CODE, TRUST AND FUTURE

THE ENGINEERING DIMENSIONS OF AI

EDITOR
JOLANA GUBALOVA

CODE, TRUST AND FUTURE: THE ENGINEERING DIMENSIONS OF AI- 2025

Edited By

Jolana GUBALOVA

ISBN: 978-625-92720-0-9

DOI: 10.5281/zenodo.17406850

October / 2025

İstanbul, Türkiye

Date: 21.10.2025

Halic Publishing House
İstanbul, Türkiye
www.halicyayinevi.com

adopted by Esma AKSAKAL
ISBN: 978-625-92720-0-9

CODE, TRUST AND FUTURE: THE ENGINEERING DIMENSIONS OF AI

EDITOR
Jolana GUBALOVA

AUTHORS
Prof. Dr. Shamim RIPON
Rubaiya RAHMAN
Moshiur Mahamud PIASH
Christian Nwankwo CHIJIOKE
Gbolahan Afeez ADIGUN
Omoboriowo Samson LOYE
Samuel Sola AKOSILE
Jolana GUBALOVA
Robert HLAVAC

TABLE OF CONTENTS

PREFACE

This volume explores how artificial intelligence is transforming healthcare, transportation, and cybersecurity. Each chapter offers a focused study on AI's practical applications, challenges, and future potential across different global contexts.

The first chapter presents a hybrid CNN–ML model for breast cancer histopathology, aiming to improve diagnostic accuracy while ensuring trustworthy AI in medical settings. The second examines Lagos State's transportation system, identifying barriers to AI adoption and proposing strategies for smarter urban mobility.

The final chapter analyzes cybersecurity practices in Slovak SMEs, highlighting vulnerabilities and offering tailored recommendations. Together, these studies underscore the need for responsible, context-aware AI to drive innovation and resilience across sectors.

Editoral Team
October 21, 2025
Türkiye

CHAPTER 1
TOWARDS TRUSTWORTHY AI IN DIGITAL PATHOLOGY: HYBRID CNN–ML MODELS FOR BREAST CANCER HISTOPATHOLOGY CLASSIFICATION

[1]Rubaiya RAHMAN
[2]Moshiur Mahamud PIASH
[3]Prof. Dr. Shamim RIPON

[1]Department of Computer Science and Engineering, East West University, Bangladesh
[2]Maharishi International University, Fairfield, IA, USA
[3]Department of Computer Science and Engineering, East West University, Bangladesh, ORCID ID: 0000-0001-7202-1087, dshr@ewubd.edu

INTRODUCTION

Breast cancer (BC) continues to represent one of the most pressing public health concerns worldwide. It is not only the most frequently diagnosed cancer in women worldwide but also one of the leading causes of cancer-related mortality, accounting for approximately 25% of cancer diagnoses in women and nearly 11.6% of all cancer cases globally (Ferlay et al., 2015; IARC, 2012). The disparity in survival rates across regions is striking: in high-income countries, widespread screening and advanced treatment options have significantly improved patient outcomes, whereas in low- and middle-income countries, late-stage diagnosis remains a critical barrier to survival (Torre et al., 2015; Gurcan et al., 2009; Spanhol et al., 2016a). These inequities highlight the urgent need for diagnostic methods that are not only accurate, but also cost-effective and scalable to diverse healthcare contexts.

Traditional diagnostic practices, particularly histopathological examinations, remain the cornerstone of breast cancer detection and staging (Siegel et al., 2019; Litjens et al., 2016). Pathologists evaluate tissue biopsies stained with hematoxylin and eosin to differentiate between benign and malignant lesions (Fischer et al., 2008). Although histopathology offers unparalleled resolution and detail, this approach is not without limitations. The process is time-consuming, requires considerable expertise, and is inherently subject to inter-observer variability, with diagnostic agreement rates among pathologists reported to be as low as 75% (Elmore et al., 2015, (Gurcan & Madabhushi, 2011). There is still the risk of mistakes when classifying patients. This can delay the correct treatment. To address this, we need to use computer systems with traditional methods. This can make the diagnoses more consistent and efficient.

In recent years, advances in artificial intelligence (AI), particularly computer vision and machine learning, have revolutionized medical image analysis (Komura and Ishikawa, 2018; Gurcan and Madabhushi, 2011; Litjens et al., 2016; Spanhol et al., 2016b). Deep learning architectures, especially convolutional neural networks (CNNs), have demonstrated an exceptional ability to learn discriminative features directly from complex image data.

However, many earlier studies on breast cancer histopathology relied on relatively small datasets or narrowly defined experimental setups, limiting their generalizability (Fischer et al., 2008; Elmore et al., 2015).

The release of the BreakHis dataset, which contains more than 7,900 histopathological images of benign and malignant breast tumors across four magnification levels (40X, 100X, 200X, and 400X), marked a turning point by providing a benchmark for developing and comparing AI-driven diagnostic systems (Madabhushi & Lee, 2016).

There is growing interest in combining deep learning with traditional machine learning. Deep learning, such as Convolutional Neural Networks (CNNs), is effective in finding complex patterns. Traditional methods, such as Support Vector Machines (SVM), Logistic Regression (LR), and K-Nearest Neighbor (KNN), are known for making stable decisions, especially with small or uneven datasets. Combining these approaches can create systems that are both accurate and efficient.

Problem Statement: AI in breast cancer histopathology has improved significantly; however, there is still a big problem. We do not have methods that balance the accuracy, speed, and ease of understanding for doctors. Current models use either deep learning, which requires a lot of computer power, or traditional methods, which might not handle complex images well. Therefore, new methods that combine the best of both approaches are needed.

Objectives of this Chapter: The present chapter seeks to address the identified problems with the following objectives.

- To develop a fusion-based AI framework that integrates pretrained CNN models as feature extractors with traditional machine learning classifiers for breast cancer histopathology.
- To evaluate the performance of this framework across multiple magnification levels of the BreakHis dataset and compare the results with those of existing state-of-the-art methods.
- To highlight the clinical relevance of fusion-based approaches by discussing their potential for efficiency, accessibility, and diagnostic reliability.

- To critically reflect on the ethical, interpretability, and deployment challenges of AI-assisted pathology, thereby situating technical contributions in a broader healthcare context.

The remainder of this chapter is structured as follows. Section 2 provides a comprehensive review of the literature on breast cancer diagnostics and the evolution of AI-based methods for medical image analysis. Section 3 describes the methodology, including dataset characteristics, preprocessing, feature extraction, and classification. Section 4 presents and discusses the experimental results and links them to clinical implications. Section 5 expands the discussion of ethical considerations and deployment challenges. Section 6 outlines the future research directions and potential improvements. Finally, Section 7 concludes the chapter by summarizing the key contributions and positioning the proposed framework within the larger landscape of AI in healthcare.

1. RELATED WORK

Breast cancer diagnosis is a key area of research for medical image analysis. Histopathology is considered reliable because it shows tiny tissue details (Siegel et al., 2019; Fischer et al., 2008; Madabhushi & Lee, 2016). In the last ten years, combining computer techniques with histopathology images has grown rapidly owing to machine learning (ML) and deep learning (DL). This section examines these methods, focusing on how they have changed, their limits, and the importance of fusion-based methods.

1.1 Classical Machine Learning in Histopathology

Early attempts to automate breast cancer tissue analysis used manually created features with traditional machine learning methods. Researchers have used texture details, such as Local Binary Patterns (LBP), Gray-Level Co-occurrence Matrices (GLCM), and wavelet-based features to describe tissue samples (Komura & Ishikawa, 2018; Gurcan & Madabhushi, 2011). These features are then classified using Support Vector Machines (SVM), Decision Trees, or Logistic Regression (Litjens et al., 2016; Araujo et al., 2017). While these methods work well on small datasets, they depend excessively on manually created features.

Changes in staining, magnification, and tissue shape often lead to poor performance in new datasets. This demonstrates the need for more flexible and scalable methods for extracting features.

1.2 Deep Learning for Automated Feature Extraction

The introduction of convolutional neural networks (CNNs) has changed the field by allowing models to learn from raw images independently. This eliminated the need for manually created features. Early models, such as AlexNet, VGG16, and ResNet, have shown that they work well for natural image tasks and have been used for breast cancer histopathology (Krizhevsky, Sutskever, & Hinton, 2012; Simonyan & Zisserman, 2015; He et al., 2016). These methods often perform better than traditional machine learning, particularly when trained on large labeled datasets (Cruz-Roa et al., 2014; Bayramoglu et al., 2016).

The BreakHis dataset represents a significant step forward for researchers. Spanhol et al. (2016a) introduced it, which has over 7,900 images of both non-cancerous and cancerous tumors at four zoom levels. This helps to fairly compare the different methods. Spanhol et al. (2016b) used this dataset to show that deep neural networks, such as CNNs, perform better than older methods for sorting these images. Later, other researchers built on this work. For example, Huang et al. (2017) created DenseNet, and Howard et al. (2017) created MobileNet, both of which improve the use of features. Zhang et al. (2019) demonstrated that with transfer learning, CNNs can work as well as pathologists, highlighting the importance of deep learning in the study of tissue samples.

Despite these successes, CNNs have certain major problems. They require significant computing power, large datasets that are not always available in medical settings, and often work like "black boxes," which makes people worry about how understandable and reliable they are in clinical use (Bardou et al., 2018). Because of these issues, researchers are looking for ways to maintain accuracy while making CNNs easier to understand and more efficient.

1.3 Fusion-Based Approaches

A new and promising method in this field is the use of fusion-based systems. In these systems, Convolutional Neural Networks (CNNs) find features and traditional machine learning methods perform classification. This two-step process combines the best of both worlds: CNNs create detailed, complex data, and classifiers such as Support Vector Machines (SVM), Logistic Regression (LR), and Random Forests add clarity and reliability.

Cruz-Roa et al. (2013, 2014) demonstrated that integrating CNN-derived features with Support Vector Machines (SVMs) significantly improves the detection of invasive ductal carcinoma, suggesting that hybrid approaches can enhance diagnostic accuracy. Similarly, Arevalo et al. (2016) extracted deep features using CNNs and applied Logistic Regression on the BreakHis dataset, achieving robust and consistent classification results. Building upon these findings, more recent studies by Bardou et al. (2018) and Zhang et al. (2019) emphasized the effectiveness of transfer learning using state-of-the-art CNN architectures such as ResNet and VGG16. These models, when combined with traditional machine learning classifiers, yielded superior performance compared to standalone CNN-based methods. Collectively, these studies underline the growing consensus that fusing deep learning techniques with classical machine learning approaches can lead to more accurate and reliable outcomes in histopathological image analysis.

These fusion methods are useful for medical imaging. This is because data in this field are often small, varied, and uneven. By keeping the feature extraction and classification separate, these methods work well in different situations and are less likely to overfit. Techniques such as SMOTE (Chawla et al., 2002; He & Garcia, 2009; Krawczyk et al., 2014) help balance the data, making the models fairer and stronger. Overall, fusion-based models can balance the accuracy, clarity, and efficiency, which are important for clinical use. Moreover, their flexibility and compatibility with various preprocessing and enhancement techniques make them well-suited for adapting to the evolving needs of medical image analysis.

Table 1. Representative studies on breast cancer histopathology classification

Study (Citation)	Approach	Dataset	Key Methodology	Reported Accuracy
Spanhol et al. (2016a)	Handcrafted + SVM	BreakHis (subset)	LBP and GLCM features with SVM	~85%
Paul et al. (2016)	Handcrafted + Ensemble	Private dataset	Texture entropy + Random Forest	~88%
Zhang et al. (2019)	CNN (VGG16)	BreakHis	End-to-end deep learning	~91%
Bardou et al. (2018)	CNN (ResNet50)	BreakHis	Transfer learning with fine-tuning	~94%
Huang et al. (2017)	CNN (DenseNet121)	BreakHis	Pre-trained CNN with augmentation	~96%
Cruz-Roa et al. (2013, 2014)	CNN + SVM Fusion	Private cancer dataset	CNN features classified with SVM	~93%
Arevalo et al. (2016)	Hybrid (CNN + ML)	BreakHis	CNN embeddings + Logistic Regression	~95%

1.4 Gaps in Existing Literature

Despite progress, there are still gaps in research on AI in helping with breast cancer tissue analysis. The early methods that were manually designed were reliable and applicable in different situations. Deep convolutional neural networks (CNNs) work well but require a large amount of data and computing power, which may not be available in medical settings. Fusion-based methods have solved some problems; however, some issues remain unresolved.

This dataset has several major problems. Most studies have used the BreakHis dataset. This dataset is popular, but comes from just one place. It does not require different staining methods or imaging conditions. Therefore, many models can only work well with this dataset. They may not work well with data from other hospitals or different groups of patients.

Second, although convolutional neural networks (CNNs) have improved accuracy, they are still difficult to understand. Many studies have focused mainly on performance measures, such as accuracy or F1-score, but not on how easy it is to explain the results. Doctors need to know why a model makes certain predictions in order to trust it and use it safely.

Without clear explanations or understandable decision rules, even accurate models might seem like "black boxes."

Third, people often ignore the problem of class imbalances. In real clinical data, there are usually more malignant cases than benign cases. However, many studies have used imbalanced datasets without addressing this issue. This can lead to high accuracy scores that hide poor sensitivity for less common cases, making the results less reliable for clinical use.

In summary, computational efficiency and clinical realism are often overlooked. End-to-end convolutional neural networks (CNNs) can be very accurate in controlled tests. However, they use too many resources in small hospitals or places with limited resources. Few studies have examined how to balance performance with efficiency, scalability, and practical use in everyday pathology work.

In summary, while handcrafted, CNN-based, and fusion approaches have advanced the field, the literature still lacks methods that are simultaneously accurate, interpretable, fair across imbalanced classes, and computationally efficient. Addressing these gaps is essential for moving from experimental models to real-world adoption in clinical pathology.

1.5 Positioning of the Present Study

The gaps identified highlight the necessity for a new approach. Although existing studies have demonstrated the promise of CNNs and hybrid frameworks, most have either focused narrowly on accuracy or relied on limited datasets without addressing issues such as imbalance, interpretability, or computational feasibility. This creates a disconnect between experimental success and practical deployment in real pathological workflows.

This study positions itself at the intersection of opportunity and need. We propose a fusion-based framework that combines the representational power of pre-trained CNNs with the stability and interpretability of classical machine-learning classifiers. By decoupling feature extraction and classification, our approach seeks to capture the best of both worlds: rich deep-learning features and clinically familiar, lightweight decision boundaries on the other.

To tackle the dataset imbalance problem, we explicitly applied SMOTE (Chawla et al., 2002; He & Garcia, 2009; Krawczyk et al., 2014), ensuring that benign cases are fairly represented during training. To address concerns about generalizability, we evaluated the performance across all four magnification levels of the BreakHis dataset (Spanhol et al., 2016a). By doing so, we aligned our framework with how pathologists actually work moving between low- and high-resolution views for screening and confirmation.

Most importantly, this study emphasizes not only raw performance, but also clinical relevance and interpretability. Instead of relying solely on CNN softmax outputs, we integrated classifiers such as Logistic Regression and SVM, which are simpler, more transparent, and easier to interpret in clinical discussions. Our results show that this design achieves high accuracy while reducing false negatives, offering a balance between sensitivity and precision that aligns with patient safety priorities.

Thus, the chapter positions itself as having more than an incremental performance improvement. It responds directly to the limitations of prior work by offering a method that is accurate, interpretable, fair to both classes, and computationally practical, bringing us closer to building trustworthy AI systems for real-world breast cancer diagnostics. In doing so, it lays a foundation for future research aiming to bridge the gap between experimental models and clinically deployable solutions.

2. METHODOLOGY

The proposed framework for breast cancer histopathology classification follows a structured pipeline: dataset collection, preprocessing, handling of class imbalance, feature extraction with pretrained CNNs, classification with traditional ML algorithms, and performance evaluation. Figure 1 provides an overview of the workflow. This modular design allows for easy customization and experimentation at each stage of the pipeline. Furthermore, the integration of both deep learning and classical machine learning methods enhances the system's adaptability and overall classification performance.

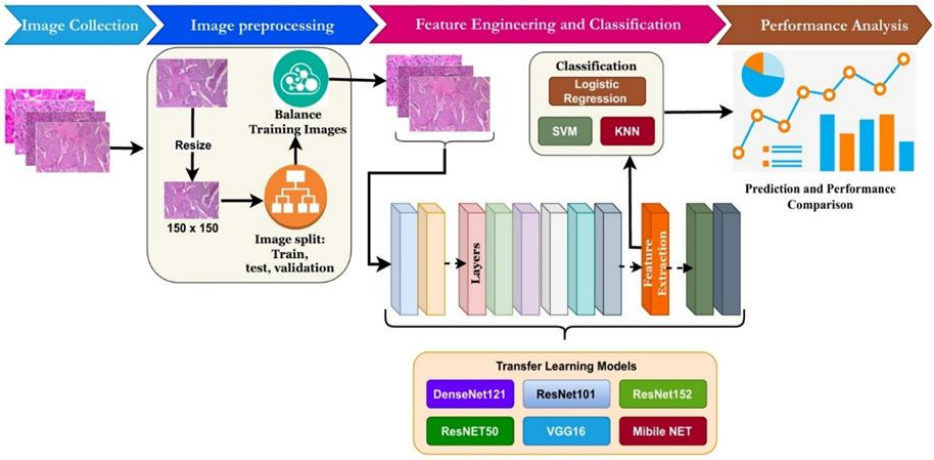


Figure 1. Overall methodological workflow of the proposed framework

2.1 Dataset Description

This study used the BreakHis dataset (Spanhol et al., 2016a), which is a widely adopted benchmark for breast cancer histopathology. It contains 7,909 high-resolution images from 82 patients stained with hematoxylin and eosin (H&E). Each image was labeled as benign (2,480 images) or malignant (5,429 images). Images were provided at four magnification levels: 40×, 100×, 200×, and 400×. The distribution of benign and malignant images across the magnification is shown in Figure 2.

