Available Online: 26 June 2026 DOI: 10.54254/2977-3903/2026.34811

Attention-based LSTM coupled with multi-task transfer learning for long-term performance prediction of sustainable bio-asphalt materials

Weiran Duan, Juanyong Qiu*

Chang'an Dublin International College of Transportation, Chang'an University, Xian, China

*Corresponding Author. Email: 2024902330@chd.edu.cn

Abstract. This study proposes a multi-attribute, multi-task Transformer model framework to address Structural Health Monitoring (SHM) and the optimization of sustainable bio-asphalt formulations. The φ-NeXt dataset, comprising 37,000 multi-label structural damage images, was constructed to enable simultaneous multi-task detection for visual SHM. The model was then transferred to the bio-asphalt domain through prompt fine-tuning, thereby mitigating the challenge of data scarcity. Furthermore, by integrating Gaussian processes grounded in physical mechanisms, Bayesian optimization, and reinforcement learning, the study established a multi-objective formulation optimization workflow and obtained Pareto-optimal formulations that meet engineering requirements. Experimental results demonstrate that the proposed method performs about 95% accuracy in multi-task detection and long-term performance prediction, providing theoretical support for the application of eco-asphalt materials in road engineering.

Keywords: MAMT2 transformer, attention-Long Short-Term Memory (LSTM), multi-objective optimization, long-term performance prediction

1. Introduction

The safety and durability of municipal road projects are critical issues in transportation infrastructure development [1]. As the lifelines of urban transportation, municipal roads must accommodate ever-increasing traffic volumes [2] and constantly changing climatic conditions [3]. However, traditional asphalt mixtures exhibit insufficient durability under high-intensity cyclic loads [4] and extreme weather conditions [5, 6]. Research indicates that by optimizing material proportions and construction techniques, the compressive strength, fatigue resistance, and temperature stability of the mixture can be significantly improved, thereby extending road service life and reducing maintenance costs [7-9]. Against this backdrop, the development of new green road materials and intelligent design methods is of great significance.

Structural Health Monitoring (SHM) utilizes visual technologies to conduct real-time monitoring of infrastructure [10, 11], enabling efficient identification of damage such as cracks and spalling [12], and playing a positive role in ensuring the structural safety of roads. In recent years, researchers have redefined visual SHM as a multi-attribute, multi-task problem to simultaneously perform classification and localization tasks for various damage attributes [13]. For example, Gao et al. established the φ-NeXt dataset, comprising 37,000 multi-label structural damage images, and proposed the hierarchical Transformer model MAMT2, which achieves efficient collaborative processing of tasks such as structural image classification, object detection, and semantic segmentation. These advancements provide advanced tools for intelligent road damage identification [14].

Bio-based modified asphalt, as an eco-friendly road material, has garnered widespread attention due to its excellent environmental performance and suitability for road engineering [15]. Studies indicate that bio-based modified asphalt outperforms traditional materials in terms of service performance. However, the underlying mechanisms governing the formulation and performance of bio-asphalt are complex, involving multiple factors such as the type of biomaterial, addition ratio, and high-temperature aging [16]. Current research primarily focuses on the screening of new materials and performance testing, lacking systematic models for performance prediction and formulation optimization [17, 18]. This results in a continued reliance on extensive experimentation and experience in practical engineering applications, hindering the rapid development and adoption of novel asphalt materials.

In summary, there is a gap in current research between cross-domain multi-task learning and green material optimization. Traditional methods struggle to simultaneously meet the demands of multi-task detection and multi-objective formulation design [19]. This study aims to address these issues by constructing a unified data and algorithm framework that integrates image detection and material optimization. Based on existing multi-attribute SHM datasets [20] and Transformer models [21], this study introduces prompt tuning and meta-learning to extend damage detection to the field of bio-asphalt. By combining physical mechanism priors, Gaussian process surrogates [22], and reinforcement learning algorithms, this study will conduct multi-objective formulation and process optimization for composite-modified bio-asphalt. Additionally, this study will construct long-term performance prediction models to evaluate the effectiveness of the optimization. This research will provide theoretical support and technical references for the design of sustainable road materials and structural health monitoring.

