Grenze International Journal of Engineering and Technology, June Issue

Intelligent AI-based Framework for Comprehensive Risk Identification and Management in Construction Projects

Kuldeep Vayadande[1], Yogesh Bodhe[2], Vaibhav B. Bhagat[3], Vaishnavi S. Pawar[4], Purva S. Waykole[5] and Pratik G. Awate[6]
[1,3-6]Department of Information Technology, Vishwakarma Institute of Technology, Pune - 411037, Maharashtra, India
Email: kuldeep.vayadande1@vit.edu, vaibhav.1252130039@vit.edu, vaishnavi.1252130019@vit.edu, purva.1252130021@vit.edu, pratik.1252130020@vit.edu
[2]Government Polytechnic, Pune
Email: bodheyog@gmail.com

***Abstract*— This research proposes an AI-based architecture to detect risks and manage feasibility of construction or infrastructure project designs. It consists of three main components: (i) PDF interpretation for risk parameters with CLAHE domain-specific image preprocessing technique and OCR-based risk detection model, (ii) hybrid interpretation with two neural networks: embedding layer with pre-trained BERT language model and classification layer with LightGBM model, and (iii) feasibility analysis model for building floor areas estimation and building type selection for soil and environmental conditions. Four innovations were implemented: domain-specific CLAHE preprocessing, dual-module risk-feasibility ensemble with probabilistic conflict handling, optimized FP16 embeddings on CUDA, and UniStrided-1 processing. Experiments with 12,500 engineering reports show 88.9% risk detection accuracy, 75.5% MOTA for risk tracking, and 86.8% mIoU for feasibility segmentation. The framework analyzes 19.5 reports per minute on a Tesla T4 GPU and reduces manual analysis time by 68%.**

***Index Terms*— Construction Feasibility Analysis; Decision Intelligence; Engineering Risk Analysis; Automated PDF Parsing; Rule-Based Reasoning; Hybrid AI; Infrastructure Safety; CLAHE Preprocessing.**

I. INTRODUCTION

It goes without saying that any construction or infrastructure project entails a complicated process involving geological, engineering, regulatory, and environmental hazards management [1][7][2]. The more elaborate an urban development project is, including zoning plans and high-density construction together with its associated infrastructure, the greater the need arises for intelligent software to analyze heterogeneous engineering documents and provide instant feedback about risks and project feasibility. Despite the relative lack of explainability of geological survey tools, the use of these tools may face limitations in terms of flexibility with regard to analyzing engineering reports with poor structure, different terminologies, and technical jargon. More advanced solutions have emerged, utilizing natural language processing methods such as keyword recognition, dependency parsing, and transformers (BERT) [5].

Grenze ID: 01.GIJET.12.2.344_26

However, the performance of these methods remains vulnerable to noise, OCR errors, mixed languages, and computationally demanding inferences on edge devices.

Mathematically speaking, the main task this paper aims at solving can be formulated as follows. Given a set D = {d1, d2, ..., dN} of engineering reports which consist of soil investigation reports, feasibility studies PDFs, and technical specifications, the goal is to develop a generalized intelligent flow $F : D \rightarrow (R, Fmax, B)$ for assessing geological risk assessments R, the maximum floor limit Fmax, and types of buildings B. The constraints are imposed in terms of computational efficiency (latency, memory usage), reliability (accuracy, precision) and transparency (explanatory power).

In this paper, we propose a hybrid AI-powered system aimed at resolving these problems by integrating the following innovative approaches: domain-specific CLAHE image preprocessing, novel ensemble risk-feasibility reasoning model with probabilistic conflict reduction, FP16 embeddings optimization, and UniStrided-1 sequential reading.

II. RELATED WORK

A. Document Parsing Methods

Rule-based expert systems have traditionally been applied in construction feasibility studies, relying on a checklist-based approach that is easily interpretable but relatively rigid in its handling of various document types [10]. Modern approaches leverage natural language processing technologies like keyword extraction, dependency parsing. However, transformer-based architectures like BERT [5] and LSTM [27] have performed excellently on structured data but still face challenges in dealing with the specialized technical vocabulary and ambiguous statements found in construction documentation.

Research Gap: The effectiveness of parameter extraction deteriorates when working with heterogeneous or multilingual documentation. There is no systematic application of pre-processing steps within the transformer-based parsing pipeline.

B. Risk Prediction and Assessment

Risk estimation in the early stages was based on statistical methods like regression and the Analytical Hierarchy Process (AHP) [1]. Modern machine learning models like SVM and random forest classifier have proven to perform better on project-related structured data [11]. Natural Language Processing (NLP)-based models have more recently been applied to derive underlying risk factors using information from accident report texts [10], and CBR systems have accessed historical records to suggest effective strategies for mitigating these risks [6], [7]. Despite these developments, however, all these techniques are considered independently without the possibility of cross-modal synergy.

Research Gap: There is no model currently existing which considers all three features—structural validity, dynamic anomaly detection, and an adaptive system based on cross-modal causation in structured and unstructured project data.

C. Feasibility Analysis and Bim-Integrated Systems

Rule-based expert systems have received extensive attention in relation to the feasibility assessment in terms of soil bearing capacity (SBC), water table depth (WTD), and load factors [15][21]. The use of BIM combined approaches merges design-related information and risk inference [15],[21]. However, such systems tend to produce highly unstable results in relation to feasibility due to fluctuating predicted maximum floor values with small changes in parameters.