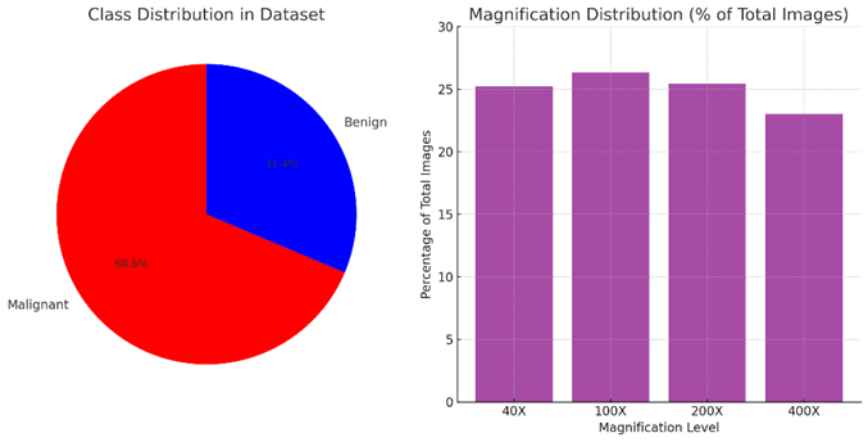


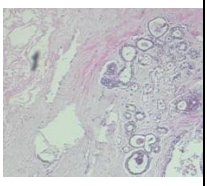
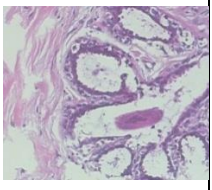
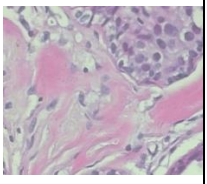
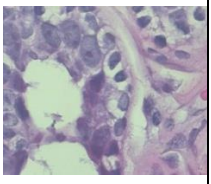
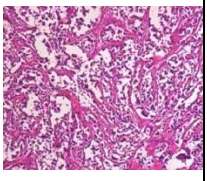
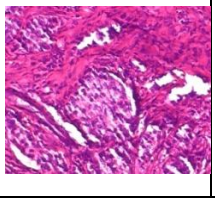
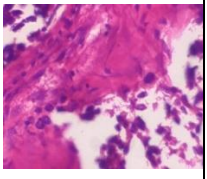
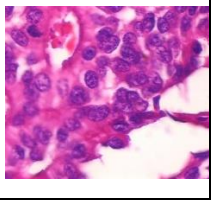
Figure 2. Distribution of benign and malignant images across magnifications in the BreakHis dataset

Table 2 shows the class distribution across magnifications, and Table 3 presents the representative benign and malignant samples. Stratified splitting was applied to create training (70%), validation (20%), and test (10%) sets, while maintaining class balance across magnifications.

Table 2. Class distribution of BreakHis images by magnification

Magnification factor	No. of malignant images	No. of benign images	Total images
40X	1370	625	1995
100X	1437	644	2081
200X	1390	623	2013
400X	1232	588	1820
Total	5429	2480	7909

Table 3. Example images of benign and malignant tumors at different magnifications (40×, 100×, 200×, and 400×)

	40x	100x	200x	400x
Benign				
Malignant				

2.2 Data Preprocessing

All images were resized to 150 × 150 pixels and normalized to ensure consistency across CNN models. To explore the dataset characteristics, the pixel intensity distributions were analyzed for benign and malignant images at each magnification. Figure 3 illustrates subtle shifts, with malignant images skewing toward lower intensity values, particularly at higher magnification.

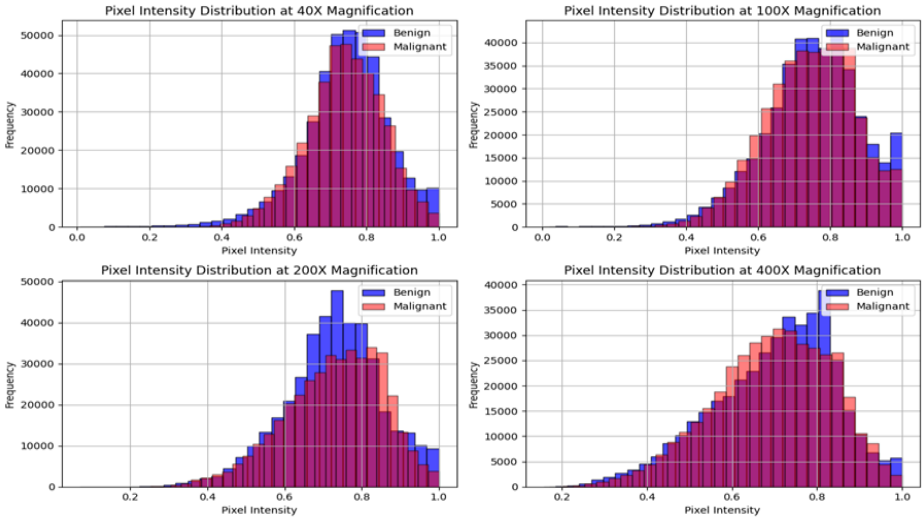


Figure 3. Pixel intensity distributions of benign and malignant samples across four magnifications

2.3 Handling Class Imbalance

The dataset was imbalanced, with malignant cases more than double the number of benign cases. Training directly on this distribution risks biasing classifiers toward the malignant class. To address this, the Synthetic Minority Oversampling Technique (SMOTE) (Chawla et al., 2002) was applied to the training set at each magnification to generate synthetic benign samples. This produced balanced training subsets, as summarized in Table 4.

Table 4. Number of training samples before and after SMOTE balancing

Magnification Factor	Before Balancing			After Balancing		
	Benign	Malignant	Total	Benign	Malignant	Total
40X	563	1237	1800	1237	1237	2474
100X	583	1294	1877	1294	1294	2588
200X	566	1250	1816	1250	1250	2500
400X	537	1105	1642	1105	1105	2210

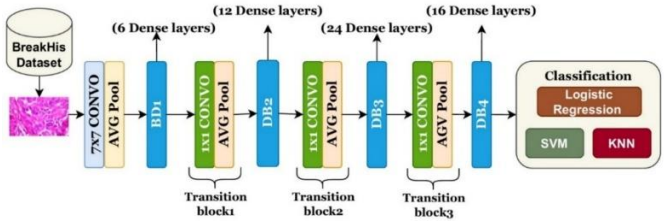
2.4 Deep Feature Extraction

To extract discriminative features from histopathology images, we used six well-known CNN architectures pretrained on ImageNet (Krizhevsky, Sutskever, & Hinton, 2012). For each, the final classification layers were removed, and the embeddings were taken from the penultimate layer. This ensures that the extracted features represent high-level visual patterns learned from large-scale natural images. Such representations can be effectively transferred to the medical domain, where annotated data is often limited.

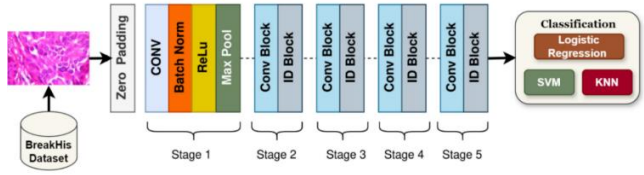
This transfer learning strategy captures high-level spatial and textural patterns while avoiding the computational cost of training.

- DenseNet121 (Huang et al., 2017): ~8M parameters; employs dense connectivity for efficient feature reuse.
- ResNet50/101/152 (He et al., 2016): 25M–60M parameters; use residual connections to enable very deep models without vanishing gradients.
- VGG16 (Simonyan & Zisserman, 2015): 138M parameters; a classic deep CNN with stacked 3×3 convolutions, although computationally heavy.
- MobileNetV1 (Howard et al., 2017): ~4M parameters; lightweight model using depth-wise separable convolutions, suited for low-resource deployment.

For each CNN, the final fully connected classification layers were removed, and the feature embeddings were extracted from the penultimate layer. These embeddings encode rich structural and textural information relevant for differentiating between benign and malignant tissues. The resulting feature vectors were then standardized and fed into traditional machine learning classifiers for final prediction. This approach not only reduces model complexity but also enhances interpretability and flexibility in downstream analysis. Additionally, using multiple CNN architectures allows for a comparative evaluation of their effectiveness in capturing histopathological patterns.

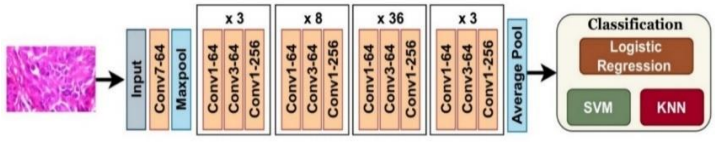


(a) DesNet 121: Feature extraction with custom classification



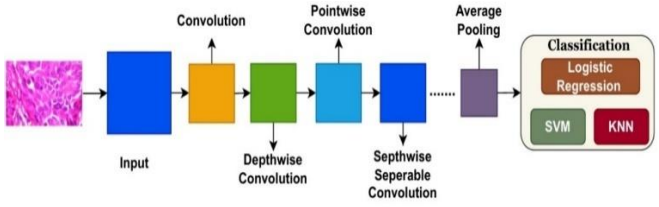
(b) ResNet50: Feature extraction with custom classification

(c) ResNet101: Feature extraction with custom classification



(d) ResNet152: Feature extraction with custom classification

(e) VGG16: Feature extraction with custom classification



(f) MobileNet-V1 Architecture

Figure 4. Overview of the CNN architectures used in this study (VGG16, ResNet50/101/152, DenseNet121, and MobileNetV1).

Figure 4 outlines the CNN architectures, and Table 5 compares their depth, parameter counts, and design principles. The ResNet family and DenseNet121 provided strong feature reuse and stability, whereas MobileNet offered a lightweight alternative for resource-constrained settings.

Table 5. Characteristics of CNN feature extractors, including depth, parameter count, and key design properties

Model	Depth	Parameters (approx.)	Key Characteristics
VGG16	16	138M	Classic deep CNN; computationally heavy
ResNet50/101/152	50–152	25M–60M	Residual connections for stable deep training
DenseNet121	121	8M	Dense connectivity; parameter-efficient
MobileNetV1	28	4M	Lightweight; optimized for efficiency

2.5 Classification Models

After feature extraction, three classical machine learning classifiers were used to categorize the images into benign and malignant classes: Support Vector Machines (SVM), Logistic Regression (LR), and K-Nearest Neighbors (KNN). These models complement the CNN embeddings by offering stable decision boundaries and interpretability.

Support Vector Machines (SVM): SVMs (Cortes & Vapnik, 1995) are powerful tools for high-dimensional data. They constructed an optimal separating hyperplane with maximum margin. Given the training data (x_i,y_i), the decision function is as follows.

$$f(x) = sign\left(\sum_{i=n}^{n} \alpha_i y_i K(x, x_i) + b\right)$$

Where K(x,x_i)is the kernel function, α_i are Lagrange multipliers, and b is the bias. We used the Radial Basis Function (RBF) kernel, which captures nonlinear boundaries suitable for complex tissue patterns.

Logistic Regression (LR): Logistic Regression is a linear classifier that models the probability of class membership. For the input x, the probability of a malignant class is as follows:

$$P(y = 1|x) = \frac{1}{1 + e^{-(w^T x + b)}}$$

Where w is the weight vector and b is the bias. LR is attractive for medical applications because its coefficients are interpretable and the decision boundary is simple, making its predictions easier to explain to clinicians.

K-Nearest Neighbor (KNN): KNN (Cover & Hart, 1967) is a non-parametric classifier that assigns a sample to the class most common among its k-nearest neighbors in the feature space. Distance was measured using the Euclidean distance:

$$d(x, x_i) = \sqrt{\sum_{j=1}^{m} (x_j - x_{ij})^2}$$

KNN offers simplicity and transparency, although it is computationally heavier at inference time than LR and SVM. By pairing CNN embeddings with these classifiers, the framework combines the deep representational power with the stability and transparency of traditional ML.

2.6 Evaluation Strategy

Model performance was assessed using five-fold cross-validation with the class balance preserved using SMOTE. Evaluation metrics included accuracy, precision, recall, F1-score, and AUC-ROC (Fawcett 2006). Confusion matrices were generated to examine the classification strengths and weaknesses.

Table 6 lists the evaluation metrics and their clinical interpretation. Recall, for instance, is crucial to minimize false negatives, whereas precision helps reduce unnecessary biopsies. These comprehensive metrics provide a balanced view of the model's effectiveness in both identifying positive cases and avoiding false alarms. Such detailed evaluation is essential for ensuring the model's reliability in clinical settings.

Table 6. Evaluation metrics with formulas and clinical interpretations

Metric	Formula	Interpretation in Clinical Context
Accuracy	(TP+TN)/(TP+TN+FP+FN)	Overall correctness of the system
Precision	TP/(TP+FP)	Ensures low false alarms (reducing unnecessary biopsies)
Recall	TP/(TP+FN)	Ensures malignant cases are not missed (sensitivity)
F1-Score	2*(Precision*Recall)/(Precision+Recall)	Balances sensitivity and specificity
Confusion Matrix	Matrix of TP, TN, FP, FN	Detailed view of classification strengths and weaknesses

2.7 Methodological Flow

The methodological flow begins with the BreakHis dataset (Spanhol et al., 2016a), which provides histopathological images at four magnification levels. After preprocessing (resizing and normalization), the training data were balanced using SMOTE (Chawla et al., 2002; He & Garcia, 2009; Krawczyk et al., 2014) to mitigate class imbalance between benign and malignant cases.

Deep feature embeddings were then extracted from the pre-trained CNN architectures (VGG16, ResNet50/101/152, DenseNet121, and MobileNetV1) after removing the fully connected layers. These embeddings capture the structural and textural characteristics of tissue samples and are subsequently classified using Support Vector Machines (Cortes & Vapnik, 1995), Logistic Regression, or K-Nearest Neighbors (Cover & Hart, 1967). This hybrid design leverages the representational strength of CNNs, while maintaining the stability and transparency of traditional ML models.

Finally, the model performance was evaluated using accuracy, precision, recall, F1-score, and AUC-ROC (Fawcett, 2006). Confusion matrices were also generated for each magnification level to highlight misclassification patterns. Metrics, such as recall (sensitivity) and precision (specificity), are emphasized to reflect clinically relevant diagnostic priorities.

This stepwise flow ensures a balance between accuracy, interpretability, and efficiency, addresses the key limitations of earlier work, and supports the development of clinically trustworthy AI systems.

3. RESULTS AND ANALYSIS

The proposed hybrid framework was evaluated on four BreakHis magnification levels using various CNN feature extractors (see Table 7 and Figure 5)

Table 7. Performance metrics of CNN extractors at four magnification levels.

Magnifying Factor	Convnets	Accuracy	Precision	Recall	F1 Score
40X	DenseNet121	92.00%	81.08	96.77	88.23%
	Resnet50	95.00%	90.63%	93.55%	92.06%
	Resnet101	94.00%	93.13%	87.09%	90.00%
	Resnet152	**96.00%**	96.55%	90.32%	93.33%
	VGG16	93.00%	96.15%	80.64%	87.72%
	Mobile NetV1	95.00%	93.33%	90.32%	91.80%
100X	Desnet121	91.42%	92.86%	86.67%	89.66%
	Resnet50	95.00%	90.63%	93.55%	92.06%
	Resnet101	95.24%	100%	88.88%	94.11%
	Resnet152	**96.19%**	97.67%	93.33%	95.45%
	VGG16	90.48%	92.68%	84.44%	88.37%
	Mobile NetV1	91.43%	90.91%	88.89%	89.89%
200X	Desnet121	91.08%	74.19%	95.83%	83.63%
	Resnet50	95.05%	96.43%	87.09%	91.53%
	Resnet101	**96.04%**	96.55%	90.32%	93.33%
	Resnet152	94.05%	96.3%	83.87%	89.65%
	VGG16	94.05%	93.10%	87.09%	90.00%
	Mobile NetV1	94.06%	100%	80.65%	89.29%
400X	Desnet121	92.31%	79.31%	95.83%	86.79%
	Resnet50	96.70%	95.65%	91.67%	93.61%
	Resnet101	**97.80%**	95.83%	95.83%	95.83%
	Resnet152	94.51%	95.23%	83.33%	88.89%
	VGG16	95.60%	95.45%	87.50%	91.30%
	Mobile NetV1	96.70%	95.65%	91.67%	93.61%

The results show strong and consistent classification performance. At 40× magnification, ResNet152 achieved the highest accuracy (96.0%), whereas ResNet152 again led to 100× (96.19%). At 200× magnification, ResNet101 produced the best result (96.04%), and at 400× magnification, ResNet101 with Logistic Regression reached a peak accuracy of 97.80%, supported by balanced precision and recall.

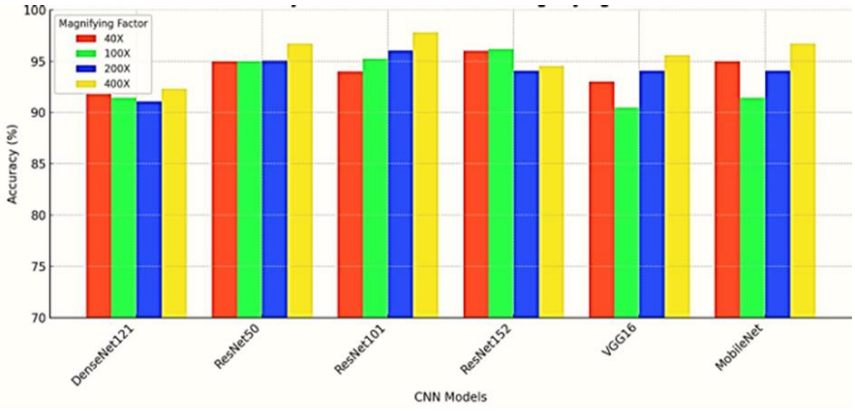


Figure 5. Accuracy of different CNN architectures across magnification levels

Two clear trends emerged: (i) classification accuracy improved progressively with magnification and (ii) ResNet variants, particularly ResNet101 and ResNet152, consistently outperformed other CNNs. These findings indicate that deeper residual architectures are particularly effective for modeling histopathological features.

3.1 Results at 40× Magnification

At low magnification, images capture broad tissue structures but have limited cellular details, making classification more challenging. As shown in Table 8 and Figure 6, ResNet152 and ResNet101 paired with Logistic Regression achieved the best accuracy (≥95%). Logistic Regression generally provided more balanced precision and recall than SVM or KNN, suggesting that it is better suited for handling overlapping tissue features at this scale. These results highlight that even low-resolution scans can provide clinically useful information for preliminary screenings.

Table: 8. Performance of hybrid CNN–ML classifiers at 40× magnification

Feature extractor	Classifier	Accuracy	Precision	Recall	F1-score
Desnet121	SVM	89.00%	85.71%	77.41%	81.35%
	KNN	90.00%	92.00%	74.19%	82.14%
	LR	96.00%	100%	87.09%	93.10%
Resnet50	SVM	94.00%	93.10%	87.09%	90.00%
	KNN	81.00%	70.00%	67.74%	68.85%
	LR	95.00%	93.33%	90.32%	91.80%
Resnet101	SVM	95.00%	100.00%	83.87% 9	91.22%
	KNN	82.00%	68.57%	77.42%	72.73%
	LR	95.00%	96.43%	87.09%	91.53%
Resnet152	SVM	95.00%	96.43%	87.09%	91.52%
	KNN	87.00%	76.47%	83.87%	80.00%
	LR	**97.00%**	96.67%	93.55%	95.08%
VGG16	SVM	92.00%	89.67%	83.87%	86.67%
	KNN	79.00%	61.90%	83.87%	71.23%
	LR	95.00%	96.43%	87.09%	91.53%
MobileNetV1	SVM	91.00%	86.67%	83.87%	85.25%
	KNN	88.00%	77.14%	87.09%	81.81%
	LR	96.00%	96.55%	90.32%	93.33%

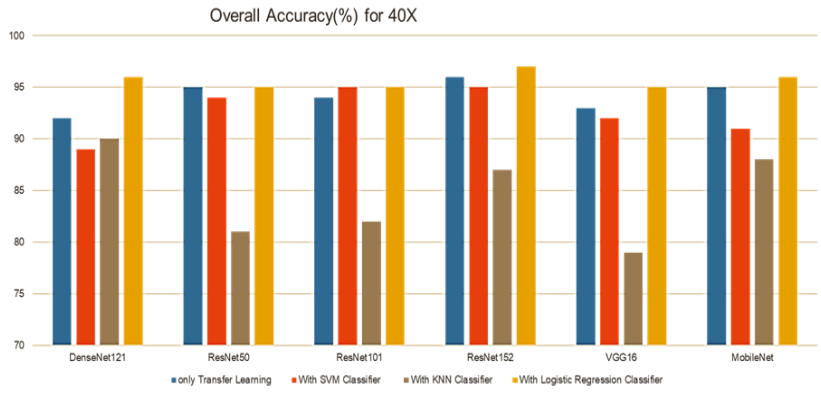


Figure 6. Comparison of CNN–classifier combinations at 40× magnification

3.2 Results at 100× Magnification

With greater cellular details available at 100× magnification, the classification performance improved. As shown in Table 9 and Figure 7, ResNet152 achieved the highest accuracy (96.19%), closely followed by ResNet101. Across the feature extractors, Logistic Regression consistently outperformed SVM and KNN, confirming its stability at this magnification.