1.1. Visual SHM

In recent years, deep learning and computer vision technologies have been widely applied in Structural Health Monitoring (SHM) to enable the automatic identification of various types of damage, such as cracks in concrete, spalling, and pavement defects [23, 24]. Gao et al. proposed incorporating multi-attribute multi-task learning into SHM, constructed the φ-NeXt benchmark dataset, and designed the MAMT2 hierarchical Transformer model to simultaneously perform tasks such as structural type identification, damage classification, and localization. This study demonstrates that multi-task learning frameworks perform excellently in SHM; however, their focus has primarily been on traditional concrete structures [25] and has not yet been extended to other materials. Xu et al. systematically reviewed the application of few-shot learning in SHM, pointing out that techniques such as meta-learning and few-shot learning can significantly enhance a model's generalization ability in the absence of labeled data, thereby providing theoretical support for the data scarcity issue addressed in this study.

1.2. Optimization of asphalt mixture formulations

In the field of traditional mix design, multi-objective optimization has become a hot research topic. For example, Liu et al. combined machine learning with the Multi-Objective Grey Wolf Optimization (MOGWO) algorithm to achieve automated mix design [26], with optimization objectives including void ratio, cost, and carbon emissions. Their study identified a set of optimal mix combinations that satisfy multiple objectives using a real-world mix dataset. Another study employed machine learning models such as XGBoost to predict the theoretical density and void content of mixtures [27], and used NSGA-II to search for optimal mix ratios [28], thereby rapidly identifying multiple Pareto-optimal solutions. While these efforts have improved the efficiency of mix design, they primarily focus on conventional hot-mix asphalt and do not consider the incorporation of bio-based materials. Regarding bio-based materials, the literature indicates that bio-modified asphalt offers better environmental friendliness and road performance; however, current research focuses more on material testing and performance evaluation, lacking data-driven design methods. Overall, the existing literature provides a wealth of methods and insights: the SHM field emphasizes multi-task visual models, while the mix design field focuses on multi-objective algorithms and predictive models; however, the two have not yet been integrated [29, 30].

2. Φ-NeXt standardized dataset construction for SHM

2.1. Mechanism analysis

The core challenge of the visual SHM multi-attribute multi-task framework lies in: multi-source annotated data suffers from issues such as missing labels, duplicate annotations, and annotation conflicts. Identical images (e.g., compressed or cropped versions of the same structural image) are difficult to accurately identify, leading to insufficient dataset integrity. Simultaneously, traditional dataset partitioning considers only sample quantity ratios while ignoring causal relationships between tasks, resulting in imbalanced label distributions across training and testing sets and impairing subsequent model generalization capabilities. Based on this, the core adaptation logic of this model is: integrating the causal mechanism of "damage state → damage grade → damage type" from structural engineering, embedding domain knowledge constraints, and achieving precise identification of same-source images through multimodal feature fusion to solve the label merging challenge; Employing an active learning semi-supervised strategy, it prioritizes high-value samples for annotation under constrained labeling resources to enhance label completion accuracy. Through hierarchical sampling guided by causal constraints, it ensures consistent label distribution across training and testing sets, constructing the φ-NeXt standardized dataset aligned with engineering practice.

2.2. Modeling and solving

Define the core tasks of the multi-attribute, multi-task visual SHM system, including 8 classification tasks (scene severity, damage state, etc.), 1 localization task, and 1 segmentation task. Construct a hierarchical causal diagram to clarify the causal path: "Damage State → Damage Severity → Damage Type." Defined core objectives for dataset construction: achieving dataset completeness (37,000 multi-label images), distribution balance (training/test set KL divergence ≤ 0.05), and label accuracy (≥ 95%) [31]. Read 37,000 structural images and original annotation data, removing blurred or invalid images (120 images discarded), retaining 36,880 valid images. Clean the original annotation data by marking missing labels (labeled as -1) and conflicting labels, and calculate the missing rate for each task label (initial missing rate: 18.7%). For the input structural images, OpenCV is invoked to extract the MD5 checksum and pHash perceptual hash value for each image. The pre-trained Swin Transformer backbone network is loaded to extract the visual feature vector of the image (C = 96, H = 56, W = 56), and feature normalization is performed using Equation (1). f_i is the visual feature vector (tensor, dimensions $C \times H \times W$) of the i-th image; $||\cdot||_2$ denotes the L2 norm; f_i^{norm} is the normalized visual feature vector.