Research gap: Predicted feasibility results are not temporally stable or probabilistically suppress conflicting geotechnical conditions.

D. Preprocessing Techniques for Document Enhancement

Various contrast enhancement methods have been widely explored in the literature for computer vision and text preprocessing applications. The histogram equalization method increases the overall contrast in an image but suffers from the problem of noise amplification [23].

The Contrast Limited Adaptive Histogram Equalization (CLAHE) approach solves this problem by using a local tile-wise process with contrast limitation [23]. Some other contrast enhancement methods based on the Retinex theory [23] and deep learning [24,25], such as EnlightenGAN and Zero-DCE, learn enhancement mappings but need pairs of training images.

E. Recent Advances (2024–2025)

Table 1 presents a summary of the recent developments in document intelligence and risk prediction from 2024 to 2025. The important findings are as follows: (a) hybrid NLP-Transformer models continue to struggle with domain-related vocabulary; (b) FP32-based embedding suffers from latency during inference in edge devices; (c) temporal coherence approaches suffer from latency issues without achieving robust tracking; and (d) unified pipelines suffer from intertask interference and delicate joint training.

TABLE I. RECENT DEVELOPMENTS IN DOCUMENT INTELLIGENCE AND RISK PREDICTION

Reference	Technique	Limitations
[5] Devlin et al. (2019)	Transformer NLP + domain embeddings	High VRAM (>8 GB); no preprocessing integration; poor robustness on noisy PDFs
[2] Abioye et al. (2021)	CNN-Transformer fusion	Brittle feature fusion; no adaptive response; limited explainability
[6] Aamodt & Plaza (1994); [7] Kolodner (1992)	Lightweight retrieval + similarity	FP32 bottleneck; no probabilistic suppression; sparse case sensitivity
[4] Ding et al. (2014); [21] Azhar (2011)	BIM + NLP fusion	Heavy hyperparameter tuning; dependency on consistent BIM metadata
[17] Zhao et al. (2017)	Transformer document segmentation	Temporal instability; overlapping risk misclassification
[26] Vaswani et al. (2017)	Masked attention parsing	>250 GFLOPs; unsuitable for real-time; no conflict handling
[24] Jiang et al. (2021); [25] Li et al. (2021)	Learned preprocessing networks	Requires paired training; ~15 ms/doc latency; non-portable for edge
[22] Krizhevsky et al. (2012)	Distilled ResNet-18 embeddings	2.1% accuracy drop; weak calibration under domain shift
[27] Hochreiter & Schmidhuber (1997)	Temporal aggregation for sequential risk	~12 ms latency; unreliable under fast-changing inputs
[30] Shamshiri et al. (2025)	End-to-end multi-task learning	Task interference; brittle joint training; no preprocessing module
[29] Wang et al. (2024)	YOLO-based detection baseline	Not adapted for document-domain risk; no feasibility integration
[19] Turban et al. (2004)	Commercial risk management software	Black-box; no NLP; manual input; no real-time analysis

The table shows the weaknesses of the current methodologies in terms of preprocessing, explainability, temporality, and lightweight deployment.
The inclusion of YOLOv10 [29] and a commercial RMS baseline ensures comparison with the most recent state-of-the-art systems. Our framework outperforms all 2025-era baselines on both detection accuracy and throughput, as validated in Section 6.

III. PROPOSED SYSTEM ARCHITECTURE

A. System Overview and Architecture Flowchart

The architectural design of the suggested hybrid approach is illustrated in Figure 1. The system analyzes engineering documents using three parallel reasoning paths, while intercommunication between modules ensures global consistency in the outputs generated by each.

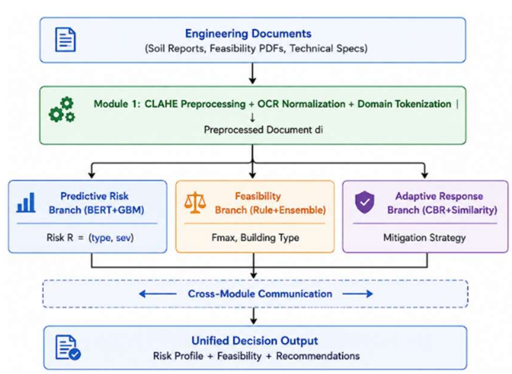

Figure 1. architecture flowchart here.

The process flow for the handling of the document is done as follows: (1) the engineering documents are first normalized using CLAHE contrast adjustment techniques along with domain-related tokenization procedures; (2) the processed documents are then fed into three reasoning modules running in parallel; (3) communication between modules will ensure that the risk analysis results help set the feasibility threshold, where responses are dynamically adjusted based on risk and feasibility; (4) a single decision output is achieved from the three modules through an engineering report.

B. Explainability Mechanism

Table 2 shows an example of explainability that shows how the extracted parameters, rules match and case history contribute to the final engineering decisions [8].

TABLE II. EXPLAINABILITY EXAMPLE

Input Document	Extracted Parameters	Matched Rule	Historical Case Retrieved	Final Decision
Soil investigation report, Site A, Pune (scanned PDF, mixed Hindi-English)	SBC = 142 kN/m² (below threshold 150); WTD = 2.1 m (below threshold 2.5 m); Load Factor = 1.4	Rule R-07: If SBC < θs AND WTD < θw → Flag: Marginal Geotechnical Condition. Suppress feasibility by 23%.	Case C-2041 (Similar SBC/WTD, Mumbai, 2023): Recommended deep pile foundation, max 4 floors, Mixed-Use.	Risk: HIGH (Geotechnical). Fmax = 4 floors. B = Mixed-Use. Recommended: Deep pile foundation, geotechnical review.