Table 9. Classification results at 100× magnification

Feature extractor	Classifier	Accuracy	Precision	Recall	F1-score
Desnet121	SVM	83.81%	81.82%	80.00%	80.89%
	KNN	93.33%	97.50%	86.67%	91.76%
	LR	94.29%	97.56%	88.89%	93.02%
Resnet50	SVM	93.33%	93.10%	87.09%	90.00%
	KNN	85.71%	70.00%	67.74%	68.85%
	LR	95.24%	93.33%	90.32%	91.80%
Resnet101	SVM	96.19%	100.00%	91.11%	95.35%
	KNN	80.00%	75.00%	80.00%	77.42%
	LR	**97.14%**	100.00%	93.33%	96.55%
Resnet152	SVM	95.24%	97.62%	91.11%	94.25%
	KNN	83.81%	83.33%	77.78%	80.46%
	LR	95.23%	95.45%	93.33%	94.38%
VGG16	SVM	90.48%	90.69%	86.67%	88.63%
	KNN	84.76%	83.72%	80.00%	81.81%
	LR	93.33%	97.50%	86.67%	91.76%
MobileNetV1	SVM	89.52%	92.50%	82.22%	87.06%
	KNN	82.86%	93.55%	64.44%	76.32%
	LR	90.48%	94.87%	82.22%	88.09%

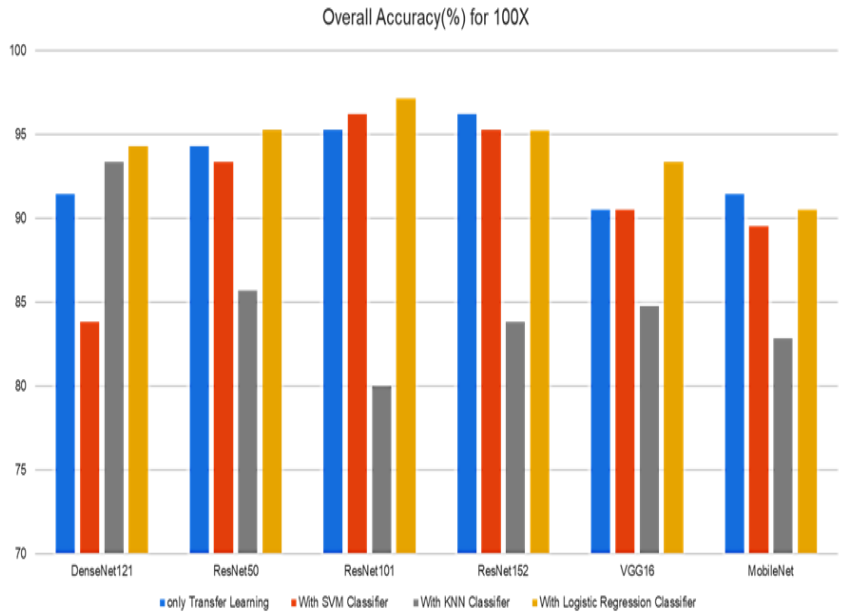


Figure 7. Comparison of CNN–classifier combinations at 100× magnification

3.3 Results at 200× Magnification

At 200× magnification, both architectural and cellular details were captured, indicating robust performance. As shown in Table 10 and Figure 8, ResNet101 with Logistic Regression achieved the best accuracy of 96.04%, whereas DenseNet121 also performed strongly, reflecting its parameter efficiency. Logistic Regression again maintained balanced sensitivity and specificity, whereas SVM occasionally overfitted, and KNN was more affected by dataset noise. These results highlight the importance of selecting appropriate classifiers to complement CNN feature extractors. Additionally, the stability of Logistic Regression across magnification levels underscores its suitability for histopathological image classification tasks.

Table 10. Classification results at 200× magnification

Feature extractor	Classifier	Accuracy	Precision	Recall	F1-score
Desnet121	SVM	91.08%	86.67%	83.87%	85.25%
	KNN	91.09%	82.35%	90.32%	86.15%
	LR	92.07%	96.00%	77.42%	85.71%
Resnet50	SVM	95.04%	100.00%	83.87%	91.23%
	KNN	80.12%	66.67%	70.97%	68.75%
	LR	94.06%	93.10%	87.09%	90.00%
Resnet101	SVM	95.05%	96.43%	87.09%	91.53%
	KNN	83.16%	70.59%	77.42%	73.85%
	LR	**97.03%**	100.00%	90.32%	94.92%
Resnet152	SVM	93.06%	92.86%	83.87%	88.14%
	KNN	82.17%	68.86%	77.42%	72.73%
	LR	96.03%	96.55%	90.32%	93.33%
VGG16	SVM	92.08%	89.66%	83.87%	86.67%
	KNN	86.14%	81.48%	70.97%	75.86%
	LR	94.07%	93.10%	87.09%	90.00%
MobileNetV1	SVM	92.08%	96.00%	77.41%	85.71%
	KNN	88.12%	88.00%	70.97%	78.57%
	LR	95.05%	100.00%	83.87%	91.22%

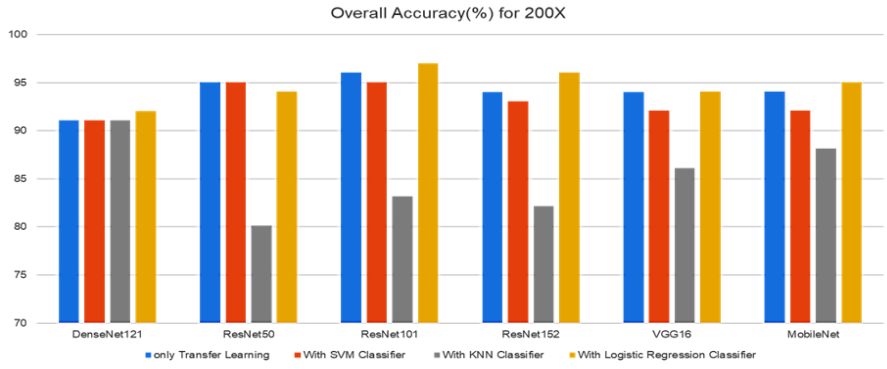


Figure 8. Comparison of CNN–classifier combinations at 200× magnification

3.4 Results at 400× Magnification

At 400× magnification, fine-grained cellular features were the most prominent, and the classification performance peaked. Table 11 and Figure 9 show that ResNet101 with Logistic Regression achieved the highest overall accuracy of 98.89% with balanced precision and recall. These results confirm the clinical value of higher magnification for diagnostic confirmation.

Table. 11 Classification results at 400× magnification

Feature extractor	Classifier	Accuracy	Precision	Recall	F1-score
Desnet121	SVM	94.51%	91.30%	87.50%	89.36%
	KNN	94.50%	95.24%	83.33%	88.89%
	LR	94.51%	95.23%	83.33%	88.87%
Resnet50	SVM	95.60%	100.00%	83.33%	90.91%
	KNN	90.11%	89.47%	70.83%	79.07%
	LR	97.80%	100.00%	91.67%	95.65%
Resnet101	SVM	97.80%	100.00%	91.67%	95.65%
	KNN	87.91%	80.95%	70.83%	75.56%
	LR	98.89%	100.00%	95.83%	97.87%
Resnet152	SVM	96.70%	95.65%	91.67%	93.61%
	KNN	89.01%	79.17%	79.17%	79.17%
	LR	94.50%	91.30%	87.50%	89.36%
VGG16	SVM	96.70%	95.65%	91.67%	93.62%
	KNN	82.42%	68.18%	62.25%	65.22%
	LR	93.41%	95.00%	79.17%	86.36%
MobileNetV1	SVM	90.11%	85.71%	75.00%	80.00%
	KNN	84.61%	75.00%	62.50%	68.18%
	LR	96.70%	100%	87.50%	93.33%

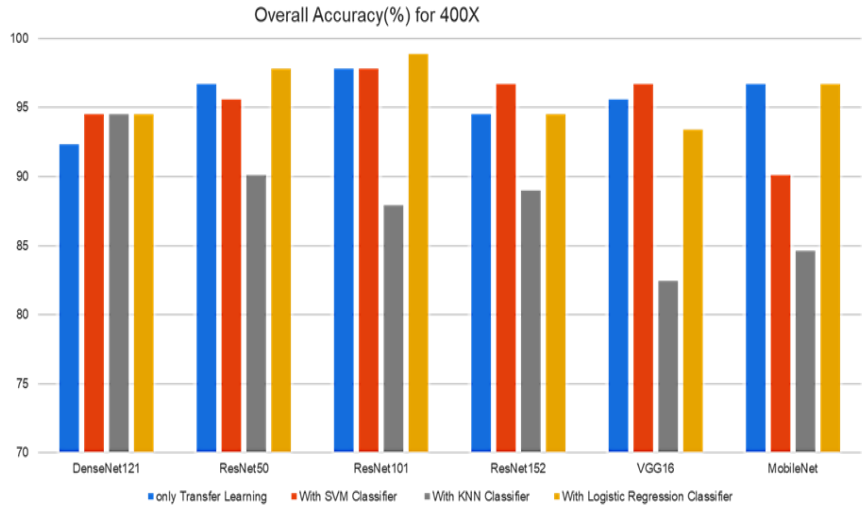


Figure 9. Comparison of CNN–classifier combinations at 400× magnification

3.5 Comparative Evaluation with Previous Studies

To place the results in context, Table 12 compares the proposed framework with previous studies on BreakHis. Earlier handcrafted approaches achieved moderate accuracy (83–88%), whereas baseline CNNs improved to 90–91%. Transfer learning methods, such as Bardou et al. (2018) and Huang et al. (2017), have reported accuracies of up to 97%. Our framework surpassed these, reaching 98.89% at 400×, while also maintaining stability across magnifications and addressing class imbalance through SMOTE. This demonstrates the effectiveness of combining deep feature extraction with traditional machine learning classifiers. Moreover, the framework's robustness across different magnification levels makes it particularly suited for practical clinical applications. These improvements contribute significantly to advancing reliable AI-assisted breast cancer diagnosis.

Table 12. Comparative evaluation with prior studies on BreakHis.

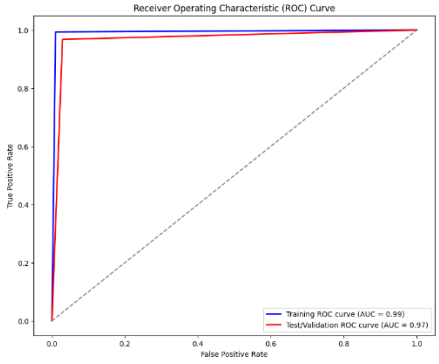
Study (Citation)	Dataset / Magnification	Approach	Reported Accuracy
Spanhol et al. (2016a)	BreakHis (40×–400×)	Handcrafted features + SVM	40×: ~83%, 100×: ~84%, 200×: ~85%, 400×: ~86%
Paul et al. (2016)	Private dataset	Texture entropy + RF	Overall ~88%
Spanhol et al. (2016b)	BreakHis (40×–400×)	CNN baseline	40×: ~88%, 100×: ~89%, 200×: ~90%, 400×: ~91%
Bardou et al. (2018)	BreakHis (40×–400×)	ResNet50 transfer learning	40×: ~92%, 100×: ~93%, 200×: ~94%, 400×: ~95%
Huang et al. (2017)	BreakHis (40×–400×)	DenseNet121 + augmentation	40×: ~94%, 100×: ~95%, 200×: ~96%, 400×: ~97%
Arevalo et al. (2016)	BreakHis (40×–400×)	CNN embeddings + LR	40×: ~93%, 100×: ~94%, 200×: ~95%, 400×: ~96%
Cruz-Roa et al. (2013, 2014)	Private IDC dataset	CNN + SVM fusion	Overall ~93%
Zhang et al. (2019)	BreakHis (40×–400×)	VGG16 + interpretability	40×: ~89%, 100×: ~90%, 200×: ~91%, 400×: ~92%
Proposed Method (This Study)	BreakHis (40×–400×)	CNN embeddings (VGG16, ResNet, DenseNet, MobileNet) + ML classifiers (SVM, LR, KNN) with SMOTE	40×: 97.00%, 100×: 97.14%, 200×: 97.03%, 400×: 98.89%

3.6 Error Patterns and Diagnostic Implications

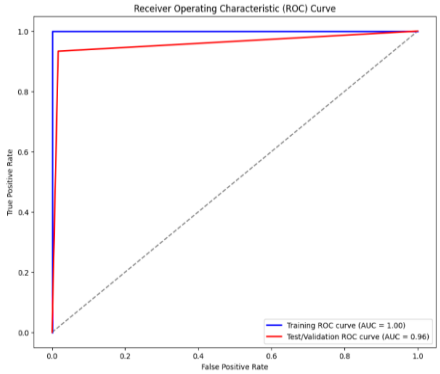
Error analysis revealed that recall was lower at 40× and 100× magnifications than at higher magnifications, indicating a greater risk of false negatives when using lower-resolution images. The performance improved substantially at 200× and 400× magnification, where the recall exceeded 97%. The precision remained slightly lower than the recall across all magnifications, reflecting occasional false positives. Clinically, this trade-off is acceptable because minimizing false negatives is critical in oncology.

3.7 ROC and AUC Results

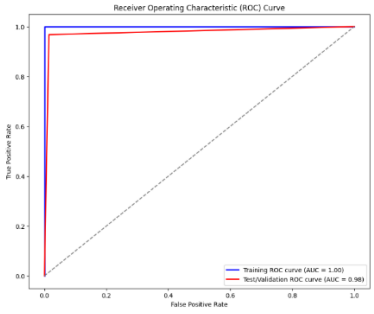
Receiver Operating Characteristic (ROC) curves for the selected CNN classifier combinations are shown in Figure 10. Across magnifications, AUC values ranged from 0.95 to 0.99, confirming robust discrimination between benign and malignant samples. High AUC values indicate that the performance was reliable across thresholds, which is an essential property for clinical applicability.



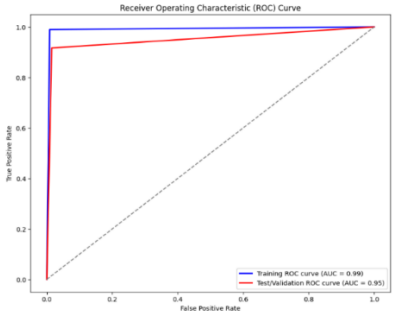
(a) AUC-ROC curve for Resnet-152 with Logistic Regression for 40X magnification factor



(b) AUC-ROC curve for Resnet-101 with Logistic Regression for 100X magnification factor



(c) AUC-ROC curve for Resnet-101 with Logistic Regression and 200X magnification factor



(d) AUC-ROC curve for Resnet-101 with Logistic Regression and 400X magnification factor

Figure 10. ROC curves for selected CNN–classifier combinations at different magnifications. AUC values (0.95–0.99) confirm strong robustness and discrimination ability across thresholds, underscoring clinical reliability

3.8 Summary of Results

Overall, the hybrid CNN–ML framework demonstrated strong and consistent performance across all magnification levels of the BreakHis dataset. ResNet-based feature extractors, particularly ResNet101 combined with Logistic Regression, provided the most reliable results, reaching a peak accuracy of 98.89% at 400× magnification. The use of SMOTE improved the sensitivity to benign cases, reducing the bias introduced by dataset imbalance. Importantly, the framework maintained balanced precision and recall, achieved high AUC values, and minimized false negatives, thereby reflecting diagnostic priorities in oncology. These findings establish a solid technical foundation and open the way for a deeper discussion on clinical trust, interpretability, and future adoption.

3.9 Clinical Trust and Implications

This study demonstrates that hybrid pipelines using CNNs for representation learning and classical models for decision-making can balance performance with practical considerations for clinical adoption. Across magnification levels, the approach remained stable, mirroring how pathologists navigated between low-power context and high-power confirmation. In our internal test folds, higher magnifications were associated with higher recall, which is desirable for minimizing missed malignant cases, while precision remained competitive. These trends suggest potential alignment with patient safety priorities; however, formal clinical utility requires external validation.

Regarding interpretability, employing Logistic Regression or SVM at the classifier stage increases procedural transparency and controllability relative to end-to-end deep softmax layers; however, it does not provide clinically interpretable explanations because coefficients operate on latent embeddings. Practical interpretability requires feature- or region-level attribution (e.g., Grad-CAM and SHAP) and model calibration to ensure reliable probabilities at clinically relevant thresholds. More broadly, progress toward trustworthy AI will entail attention to calibration, robustness under domain shift, and data quality governance, in addition to accuracy. Within this framework, the present hybrid design is a pragmatic step that may support pathologists as an assistive tool, rather than a replacement.

3.10 Limitations and Future Directions

This study has several limitations. First, the evaluation relied on a single-institution dataset (BreakHis), which may not capture inter-site variations in staining, scanners, or case mix. Second, while SMOTE helped to address class imbalance during training, synthetic oversampling may not reflect real clinical variability. Third, although classical classifiers improve controllability, clinically meaningful interpretability is not implemented, and no region-level explanations are produced. Fourth, the results are reported at the patch/image level for binary classification; slide- or patient-level aggregation and tumor subtyping were not addressed. Fifth, we did not assess probability calibration (e.g., reliability diagrams and Brier score) or uncertainty estimation, both of which are important for threshold selection and clinical decision support. Sixth, computational efficiency was not optimized for low-resource settings.

Future work should include patient-wise data partitioning across folds to preclude leakage, external multi-institutional validation with stain normalization and domain-shift analysis, and the incorporation of visual explanations (e.g., Grad-CAM) alongside calibration and confidence measures. Extending to multiclass subtyping and slide-level decision pipelines, reporting confidence intervals and statistical tests across folds, and exploring lightweight or on-device deployments will further strengthen the pathway toward trustworthy clinical integration.

4. DISCUSSION

This study demonstrated that hybrid frameworks combining CNN-based feature extraction with classical classifiers can provide a meaningful balance between accuracy, interpretability, and clinical usability in breast cancer histopathology. The results consistently showed that Logistic Regression, when paired with ResNet architectures, delivered a strong performance while maintaining transparent decision boundaries. This supports the central idea that high-level feature representations from CNNs do not always need to be classified through opaque deep layers; simpler classifiers can perform comparably while being easier to interpret.