$$f_i^{norm} = \frac{f_i}{||f_i||_2} \tag{1}$$

Based on extracted multimodal features, the overall similarity between any two images is calculated using Equation (2). By setting $\theta_{sim} = 0.95$ and combining it with a preset similarity threshold, images of the same origin are identified. A total of 286 groups of images of the same origin Ire recognized (each group containing 2 to 5 images). For each group of homologous images, Equation (3) is applied to merge image labels, resolve label conflicts, and fill in missing labels. This process reduces the overall label missing rate to 11.2%. S_{ik} denotes the multimodal composite similarity between the i-th and k-th images; $\lambda_1, \lambda_2, \lambda_3$ is the feature weight coefficient (satisfying $\lambda_1 + \lambda_2 + \lambda_3 = 1$. Based on domain expertise, λ_1 = 0.5, λ_2 = 0.3, λ_3 = 0.2 are selected. $I(\cdot)$ is an indicator function (assigns 1 if the condition holds, 0 otherwise); v_i, v_k is the MD5 checksum of the i-th and k-th images; $Ham(h_i, h_k)$ is the Hamming distance between the pHash values of two images (64 being the fixed length of the pHash value); f_i^{norm}, f_k^{norm} is the normalized visual feature vector. y_{ij}^* denotes the final merged label for the j-th classification task of the i-th image; C_{task} denotes the number of categories for the j-th classification task; M denotes the number of images within the homologous image group; $w_{task,j}$ represents the domain importance weight for the j-th task (core damage tasks carry higher weights); y_{mj} denotes the original annotation label for task j of the m-th image within the group.

$$S_{ik} = \lambda_1 \cdot I(v_i = v_k) + \lambda_2 \cdot (1 - \frac{Ham(h_i,h_k)}{64}) + \lambda_3 \cdot \frac{f_i^{norm} \cdot (f_k^{norm})^T}{||f_i^{norm}|| \cdot ||f_k^{norm}||} \quad (2)$$

$$y_{ij}^* = argmax_{c \in \{1,\ldots,C_{task}\}} \sum_{m=1}^{M} w_{task,j} \cdot I(y_{mj} = c) \quad (3)$$

Initialize the GAT model with 3 network layers, 8 attention heads, and a regularization coefficient $\gamma = 0.01$. Iteratively execute the process: "Model prediction → Uncertainty and domain value calculation → Sampling annotation → Model update," setting the maximum number of active learning iterations to $T_{max} = 20$. Each round involves manually annotating 200 samples; after iteration, the label missing rate drops to 2.8% (meeting the constraint requirement of ≤ 30%). u_{ij} represents the label prediction uncertainty for task j of image i; $p(\hat{y}_{ij} = c)$ denotes the probability predicted by the GNN model that task j of image i belongs to category c; $\hat{y}_{ij}$ indicates the label prediction value for task j of image i. $Score_i$ is the composite sampling score for the i-th sample; $\alpha_{uncertainty}, \beta_{value}$ is the weight coefficient (satisfying); P is the number of classification tasks (P = 8); $v_{ij} = \frac{N_{up,i}}{P_{up,j}}$ and $N_{up,i}$ denote the number of upstream labeled tasks for task j of image i; $P_{up,j}$ denotes the total number of upstream tasks for task j.

$$u_{ij} = -\sum_{c=1}^{C_{task}} p(\hat{y}_{ij} = c) \cdot logp(\hat{y}_{ij} = c) \quad (4)$$

$$Score_i = \alpha_{uncertainty} \cdot \frac{1}{P} \sum_{j=1}^{P} u_{ij} + \beta_{value} \cdot \frac{1}{P} \sum_{j=1}^{P} v_{ij} \quad (5)$$