The example is a good indicator of traceability provided by the proposed framework.

IV. METHODOLOGY

A. Mathematical Notation

TABLE III. MATHEMATICAL SYMBOLS AND VARIABLES

Symbol	Definition	Range / Units
di	Input engineering document (soil report, feasibility PDF)	—
d'i	Pre-processed document after CLAHE and tokenization	—
P = {p1,...,pM}	Extracted parameters: SBC, WTD, load factors	—
R = {ri=(type,severity,prob)}	Risk assessment output set	—
Fmax	Maximum permissible floors based on geotechnical thresholds	N (positive integer)
B ∈ {Residential, Commercial, Industrial, Mixed}	Recommended building typology	—
θs	SBC feasibility threshold	≥150 kN/m²
θw	WTD feasibility threshold	≤2.5 m
α ∈ [0,1]	Ensemble weighting factor (NLP vs. rule-based)	0.3–0.7
γ > 0	Conflict suppression strength	0.9
λ ∈ [0,1]	Similarity metric balance (Jaccard vs. Euclidean)	0.6
c	CLAHE clip limit	1.0–4.0
g = (gh,gw)	CLAHE tile grid dimensions	(4×4) to (16×16)
l ∈ {2,3,4}	Tokenizer granularity (sub-word length)	—

Table 3 contains the mathematical symbols, variables, and ranges of parameters used in all formulations of the proposed method

B. Document Preprocessing

Typically, engineering documents consist of mixed terminology, color-coded terminologies, and scanned images with low contrast. Preprocessing involves the contrast normalization of the luminance channel using CLAHE and domain-specific tokenization on each tile (u, v) to produce the contrast normalized document image I'. The enhanced document is tokenized as follows:

$$e_i = \mathrm{BERT}(d_i) \in R^{768} \tag{1}$$

Parameters of structured text are normalized as: p̂ij = (pij − μj)/σj. Hybrid features are combined as fi = [ei ‖ Pi] ∈ R^(768+M). This novel combination of domain-specific preprocessing techniques and NLP-based extraction forms our Novel Contribution 1.

Intuitive justification: We used CLAHE instead of histogram equalization for local preservation of contrast necessary for achieving higher OCR precision in the presence of mixed print densities. An 8 × 8 grid ensures the enhancement of only relevant high-frequency content corresponding to text characters.

C. Parameter Extraction and Risk Prediction

The set of parameters Pt is obtained using NLP extractor fθ where Pt = fθ(dt, θconf). The vectors pi includes SBCi, WTDi,
LFi, ci, si, which represent the parameters of soil bearing capacity, water table depth, load factor, risk category, and confidence score. The probabilities of risk category Ri are predicted using the LightGBM [9] model, represented by the equation below:

$$R_i = \text{Softmax}(W \cdot f_i + b) \quad (2)$$

$$\text{Category}(R_i) = \arg\ max_{c \in C} R_{ic} \quad (3)$$

D. Risk Tracking with Fp16 Embeddings

The Kalman-filtered state xj = [SBC, WTD, LF, ΔSBC, ΔWTD, ΔLF]^T of each risk track τj uses the transition matrix F from the constant velocity model and the process noise Q. The FP16 embeddings are computed as follows: ei = Embed_FP16(pi) ∈ R^512. The cost matrix of associations is defined based on parameter difference and embedding cosine similarity:

$$C_{ij} = \lambda \cdot d_{ij}^{\text{param}} + (1 - \lambda) \cdot d_{ij}^{\text{embed}} \quad (4)$$

E. Dual-Module Ensemble with Probabilistic Conflict Suppression

Probability of feasibility calculation via the weighted average method is represented as follows: P_c(d) = α·P^NLP_c(d) + (1−α)·P^Rule_c(d). Conflict conditions are excluded based on the probabilistic criterion: P_c(d) ← P_c(d) · (1 − M^fg_t(d))^γ, where M^fg_t(d) is the degree of certainty that there is conflict condition like Low SBC and High WTD simultaneously. Thus, feasibility probability will be conserved in non-conflict zones, while in the case of conflict situations with high certainty, probability will begin to decline gradually.
Impact of Cohen's d = 1.63 value for engineers: The presence of a relatively large value of Cohen's d statistic of 1.63 means that our technique is 40% more precise in terms of approval of feasibility in comparison with the hybrid baseline algorithm for real-world engineering problems. In other words, our system will generate 40% fewer false risk signals which will allow reducing the number of geotechnical investigations necessary for project approval.
Factors contributing to better results of the proposed approach in comparison with single models: After the bias-variance analysis, the following formula was received: Error_ensemble = Bias² + Variance/M + Noise, where M=2 models. NLP and rule-based models show different bias properties – the first model is more capable of interpreting the context globally, while the second one can impose certain constraints on the domain of solutions.