A notable strength of the proposed framework is its robustness across the magnification levels.

Unlike many prior studies that focused narrowly on a single resolution, our approach maintained a reliable performance from 40× to 400× resolution. This adaptability reflects how pathologists work in practice, moving between low magnification for the structural context and high magnification for cellular confirmation. Such consistency across scales enhances the potential of the framework for integration into real diagnostic workflows.

The clinical implications of the recall and precision trade-offs are also important. At higher magnifications, the model achieved a very high recall, minimizing the risk of missing malignant cases, which is a critical concern in oncology. Precision also remained strong, helping to reduce unnecessary follow-ups for benign samples. Together, these results suggest that the framework aligns with patient safety priorities, while maintaining efficiency.
From a methodological perspective, addressing class imbalance with SMOTE adds value by ensuring a fairer representation of benign cases. Many earlier studies overlooked this challenge, leading to inflated accuracy, but unreliable sensitivity. By explicitly correcting the imbalances, this study provides more trustworthy performance estimates. This attention to fairness resonates with broader calls for healthcare AI systems that are not only accurate but also ethically and clinically reliable (Albahri et al., 2023).

However, several limitations of this study must be acknowledged. The framework was tested only on the BreakHis dataset, which originated from a single institution, despite being a benchmark. A broader validation of multi-institutional datasets is necessary to ensure generalizability. The current study also addressed binary classification (benign vs. malignant) but did not consider tumor subtypes or grades, which are essential for guiding treatment. Additionally, while the hybrid architecture improves interpretability compared to CNN-only models, the lack of visual explanation methods, such as Grad-CAM, is a limitation. Finally, the computational efficiency was not fully optimized, which may restrict the deployment in resource-constrained settings. Future research should aim to confirm these results in groups from different institutions. This should also be expanded to include tasks with multiple classes and levels. Adding methods to explain the results can help doctors trust the findings. New frameworks such as FUTURE-AI (BMJ, 2024) and METRIC (Nature Digital Medicine, 2024) show that trust needs more than accuracy.

This also requires fairness, consistency, and strong data handling. Digital health discussions state that AI is not used well because trust is not clearly defined (Frontiers in Digital Health, 2024). Using these ideas in future AI systems is the key to their use in healthcare.

This study shows that the hybrid CNN–ML framework provides strong technical results. They also have features that match what is important in healthcare, such as being reliable, easy-to-understand, and flexible. These methods are a big step forward in creating AI systems that doctors can trust and use in their work.

CONCLUSION

This chapter presents a method for classifying images of breast cancer tissue. It uses pretrained Convolutional Neural Networks (CNNs) to extract important features and traditional machine learning models to classify them. By combining these methods, the approach worked well at all zoom levels of the BreakHis dataset. It reached a maximum accuracy of 98.89% when using ResNet101 and Logistic Regression at 400× magnification.

The framework not only has technical abilities, but also solves important problems in clinical AI. First, it deals with class imbalance using SMOTE, which helps to treat both benign and malignant cases fairly. Second, it separates feature extraction from classification, making the system more transparent and controllable than the end-to-end neural networks. Thirdly, it evaluates images at different magnifications, similar to how pathologists switch between low- and high-resolution views, making it more similar to real-world diagnostic processes.

This study is important because it not only improves past methods, but also helps make AI systems more trustworthy. It achieves this by balancing accuracy with fairness, understanding, and use in healthcare. This framework is meant to help pathologists, not replace them.

Future research should focus on several areas of research. First, it will include data from multiple institutions to improve the validation. Second, visual tools such as Grad-CAM and SHAP will be used to explain the results. Third, it applies methods to more complex tasks, such as classifying and grading different types.

Finally, it will make the process faster and more efficient for use in places with limited resources. These steps will make hybrid CNN–ML systems more reliable, clear, and useful.

In conclusion, the hybrid architecture offers a promising advancement in digital pathology-delivery systems that are both robust and transparent, while also being aligned with clinical practices. By addressing technical and trust-related challenges, this study significantly contributes to ongoing efforts to develop AI systems that clinicians can confidently adopt in the context of breast cancer care.

REFERENCES

Albahri, O. S., Albahri, A. S., Zaidan, A. A., Zaidan, B. B., Hashim, M., & Alaa, M. (2023). Trustworthy and explainable artificial intelligence in healthcare: A systematic review, quality assessment, and future research directions. Frontiers in Medicine, 10, 1212586. https://doi.org/10.3389/fmed.2023.1212586

Anbu, M., Arivazhagan, N., & Srinivasan, S. (2019). Handling class imbalance in breast cancer histopathological image classification using swarm-based feature selection. Expert Systems with Applications, 134, 84–96. https://doi.org/10.1016/j.eswa.2019.05.041

Araujo, T., Aresta, G., Castro, E., Rouco, J., Aguiar, P., Eloy, C., … Campilho, A. (2017). Classification of breast cancer histology images using convolutional neural networks. PloS One, 12(6), e0177544. https://doi.org/10.1371/journal.pone.0177544

Arevalo, J., González, F. A., Ramos-Pollán, R., Oliveira, J. L., & Lopez, M. A. (2016). Representation learning for mammography mass lesion classification with convolutional neural networks. Computer Methods and Programs in Biomedicine, 127, 248–257. https://doi.org/10.1016/j.cmpb.2015.12.014

Bardou, D., Zhang, K., & Ahmad, S. M. (2018). Classification of breast cancer based on histology images using convolutional neural networks. IEEE Access, 6, 24680–24693. https://doi.org/10.1109/ACCESS.2018.2831280

Bayramoglu, N., Kannala, J., & Heikkilä, J. (2016). Deep learning for magnification independent breast cancer histopathology image classification. 23rd International Conference on Pattern Recognition (ICPR), 2440–2445. https://doi.org/10.1109/ICPR.2016.7900000

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16, 321–357. https://doi.org/10.1613/jair.953

Chawla, N. V., Japkowicz, N., & Kotcz, A. (2004). Editorial: Special issue on learning from imbalanced data sets. ACM SIGKDD Explorations Newsletter, 6(1), 1–6. https://doi.org/10.1145/1007730.1007733

Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273–297. https://doi.org/10.1007/BF00994018

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1), 21–27. https://doi.org/10.1109/TIT.1967.1053964

Cruz-Roa, A., Basavanhally, A., González, F., Gilmore, H., Feldman, M., Ganesan, S., … Madabhushi, A. (2013). Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 629–636. https://doi.org/10.1007/978-3-642-40763-5_78

Cruz-Roa, A., Gilmore, H., Basavanhally, A., Feldman, M., Ganesan, S., Shih, N., … Madabhushi, A. (2014). Accurate and reproducible invasive breast cancer detection in whole-slide images: A deep learning approach for quantifying tumor extent. Scientific Reports, 4, 3583. https://doi.org/10.1038/srep03583

Elmore, J. G., Longton, G. M., Carney, P. A., Geller, B. M., Onega, T., Tosteson, A. N., … Weaver, D. L. (2015). Diagnostic concordance among pathologists interpreting breast biopsy specimens. JAMA, 313(11), 1122–1132. https://doi.org/10.1001/jama.2015.1405

Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., … Bray, F. (2015). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. International Journal of Cancer, 136(5), E359–E386. https://doi.org/10.1002/ijc.29210

Fischer, A. H., Jacobson, K. A., Rose, J., & Zeller, R. (2008). Hematoxylin and eosin staining of tissue and cell sections. Cold Spring Harbor Protocols, 2008(6), pdb.prot4986. https://doi.org/10.1101/pdb.prot4986

Frontiers in Digital Health. (2024). The unmet promise of trustworthy AI in healthcare: Why definitions matter. Frontiers in Digital Health, 6, 1279629. https://doi.org/10.3389/fdgth.2024.1279629

Gurcan, M. N., Boucheron, L. E., Can, A., Madabhushi, A., Rajpoot, N. M., & Bulent, Y. (2009). Histopathological image analysis: A review. IEEE

Reviews in Biomedical Engineering, 2, 147–171. https://doi.org/10.1109/RBME.2009.2034865

Gurcan, M. N., & Madabhushi, A. (2011). Digital pathology: Historical perspectives, current status, and future directions. Annual Review of Biomedical Engineering, 13, 331–368. https://doi.org/10.1146/annurev-bioeng-071910-124615

He, H., & Garcia, E. A. (2009). Learning from imbalanced data. IEEE Transactions on Knowledge and Data Engineering, 21(9), 1263–1284. https://doi.org/10.1109/TKDE.2008.239

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778. https://doi.org/10.1109/CVPR.2016.90

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4700–4708. https://doi.org/10.1109/CVPR.2017.243

International Agency for Research on Cancer (IARC), World Health Organization. (2012). GLOBOCAN 2012: Estimated cancer incidence, mortality and prevalence worldwide in 2012. Lyon, France: IARC.

Komura, D., & Ishikawa, S. (2018). Machine learning methods for histopathological image analysis. Computational and Structural Biotechnology Journal, 16, 34–42. https://doi.org/10.1016/j.csbj.2018.05.004

Khoshgoftaar, T. M., Gao, K., & Napolitano, A. (2013). Software quality modeling with imbalanced class data. International Journal of Software Engineering and Knowledge Engineering, 23(8), 1141–1165. https://doi.org/10.1142/S0218194013500401

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 1097–1105.

Krawczyk, B. (2016). Learning from imbalanced data: Open challenges and future directions. Progress in Artificial Intelligence, 5(4), 221–232. https://doi.org/10.1007/s13748-016-0094-0

Krawczyk, B., Woźniak, M., & Herrera, F. (2014). One-class classifiers for imbalanced data: Applications in medicine and bioinformatics. Computational Statistics & Data Analysis, 69, 1–17. https://doi.org/10.1016/j.csda.2013.07.018

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., … van Ginneken, B. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60–88. https://doi.org/10.1016/j.media.2017.07.005

Litjens, G., Sánchez, C. I., Timofeeva, N., Hermsen, M., Nagtegaal, I., Kovacs, I., … van Ginneken, B. (2016). Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Scientific Reports, 6, 26286. https://doi.org/10.1038/srep26286

Madabhushi, A., & Lee, G. (2016). Image analysis and machine learning in digital pathology: Challenges and opportunities. Medical Image Analysis, 33, 170–175. https://doi.org/10.1016/j.media.2016.06.037

Nature Digital Medicine. (2024). METRIC: A framework for evaluating trustworthy AI data quality in healthcare. npj Digital Medicine, 7, 196. https://doi.org/10.1038/s41746-024-01196-4

Paul, A., Mukherjee, D. P., & Das, A. (2016). Classification of histopathological images using texture entropy and supervised classifiers. Biocybernetics and Biomedical Engineering, 36(4), 731–742. https://doi.org/10.1016/j.bbe.2016.06.002

Siegel, R. L., Miller, K. D., & Jemal, A. (2019). Cancer statistics, 2019. CA: A Cancer Journal for Clinicians, 69(1), 7–34. https://doi.org/10.3322/caac.21551

Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. International Conference on Learning Representations (ICLR).

Spanhol, F. A., Oliveira, L. S., Petitjean, C., & Heutte, L. (2016a). A dataset for breast cancer histopathological image classification. IEEE Transactions on Biomedical Engineering, 63(7), 1455–1462. https://doi.org/10.1109/TBME.2015.2496264

Spanhol, F. A., Oliveira, L. S., Petitjean, C., & Heutte, L. (2016b). Breast cancer histopathological image classification using convolutional neural networks. International Joint Conference on Neural Networks (IJCNN), 2560–2567. https://doi.org/10.1109/IJCNN.2016.7727519

Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., & Jemal, A. (2015). Global cancer statistics, 2012. CA: A Cancer Journal for Clinicians, 65(2), 87–108. https://doi.org/10.3322/caac.21262

Vapnik, V. (1998). Statistical learning theory. Wiley.

Wang, X., Yang, W., Weinreb, J., Han, J., Li, Q., Kong, J., & Yan, C. (2017). Searching for prostate cancer by fully automated magnetic resonance imaging–based deep learning detection. Journal of Clinical Oncology, 35(27), 3100–3108. https://doi.org/10.1200/JCO.2017.72.1442

Xu, J., Luo, X., Wang, G., Gilmore, H., & Madabhushi, A. (2016). A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images. Neurocomputing, 191, 214–223. https://doi.org/10.1016/j.neucom.2016.01.034

Zhang, Z., Chen, P., McGough, M., Xing, F., Wang, C., Bui, M., & Yang, L. (2019). Pathologist-level interpretable whole-slide cancer diagnosis with deep learning. Nature Machine Intelligence, 1(5), 236–245. https://doi.org/10.1038/s42256-019-0052-1

CHAPTER 2
AI IN LAGOS STATE TRANSPORTATION SYSTEM: BARRIERS, CURRENT STATE AND FUTURE POTENTIAL

[1]Christian Nwankwo CHIJIOKE
[2]Gbolahan Afeez ADIGUN
[3]Omoboriowo Samson LOYE
[4]Samuel Sola AKOSILE

[1]National Centre for Artificial Intelligence and Robotics, National Information Technology Development Agency, Abuja, Nigeria, christianchijioke001@gmail.com, ORCID ID: 0009-0009-3413-3219
[2]Obafemi Awolowo University, Osun, Nigeria, gbolahanadigun1606@gmail.com, ORCID ID: 0009-0009-7136-9431
[3]Obafemi Awolowo University, Osun, Nigeria, loyeomoboriowo@gmail.com, ORCID ID: 0009-0007-2979-0295
[4]Morgan State University, Maryland, United States of America, saako2@morgan.edu, ORCID ID: 0009-0002-6275-8768

INTRODUCTION

The economy of any nation or region depends heavily on the state and operation of its transport systems. The transportation sector is an integral part of society, needed for social and economic growth (Rodrigue, 2024). Enhancing the movement of people and goods, transportation plays an indispensable role in every other sector (Dostál and Adamec, 2011). Increasing transportation demands due to global population growth imply more serious and novel challenges. The unpredictability in users' behavior complicates the matter further. Commuting is therefore becoming increasingly complex by the day, and it requires new approaches to counter the challenges, including traffic congestion, unpredictability in commuting, safety issues, energy use, and carbon emissions associated with modern mobility (Abduljabbar et al., 2019). Technology has improved drastically and exponentially, making daily operations easier. Artificial Intelligence (AI), including machine learning (ML) and deep learning (DL), has become a tool to solve problems smartly and effectively in diverse fields, including Fintech, medicine, agriculture, transportation, etc. (Bharadiya, 2023; Chakrabarty et al., 2026; Yu et al., 2018). The application of AI is based on the underlying belief that machines (and software) can have the same level of intelligence that humans have and can therefore make intelligible and evidence-based decisions (Jiang et al., 2022). Considering the traffic infrastructure, vehicles, and users, AI is currently deployed in the transportation sector, transitioning from the conventional operation of transport systems to a more intelligent, safer, smarter, and more effective operation (Abduljabbar et al., 2019). Recent technologies in transportation include autonomous driving, connected vehicles, vehicle-to-environment (V2X) communication, shared mobility and smart traffic management (Ji et al., 2025). These technologies have enhanced predictive maintenance of transport infrastructure, traffic management, commuting demand forecasting, and safety (Cohen, 2024).

1. PROBLEM STATEMENT

Lagos state, being the most populous state in Nigeria and Africa as a continent, is one whose operations have a very significant influence on the development of the continent.

Recently, the state's transport systems have undergone significant improvements, with the construction of advanced rail projects and the introduction of intelligent systems at strategic locations. However, commuters in the state still struggle with incessant traffic congestion, long waiting and commuting hours, exorbitant transport costs and a high rate of accidents. The purpose of this study is to assess the current state of the state's transportation and explore the maximization of artificial intelligence (AI) in the more effective management of the transport systems.

1.1 Research Aims and Objectives

The primary objective of this research is to evaluate the current application of AI in transport systems in Lagos State and explore ways to optimize it for more effective transportation. The specific objectives are:

a. To examine how AI is currently being used in the Lagos State transport system.
b. To identify key barriers to integrating AI into the Lagos State transport system.
c. To propose ways AI could improve traffic, safety, and transport planning.

1.2 Structure of Paper

This paper describes, in detail, the current state of the transport systems in Lagos State. The research then delves into the current use of AI in managing Lagos' transport systems, identifies barriers limiting the broader integration of AI, and suggests solutions to tackle these barriers, ultimately aiming to optimize AI in the management and operation of the transport systems.

1.3 Definition of Key Terms

a. Transportation: Transportation refers to the mobility of people and goods from one geographical location to another by land, water, or air. With each mode of transportation offering its peculiar advantages over others, a vibrant economy is heavily invested in the transport networks via these three modes. Transportation by road involves the use of vehicles, bikes, tricycles on roads and bridges, and trains on railways. It also involves the movement of fluids like crude oil and natural gas through pipelines (Johnstone, 2015).

Suitable for the transportation of heavy goods and cargoes, water transportation makes use of boats, ships, and canoes on waterways such as oceans, lagoons, and rivers. Air transportation, the most expensive and fastest means of travel, utilizes airplanes, jets, helicopters, and other aircraft. Beyond the movement of people and goods, transportation facilitates the delivery of services, creates employment opportunities, promotes international trade, and fosters development, amongst other indispensable benefits. Conventionally, effective transportation has focused on the construction and maintenance of transportation networks; however, the human factor in transportation—the operation of commuters along these networks—is now also being prioritized.

Traffic basically refers to the mobility of humans, vehicles, and goods measured along a particular route. Congestion occurs when the number of vehicles plying a road exceeds the road capacity, often characterized by low speeds and elongated travel times (Gladys Chidinma Oweisana & Ndubuisi Ordua, 2022). Usually, commuters experience heightened traffic congestion during rush hours - in the morning and evening, due to mobility to and from their workplaces. Traffic congestion has a negative impact on the productivity of commuters and the economy of the region, at large (Akhtar & Moridpour, 2021). Other effects include stress and discomfort, noise and air pollution, and increased fuel usage (Habiba & Talukdar, 2025; Xu et al., 2024).

b. Public Transit: Also known as mass transit, urban transit, or public transportation, this is any form of transportation, whether privately or government-owned, that is available for the public to use and is typically paid for. This includes all modes of transportation – road, rail, air, and water transportation (Daganzo & Ouyang, 2019; Victoria Transport Policy Institute, 2025). Public transit is an important part of the transportation system of any region - it significantly reduces the usage of individual cars, thereby reducing environmental pollution, traffic congestion, and wear on transport networks (Montalvo-Martel et al., 2022). Many public transit services offer a much cheaper means of transportation than riding personal vehicles, while others, like chaffering and taxi-riding, charge higher fares.

c. Artificial Intelligence (AI): Artificial Intelligence is the term given to the ability of computer programs to analyse situations and solve problems, just as a human mind would (Bharadiya, 2023).