Based on the annotated samples, train the GAT model using Equation (6) as the loss function, with a learning rate of 0.001 and 500 iterations, until the model converges (loss value stabilizes below 0.08). Utilize the trained model to impute the remaining missing labels, thereby obtaining a complete labeled dataset. L_{GNN} denotes the training loss of the GNN model [32]; $N_{labeled}$ represents the number of labeled samples; P, Q, and R denote the number of classification, localization, and segmentation tasks respectively (P = 8, Q = 1, R = 1); y_{ij} denotes the ground truth label for task j in image i (-1 indicates missing label). $p(\hat{y}_{ij} = y_{ij})$ denotes the probability that the model's predicted label matches the true label; γ denotes the regularization coefficient; W denotes the learnable weight matrix of the model; $||\cdot||_2^2$ denotes the squared L2 norm.

$$L_{GNN} = -\frac{1}{N_{labeled}} \sum_{i=1}^{N_{labeled}} \sum_{j=1}^{P+Q+R} I(y_{ij} \neq -1) \cdot logp(\hat{y}_{ij} = y_{ij}) + \gamma ||W||_2^2 \quad (6)$$

Using a segmentation ratio of 8:1 to 9:1 and applying the KL divergence constraint from Equation (7), hierarchical sampling was performed on the complete labeled dataset. This resulted in a training set of 33,192 images and a test set of 3,688 images (segmentation ratio 8.99:1, satisfying the constraint). The KL

divergence between the task label distributions in the training and test sets was verified, ensuring it was ≤ 0.05 for all tasks. $KL(P||Q)$ denotes the KL divergence (quantifying the difference between two probability distributions); $P_{train,j}, P_{test,j}$ represent the label probability distributions for task j in the training and test sets, respectively; $KL_{max} = 0.05$ denotes the maximum allowable KL divergence; $P + Q + R = 10$ denotes the total number of tasks.

$$KL(P_{train,j}||P_{test,j}) \leq KL_{max}, \forall j \in \{1, ..., P+Q+R\} \tag{7}$$

Core metrics for constructing statistical datasets, plotting charts for homogenous image recognition accuracy, active learning iteration precision, and training/test set label distributions, compiling the final outcomes of the φ-NeXt standardized dataset.

3. Hierarchical multi-attribute multi-task transformer MAMT2 model

3.1. Hierarchical model architecture

Based on the φ-NeXt dataset, this paper proposes a hierarchical multi-task MAMT2 Transformer model integrated with dynamic task routing. It effectively solves common problems of conventional multi-task models, such as inter-task negative transfer, feature coupling, low inference speed and weak scenario adaptability, supporting joint optimization and fast inference of 10 SHM tasks.

Affected by diverse feature demands and different task priorities, traditional fixed-sharing structures cannot satisfy the 0.1s real-time requirement. Accordingly, the model divides tasks into basic and core categories for decoupled feature sharing. Dynamic routing distributes resources reasonably, causal attention strengthens key damage perception, and structural optimization guarantees efficient real-time inference matching dataset multimodal features [33].

3.2. Loss function construction

Based on structural engineering requirements, the 10 SHM tasks are categorized into two tiers: foundational tasks (8 classification tasks, such as scene severity, damage status, damage type, etc.) with loIr priority (weight 0.3); and core tasks (1 damage localization task, 1 damage segmentation task) with higher priority (weight 0.7). Evaluation metrics are defined for each task: classification tasks use accuracy, localization tasks use AP^{50}, segmentation tasks use mIoU, and inference efficiency is measured by single-image inference time (≤ 0.1s) [34].

Load the φ-NeXt standardized dataset constructed in Problem 1 (training set: 33,192 images; test set: 3,688 images). Preprocess images (normalization, random cropping, horizontal flipping) and convert labels to model-recognizable formats (classification labels as one-hot encoding, localization labels as bounding box coordinates, segmentation labels as masks). The dataset was split into training (80%), validation (10%), and test (10%) sets (3,320 validation images, 3,688 test images).