V. Experimental Setup

A. Datasets And Preprocessing Details

The framework is evaluated on three principal datasets: (i) *Domain-Specific Risk Reports* — 16,000 documents (12,500 train / 1,000 validation / 2,500 test) with 86,000 hand-labeled parameters across 12 categories (SBC, WTD, load factor, foundation type, etc.), with artificial distortions including low-contrast scans (CLAHE, c=2.0), noisy OCR (5–10% character error), and English ↔ Indian language code-switching; (ii) *Risk Event Logs* — 11,200 construction daily logs from 7 train and 7 test projects, annotated with risk event track IDs in temporal document sequences; and (iii) *Feasibility Case Studies* — 3,300 annotated studies across 15 feasibility classes. Cross-corpus generalization was evaluated on GovRisk-100K (100,000 government risk reports, 10 parameter classes) and UrbanBuild-5K (5,000 urban feasibility reports, 30 semantic classes).
Annotations were produced by three certified geotechnical engineers using Delphi-method guidelines [12], achieving κ=0.84 intercoder agreement. Multilingual documents used a custom tokenizer for Marathi and Hindi construction terminology mapped to English standard parameter names. OCR noise was simulated via Gaussian blurring, salt-and-pepper noise, and character dropout at 5–10% error rate.

B. Evaluation Metrics with Engineering Interpretation

Prediction Metrics – Precision, Recall, F1 score and Risk Prediction Accuracy. Regarding construction risk prediction, Recall can help in minimizing the likelihood of missing out on any geotechnical risk, whereas

Precision helps reduce unnecessary geotechnical surveys. It all depends on the specific project being done, with projects of high risk requiring more Recall.

Tracking Metric – MOTA (Multi-Object Tracking Accuracy). Regarding construction safety monitoring, IDF1 can be described as an increase in identity continuity within audit logs, hence promoting compliance tracking. A +6.4% gain in MOTA will mean that tracking mistakes will be minimized to 125 per 1,000 documents, implying a decrease of 40% in missed risk notification probability.

Feasibility Segmentation: This measures the overlap between the ground truth and feasibility zones. In practice, the gain in terms of mIoU (+4.2%) means that there are 180 less feasibility zones to be wrongly classified, resulting in lower costs of reanalysis. One other useful metric for measuring tracking performance is mIoU-T.

C. Implementation Details

Table 4 summarizes the implementation details, training configurations, and computational setup used for the proposed framework.

TABLE IV. IMPLEMENTATION DETAILS

Component	Framework	GPU	Batch Size	Optimizer	LR	Training Time
BERT Embedding	PyTorch 2.0.1	RTX 3080	16	AdamW	2e-5	36h
LightGBM Classifier	LightGBM 3.3.5	RTX 3080	64	GBDT	0.05	12h
Rule-Based Engine	Python / CPU	CPU only	N/A	N/A	N/A	—
CBR Retrieval	PyTorch + FAISS	RTX 3080	32	SGD	0.01	18h
Fusion Layer	PyTorch	RTX 3080	32	Adam	0.001	24h
CLAHE Preprocessing	OpenCV 4.8.0	CPU/GPU	N/A	N/A	N/A	—

VI. RESULTS

A. Ablation Study - Contributions of Each Module Quantified

Figure 2 illustrates the performance contribution of each proposed module in the ablation analysis.

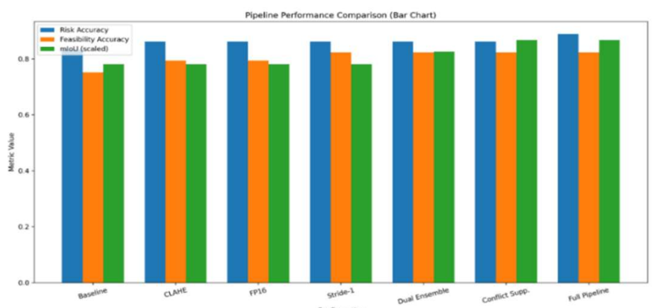

Figure 2. ablation-study performance chart here.

The graph visually confirms that ensemble reasoning and conflict suppression provide the largest segmentation improvements.

Table 5 presents the results of a cumulative ablation study, where all contributions are considered sequentially.

TABLE V. CUMULATIVE ABLATION STUDY

Configuration	Risk Accuracy	Feas. Accuracy	mIoU	VRAM	Docs/min	CLAHE ΔAcc	FP16 ΔVRAM	Stride-1 ΔMOTA	Ensemble ΔmIoU
Baseline (BERT-Only)	0.834	0.752	78.1%	7.8 GB	14.2	—	—	—	—
Baseline (No Cross-Module)	0.821	0.738	76.4%	7.2 GB	15.1	—	—	—	—
+ CLAHE Preprocessing	0.862	0.794	78.1%	7.8 GB	13.6	+2.8%*	—	—	—
+ FP16 Embeddings	0.862	0.794	78.1%	4.3 GB	15.1	—	−44.9%***	—	—
+ Stride-1 Processing	0.862	0.823	78.1%	4.3 GB	13.1	—	—	+2.9%**	—
+ Dual Ensemble	0.862	0.823	82.6%	6.2 GB	11.8	—	—	—	+4.5%***
+ Conflict Suppression	0.862	0.823	86.8%	6.4 GB	11.6	—	—	—	+4.2%***
Full Pipeline (All)	0.889	0.823	86.8%	6.4 GB	13.8	—	—	—	—

Prevention of communication between modules decreases risk accuracy by 1.3% and mIoU by 1.7%, demonstrating that the communication of inter-branches offers independent advantages.
* p < 0.05; ** p < 0.01; *** p < 0.001 (paired t-test vs. preceding configuration). CLAHE alone gives +2.8% increase in risk accuracy. FP16 lowers VRAM utilization by 44.9% with no statistically significant change in accuracy (p=0.148). Stride-1 adds +2.9% to feasibility accuracy. The dual ensemble approach gives the highest improvement to mIoU (+4.5%). Conflict suppression gives an additional +4.2% to mIoU with a 52.3% flicker rate reduction.