AI has many applications, like natural language processing (NLP), image processing, prediction making, autonomy, amongst others, in different fields of study, including medicine, law, engineering, energy, and manufacturing (Buch et al., 2018; Lyu & Liu, 2021). With increasing levels of challenges arising by the day, AI is now a household name and has helped to achieve tremendous progress with its human-level performance (Jiang et al., 2022). The performance of Machine Learning (ML), the most popular subset of AI, relies heavily on the quality and quantity of training data used to make predictions (Zhang & Lu, 2021). ML also has a subset, called Deep Learning (DL), in which multilayered neural networks learn from big data (vast amounts of data) (Holzinger et al., 2019).

d. Intelligent Transportation Systems (ITS): ITS are facilities with in-built computer programs that improve transport network optimization as well as commuters' satisfaction. ITS-based devices are essential for developing smart cities (Elassy et al., 2024). They provide and analyze a large amount of real-time traffic data to enhance traffic management, traffic law enforcement, and safety (Zhu et al., 2019). They also provide a range of other functions, including route suggestion, autonomous driving, vehicle-to-vehicle (V2V) communication, and vehicle-to-infrastructure (V2I) communication. ITS applications are mainly found at strategic locations, such as intersections (Coogan & Arcak, 2015; Ganin et al., 2019). These systems ensure there is a free flow of traffic as much as possible, thereby reducing transit time, accident rate, occurrence of congestion, and carbon emissions (Elassy et al., 2024).

2. METHODOLOGY

This paper adopts a traditional review approach to construct a narrative from existing studies on AI in Lagos State transportation. Most of the sources were obtained from peer-reviewed journal articles accessed through Google Scholar and ScienceDirect, as well as preprints available on arXiv. The study focuses on studies published between 2015 and 2025, which highlight the latest developments in the field. Literature was selected based on its relevance to AI in Lagos State transportation. Additionally, the paper reviews literature on AI in transportation more broadly. This was done to identify ways Lagos State could benefit from such studies.

Lagos, the study area, is located at approximately latitude 6°22′ North to 6°42′ and longitude 2°45′ East to 4°20′ (Koko et al., 2021). As mentioned earlier, it is the most populous city in Nigeria and is ranked among the largest metropolitan areas in Africa (Sasu, 2024). The city served as the capital of Nigeria until 1991, when the federal capital was officially relocated to Abuja (Takyi, 2016). Due to its former status and location, Lagos has retained substantial infrastructure and opportunities (Adewolu, 2023). This has attracted many Nigerians and people from abroad (Lawanson, 2023).

Lagos was selected as the study area because, as a former federal capital of Nigeria with a dense population, it reflects the characteristics of a national hub. We assume that achievements in Lagos could potentially be replicated in other parts of Nigeria and Africa.

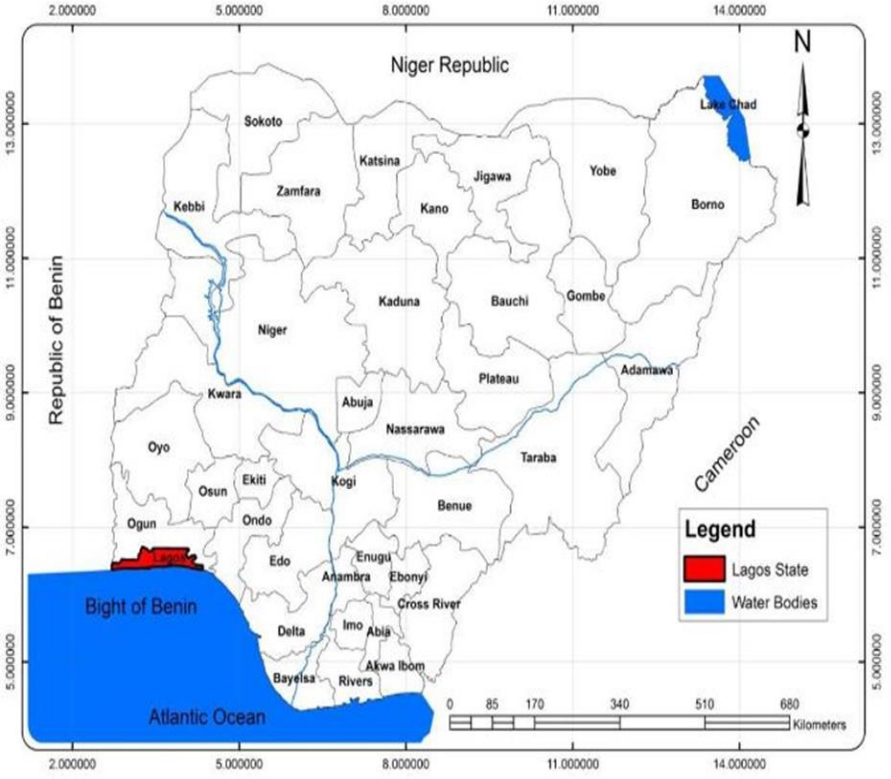


Figure 1. Geographical Location of Lagos State in Nigeria

3. CURRENT CONDITION OF AI IN LAGOS STATE TRANSPORTATION SYSTEMS

The road network forms a major part of the transportation networks in Lagos State, accounting for over 95% of all trips in the state (Atubi, 2013). The road networks run through the two parts of the state, the Mainland and the Island, with the Eko Third Mainland (also called the Ibrahim Babaginda Boulevard), and Carter bridges linking the two (Institution of Civil Engineers (ICE), 2025). However, the existing roads are not in excellent condition due to a lack of adequate maintenance and funding. Additionally, the required number of new roads to accommodate the increasing number of commuters has not been constructed over the last twenty years (LAMATA, 2024).

In addition to these, the state houses Nigeria's busiest railways, airports, and seaports. The state rail network systems, collectively known as the Lagos Rail Mass Transit (LRMT), cover strategic locations within and outside the state. The state's existing rail lines include the Blue Line and Red Line, with proposals in place for expanding existing rail lines and constructing new ones. The Blue Line records over 40,000 commuters daily (Salami, 2025). The busiest of the five international airports in Nigeria is located in Lagos State. The Murtala Muhammed International Airport (MMIA) serves approximately 15 million passengers annually (Okeke-Korieocha, 2025). The airport also features a dedicated terminal, known as the Murtala Muhammed Airport General Aviation Terminal (GAT), for domestic flights. Lagos State is also home to three of the seven major seaports in Nigeria – The Lagos Port Complex (Premiere Port/Apapa Quays), Tin Can Island Port Complex, and Lekki Deep Sea Port. The seaports, which account for above 97% and 80% of Nigeria's exports and imports, respectively, are connected to rail and road networks for the efficient transportation of cargoes (Anagor-Ewuzie, 2024).

With the state being home to more than 22 million people, the importance of these transport networks cannot be overemphasized (LAMATA, 2025).

3.1 Current use of AI in Lagos State Transport Systems

Although the incorporation of AI in Lagos State's Transport Systems is in its infancy, remarkable progress has been made (Taiwo, 2022).

As of the end of 2025 first quarter, the Lagos State government has installed functional Intelligent Transport Systems (ITS) at eleven strategic locations in the state, including the Murtala Mohammed International Airport Road, Third Mainland Bridge, Ojota Bridge, Lekki-Ikoyi Bridge, Ikorodu Road, Nurudeen Olowopopo Road, Oshodi-Apapa Expressway, Allen Avenue Road, Mobolaji Bank Anthony Way, Ogudu Road, and Epe Expressway. The locations, chosen mainly due to their high traffic density and accident rates, utilize technologies such as speed cameras, electronic police systems, and traffic control systems. These technologies monitor traffic violations, record vehicle speeds, and capture license plates, thereby ensuring compliance with speed limits, aiding traffic enforcement, and tracking traffic conditions. The government has also launched a public awareness campaign to educate commuters about the technologies and penalties for breaking traffic laws (Simon, 2025).

Figure 2 . Automatic Number Plate Recognition (ANPR) cameras in Lagos State (Olagunju, 2023)

3.2 Government Future Initiatives and Private Sector Contributions

As part of its plan for a smart city, the Lagos State Government is set to construct four new ITS sites at strategic points in the state. This AI-enhanced plan is to enhance traffic control, improve road safety, and reduce congestion by providing real-time traffic conditions. The technologies to be deployed by Huawei Technologies include Automatic Number Plate Recognition (ANPR) cameras and Traffic Management Solution (TMS) devices (African Review, 2025).

Additionally, the state government plans to integrate AI into the operation of its public transportation system, specifically the Bus Rapid Transit (BRT). Before the introduction of the BRT scheme in 2008, Lagosians only had access to public transportation by minibuses (danfo), midi-buses (molue), cabs (kabu-kabu), and motor bikes (okada). These means of public transportation are often disorderly, expensive, and uncomfortable, with passengers often hustling to find a seat or standing in them. Being the first of its kind in Africa, the BRT scheme started operations to provide a fast, affordable, and comfortable means of public transport. To optimize the current operations of the BRT scheme and upgrade it to an AI-powered system, the state government has partnered with CapitalCore and Optibus. This plan is to prepare the BRT scheme for the anticipated increased ridership by providing more new buses and adopting a new scheduling software (Levner, 2024). With proven testimonies of Optibus' intervention in Brazil's bus transit, the Lagos state government hopes to improve the passenger experience and quality of transportation services rendered by its BRT system (Levner, 2022).

4. BARRIERS TO AI INTEGRATION IN LAGOS STATE TRANSPORTATION SYSTEM

The implementation of AI solutions in transportation varies among countries due to differences in the state of infrastructure, research, and the availability of skilled personnel in each region (Conde and Twinn, 2019). In Nigeria, the infrastructure required for the implementation faces challenges, including poor internet connectivity, inconsistent electricity, and inadequate data centers (Oyeyemi et al., 2025).

Hence, the integration of AI in Lagos State's transport system has been slow due to some of these systemic challenges. A high-speed internet connection is essential for deploying AI solutions, such as smart traffic management and autonomous vehicles (Shankar Iyer, 2021). Unlike other advanced countries where AI systems have been fully deployed, Nigeria lacks advanced digital infrastructure, such as widespread 5G coverage and high-capacity data centers. Lagos, being a state in Nigeria, lacks high-speed internet connectivity, which limits the transfer of real-time data required for AI-powered solutions in the state's transportation system.

Unreliable power supply in Lagos also hinders the smooth deployment of AI systems in transportation. The continuous operation of AI-powered systems relies on the reliable operation of sensors, communication networks, and servers, all of which are dependent on a consistent electricity supply. Additionally, the scarcity of large data centers affects the processing and storage of transportation datasets required for building AI models. The lack of centralized data management obstructs data retrieval, access, and integration for AI applications (Martin et al., 2025). AI systems are fundamentally data-driven; therefore, incomplete, low-quality, or biased data increases the risk that the system will make unreliable or unfair decisions (Atubi, 2025). “Data is the lifeblood of AI systems,” but in Nigeria, the scarcity of structured and robust datasets hinders accurate model training and deployment (Oyeyemi et al., 2025). Lagos has limited access to structured data and lacks a centralized, standardized transportation database. Fragmented data collection across agencies in Lagos, combined with limited access to open data sources, limits the development of AI models. In view of all these, it is challenging for Lagos to build the foundational technology required for broad AI adoption in its transport sector.

4.1 Socio-Economic and Policy Constraints

A major barrier to the development of AI in transportation is the potentially high cost of AI systems (software and hardware) (Conde and Twinn, 2019). Integrating AI into transportation systems requires a huge investment in sensors, smart cameras, servers, software platforms, and skilled personnel.

Although AI-driven predictive maintenance reduces operational cost in multimodal transport, its high initial capital demands and technical hurdles necessitate innovative and tailored financing frameworks (Wiese, 2024). Transport agencies in Lagos are already facing funding issues to cater to the continually deteriorating roads, stops, and transit systems, among other infrastructure needs. Therefore, the high financial burden required to integrate AI into the transport system makes adoption difficult.

Additionally, the skill gap and lack of human capital have been constraints to the rapid adoption of AI in Lagos' transport sector. Because the transport sector is traditionally built on manual expertise and operations, it lacks the readiness for the data-centric nature of AI technologies (Martin et al., 2025). The lack of advanced expertise poses a major barrier to bridging the gap between Nigeria citizens and AI technology (Robinson, 2018). A shortage of local experts in AI, machine learning, and data science in Lagos affects the design and implementation of AI-powered systems, giving rise to a reliance on foreign expertise and solutions.

Another major barrier hindering the integration of AI in Lagos State's transportation sector is the lack of clear and comprehensive regulations and policies that guide the integration of AI systems in transport. The lack of unified governance frameworks results in a regulatory vacuum, leading to significant uncertainties in compliance obligations by stakeholders (Floridi et al., 2018). Policies guiding data privacy and ethical standards are necessary to encourage the advancement of AI integration in Lagos' transport sectors by both the investors and operators. The lack of a clear regulatory framework in Lagos makes the broad adoption of AI systems challenging. Therefore, the combination of the huge investment required, limited skilled personnel, and a lack of clear policies hinders the broad deployment of AI in Lagos' transport systems.

4.2 Cultural and Behavioral Resistance

As the adoption of AI expands, the transition from the traditional economic model towards a service-oriented economy accelerates while simultaneously increasing the risk of job losses among low-skilled workers, including those in the transportation sector (Meltzer, 2018).

The advancement of automation technologies in the transport sector is expected to reduce manual labor demands, which may result in job losses (Atubi, 2025). Informal transport workers in Lagos, such as danfo drivers and conductors, view the AI automation system in transportation as a threat to their daily jobs. As a result of this fear, workers in the sector are not open to reforms that would facilitate the implementation of AI and intelligent systems in Lagos' transport services. For instance, the already existing smart cameras that capture traffic offenders on the road are viewed by the traffic/road taskforce workers in Lagos as a threat to their livelihood.

A lack of public trust has also contributed to the low adoption of AI in the Lagos transport sector. Locals, both operators and commuters, find it difficult to believe in the reliability, safety, and inclusiveness of AI solutions in transportation problems (Martin et al., 2025). The widespread belief among these locals is that AI-driven solutions are foreign, non-inclusive, unsafe, and elitist. A lack of understanding of the process is another factor that contributes to a lack of trust among the people. The government must make considerable efforts to earn the trust of the people in AI, as adoption in the transport sector will remain low if this is not achieved. Public fear must be addressed through awareness, engagement, and inclusivity for AI-driven innovation in Lagos transport to flourish.

5. FUTURE POTENTIAL OF AI IN LAGOS STATE TRANSPORTATION

AI has benefited nearly every aspect of human life. In healthcare, it supports doctors in diagnosing diseases and recommending treatments (Alowais et al., 2023). In the music industry, it enables musicians and songwriters to generate content in seconds (Semancik, 2024). Perhaps most remarkably, AI extends its influence beyond Earth. It has powered space missions and operations (Halloran, 2025). Building on these successes across numerous fields, this section reviews previous studies that explore the future potential of AI in shaping transportation in Lagos State.

5.1 Improving Traffic Flow and Congestion

Paul and McSharry (2021) noted that the primary reason for traffic congestion is an excess of vehicles on roadways. Focusing on the Lagos State Bus Rapid Transit (BRT) system, they analyzed urban travel demand using individual commuter trips. Their goal was to make the BRT more available to increase its share of urban trips. At the time, the BRT served fewer than 5% of Lagos passengers. The researchers believed that increased BRT patronage would cut down private car usage. The study's primary objective was to compare the average waiting times for commuters under both fixed and dynamic bus scheduling systems. They applied a K-means clustering model to segment stations by demand patterns. Using these clusters, they built a simulation in MATLAB to replicate observed demand and supply. The results showed that dynamic scheduling could reduce average waiting times by up to 80% compared to fixed scheduling. Dynamic scheduling, as an element of demand-responsive transit, requires analyzing real-time data on passenger requests, traffic, and other factors to set routes and schedules dynamically. AI-powered systems analyze this data and create a responsive and efficient transit service.

Udoh et al. (2025) found that traditional road expansion methods were insufficient in improving traffic flow due to a rapid increase in the number of vehicles on roadways. The researchers observed that developed countries already use AI-based traffic management systems and decided to adapt these for Lagos. They proposed two AI models. An Artificial Neural Network (ANN) and a Fuzzy Inference System (FIS). Both models were trained on field data from traffic light timing and vehicle distances collected in the Abule Egba and Ikeja areas of Lagos State. The researchers integrated both AI models into a SIMULINK traffic management simulation and evaluated the performance using R-squared values. While ANN showed superior accuracy, both models achieved prediction accuracies above 99%. The study recommended adopting AI models like ANN for real-time traffic management to improve traffic flow.

Gunarathna et al. (2019) argued that the number of lanes in each direction of a road network should be dynamic. The researchers proposed a multi-agent AI architecture that uses Reinforcement Learning (RL). The system dynamically allocated lanes per direction based on traffic demand.

They evaluated the methodology using a traffic simulation based on data from New York City. Their RL model outperformed all other solutions. This study holds significant potential for Lagos State. Instead of the costly and disruptive process of building new roads, the number of lanes per direction could be changed based on demand. Lagos traffic is also highly unpredictable due to various factors (Atakiti et al., 2016). An AI system that dynamically adjusts the number of lanes per direction in real-time would be more resilient than static systems (Gunarathna et al., 2019).

Numerous studies have shown that AI can reduce congestion. However, most focus on only vehicles. A system that minimizes car delays can cause pedestrians to wait for a long time (Yazdani et al., 2023). To solve this, Yazdani et al. (2023) proposed the Intelligent Vehicle Pedestrian Light (IVPL). Their approach regulated traffic signal for both. Its goal was to share green time fairly between all road users. Their results proved IVPL outperformed a standard smart traffic system. Lagos State's large pedestrian population would greatly benefit from this approach.

5.2 Enhancing Safety and Security

Li et al. (2024) introduced a paradigm shift from reactive to predictive transportation management. The researchers developed a new system called Digital Twin-based Driver Risk-Aware Intelligent Mobility Analytics (DT-DIMA). It used existing Pan-Tilt Cameras (PTCs) to collect data from road networks. This data feeds a digital twin, which is then used to predict where traffic jams and hazardous driving situations are likely to occur. Adopting such a safety-focused approach in Lagos would reduce accidents and save lives.

Several studies have examined the use of Automatic Incident Detection (AID) systems in traffic management. AID systems are designed to detect incidents automatically with speed and accuracy. They monitor traffic parameters and trigger alarms when deviations from normal conditions exceed predefined thresholds (ElSahly & Abdelfatah, 2022). In their systematic review, ElSahly and Abdelfatah (2022) categorized and compared the major algorithms developed for incident detection. They concluded that AI- and ML-based algorithms show the greatest promise compared to traditional approaches.

For Lagos State, where congestion and frequent road incidents disrupt mobility, adopting AID systems could increase safety by improving overall transport efficiency.

5.3 Sustainability

Traffic optimization leads to shorter travel times and reduced journey duration directly decreases pollutant emissions (Mirindi, 2024).

Dhulasi and Saranya (2023) conducted a narrative review on AI in sustainable transportation. They found that AI technologies can enhance eco-friendly transport solutions across multiple sectors. Their research connects these capabilities to specific UN Sustainable Development Goals. The relevant goals include health, infrastructure development, sustainable urban planning, and climate action.

For roadway systems, the authors identify several AI applications. These include intelligent traffic management, computer vision for accident detection, delay identification systems, and coordinated vehicle grouping. Railway transportation benefits from AI through maintenance prediction and automated operations. Maritime shipping gains advantages from fuel efficiency algorithms, optimal route selection and intelligent port management.