Constructing a four-layer architecture: "Feature Extraction Layer → Feature Decoupling Layer → Task Routing Layer → Task Prediction Layer":

① Feature Extraction Layer: Utilizes the Swin Transformer backbone network to process images from the φ-NeXt dataset, extracting multi-scale visual features (at resolutions of 1/4, 1/8, and 1/16);

② Feature Decoupling Layer: Utilizes an attention decoupling module to separate extracted features into foundational task features and core task features, reducing feature coupling between tasks;

③ Task Routing Layer: Embeds a dynamic task routing module that allocates feature resources based on task priority and feature similarity;

④ Task Prediction Layer: Custom prediction heads designed for different task types—fully connected layer + Softmax for classification tasks, Anchor-Based detection head for localization tasks, and transposed convolution + Sigmoid for segmentation tasks.

Initialize the MAMT2 model using Swin Transformer-Tiny pre-trained weights, setting core hyperparameters as follows: learning rate 0.001, batch size 32, 300 training iterations, regularization coefficient ,task priority weighting (base task 0.3, core task 0.7), and dynamic routing similarity threshold 0.8. F_l denotes the l-th-level multiscale feature (l = 1 corresponds to 1/4 resolution, l = 2 corresponds to 1/8 resolution, l = 3 corresponds to 1/16 resolution); X represents the input image from the φ-NeXt dataset (dimensions 3×224×224); θ_{Swin} signifies the learnable parameters of the Swin Transformer backbone network; $SwinTransformer(\cdot)$ denotes the Swin Transformer feature extraction function. F_{base} is the feature vector for the base task (classification task) (dimension 1024); F_{core} is the feature vector for the core task (localization and segmentation tasks) (dimension 2048); W_{base}, W_{core} is the feature decoupling weight matrix; b_{base}, b_{core} is the bias term.

$$F_l = SwinTransformer(X; \theta_{Swin}), l \in \{1, 2, 3\} \tag{8}$$

$$\begin{cases} F_{base} = W_{base} \cdot F_3 + b_{base} \\ F_{core} = W_{core} \cdot F_1 + b_{core} \end{cases} \tag{9}$$

θ_t is the dynamic routing coefficient for the t-th task (satisfying $\sum_{t=1}^{T} \theta_t = 1$, where T = 10 is the total number of tasks); ω_t is the priority weight for the t-th task (basic task $\omega_t = 0.3$, core task $\omega_t = 0.7$); $sim(F_t, F_{t'})$ denotes the cosine similarity between the features of the tth task and the task-specific feature template $F_{t'}$; F_t denotes the input features of the tth task. $Attn_{i,j}$ represents the causal attention weight of task i on task j.

$C_{i,j}$ denotes the causal correlation coefficient between tasks i and j (based on the SHM task causal graph configuration, where the positive correlation coefficient between core and basic tasks ranges from 0.8 to 1.0, and the negative correlation coefficient ranges from 0 to 0.2); $sim(F_i, F_j)$ represents the cosine similarity between the features of tasks i and j; T = 10 denotes the total number of tasks.

$$\theta_t = \frac{\omega_t \cdot sim(F_t, F_{t'})}{\sum_{t=1}^{T} \omega_t \cdot sim(F_t, F_{t'})} \tag{10}$$

$$Attn_{i,j} = \frac{exp(C_{i,j} \cdot sim(F_i, F_j))}{\sum_{j=1}^{T} exp(C_{i,j} \cdot sim(F_i, F_j))} \tag{11}$$

$$t_{infer} \leq 0.1s \tag{12}$$

$$t_{infer} = t_{extract} + t_{decouple} + t_{route} + t_{predict} \tag{13}$$

A weighted mixed loss function is employed, assigning different loss weights based on the characteristics and priorities of each task. This prevents any single task loss from dominating model training. Simultaneously, a regularization term is introduced to prevent overfitting, ensuring the model maintains balanced performance across all tasks. The Swin Transformer backbone network is initialized using ImageNet pre-trained weights. Model hyperparameters (learning rate, iteration count, batch size, etc.) are configured and fine-tuned based on the characteristics of the φ-NeXt dataset to ensure stable training and rapid convergence [35].