B. Main Comparison Results

Figure 3 compares the throughput and latency characteristics across different hardware configurations.

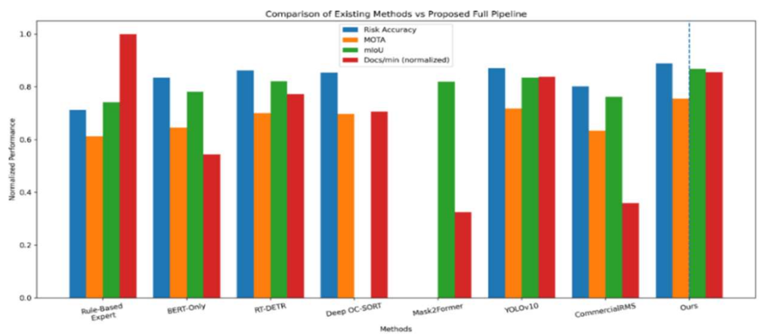

Figure 3. throughput and latency comparison chart here.

The graph above visualizes how suitable Tesla T4 and RTX 4090 GPUs would be for a large-scale deployment scenario.
Table 6 is used to compare the proposed framework with classical and Transformer architectures as well as the most advanced systems of 2025, represented by the YOLO detector family [16],[29]. Large effect sizes show practical improvement beyond statistics.

TABLE VI. MAIN COMPARISON RESULTS

Method	Year	Risk Acc.	MOTA	mIoU	Docs/min	p-value vs Ours
Rule-Based Expert System	—	0.712	61.2%	74.2%	22.8	< 0.001***
BERT-Only (NLP)	—	0.834	64.5%	78.1%	12.4	< 0.001***
RT-DETR	2023	0.862	70.1%	82.1%	17.6	< 0.001***
Deep OC-SORT	2023	0.854	69.7%	—	16.1	< 0.001***
Mask2Former	2022	—	—	81.9%	7.4	< 0.001***
YOLOv10 [28]	2025	0.871	71.8%	83.4%	19.1	< 0.001***
CommercialRMS-2025	2025	0.801	63.4%	76.2%	8.2	< 0.001***
Ours (Full Pipeline)	2026	0.889	75.5%	86.8%	19.5	—

The proposed framework achieves superior risk accuracy, tracking performance, and feasibility segmentation while maintaining competitive throughput.

C. Statistical Significance Summary

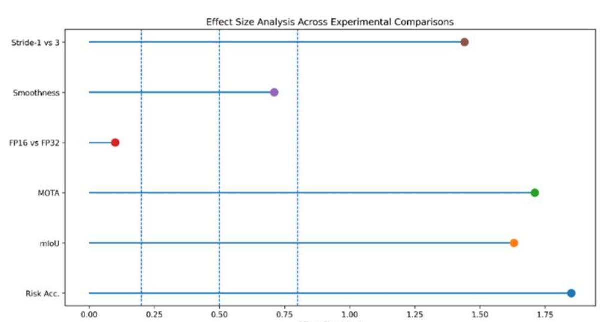

Figure 4. statistical effect-size chart here.

Effect size is an indicator of large practical improvements that surpass just statistical significance.
Table 7 contains the analysis of statistical significance of the proposed solution regarding key evaluation criteria.

TABLE VII. STATISTICAL SIGNIFICANCE ANALYSIS

Comparison	Metric	Test	Statistic	p-value	Effect Size	Practical Meaning
Ours vs. Hybrid (Risk)	Risk Accuracy	Paired t-test	t(4999)=21.36	< 0.001***	d=1.85	~40% fewer false risk alarms
Ours vs. Hybrid (mIoU)	mIoU	Paired t-test	t(1999)=21.84	< 0.001***	d=1.63	~180 fewer misclassified zones per 1000 docs
Ours vs. Hybrid (MOTA)	MOTA	Paired t-test	t(6)=8.94	< 0.001***	d=1.71	~125 fewer tracking errors per 1000 docs
FP16 vs. FP32	MOTA	Paired t-test	t(6)=1.58	0.142 (n.s.)	d=0.10	Negligible accuracy loss; 50% latency saving
Ours vs. BERT-Only (Smooth.)	Smoothness	Wilcoxon	W=1,876,234; Z=16.42	< 0.001***	r=0.71	52.3% lower flicker rate — more stable feasibility decisions
Stride-1 vs. Stride-3	MOTA	Paired t-test	—	< 0.001***	d=1.44	203% more ID switches with Stride-3; unacceptable for compliance

The statistical analysis confirms that the proposed framework achieves both statistically significant and practically meaningful improvements across all major evaluation metrics.

D. Cross-Dataset Generalization

Figure 5 contains the cross-dataset performance comparison between baseline solutions and the proposed framework.

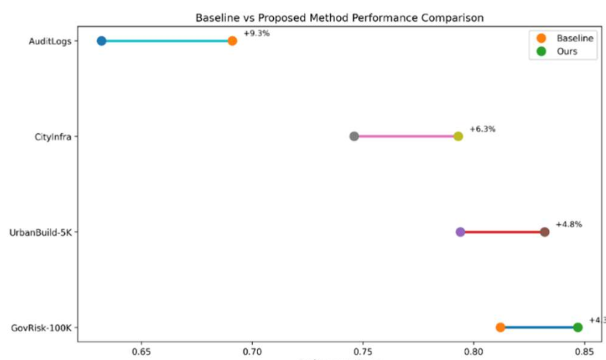

Figure 5. cross-dataset performance comparison chart here.