The study did not focus on Lagos State. However, the research findings remain valuable for AI in Lagos State transportation. The identified technologies could benefit Lagos's diverse transport network. This includes Bus Rapid Transit systems, commercial minibus services, railway lines, water ferries, and cargo port operations diverse transport network.

Akter (2024) identified AI as a key tool for addressing climate change. The study showed how AI-driven optimization supports intelligent route planning for public transit and electric vehicles (EVs). It provided empirical evidence, noting that AI-optimized buses could reduce emissions by up to 50% by 2050.

CONCLUSION

Lagos, being a rapidly growing megacity, faces numerous transportation challenges that hinder its economic growth, productivity, and quality of life. This paper has assessed the current state of AI applications in Lagos State's transportation, highlighting the Intelligent Transportation Systems deployed by the state government and the innovations in public transit by the private sector. It has also identified major barriers, including infrastructure gaps, high initial deployment costs, limited expertise, regulatory gaps, and a lack of public trust.

Despite all these challenges, the paper highlights the transformative potential of AI-driven systems in Lagos State's transportation. AI can potentially provide innovative solutions to transportation issues, which include congestion, unsafe intersection traffic flow, unpredictability in commuting, and carbon emissions in Lagos. AI innovative solutions, including adaptive traffic signal control, predictive modeling, and demand-responsive public transit, will improve the efficiency, safety, and sustainability of Lagos State's transportation. Although the journey to a fully smart transport system driven by AI systems is complex, it can be achieved through strategic investments and financing, the harmonization of policies, capacity building, stakeholder engagement, and a commitment to utilizing local innovations.

In conclusion, AI is the key to unlocking and reshaping the potential of Lagos State's smart transportation system. By addressing the current barriers and capitalizing on the future opportunities identified in this paper, Lagos State can develop a smart, efficient, sustainable, and equitable transportation system powered by AI.

Recommendation

To accelerate AI adoption and harness its potential, all stakeholders must be involved. These stakeholders include the government, the private sector, and academia. The government must invest in digital infrastructure, including broadband connectivity, power supply, and centralized data centers. Collaboration with the private sector to promote public-private partnerships will help bridge the infrastructure gap and mitigate the high initial costs associated with deploying AI systems.

The government is also responsible for developing a comprehensive regulatory policy that guides AI governance and data protection. The government must also create public awareness to alleviate the fear and skepticism of stakeholders, including both commuters and operators. This public awareness will reduce resistance and gain the trust of the people, which will, in turn, lead to a high rate of acceptance of AI integration in Lagos State's transportation sector.

The private sector must collaborate with the government in employing local AI solutions tailored to Lagos State, rather than importing foreign technologies to address the transportation issues in Lagos. The private sector must also collaborate with academia and invest in local research that advances AI in transportation. Public-private partnerships should be encouraged to facilitate the sharing of costs and risks associated with deploying large-scale AI solutions.

Finally, universities and institutions should establish dedicated research data centers and develop specialized training programs in AI, data science, and Intelligent Transportation Systems to train the next generation of engineers, data scientists, and planners who are well-equipped to develop and maintain these AI systems locally. With this, academia, in collaboration with the government, can help bridge the existing skill gap. Academia must also play a significant role in generating locally sourced datasets and developing AI systems to address Lagos' unique transport challenges.

REFERENCES

Abduljabbar, R., Dia, H., Liyanage, S., & Bagloee, S. A. (2019). Applications of artificial intelligence in transport: An overview. Sustainability (Switzerland), 11(1). https://doi.org/10.3390/SU11010189

Adeniran, I., Otokiti, K. V., & Durojaye, P. (2020). Climate change impacts in a rapidly growing urban region – A case study of Ikeja, Lagos, Nigeria. International Journal of Environmental Planning and Management, 6(1), 13–23. American Institute of Science.

Adewolu, A. O. (2023). Infrastructure growth and sustainable development: Review of Lagos city profile. International Journal of Engineering Inventions, 12(5), 253–276.

African Review. (2025). Lagos, Huawei advance smart transport systems - African Review. https://africanreview.com/ict/lagos-huawei-advance-smart-transport-systems

Akhtar, M., & Moridpour, S. (2021). A Review of Traffic Congestion Prediction Using Artificial Intelligence. Journal of Advanced Transportation, 2021(1), 8878011. https://doi.org/10.1155/2021/8878011

Akter, M. S. (2024). Harnessing technology for environmental sustainability: utilizing AI to tackle global ecological challenge. Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023, 2(1), 61-70.

Alowais, S. A., Alghamdi, S. S., Alsuhebany, N., Alqahtani, T., Alshaya, A. I., Almohareb, S. N., Aldairem, A., Alrashed, M., Bin Saleh, K., Badreldin, H. A., Al Yami, M. S., Al Harbi, S., & Albekairy, A. M. (2023). Revolutionizing healthcare: The role of artificial intelligence in clinical practice. BMC Medical Education, 23, Article 689. https://doi.org/10.1186/s12909-023-04698-z

Anagor-Ewuzie, A. (2024). Here are top four Nigerian ports by operational value - Businessday NG.

Atakiti, I. O., Ogunwemimo, T. A., Alao, O. O., Chioma, P. E., & Ofurum, I. (2016). The role of Lagos traffic radio in educating road users on traffic management in Lagos State, Nigeria. International Journal of Current Research, 8(1), 25119-25125.

Atubi, A. O. (2013). An Evaluation of Transport Infrastructure in Lagos State, Nigeria. Journal of Geography and Earth Science, 1(1), 9–18. www.aripd.org/jges

Atubi, O. A. (2025). Artificial Intelligence: Exploring its Application in Transportation Industry. Social Sciences and Education Research Review, 12(1), 179 – 186. https://doi.org/10.5281/zenodo.15804536

Bharadiya, J. (2023). Artificial Intelligence in Transportation Systems A Critical Review. American Journal of Computing and Engineering, 6(1), 34–45. https://doi.org/10.47672/AJCE.1487

Buch, V. H., Ahmed, I., & Maruthappu, M. (2018). Artificial intelligence in medicine: current trends and future possibilities. The British Journal of General Practice, 68(668), 143. https://doi.org/10.3399/BJGP18X695213

Chakrabarty, S., Deb, C. K., Marwaha, S., Haque, M. A., Kamil, D., Bheemanahalli, R., & Shashank, P. R. (2026). Application of artificial intelligence in insect pest identification - A review. Artificial Intelligence in Agriculture, 16(1), 44–61. https://doi.org/10.1016/J.AIIA.2025.06.005

Cohen, A. (2024). The Role of Artificial Intelligence in Transportation Introduction and Background.

Conde, M. L. and Twinn, I. (2019). How Artificial Intelligence is Making Transport Safer, Cleaner, More Reliable and Efficient in Emerging Markets. EMCompass; Note 75. International Finance Corporation. http://dx.doi.org/10.1596/33387

Coogan, S., & Arcak, M. (2015). A Compartmental Model for Traffic Networks and Its Dynamical Behavior. IEEE Transactions on Automatic Control, 60(10), 2698–2703. https://doi.org/10.1109/TAC.2015.2411916

Daganzo, C. F., & Ouyang, Y. (2019). Transit Basics. Public Transportation Systems, 1–31. https://doi.org/10.1142/9789813224100_0001

Dhulasi, P. S., & Saranya, K. G. (2023). Significance of artificial intelligence in the development of sustainable transportation. The Scientific Temper, 14(2), 418–425.

Dostál, I., & Adamec, V. (2011). Transport and its Role in the Society. Transactions on Transport Sciences, 4(2), 43–56.

Elassy, M., Al-Hattab, M., Takruri, M., & Badawi, S. (2024). Intelligent transportation systems for sustainable smart cities. Transportation Engineering, 16, 100252. https://doi.org/10.1016/J.TRENG.2024.100252

ElSahly, O., & Abdelfatah, A. (2022). A systematic review of traffic incident detection algorithms. Sustainability, 14(22), 14859. https://doi.org/10.3390/su142214859

Floridi, L., Cowls, J., Beltrametti, M. et al. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. Minds & Machines, 28, 689–707. https://doi.org/10.1007/s11023-018-9482-5

Ganin, A. A., Mersky, A. C., Jin, A. S., Kitsak, M., Keisler, J. M., & Linkov, I. (2019). Resilience in Intelligent Transportation Systems (ITS). Transportation Research Part C: Emerging Technologies, 100, 318–329. https://doi.org/10.1016/J.TRC.2019.01.014

Gladys Chidinma Oweisana, B., & Ndubuisi Ordua, V. (2022). Influence of Traffic Congestion on Psychological Stress and Pro-Social Behaviour among Commuters in Port-Harcourt Metropolis.

Gunarathna, U., Xie, H., Tanin, E., Karunasekara, S., & Borovica-Gajic, R. (2019). Dynamic graph configuration with reinforcement learning for connected autonomous vehicle trajectories. arXiv.

Habiba, U., & Talukdar, S. (2025). The impact of traffic congestion, aggression and driving anger on driver stress: A structural equation modelling approach. Journal of Transportation Safety and Security, 17(8), 901–922. https://doi.org/10.1080/19439962.2025.2471290

Halloran, K. (2025, January 7). NASA's AI use cases: Advancing space exploration with responsibility. NASA.

Holzinger, A., Langs, G., Denk, H., Zatloukal, K., & Müller, H. (2019). Causability and explainability of artificial intelligence in medicine. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 9(4), e1312. https://doi.org/10.1002/WIDM.1312

Institution of Civil Engineers (ICE). (2025). Third Mainland Bridge Lagos: History And Engineering Facts . https://www.ice.org.uk/what-is-civil-engineering/infrastructure-projects/third-mainland-bridge-lagos

Ji, A., Huang, J., Qin, Z., Sun, Z., Zhao, R., & Zheng, G. (2025). Cooperative merging for connected automated vehicles in mixed traffic: A multi-agent reinforcement learning approach. Artificial Intelligence for Transportation, 2, 100007. https://doi.org/10.1016/J.AIT.2025.100007

Jiang, Y., Li, X., Luo, H., Yin, S., & Kaynak, O. (123 C.E.). Discover Artificial Intelligence Quo vadis artificial intelligence? Discover Artificial Intelligence, 2, 4. https://doi.org/10.1007/s44163-022-00022-8

Jiang, Y., Li, X., Luo, H., Yin, S., & Kaynak, O. (2022). Quo vadis artificial intelligence? Discover Artificial Intelligence 2022 2:1, 2(1), 1–19. https://doi.org/10.1007/S44163-022-00022-8

Johnstone, R. W. (2015). Transportation Systems and Security Risks. Protecting Transportation, 73–106. https://doi.org/10.1016/B978-0-12-408101-7.00003-9

Koko, A. F., Yue, W., Abubakar, G. A., Roknisadeh, H., & Alabsi, A. A. N. (2021). Analyzing urban growth and land cover change scenario in Lagos, Nigeria using multi-temporal remote sensing data and GIS to mitigate flooding. Geomatics, Natural Hazards and Risk, 12(1), 631–652.

LAMATA. (2024). Lagos State Transport Policy. https://www.lamata-ng.com/wp-content/uploads/2025/04/LAGOS-STATE-TRANSPORT-POLICY-.pdf

Lawanson, L. (2023, August 18). Lagos from its margins: Land and insecurity in an African megacity. Heinrich Böll Foundation. https://hoa.boell.org/en/2023/08/18/lagos-its-margins-land-and-insecurity-african-megacity?utm_source

Levner, A. (2022). Optimizing BRT Sorocaba for 1.6 million monthly passengers. https://blog.optibus.com/optimizing-brt-sorocaba-for-1.6-million-monthly-passengers

Levner, A. (2025). Nigeria's first AI-powered BRT system will be powered by Optibus and CapitalCore. https://blog.optibus.com/nigerias-first-ai-powered-brt-system-will-be-powered-by-optibus-and-capitalcore

Li, T., Bian, Z., Lei, H., Zuo, F., Yang, Y.-T., Zhu, Q., Li, Z., Chen, Z., & Ozbay, K. (2024). Digital Twin-based Driver Risk-Aware Intelligent Mobility

Analytics for Urban Transportation Management. arXiv. https://doi.org/10.48550/arXiv.2407.15025

Lyu, W., & Liu, J. (2021). Artificial Intelligence and emerging digital technologies in the energy sector. Applied Energy, 303. https://doi.org/10.1016/J.APENERGY.2021.117615

Martin, P., Jamal, O., Bozorg, S., Khullar, T., and Workman. R. (2025). Bridging the Gap: Overcoming the barriers to AI adoption in transport. Published Project Report PPR2060. TRL Limited, Crowthorne. https://doi.org/10.58446/ansu5904

Meltzer, J. (2018). The Impact of Artificial Intelligence on International Trade. Center for technology Innovation at Brookings, 9.

Mirindi, D. (2024). A review of the advances in artificial intelligence in transportation system development. Journal of Civil, Construction and Environmental Engineering, 9(3), 72–83.

Montalvo-Martel, M., Ochoa-Zezzatti, A., Carrum, E., & Perez, P. (2022). Proposal of a smart framework for a transportation system in a smart city. Artificial Intelligence and Industry 4.0, 1–2, 143–174. https://doi.org/10.1016/B978-0-323-88468-6.00007-3

Okeke-Korieocha, I. (2025). FG to rebuild old MMIA terminal - Businessday NG. https://businessday.ng/aviation/article/fg-to-rebuild-old-mmia-terminal/

Olagunju, K. (2023). Transportation technology and security: the hub for national and sub-national development and safety. https://frsc.gov.ng/wp-content/uploads/2023/09/OLAGUNJUS-PRESENTATION.pdf

Oyeyemi, B. B., John, A. O., and Awodola, M. (2025). Infrastructure and Regulatory Barriers to AI Supply Chain Systems in Nigeria vs. the U.S. Engineering Science & Technology, 6(4), 155-172.

Paul, O., & McSharry, P. E. (2021, May 25). Public transportation demand analysis: A case study of metropolitan Lagos. SSRN. https://doi.org/10.2139/ssrn.3852747

Robinson, R. N. (2018). Artificial Intelligence: Its Importance, Challenges and Applications in Nigeria. Direct Research Journal of Engineering and

Information Technology, 5(5), 36-41. https://doi.org/10.26765/DRJEIT.2018.4780

Rodrigue, J. P. (2024). The geography of transport systems. The Geography of Transport Systems, 1–402.

Salami, U. (2025). Blue Line Rail records over 2 million passengers – Govt - Punch Newspapers. https://punchng.com/blue-line-rail-records-over-2-million-passengers-govt/

Sasu, D. D. (2024, August 16). Largest cities in Nigeria in 2024 (in 1,000 individuals). Statista. https://www.statista.com/statistics/1121444/largest-cities-in-nigeria/

Semancik, A. (2024, April 3). How AI is transforming the creative economy and music industry. Ohio University. https://www.ohio.edu/news/2024/04/how-ai-transforming-creative-economy-music-industry

Shankar Iyer, L. (2021). AI enabled applications towards intelligent transportation. Transportation Engineering, 5, 1-11.

Simon, E. (2025). Lagos State's Smart Traffic Management Overhaul: A Comprehensive Look at the ITS Rollout – CLMI. https://clmi.ng/lagos-states-smart-traffic-management-overhaul-a-comprehensive-look-at-the-its-rollout/

Taiwo, F. (2022). The impact of artificial intelligence (AI) innovations on public service delivery in Nigeria.

Takyi, S. A. (2016). Comparative study of capital city elements: the case of Ghana and Nigeria. African Geographical Review, 35(2), 168–191. https://doi.org/10.1080/19376812.2015.1134335

Udoh, U.-A. A., Itaketo, U. T., Udofia, K. M., & Joe, O. I. (2025). Modeling and control of vehicular traffic systems using artificial intelligent network. International Journal of Real-Time Applications and Computing Systems (IJORTACS), 4(4), 840–852.

Wiese, T. (2024). Predictive Maintenance Using Artificial Intelligence in Critical Infrastructure: A Decision-Making Framework. International Journal of Engineering, Business and Management (IJEBM), 8(4), 1-4.

Xu, S., Sun, C., & Liu, N. (2024). Road congestion and air pollution -Analysis of spatial and temporal congestion effects. Science of the Total Environment, 945. https://doi.org/10.1016/j.scitotenv.2024.173896

Yazdani, M., Sarvi, M., Bagloee, S. A., Nassir, N., Price, J., & Parineh, H. (2023). Intelligent vehicle pedestrian light (IVPL): A deep reinforcement learning approach for traffic signal control. Transportation Research Part C: Emerging Technologies, 149, 103991.

Yu, K. H., Beam, A. L., & Kohane, I. S. (2018). Artificial intelligence in healthcare. Nature Biomedical Engineering, 2(10), 719–731. https://doi.org/10.1038/S41551-018-0305-Z,

Zhang, C., & Lu, Y. (2021). Study on artificial intelligence: The state of the art and future prospects. Journal of Industrial Information Integration, 23, 100224. https://doi.org/10.1016/J.JII.2021.100224

Zhu, L., Yu, F. R., Wang, Y., Ning, B., & Tang, T. (2019). Big Data Analytics in Intelligent Transportation Systems: A Survey. IEEE Transactions on Intelligent Transportation Systems, 20(1), 383–398. https://doi.org/10.1109/TITS.2018.2815678

CHAPTER 3
CYBERSECURITY PRACTICES IN SLOVAK SMALL AND MEDIUM-SIZED ENTERPRISES: ANALYTICAL STUDY WITH RECOMMENDATIONS

[1]Jolana GUBALOVA
[2]Robert HLAVAC

[1]Faculty of Economics, Matej Bel University, jolana.gubalova@umb.sk, ORCID ID: 0000-0003-4901-9207
[2]Faculty of Economics, Matej Bel University, robohlavac5@gmail.com

INTRODUCTION

In today's global business environment, small and medium-sized enterprises (SMEs) face increasing pressure to maintain their competitive advantage. This often requires continuous adaptation of business processes and a willingness to embrace change. In the digital age, where technologies play an ever more important role in daily operations, information systems have become a fundamental tool SMEs use to improve efficiency. Although these systems bring many benefits, their deployment also introduces new risks that must not be underestimated. One of these risks is cyber attacks, which can negatively disrupt the operation of an entire company. While large corporations were historically the primary targets of attacks, today virtually any business that uses the internet and digital technologies is at risk. SMEs are therefore becoming increasingly attractive targets for attackers, especially due to often weaker security and lower awareness of cybersecurity.

The chapter structure consists of two main subchapters that follow thematically. The first subchapter focuses on the theoretical foundations of cybersecurity, in which we define and explain basic concepts. In addition to clarifying the importance of cybersecurity, we will examine its main principles, relevant legal regulations in the field of information security applicable in the European Union, typical forms of cyber attacks, and possibilities for their mitigation through appropriate technical and organizational measures.

The second chapter presents the analytical part, which assesses the level of cybersecurity in the SME environment in Slovakia. This part is based on quantitative research conducted by means of a questionnaire survey, which provides insight into how SMEs respond to challenges associated with cyberspace, what strategies and measures they adopt to protect themselves against cyber threats, and what factors influence their ability to build effective security mechanisms. The analysis also identifies areas where SMEs experience shortcomings or constraints that hinder further development of their cyber resilience.

