–2734, Apr. 2021.
[82] D. Richard Tom Hax, P. Penava, S. Krodel, L. Razova, and R. Buettner, "A novel hybrid deep learning architecture for dynamic hand gesture recognition," *IEEE Access*, vol. 12, pp. 28761–28774, 2024.
[83] A. Aboah, B. Wang, U. Bagci, and Y. Adu-Gyamfi, "Real-time multi-class helmet violation detection using few-shot data sampling technique and YOLOv8," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2023, pp. 5350–5358.

DANIEL HEUSCHMID received the B.Sc. degree in business administration with a major in business information systems and strategic management and organization from the University of Bayreuth, Germany, in 2023, where he is currently pursuing the M.Sc. degree in digitalization and entrepreneurship. His research interests include digital entrepreneurship, business process management, and applications of machine learning in strategic and defense-related contexts.

OLIVER WACKER received the B.Sc. degree in business administration from the University of Bamberg, Germany, in 2022, with a focus on banking and taxation. He is currently pursuing the master's degree in digitalization and entrepreneurship with the University of Bayreuth. He embarked on his educational journey by completing training as a bank clerk, followed by further training as a certified banking specialist at the Frankfurt School of Finance and Management. This foundational experience in finance laid the groundwork for his academic pursuits. His studies are centered around the fusion of innovation through business models, reflecting his keen interest in digitalization for entrepreneurial success.



YANNICK ZIMMERMANN received the Bachelor of Science degree in business administration from Technische Hochschule Mittelhessen (THM), Giessen, Germany, in 2022. He is currently pursuing the master's degree in digitalization and entrepreneurship with the University of Bayreuth, focusing on the intersection of innovation and social impact through business models.



PASCAL PENAVA received the B.Sc. degree in information systems from Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany, in 2021, and the M.Sc. degree in digitalization and entrepreneurship from the University of Bayreuth, Bayreuth, Germany, in 2023. He is currently pursuing the Ph.D. degree with the Chair of Hybrid Intelligence, Helmut-Schmidt-University/University of the Federal Armed Forces Hamburg, Germany. His research interests include the development of EEG-based BCIs and machine learning-based analyses of time-series data.

RICARDO BUETTNER (Senior Member, IEEE) received the Dipl.-Inf. degree in computer science and the Dipl.-Wirtsch.-Ing. degree in industrial engineering and management from the Technical University of Ilmenau, Germany, the Dipl.-Kfm. degree in business administration from the University of Hagen, Germany, the Ph.D. degree in information systems from the University of Hohenheim, Germany, and the Habilitation (venia legendi) degree in information systems from the University of Trier, Germany. He is currently a Chaired Professor of hybrid intelligence with the Helmut-Schmidt-University/University of the Federal Armed Forces Hamburg, Germany. He has published over 160 peer-reviewed articles, including articles in *Electronic Markets*, *AIS Transactions on Human-Computer Interaction*, *Personality and Individual Differences*, *European Journal of Psychological Assessment*, *PLOS One*, and IEEE Access. He has received 18 international best paper, best reviewer, and service awards and award nominations, including best paper awards by *AIS Transactions on Human-Computer Interaction*, *Electronic Markets*, and HICSS, for his work.