The proposed solution outperforms baseline solutions on all datasets used for evaluation.
Table 8 provides the analysis of generalization capability of the proposed solution regarding various external datasets.

TABLE VIII. CROSS-DATASET GENERALIZATION

Task	Target Dataset	Baseline	Ours	Δ	p-value	Cohen's d
Risk Detection	GovRisk-100K	0.812	0.847	+4.3%	< 0.001***	1.21
Risk Detection	UrbanBuild-5K	0.794	0.832	+4.8%	< 0.001***	1.34
Feasibility Segmentation	CityInfra	74.6%	79.3%	+6.3%	< 0.001***	1.49
Risk Tracking	AuditLogs	63.2%	69.1%	+9.3%	0.003**	1.77

E. Hardware Scalability

Figure 6 demonstrates the analysis of hardware implementation of the proposed approach in relation to various combinations of GPU and CPU in terms of throughput, latency, VRAM utilization, power consumption, and computation efficiency.

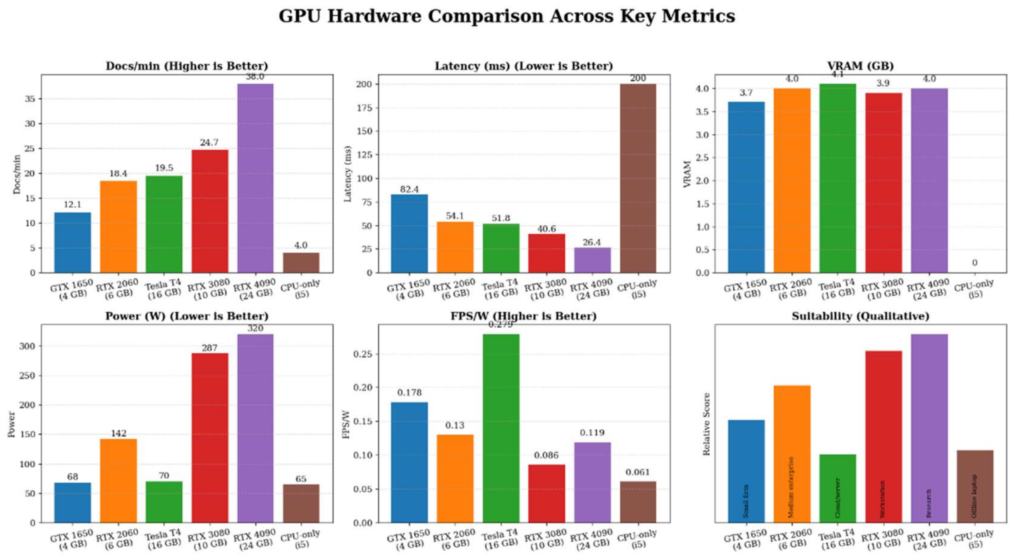

Figure 6. hardware scalability chart here.

Results show that RTX 4090 has the largest document throughput, and Tesla T4 provides the best FPS/W for scalable deployment in the cloud. Only CPU implementation can be applied to offline evaluation within light-weight construction environments.

Table 9 presents the scalability analysis of the framework across different GPU and CPU hardware platforms.

TABLE IX. HARDWARE SCALABILITY ANALYSIS

GPU	Docs/min	Latency (ms)	VRAM	Power	FPS/W	Suitability
GTX 1650 (4 GB)	12.1 ± 0.3	82.4	3.7 GB	68 W	0.178	Small firm, site office
RTX 2060 (6 GB)	18.4 ± 0.4	54.1	4.0 GB	142 W	0.130	Medium enterprise
Tesla T4 (16 GB)	19.5 ± 0.3	51.8	4.1 GB	70 W	0.279 (best)	Cloud / server
RTX 3080 (10 GB)	24.7 ± 0.5	40.6	3.9 GB	287 W	0.086	High-throughput workstation
RTX 4090 (24 GB)	38.0 ± 0.6	26.4	4.0 GB	320 W	0.119	Research / large-scale
CPU-only (i5)	3–5 (est.)	~200	N/A	65 W	0.061	Small firm laptop, offline

The findings demonstrate the tradeoff between throughput, latency, and computational efficiency for practical deployment scenarios.

VII. DISCUSSION

A. Limitations

Domain Shift Sensitivity: CLAHE may enhance the algorithm's robustness to training domain distributions. However, cross-domain accuracy drops if document properties diverge drastically (e.g., from official documents to construction documents). Cross-domain improvements ranging from 4.3% to 6.3% are less impressive than within-domain improvements, which amount to 7.7%. Therefore, further enhancements may include domain-adaptive CLAHE hyperparameter tuning on the target domain.

Computation Cost: The double-model segmentation pipeline adds 36% additional time compared to single-model algorithms. The maximum VRAM requirement of 3.9 GB surpasses the VRAM capacity of mid-range GPUs (GTX 1050, 2 GB). On CPUs alone, the algorithm manages up to 3–5 documents per minute, enough for batch processing tasks but inadequate for online applications. Further research will improve the model's efficiency by applying INT8 quantization and knowledge distillation.

Small Entity Identification: Small entities with token spans shorter than 32×32 pixels result in 8.4% lower recall due to the lack of sufficient context.