The main objective of the chapter is to, based on theoretical knowledge and the results of the questionnaire survey, provide an overview of the current state of security awareness in SMEs and propose a list of recommendations, rules and procedures to raise the level of cybersecurity in this segment.

1. THEORETICAL BACKGROUND

The prefix "cyber-" is today associated with almost every aspect of the digital world. Cyberspace is an area undergoing extremely rapid development and has become a familiar part of everyday life for individuals, organizations and society as a whole. However, it is increasingly discussed in connection with security threats that endanger information systems (Sedlak et al., 2021). The term "cyberspace", often used interchangeably with "cyber space", first appeared in 1984 when science fiction author William Gibson introduced and described it in his novel Neuromancer. In the work it is defined as "a consensual hallucination experienced daily by billions of legitimate users in every nation... an unimaginable complexity" (Gibson, 1984, p. 69). From this originally futuristic idea it follows that cyberspace is primarily a subjectively perceived world that each user experiences based on their own experience. At the same time, the work emphasizes its complexity as one of the main and defining features of cyberspace.

Given the extensive historical development of the concept, there is still no single globally accepted definition. In the academic literature it is often equated with the Internet or the virtual space. According to the U.S. National Institute of Standards and Technology (NIST), cyberspace is understood as a global digital environment that arises from the interconnection of information system infrastructures. The basis of this infrastructure includes elements such as the Internet, computers, communication networks, as well as embedded processors and control units (NIST, 2012).

Choucri (2013) expands this definition by describing cyberspace as a borderless space created by connecting computers via a global network such as the Internet. This network is built on a layered structure where physical elements such as computers, servers and networking devices enable logical connections. These in turn support the processing, modification and dissemination of information, as well as interaction between people and data, regardless of geographic location. The user plays a significant role in cyberspace. According to Sedlak et al. (2021), it is an artificial space created by humans for human purposes in which no fixed rules apply. Its form constantly changes according to how people use it.

If cyberspace were not actively used by users, it would lose its significance, cease to develop and gradually disappear. Ottis and Lorents (2010) emphasize the temporal dimension of cyberspace, describing it as an interconnected and constantly changing environment of information systems and users who actively use them. This interconnection between technology and people causes changes in cyberspace to occur over very short time periods and potentially affect a huge number of users. If one system is threatened, the problem can spread within seconds to other connected systems and negatively affect all users.

Based on multiple definitions, cyberspace can be understood as an electronic environment that enables mutual interconnection of people, devices and systems regardless of physical distance or time. It allows communication and real-time sharing of information from various corners of the world. Cyberspace includes the Internet and the World Wide Web, which are often perceived as its synonyms. It is important to emphasize that this space is intangible, existing in virtual form, yet it is very dynamic and at the same time stable. It is full of information and enables access to it by any user who has a connection.

Digital transformation has significantly affected SMEs, which were forced to move from traditional ways of doing business to digital approaches, especially during the COVID-19 pandemic. To maintain competitiveness, SMEs began to use modern technologies more intensively and move their operations into cyberspace. Many invested substantial financial resources to move closer to the Industry 4.0 concept, especially through technologies such as cloud computing, the Internet of Things and artificial intelligence. While this shift allowed the modernization of business processes, it also brought challenges around cybersecurity (Masood, Sonntag, 2020).

Cybersecurity

Cybersecurity is today one of the most important topics in the digital world. Its importance is constantly growing because, with the rapid development of technologies, it becomes increasingly necessary to protect sensitive data and systems from various cyber threats. Without adequate protection, not only individuals but also businesses may be at risk.

According to Sedlak et al. (2021, p. 15), the term cybersecurity can be defined as "a set of legal, organizational, technical and educational measures aimed at ensuring the protection of cyberspace." It is a comprehensive approach that requires coordinated use of various mechanisms and procedures that together contribute to effective protection of information systems. Measures are designed to protect systems and information from threats that may come from both external and internal sources. The aim is to minimize the risk of damage or misuse and to ensure reliable and secure operation of technological infrastructure.

Cybersecurity represents an ongoing process that requires regular monitoring, assessment of potential threats and updating of security measures. Reliable protection is based on a multi-layered approach that focuses on securing software, hardware, networks and the data that a user wants to protect. Equally important is the integrated cooperation between technologies, people and processes, which together form a single comprehensive security system capable of responding to various types of cyber threats (Abomhara, Koien, 2015).In the Slovak Republic, cybersecurity is regulated by Act No. 69/2018 Coll. on Cybersecurity. The Act defines cybersecurity as "a state in which networks and information systems are capable, to a certain degree of reliability, of withstanding any action that threatens the availability, authenticity, integrity or confidentiality of stored, transmitted or processed data or related services provided or accessible through those networks and information systems" (§ 3 of Act No. 69/2018 Coll.).

From the legislative definition it follows that cybersecurity not only covers the protection of information systems themselves but also focuses on protecting the data processed by these systems. Key principles of cybersecurity therefore include confidentiality, integrity and availability, which together form the basic model of information security commonly referred to in practice as the CIA triad.Confidentiality - protection of confidentiality means the ability to protect data from unauthorized access, leakage or misuse. Sensitive information should only be made available to persons who are authorized to access it or who have obtained consent from the data owner. A breach of confidentiality may result in financial loss, damage to an organization's reputation or even legal sanctions.

Integrity - focuses on ensuring the accuracy and completeness of data. The goal is to prevent unauthorized interventions that could modify, damage or delete data. Ensuring integrity guarantees that data remains trustworthy, accurate and reliable throughout its lifecycle. Availability - ensuring availability refers to making sure that data and systems are always ready for use and accessible to authorized users. This includes ensuring the continuous operation of key components such as servers, databases and applications so that data is available when the user needs it (Kolouch, Basta, 2019). An active approach to addressing cybersecurity issues should be a natural part of every organization's strategy and management, including SMEs. Investments in security should not be viewed merely as a cost but as an essential element of protecting business interests (Corallo, Lazoi, Lezzi, 2019).

Regulatory Requirements

Given the growing importance of cybersecurity, governments and regulatory bodies are increasingly involved in protecting information systems through the adoption of relevant legal regulations. These regulations play a decisive role in shaping a secure environment for SMEs. Their aim is to protect sensitive data, support uninterrupted business operations and maintain customer trust.

NIS 2

The NIS 2 Directive, which entered into force on 16 January 2023, represents a significant step in cybersecurity within the European Union. It replaces the original NIS Directive from 2016 and broadens its scope, especially towards a larger number of regulated entities and stricter requirements for their cyber resilience. Unlike the GDPR, which primarily focuses on personal data protection, NIS 2 targets organizations that are critical to supply chains and essential for societal and economic functioning. These include sectors such as healthcare, transport, energy, food supply, manufacturing and public administration. NIS 2 covers a wide range of measures in cybersecurity.

It requires organizations to adopt a comprehensive set of risk management measures and protect their network and information systems (European Parliament, Council of the EU, 2022). In the Slovak Republic, the directive was implemented through an amendment to Act No. 69/2018 Coll. on Cybersecurity, which took effect on 1 January 2025. Within this legal framework, a distinction is made between two types of regulated entities: operators of essential services (classified as important entities) and operators of critical essential services (classified as key entities).

In accordance with NIS 2, the amendment also introduces a rule according to which the requirements mainly apply to medium-sized and large enterprises. Micro and small enterprises are generally exempt unless they provide services that are essential to the functioning of society or the state (National Security Authority, 2024). One of the main requirements of the directive is the obligation to report a cyber incident to the relevant authorities within 24 hours of its detection. Such a rapid notification helps limit damage and improve protection of critical systems (European Parliament, Council of the EU, 2022). If an organization fails to comply with statutory obligations, it may face severe financial penalties:

- key entities: up to EUR 10 million or 2% of their global annual turnover,
- important entities: up to EUR 7 million or 1.4% of their global annual turnover whichever amount is higher.

The NIS 2 Directive also supports active engagement of SMEs in information-sharing networks that exchange knowledge about current threats and best security practices between businesses and state authorities (European Parliament, Council of the EU, 2022). Through this cooperation, SMEs can strengthen their cyber defense capabilities and contribute to a more resilient digital ecosystem.

Cyber Attacks

A cyber attack can be defined as "an intentional act by an attacker in cyberspace that uses information and communication technologies to attack another information and communication infrastructure, whether with the aim of disrupting availability, confidentiality or integrity of data" (Kolouch, Basta, 2019, p. 83).

Closely related to this concept is a cyber threat, understood as any event or circumstance that may negatively affect the organization's operations, mission, functions, image, reputation, assets or individuals through the misuse of information systems (NIST, 2012). According to the Global Risks Report published by the World Economic Forum, cybercrime and low levels of cybersecurity rank among the eight most serious global risks expected over the next 2 to 10 years (World Economic Forum, 2023).

From the attackers' perspective, SMEs are often seen as easier and more accessible targets, primarily due to their smaller size. This assumption stems from the fact that smaller businesses generally have weaker security measures and lack sufficient financial resources to deal with the consequences of cyber attacks. These factors increase their vulnerability and make them attractive targets. Attackers are also motivated by the possibility of obtaining financial gain by attacking a larger number of smaller entities at once, while the risk of detection and media attention is significantly lower compared to attacks on large corporations. The response by state and security authorities to incidents in the SME sector is often less intense, which further increases their attractiveness to cybercriminals (Rahmonbek, 2023).

2. RESEARCH AND METHODOLOGY

At the outset of the study we defined the main research objective, which we then divided into three sub-objectives. Main objective: To propose a list of recommendations, rules and procedures to increase the level of cybersecurity in the SME environment.

Sub-objectives:

C1: To examine existing measures and tools for protecting information systems and data against cyber threats.

C2: To carry out a questionnaire survey among SMEs to determine their current situation and level of cybersecurity awareness.

C3: Based on the obtained findings, to develop a list of recommendations and proposals for improving security practices in SMEs.

The theoretical part of the thesis provided expert information related to sub-objective C1, which then served as the basis for processing the practical part of the thesis. Sub-objective C2 is based on quantitative research conducted in the form of a questionnaire survey.

The population consisted of all SMEs in the Slovak Republic. The sample consisted of the individual companies that participated in the survey. The sample was selected by random sampling. Respondents were managerial staff and employees responsible for IT security who have an overview of security processes in their companies. Data collection took place during April and May 2025 via an online questionnaire on the Google Forms platform, and distribution was carried out using contact data of companies obtained from the Finstat.sk database. The questionnaire contained a total of 16 questions and focused on specific aspects of cybersecurity. Emphasis was placed on preserving respondent anonymity during data collection. No personal data that could identify respondents were collected, and respondents were informed in advance that the survey is anonymous and serves exclusively academic purposes. Due to the sensitivity of the topic, 52 respondents participated out of a total of 160 companies contacted. The collected data were processed using MS Excel. Results are presented in the form of tables and charts that visualize the most commonly implemented security measures as well as areas where SMEs show potential shortcomings.

3. RESEARCH RESULTS

The initial questions of the questionnaire focused on basic identification of the respondents. The size of the company in which the respondent works and the sector of activity were determined. The following part of the questionnaire focused on a deeper analysis of cybersecurity practices — we investigated whether SMEs have experience with cyber attacks, what security measures they use, what barriers they perceive in implementing them, and how prepared they are in case an incident occurs.

Regarding the size categories of companies participating in the survey, small enterprises predominated. In these group 24 respondents replied. Medium-sized enterprises were represented by 18 respondents and micro-enterprises by 10 respondents.

Respondents were also categorized by the sector in which their company operates. The aim was to obtain responses from companies operating in different sectors to better compare differences in approaches to cybersecurity. Most responses came from companies in trade, industry and business services. Other sectors represented included transport, construction and agriculture. This distribution was intended to ensure that the survey results are balanced and do not reflect only the experiences and opinions of SMEs from a single sector.

The survey showed that 42 companies, representing 80.8% of all respondents, use cyberspace on a daily basis. Another 13.5% of respondents reported using it several times a week and 5.8% answered that they use it several times a month. None of the surveyed companies selected the response "only occasionally" or "never," which confirms that digital technologies and online tools are now a natural part of everyday operations in SMEs.
Table 1 provides an overview of how SMEs staff the area of cybersecurity.

Table 1. Responsible person for cybersecurity in the company(Own processing)

IT department	Micro Ent.	Small Ent.	SMEs	Total
Yes	20.0 %	79.2 %	100.0 %	75.0 %
No	80.0 %	20.8 %	-	25.0 %
Absolute number of responses	10	24	18	52

The questionnaire results indicate a strong correlation between the size of the enterprise and whether there is a dedicated IT department or specific employee dealing with cybersecurity. In the category of medium-sized enterprises, all 18 respondents (100%) stated that their company has a dedicated person or team for this area. Similarly, in small enterprises, 79.2% indicated that their company has a person responsible for IT security. The most pronounced difference was in micro-enterprises, where 80% of respondents answered that this area is not covered personnel-wise, while only 20% employ such a person. Overall, cybersecurity is actively handled by 75% of all surveyed companies, which in absolute terms amounts to 39 out of the total 52 respondents.

Based on the obtained data, it can be concluded that the smallest companies often underestimate the personnel coverage of cybersecurity. Conversely, larger companies approach it more responsibly and are better prepared for the risks associated with using cyberspace.

Table 2 shows results concerning SMEs' experience with cyber attacks during 2023–2025. This time interval was chosen intentionally to take into account the consequences of the COVID-19 pandemic during which many work activities moved online.

Table 2. Experience with cyber attacks over the past three years (Own processing)

Experience with an attack	Micro Ent.	Small Ent.	SMEs	Total
Yes, once	30.0 %	50.0 %	5.6 %	30.8 %
Yes, several times	10.0 %	25.0 %	66.7 %	36.5 %
No, we have not experienced any attacks	60.0 %	25.0 %	27.8 %	32.7 %
Absolute number of responses	10	24	18	52

Based on the data in Table 2, medium-sized enterprises were most frequently the target of repeated cyber attacks during this period. As many as 12 out of 18 respondents, representing 66.7% of this category, reported having been attacked multiple times. A small percentage reported having been attacked only once or not at all. In the case of small enterprises, 12 out of 24 companies (50%) reported experiencing a cyber attack once. The responses "multiple times" and "we did not record any attack" were more evenly split at 25% each. Micro-enterprises showed the most favorable results: 60% of respondents in this category stated that they had not experienced a cyber attack during the monitored period, 30% experienced an attack once and only 10% multiple times. From these data it can be inferred that medium-sized enterprises tend to be subject to repeated attacks more often than micro and small enterprises. A possible reason may be the larger volume of processed data and more extensive technologies.

Out of all 52 respondents, 35 companies reported having experienced at least one cyber attack during 2023–2025. Since SMEs may have faced multiple forms of cyber attacks, respondents were allowed to choose more than one answer.

Among the 35 companies that confirmed an attack, a total of 64 responses were recorded. The most frequent form was phishing, in the form of emails or SMS messages. This type of attack was reported by 31 out of 35 companies, which represents 88.6% of those who reported an attack. Unauthorized access to corporate networks was second with 12 responses. Technical attacks such as malware were recorded in eight responses, while DDoS attacks were reported by seven respondents. Insider threats caused by employees were mentioned in five cases. These data confirm that social engineering represents the greatest risk for SMEs, although technical attacks are not uncommon. Table 3 provides a more detailed look at the types of cyber attacks recorded by each SME size category.

Table 3. Type of attack by company size (Own processing)

Attack type	Micro Ent.	Small Ent.	SMEs	Total
Phishing (fraudulent emails, SMS)	4	14	13	31
Unauthorized access to corporate networks	-	9	3	12
Malware (virus, worm, trojan,ransomware)	-	4	4	8
DDoS (website / service disruption)	-	4	3	7
Attack caused by own employee	-	3	2	5
Other	-	-	1	1
Absolute number of responses	4	34	26	64
Number of respondents	4	18	13	35

Because companies could report multiple types of impacts, respondents were also allowed to provide multiple answers for the consequences question. A total of 58 responses were recorded. The most common consequence of a cyber attack was temporary disruption of operations, indicated by 19 respondents, representing 54.3% of those who experienced an attack. Loss or leakage of sensitive data was reported by 18 companies, while financial losses were reported by 11 respondents.

Damage to reputation was recorded in three cases. In seven cases companies stated that the attack had no impact, which suggests that it either did not cause serious consequences or they were able to quickly resolve it. The results show that the most common risks associated with cyber attacks are operational interruptions and data loss, which can significantly affect business operations. Since SMEs also play an important role in the national economy, such consequences can negatively affect the entire economy. The findings also confirm that it is important to have an incident response plan to help SMEs know what to do before, during and after an attack. Table 4 shows the results for the question of whether SMEs have an incident response plan.

Table 4. Incident response plan (Own processing)

Processing level	Micro Ent.	Small Ent.	SMEs	Total
Yes, detailed and regularly updated	10.0 %	50.0 %	44.4 %	40.4 %
No	70.0 %	20.8 %	22.2 %	30.8 %
No, but we plan to implement it	20.0 %	29.2 %	33.3 %	28.9 %
Absolute number of responses	10	24	18	52

The results indicate that micro-enterprises show the lowest level of preparedness. Out of ten respondents in this category, only 10% stated they have a detailed and regularly updated plan. As many as 70% do not have such a plan at all and the remaining 20% only plan to implement it. The situation in small enterprises is significantly more favorable: of 24 responses, half (50%) reported having a plan and updating it regularly. Approximately 21% do not have such a plan at all and the remaining 29.2% plan to implement it. Medium-sized enterprises showed a very similar level of readiness to small enterprises. A detailed and updated plan is present in 44.4% of respondents, 33.3% plan to implement it and 22.2% do not have such a plan at all. Overall, 21 out of 52 companies, or 40.4% of all respondents, have an incident response plan. Although a significant portion of companies already actively address their cyber readiness, there remains a notable share of entities that might not know how to respond properly in the event of a cyber incident.

Table 5 provides deeper analysis of how experience with an attack affects preparedness and whether companies have an incident response plan.

Table 5. Impact of attack experience on incident response plan(Own processing)

Response plan	Yes, once	Yes, many times	Unrecorded attack	Total
Yes, detailed and regularly updated	37.5 %	68.4 %	11.8 %	40.4 %
No	25.0 %	10.5 %	58.8 %	30.8 %
No, but we plan to implement it	37.5 %	21.1 %	29.4 %	28.9 %
Absolute number of responses	16	19	17	52

The results show that SMEs that have already experienced at least one attack approach cybersecurity more responsibly. As many as 68.4% of companies that reported repeated attacks have a detailed and regularly updated incident response plan, while only 10.5% still do not have such a plan. In the group of companies that were attacked only once, 37.5% have a plan, while 25% do not. Significant differences were found among companies that have not experienced an attack; only 11.8% of this group have a plan and 58.8% stated that they do not have one. These findings indicate that SMEs often begin to recognize the need for an incident response plan only after experiencing a negative event themselves.

The research showed that SMEs most frequently rely on basic technical solutions such as antivirus software, used by 51 out of 52 companies, representing 98.1% of all respondents. The second most used measure is a firewall, indicated by 35 respondents. Data backup, which is important particularly for system recovery after attacks or failures, is used by 26 companies. Multi-factor authentication was reported by 20 respondents. Conversely, employee training in cybersecurity is offered by significantly fewer companies; only 13 respondents reported this measure, while only one company admitted to using no security measures at all. From the questionnaire results it can be confirmed that while SMEs try to protect their systems, they often focus only on the technical side of protection. Less emphasis is placed on building cybersecurity awareness and employee skills, which can be problematic since many attacks exploit user ignorance or mistakes.