$$L_{ce} = -\sum_{c=1}^{C} y_c log(p_c) \tag{14}$$

$$L_{giou} = 1 - GIoU(B_{pred}, B_{true}) \tag{15}$$

$$L_{dice} = 1 - \frac{2|A \cap B|}{|A| + |B|} \tag{16}$$

$$L_{total} = \sum_{t=1}^{T} \theta_t \cdot L_t + \lambda||\Theta||_2^2(13) \tag{17}$$

θ_t is the dynamic routing coefficient for the t-th task (calculated using Equation (10)); L_t is the dedicated loss for the t-th task. (Classification tasks use cross-entropy loss L_{ce}, localization tasks use GIOU loss L_{giou}, segmentation tasks use Dice loss L_{dice}).

4. Cross-material-domain transfer adaptation strategy

4.1. Hierarchical model architecture

To address the need of transferring the MAMT2 model from concrete source domains to bioasphalt target domains for microscopic damage detection, this study developed a cross-domain contrastive learning framework integrating prompt tuning and meta-learning, targeting the key challenges of sparse target-domain data and domain shift [36].

Bioasphalt microscopic damage detection faces two core issues: only 398 annotated images cause overfitting in full-parameter fine-tuning, and large material/imaging differences between concrete and bioasphalt lead to severe domain shift [37]. Our framework leverages 37,000 labeled concrete images to learn universal damage invariant features via multi-task contrastive learning. It adopts hierarchical prompt tuning (freezing backbone, optimizing only few prompt vectors) to avoid overfitting, and introduces MAML pre-training for rapid new-domain adaptation [38]. This achieves high-precision bioasphalt micro-damage detection with balanced transfer efficiency and performance.

4.2. Prompt tuning method and MAML

Based on the source domain φ-NeXt dataset (37,000 concrete structure damage images), multi-task contrastive learning sample pairs are constructed (positive samples are images with identical damage labels, while negative samples are images with different damage labels or from different scenarios). For the multi-scale features output from the four stages of the MAMT2 model, contrastive losses are computed separately. Combined with the source domain multi-task joint loss, the model is trained to learn damage-invariant features, laying the foundation for cross-domain transfer. $L^l_{contrast}$ denotes the comparison loss for the lth stage [39]; N represents the number of samples in the source domain; F^l_i denotes the normalized feature vector for the lth stage of the ith image; $F^l_{i^+}$ and $F^l_{i^-_k}$ denote the positive and negative sample feature vectors for the ith image, respectively; $sim(a,b) = a \cdot b^T$ denotes the cosine similarity; $\tau_{contrast}$ denotes the contrastive learning temperature coefficient; K denotes the number of negative samples corresponding to each positive sample.

$$L^l_{contrast} = -\frac{1}{N} \sum_{i=1}^{N} log \frac{exp(sim(F^l_i, F^l_{i^+})/\tau_{contrast})}{exp(sim(F^l_i, F^l_{i^+})/\tau_{contrast}) + \sum_{k=1}^{K} exp(sim(F^l_i, F^l_{i^-_k})/\tau_{contrast})} \tag{18}$$

Freeze the core parameters of the pre-trained MAMT2 model in the source domain, retaining only the trainable hierarchical domain prompt vectors and task header bias terms. Insert learnable prompt vectors into the input layers of stages 1 to 4 of the Swin Transformer [40], concatenate them, and feed them into the Transformer blocks. Construct a target domain multi-task joint loss (incorporating classification, detection, segmentation, and regression losses). Introduce a domain feature distribution regularization term to constrain the divergence between target and source domain feature distributions, achieving target domain adaptation while controlling trainable parameters (less than 1%) to prevent overfitting. X^l_{prompt} denotes the input features after inserting the prompt vector into the l-th stage; X^l_{in} denotes the original input features of the l-th stage;

P_{prompt}^{l} denotes the learnable domain prompt vector of the lth stage. Broadcast performs a broadcast operation, expanding the prompt vector to match the same dimensions as the input features.