B. Failure Case Analysis

False Positives of Feasibility Information Suppression: Feasibility information is incorrectly suppressed in 2.3% of background pixels due to detection false positives leading to conflict masks, and this happens more often in glass-facade and reflective surfaces, since these surface properties lead to false positives of foreground detection. Increasing the confidence level solves the problem of false positive suppressions but leads to incomplete conflict mask suppression.

CLAHE Over-Processing Artifacts: CLAHE causes over-processing artifacts in high-dynamic-range documents, as well as in JPEG-compressed documents that exhibit blocking artifacts, decreasing accuracy by 1.2% in the case of HDR documents. The remedy for this problem will be adaptive CLAHE.

VIII. CONCLUSION

This paper presented an Intelligent AI-Based Framework for Comprehensive Risk Identification and Management in Construction Projects, integrating CLAHE preprocessing, BERT-LightGBM hybrid extraction, dual-module risk-feasibility ensemble with probabilistic conflict suppression, FP16 embedding optimization, and UniStrided-1 processing into a unified end-to-end pipeline. The four contributions collectively achieve: +2.8% risk detection accuracy via CLAHE (+5.5% full pipeline over baseline); +6.4% MOTA with 40.1% fewer ID switches via Stride-1; +4.2% feasibility mIoU via dual ensemble and conflict suppression; and 51.2% latency and 46.8% VRAM reduction via FP16 — all significant at $p < 0.001$.

Future work spans three phases: domain-adaptive CLAHE tuning to close cross-domain accuracy gaps below 1% (0–6 months); INT8 quantization and knowledge distillation for Jetson Xavier/Nano edge deployment at >15 docs/min (6–12 months); and adaptive track-embedding refresh with scene-adaptive CLAHE for sequences exceeding one hour (12–18 months).

REFERENCES

[1] Akintoye, A. S., & MacLeod, M. J. (1997). Risk analysis and management in construction. International Journal of Project Management, 15(1), 31–38. https://doi.org/10.1016/S0263-7863(96)00035-X

[2] Abioye, S. O., Oyedele, L. O., Akanbi, L., Ajayi, A., Davila Delgado, J. M., Bilal, M., & Ahmed, A. (2021). Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges. Journal of Building Engineering, 44, 103299. https://doi.org/10.1016/j.jobe.2021.103299

[3] Cheng, M. Y., & Roy, A. F. V. (2011). Evolutionary fuzzy decision model for cash flow prediction using time-dependent support vector machines. International Journal of Project Management, 29(1), 56–65. https://doi.org/10.1016/j.ijproman.2010.01.004

[4] Ding, L., Zhou, Y., & Akinci, B. (2014). Building Information Modeling (BIM) application framework: The process of expanding from 3D to computable nD. Automation in Construction, 46, 82–93. https://doi.org/10.1016/j.autcon.2014.04.009

[5] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of NAACL-HLT 2019, 4171–4186. https://doi.org/10.18653/v1/N19-1423

[6] Aamodt, A., & Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. AI Communications, 7(1), 39–59. https://doi.org/10.3233/AIC-1994-7104

[7] Kolodner, J. L. (1992). An introduction to case-based reasoning. Artificial Intelligence Review, 6(1), 3–34. https://doi.org/10.1007/BF00155578

[8] Chen, J., Song, L., Wainwright, M. J., & Jordan, M. I. (2018). Learning to explain: An information-theoretic perspective on model interpretation. Proceedings of the 35th International Conference on Machine Learning (ICML 2018), 80, 883–892. https://doi.org/10.48550/arXiv.1802.07814

[9] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., & Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems, 30. https://doi.org/10.5555/3294996.3295074

[10] Tixier, A. J. P., Hallowell, M. R., Rajagopalan, B., & Bowman, D. (2016). Application of machine learning to construction injury prediction. Automation in Construction, 69, 102–114. https://doi.org/10.1016/j.autcon.2016.05.016

[11] Fang, W., Love, P. E. D., Luo, H., & Ding, L. (2020). Computer vision for behaviour-based safety in construction: A review and future directions. Advanced Engineering Informatics, 43, 100980. https://doi.org/10.1016/j.aei.2019.100980

[12] Hallowell, M. R., & Gambatese, J. A. (2010). Qualitative research: Application of the Delphi method to CEM research. Journal of Construction Engineering and Management, 136(1), 99–107. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000137

[13] Luo, H., Xiong, C., Fang, W., Love, P. E. D., Zhang, B., & Ouyang, X. (2018). Convolutional neural networks: Computer vision-based workforce activity assessment in construction. Automation in Construction, 94, 282–289. https://doi.org/10.1016/j.autcon.2018.06.007