Table 6 shows the main barriers SMEs face when implementing security measures. Respondents were able to select the single option that best reflected their company's situation.

Table 6. Constraints in implementing security measures (Own processing)

Type of constraint	Micro	Small Ent.	SMEs	Total
Lack of funds	60.0 %	70.8 %	61.1 %	65.4 %
Lack of expertise	30.0 %	29.2 %	22.2 %	26.9 %
Disinterest, we consider it an insignificant	10.0 %	-	5.6 %	3.9 %
Other	-	-	11.1 %	3.9 %
Absolute number of responses	10	24	18	52

The results indicate that the most significant barrier is lack of financial resources. This problem was indicated by the majority of respondents across all company categories, representing 65.4% of the total. The second most frequently cited barrier was lack of expertise, noted by an average of 26.9% of companies. This suggests that many SMEs lack internal know-how in cybersecurity. A positive finding is that only 3.9% of companies consider security measures unimportant, confirming that most respondents recognize their importance and necessity.

Data show that SMEs would most frequently invest in services of external experts, which 55.8% of respondents indicated. This preference corresponds with previous findings where many companies see the lack of internal expertise as one of the main obstacles to implementing security measures. The second most frequently mentioned priority was investment in modern antivirus programs, chosen by 21.2% of companies. Employee training was prioritized by 9.6% of companies, while the use of artificial intelligence for threat detection was selected by 8% of respondents. The remaining 5.8% of respondents indicated other forms of protection, including advanced solutions such as EDR, IPS, SIEM or SOC, which suggests a growing interest among some companies in more comprehensive security technologies. Table 7 shows estimated financial losses that SMEs would incur in the event of a one-day operational outage caused by a cyber attack.

Table 7. Financial impact of an attack with one-day outage (Own processing)

Financial intervals	Micro Ent.	Small Ent.	SMEs	Total
Less than 1 000 €	80.0 %	54.2 %	11.1%	44.2 %
1 000 – 5 000 €	10.0 %	37.5 %	33.3 %	30.8 %
5 000 – 10 000 €	10.0 %	4.2 %	5.6 %	5.8 %
Greater than 10 000 €	-	4.2 %	50.0 %	19.2 %
Absolute number of responses	10	24	18	52

In the micro-enterprise category, 80% of respondents stated that such an outage would cost them less than €1,000. A similar trend was seen among small enterprises, where more than half (54.2%) chose the same interval, while another 37.5% estimated losses in the €1,000–€5,000 bracket. The situation is markedly different for medium-sized enterprises: half of medium-sized companies (50%) stated that a one-day outage would cost them more than €10,000. These results clearly confirm that as company size increases, the scale of potential financial losses due to downtime also increases, even for an outage lasting only one day. Table 8 shows how frequently SMEs update their software and systems.

Table 8. Software and system updates (Own processing)

Update frekvency	Micro Ent.	Small Ent.	SMEs	Total
Immediately after update release	70.0 %	79.2 %	66.6 %	73.1 %
Regularly (monthly/yearly)	20.0 %	20.8 %	22.2 %	21.2 %
Only when necessary	10.0 %	-	5.6 %	3.9 %
Other	-	-	5.6 %	1.9 %
Absolute number of responses	10	24	18	52

Most SMEs approach updates responsibly. As many as 73.1% of respondents (38 out of 52 companies) reported updating their systems immediately after a new version is released. This approach is most common among small enterprises (79.2%), followed by micro-enterprises (70.0%) and medium-sized enterprises (66.6%). Regular updates (e.g., monthly or yearly) are performed by 21.2% of companies.

Although this is not a neglect of duty, such an approach may lead to delayed responses to new forms of attack. Only a small share of companies (3.9%) admitted that they update only when necessary, which presents a potential security risk. Table 9 shows analysis of whether the presence of an IT department affects the approach to software and system updates.

Table 9. Impact of having an IT department on updates (Own processing)

Update frekvency	Yes	No	Total
Immediately after update release	76.9 %	61.5 %	73.1 %
Regularly (monthly/yearly)	17.9 %	30.8 %	21.2 %
Only when necessary	2.6 %	7.7 %	3.9 %
Other	2.6 %	-	1.9 %
Absolute number of responses	39	13	52

The results confirm that SMEs with their own IT department react to updates more promptly — 76.9% of them update their systems and software immediately after a release. By contrast, among companies without an IT department this share is lower at 61.5%. Based on these findings, it can be concluded that the presence of an IT department has a positive effect on the efficiency and regularity of updates. Table 10 shows the intervals in which SMEs back up their data, aiming to determine the extent to which companies have assured data recovery in case of an incident.

Table 10. Data backup (Own processing)

Update frekvency	Micro Ent.	Small Ent.	SMEs	Total
Daily	10.0 %	79.2 %	83.3 %	67.3 %
Weekly	20.0 %	4.2 %	16.7 %	11.5 %
Monthly	-	8.3 %	-	3.9 %
Never	70.0 %	8.3 %	-	17.3 %
Absolute number of responses	10	24	18	52

The results show significant differences between categories regarding data backup. Medium-sized enterprises take this measure most responsibly, with 83.3% performing daily backups. A similarly high proportion was recorded among small enterprises (79.2%). Micro-enterprises differ starkly: only 10% perform daily backups, while 70% stated they never back up data, representing a serious security risk. In case of technical failure or cyber attack, irreversible loss of important data may occur without the possibility of recovery. Overall, 35 out of 52 respondents (67.3%) back up their data daily, reflecting a responsible approach by most small and medium enterprises. Conversely, the results indicate a much weaker level of preparedness among micro-enterprises, which largely neglect this important security measure. Table 11 shows responses to the question of whether employees with access to IT systems receive cybersecurity training.

Table 11. Employee training (Own processing)

Access to course	Micro Ent.	Small Ent.	SMEs	Total
Yes, at least once a year	10.0 %	25.0 %	33.3 %	25.0 %
Yes, only when the need arises	-	54.2 %	44.4 %	40.4 %
No, we do not provide this type of training	90.0 %	20.8 %	22.2 %	34.6 %
Absolute number of responses	10	24	18	52

The survey revealed that only 25% of respondents provide regular training at least once a year. The most critical situation was observed among micro-enterprises, where 90% of respondents stated they provide no training at all, and only 10% conduct training once a year. Small and medium enterprises fared somewhat better: 54.2% of small enterprises provide training only when needed, and 44.4% of medium enterprises said the same. Despite slightly better results, this still indicates a lack of systematic employee education, which is often carried out reactively after an incident rather than proactively. These findings confirm a significant gap in employee education and that many SMEs still underestimate the importance of prevention. Insufficient employee awareness significantly increases the risk of human-factor failures.

Table 12 presents responses on the extent to which SMEs use or consider cyber insurance as a form of protection against the consequences of cyber incidents.

Table 12. Cyber insurance (Own processing)

Respondent's attitude	Micro Ent.	Small Ent.	SMEs	Total
Yes	10.0 %	37.5 %	50.0 %	36.5 %
No, but we are considering it	40.0 %	16.7 %	27.8 %	25.0 %
No, we were not aware of the possibility of such insurance	50.0 %	45.8 %	22.2 %	38.5%
Absolute number of responses	10	24	18	52

The survey results showed that cyber insurance is currently held by only 36.5% of respondents (19 out of 52 companies). The highest rate of insurance was recorded among medium-sized enterprises, where half of respondents use this protection. Among small enterprises it is 37.5% and among micro-enterprises only 10%, indicating significant differences in approach depending on company size. A grave finding is that 38.5% of companies were not aware of the possibility of cyber insurance. This lack of awareness is most pronounced among micro-enterprises (50%).

4. RECOMMENDATION

Based on the questionnaire results, it was confirmed that SMEs in Slovakia face several challenges in cybersecurity. Their ability to protect digital assets closely relates to their size, available resources, level of expertise and perception of cyber threats. Company size emerged as an important factor in implementing security measures. While small and medium enterprises have fairly advanced measures in place, micro-enterprises often neglect even basic security standards. Therefore, we decided to formulate recommendations separately for each category, aiming to provide concrete and practically usable proposals for improving security procedures.

For micro-enterprises, it was confirmed that they are the group with the lowest level of cyber preparedness. Most of these entities do not have a designated person responsible for IT security, often lack an incident response plan and a large share do not perform data backups at all.

In this category it is therefore necessary to focus on implementing a minimum set of measures that do not require high financial or technical costs but can significantly raise the basic level of security. As a first step we recommend developing a simple and clear incident response plan that will serve as a practical guide in the event of a cyber attack, technical failure or data loss. The plan should be understandable for every employee and describe how to proceed when a cyber problem occurs. In the final phase it is necessary to analyse the cause of the incident and, based on the findings, adopt appropriate measures to prevent its recurrence. Another recommended measure is implementing regular data backups. Ideally, backups should be performed automatically without employee intervention, minimizing the risk of human error. In this context, micro-enterprises can use affordable solutions such as the Backup Solution service provided by Slovak Telekom in cooperation with the technological partner Xopero. The service allows automatic encrypted backup to the cloud without the need to invest in expensive hardware solutions. Data are additionally protected by strong encryption and are accessible from any authorized device.

Small enterprises showed a higher degree of security measure implementation, but the survey highlighted that many measures are reactive, meaning they are implemented only after an incident occurs. System updates and data backups are not always set to run automatically and are not always regularly tested. Employee training is often conducted only when needed. The survey showed that small enterprises often lack systematization and prevention in cybersecurity. Therefore, we recommend introducing regular employee training at least once a year. Training should not focus solely on technical threats but also on building employee security awareness. Very effective are practical phishing exercises in which employees receive fake emails that mimic real phishing messages. The exercise monitors how employees respond — whether they report them as suspicious, ignore them, or click on them. Such exercises help reveal weak points in employee behaviour.

Employees should also understand the basics of safe online behaviour. They should be able to recognize suspicious websites, links or pop-ups and be aware of the risks associated with downloading files or programs from untrusted sources.

It is important to provide practical advice they can easily apply in daily work. The goal is to enable employees to behave responsibly and contribute to overall company security. Small enterprises can also use free online tools, such as the Kyberarena platform operated by the Government CSIRT unit, which offers various simulated tests and exercises focused on cybersecurity and can help employees improve their skills and preparedness for potential threats. It is also important to create and implement an internal security policy that clearly defines rules for system access, password management, use of company devices and how employees should report security incidents. Attention should be paid to the management of credentials, which are among the common targets of cyber-attacks. Weak passwords, password reuse and insufficient control over access increase the risk of compromise. As a practical, free and easily accessible solution we recommend using KeePassXC for secure password management. Passwords, access credentials and notes are stored in an encrypted database on the user's local device protected by a master password known only to the user. The database can be securely shared with authorized colleagues while remaining encrypted and protected against unauthorized access. Key advantages include the ability to generate strong passwords and auto-fill credentials in web browsers. The tool does not require an internet connection, which reduces the risk of external compromise. This tool is recommended especially for micro and small enterprises that lack an IT department. It is a simple but effective step toward improving cyber hygiene and resilience to cyber threats.

Medium-sized enterprises already have several important security measures in place and are overall the best prepared among all categories. Because they have already faced various types of attacks, they should further develop their security strategy and focus on more modern and advanced technologies. We recommend cooperation with external specialists who can help implement continuous system monitoring, timely detection of threats and regular security assessments. Across all categories, SMEs should work on building a security culture in which management leads by example, employees are motivated, and cybersecurity is perceived as a natural part of everyday operations. Employers should support open and trusting communication and be helpful to employees in this regard.

In cases of human error, it is important to provide support rather than punishment. Employees should feel safe and confident to express when they do not understand something or have doubts. Such an approach builds trust within the company and creates space for open discussion about cybersecurity. The common goal of all measures should be building a resilient environment capable of effectively responding to challenges from cyberspace. It is important to emphasize that cyber risk management is not a one-off process but a continuous activity that requires a systematic approach. SMEs should continually analyze historical data, regularly reassess the current state, monitor new threats and trends and learn from incidents experienced by other companies in order to create a comprehensive overview of possible risks and continuously increase their level of cyber readiness.

CONCLUSION

Cybersecurity has become increasingly topical in recent years, especially given the growing use of modern technologies and online tools across all areas of business. The main goal of the research was to propose a list of recommendations, rules and procedures to increase the level of cybersecurity in SMEs, formulated on the basis of the questionnaire analysis. Through a combination of theoretical analysis and empirical research, we managed to create a more comprehensive view of the current situation of cybersecurity in the Slovak business environment.

The questionnaire results pointed to growing awareness of the importance of protecting data and systems. At the same time, it was confirmed that a systematic approach to security remains insufficient in many cases. The level of preparedness varies by company size, technical capabilities and available resources. Most SMEs in Slovakia rely on basic technical measures but lack employee security awareness. In many cases formal security policies are absent, regular training does not take place, and a clearly defined incident response plan is missing. Nevertheless, there is tangible potential for improvement, especially through simple, financially accessible measures and incorporating security awareness into corporate culture.

The recommendations in the thesis are based on the survey findings and were created as practical outputs grounded in real SME environments.

They target not only technical solutions but broader contexts. The proposed framework is therefore not just a set of technologies; it emphasizes linking technical tools, organizational processes and the human factor. It is important not only to implement protective elements but especially to maintain them in the long term, understand risks and have active leadership. This creates a safer digital environment for the enterprise, its customers and other stakeholders. The future of cybersecurity in SMEs will require a systematic approach. Threats will evolve, and it will be necessary to continuously update security measures. Key will be the integration of technical solutions with organizational procedures and increasing employee awareness. Security must become part of corporate culture and daily practice.

REFERENCES

Abomhara, M., Koıen, G. M. (2015). Cyber security and the internet of things: vulnerabilities, threats, intruders and attacks. In Journal of Cyber Security and Mobility, volume 4, n. 1, pp. 65–88. 2015. ISSN 2245-1439.

Brezınova, M. (2021). Basic Characteristics of Small and Medium-sized Enterprises in Terms of Their Strategic Management. In International journal of systems applications, engineering & development, volume 15, n. 15, pp. 84-87. ISSN 2074-1308.

Choucrı, N. (2013). Co-Evolution of Cyberspace and International Relations: New Challenges for the Social Sciences [online]. In Political Science Department Research Paper, volume 29. [cit. 2025-05-11]. Retrieved from: http://dx.doi.org/10.2139/ssrn.2514532.

Cısa. (2024). Understanding and Responding to Distributed Denial of Service Attacks [online]. [cit. 2025-01-25]. Retrieved from: https://www.cisa.gov/sites/default/files/2024-03/understanding-and-responding-to- distributed-denial-of-service-attacks_508c.pdf.

Corallo, A., Lazoı, M., Lezzı, L. (2019). Cybersecurity in the context of industry 4.0: A structured classification of critical assets and business impacts [online]. In Computers in Industry, volume 114. [cit. 2025-01-15]. Retrieved from: https://doi.org/10.1016/j.compind.2019.103165.

Csırt. (2024). Social engineering [online]. [cit. 2024-12-23]. Retrieved from: https://www.csirt.gov.sk/socialne-inzinierstvo.html.

DataGuard. (2022). GDPR for Small Businesses: Your All-in-One GDPR Guide [online]. [cit. 2025-05-16]. Retrieved from: https://www.dataguard.com/blog/gdpr-for- small-businesses.

Deshpande, S. (2023). Cyber insurance for small and medium businesses [online]. [cit. 2025-04-13]. Retrieved from: https://www.tcs.com/insights/blogs/cyber- insurance-small-medium-businesses.

Enısa. (2021). Cybersecurity for SMEs: Challenges and Recommendations [online]. [cit. 2025-01-26]. Retrieved from:

European Commission. (2014). EU Commission Regulation n. 651/2014 [online]. [cit. 2025-03-12]. Retrieved from: https://eur-lex.europa.eu/legal- content/SK/TXT/PDF/?uri=CELEX:32014R0651.

European Parliament, Council of the European Union. (2016). Regulation of the European Parliament and of the Council (EU) 2016/679 (GDPR) [online]. [cit. 2025-03-12] Retrieved from: https://eur-lex.europa.eu/legal-content/sk/TXT/?uri=CELEX%3A32016R0679.

European Parliament, Council of the European Union. (2022). Directive of the European Parliament and of the Council (EU) 2022/2555 (directive NIS 2) [online]. [cit. 2025-02-22]. Retrieved from: https://eur-lex.europa.eu/legal-content/SK/TXT/?uri=OJ:L:2022:333:TOC.

Gıbson, W. (1984). Neuromancer. New York: Ace Science Fiction. 279 s. ISBN 978-0- 441-56959-5.

IBM. (2024). What is incident response? [online]. [cit. 2025-04-13]. Retrieved from:https://www.ibm.com/think/topics/incident-response.

Kolouch, J., Basta, P. (2019). Cybersecurity. Praha: CZ NIC. 560 s. ISBN 978-80- 88168-34-8.

Martıns, A. (2023). Cybersecurity: A Small Business Guide [online]. [cit. 2025-04-22]. Retrieved from: https://www.businessnewsdaily.com/8231-small-business- cybersecurity-guide.html.

Masood, T., Sonntag, P. (2020). Industry 4.0: Adoption challenges and benefits for SMEs [online]. In Computers in Industry, volume 121. [cit. 2025-02-22]. Retrieved from: https://doi.org/10.1016/j.compind.2020.103261.

NBU. 2024. What did NIS2 bring? [online]. [cit. 2025-05-27]. Retrieved from: https://nis2.nbu.gov.sk/co-priniesla-nis2/.

Nextech. (2023). Cybersecurity for businesses 2023. Bratislava: Digital Visions, spol. s r.o., 72 s. ISBN 978-80-974206-6-6.

NIST. (2012). Guide for Conducting Risk Assessments [online]. [cit. 2025-05-18].

Ottıs, R., Lorents, P. (2010). Cyberspace: Definition and Implications. In Proceedings of the 5th International Conference on Information Warfare and Security: April 08-09,2010. Dayton: OH, US, s. 267–270. ISBN 97-1-906638-60-3.

Rahmonbek, K. (2023). Alarming Small Business Cybersecurity Statistics for 2023 [online]. [cit. 2025-02-22].

Sedlak, P. A Kol. (2021). Cyber (in)security Issues of security in cyberspace. Brno: Akademicke nakladatelstvi CERM. 440 p. ISBN 978-80-7623-068-2.

Segal, E. (2022). Small Businesses Are More Frequent Targets Of Cyberattacks Than Larger Companies: New Report. Forbes [online]. [cit. 2025-01-19].

Slov-Lex. (2025). Act No. 69/2018 Coll. on Cybersecurity [online]. [cit. 2025-03-13]. Retrieved from: https://www.slov-lex.sk/ezbierky/pravne-predpisy/SK/ZZ/2018/69/.

Sulc, V. (2018). Cybersecurity. Plzen: Aleš Čeněk. 147 s. ISBN: 978-80-7380-737-5.

ISBN: 978-625-92720-0-9