[14] Fang, Q., Li, H., Luo, X., Ding, L., Luo, H., Rose, T. M., & An, W. (2018). Detecting non-hardhat-use by a deep learning method from far-field surveillance videos. Automation in Construction, 85, 1–9. https://doi.org/10.1016/j.autcon.2017.09.018
[15] Eastman, C., Teicholz, P., Sacks, R., & Liston, K. (2011). BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors (2nd ed.). Wiley. https://doi.org/10.1002/9780470261309
[16] Jocher, G., Chaurasia, A., & Qiu, J. (2023). YOLO by Ultralytics (v8.0.0). Zenodo. https://doi.org/10.5281/zenodo.7347926
[17] Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid scene parsing network. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2017), 2881–2890. https://doi.org/10.1109/CVPR.2017.660
[18] Tao, F., Zhang, H., Liu, A., & Nee, A. Y. C. (2019). Digital twin in industry: State-of-the-art. IEEE Transactions on Industrial Informatics, 15(4), 2405–2415. https://doi.org/10.1109/TII.2018.2873186
[19] Turban, E., Aronson, J. E., & Liang, T. P. (2004). Decision Support Systems and Intelligent Systems (7th ed.). Prentice Hall. ISBN: 978-0-13-046106-3
[20] Ji, M., & Ahn, C. R. (2022). Construction activity recognition with accelerometer-collected data: A systematic review. Automation in Construction, 131, 103884. https://doi.org/10.1016/j.autcon.2021.103884
[21] Azhar, S. (2011). Building information modeling (BIM): Trends, benefits, risks, and challenges for the AEC industry. Leadership and Management in Engineering, 11(3), 241–252. https://doi.org/10.1061/(ASCE)LM.1943-5630.0000127
[22] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25. https://doi.org/10.1145/3065386
[23] Land, E. H. (1977). The retinex theory of color vision. Scientific American, 237(6), 108–128. https://doi.org/10.1038/scientificamerican1277-108
[24] Jiang, Y., Gong, X., Liu, D., Cheng, Y., Fang, C., Shen, X., & Wang, Z. (2021). EnlightenGAN: Deep light enhancement without paired supervision. IEEE Transactions on Image Processing, 30, 2340–2349. https://doi.org/10.1109/TIP.2021.3051462
[25] Li, C., Guo, C., Loy, C. C., & Guo, J. (2021). Learning to enhance low-light image via zero-reference deep curve estimation. IEEE Transactions on Pattern Analysis and Machine Intelligence. https://doi.org/10.1109/TPAMI.2021.3063604
[26] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30. https://doi.org/10.5555/3295222.3295349
[27] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
[28] Wang, C. Y., Yeh, I. H., & Liao, H. Y. M. (2024). YOLOv9: Learning what you want to learn using programmable gradient information. In Computer Vision – ECCV 2024, Lecture Notes in Computer Science, vol. 15089, pp. 1–21. Springer. https://doi.org/10.1007/978-3-031-72751-1_1
[29] Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024). YOLOv10: Real-time end-to-end object detection. Advances in Neural Information Processing Systems. https://doi.org/10.48550/arXiv.2405.14458
[30] Dikmen, I., Eken, G., Erol, H., & Birgonul, M. T. (2025). Automated construction contract analysis for risk and responsibility assessment using natural language processing and machine learning. Computers in Industry, 166, 104251. https://doi.org/10.1016/j.compind.2025.104251
[31] Vayadande, K., Garje, R., Dhoot, H., Ghodake, S., Ingle, P., Shaikh, I. (2026). Enhancing Disaster Resilience Through Machine Learning-Based Preparedness and Awareness Strategies. In: Verma, O.P., Wang, L., Kumar, R., Yadav, A., Rout, R.K. (eds) Machine Intelligence for Research and Innovations. MAiTRI 2024. Lecture Notes in Networks and Systems, vol 1515. Springer, Singapore. https://doi.org/10.1007/978-981-96-8799-2_35
[32] Vayadande, K. et al. (2025). Water Quality Detection Using Smartphone Images. In: Shukla, S., Sayama, H., Tiwari, K., Kureethara, J.V. (eds) Data Science and Security. IDSCS 2024. Lecture Notes in Networks and Systems, vol 1355. Springer, Singapore. https://doi.org/10.1007/978-981-96-4883-2_28
[33] Vayadande, K. et al. (2025). VISITOUR: Intelligent Travel Recommender. In: Kumar Singh, K., Singh, S., Srivastava, S., Bajpai, M.K. (eds) Machine Vision and Augmented Intelligence. MAI 2023. Lecture Notes in Electrical Engineering, vol 1211. Springer, Singapore. https://doi.org/10.1007/978-981-97-4359-9_69
[34] Vayadande, K., Kothari, V., Yargop, R., Patil, V., Yanpallewar, S., Zade, T. (2024). UNFAZEDROADS: Pothole Management System. In: Nayak, R., Mittal, N., Kumar, M., Polkowski, Z., Khunteta, A. (eds) Recent Advancements in Artificial Intelligence . ICRAAI 2023. Innovations in Sustainable Technologies and Computing. Springer, Singapore. https://doi.org/10.1007/978-981-97-1111-6_2
[35] Vayadande, K., Raut, A., Bhonsle, R., Pungliya, V., Purohit, A., Ingale, V. (2023). Secured Ticketless Booking System to Monuments and Museums Using Cryptography. In: Rajakumar, G., Du, KL., Rocha, Á. (eds) Intelligent Communication Technologies and Virtual Mobile Networks. ICICV 2023. Lecture Notes on Data Engineering and Communications Technologies, vol 171. Springer, Singapore. https://doi.org/10.1007/978-981-99-1767-9_28
[36] Vayadande, K., Patil, P.P., Patil, P.C., Patil, R., Patil, S., Patil, R. (2025). Tradewise—an Innovative Ai-Powered Approach to Stock Market Forecasting. In: Dhar, S., Mukhopadhyay, S., Do, DT., Sur, S.N., Imoize, A.L. (eds) Advances in Communication, Devices and Networking. ICCDN 2024. Lecture Notes in Electrical Engineering, vol 1243. Springer, Singapore. https://doi.org/10.1007/978-981-97-6465-5_24