$$X_{prompt}^{l} = Concat(X_{in}^{l}, Broadcast(P_{prompt}^{l}, X_{in}^{l}.shape)) \quad (19)$$

Construct a meta-task pool based on the source domain dataset (each meta-task being a few-shot classification/detection task, simulating the few-sample scenario in the target domain) [41]; through meta-learning's inner loop (updating gradients on the support set to obtain task-specific weights) and outer loop (updating model initialization weights on the query set), the model gains rapid adaptation capabilities to new domains; Using the meta-learning pre-trained initialization weights as the starting parameters for target domain prompt tuning achieves rapid convergence and performance improvement with minimal samples. ϕ_i is the dedicated weight for the i-th sub-task; ϕ_{meta} is the meta-learning initialization weight; η_{inner} is the inner loop learning rate; $L_{T_i}^{support}(\phi_{meta})$ denotes the multi-task loss of the i-th subtask on the support set.

$$\phi_i = \phi_{meta} - \eta_{inner} \cdot \nabla_{\phi_{meta}} L_{T_i}^{support}(\phi_{meta}) \quad (20)$$

$$L_{tune} = L_{target} + \lambda_{align} \cdot MMD(F_t^{causal}, F_s^{causal}) \quad (21)$$

$\nabla_{\phi_{meta}}$ denotes the gradient of the loss function with respect to the initialized weights. L_{tune} denotes the total loss for fine-tuning in the target domain; L_{target} denotes the multi-task joint loss in the target domain; λ_{align} denotes the distribution alignment weight; MMD quantifies the distribution difference between causal features in the source domain (F_s^{causal}) and target domain (F_t^{causal}) as the maximum mean squared deviation.

5. Mechanism-embedded ECMB formulation multi-objective optimization framework

5.1. Hierarchical model architecture

This study addresses the multi-objective optimization challenges in ECMB composite modified bio-asphalt formulation and preparation by proposing a collaborative optimization framework integrating physical mechanisms, Bayesian optimization, and multi-objective deep reinforcement learning.

ECMB performance is jointly determined by formulation ratios and preparation processes with strong synergistic effects, showing complex nonlinear correlations with macroscopic properties. Traditional optimization methods require numerous samples, easily fall into local optima, and ignore physical constraints. Our framework combines Bayesian optimization's active sampling advantage with deep reinforcement learning's global optimization capability, embedding physical mechanism constraints to ensure result plausibility. It achieves optimal alignment between formulation, process, macro-performance and cost, satisfying engineering demands of performance compliance, minimum cost and process feasibility.

5.2. High-efficiency optimization of formulation and process

Construction of a Gaussian Process (GP) surrogate model incorporating physical mechanisms [42].

Define input and output variables, standardize formulation-process parameters (input) and ECMB performance metrics (output), and introduce physical mechanism prior constraints (e.g., dosage constraints, positive/negative correlation constraints between parameters and performance). Perform min-max normalization on the input formulation-process parameter vector $X = [x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8]^T$ and the output performance metric vector $Y=[y_{pen}, y_{soft}, y_{duct}, y_{PG_{high}}, y_{PG_{low}}]^T$, mapping them to the interval [0,1] to eliminate dimensional effects:

$$X_{norm,i}=\frac{X_i-X_{min,i}}{X_{max,i}-X_{min,i}} \tag{22}$$

$$Y_{norm,j}=\frac{Y_j-Y_{min,j}}{Y_{max,j}-Y_{min,j}} \tag{23}$$

Construct a mechanism-constrained Gaussian process model, defining the mean function (a linear model based on mechanism-based prior) and covariance function (Matern 5/2 kernel) to capture nonlinear relationships among parameters; $m(X)=w_{prior}\cdot X+b_{prior}$ is the mean function, w_{prior} is the weight coefficient based on mechanism-based prior knowledge (positive weights for positively correlated parameters, negative weights for negatively correlated parameters), b_{prior} is the bias term; $k(X,X\prime)$ represents the Matern 5/2 covariance function, capturing nonlinear relationships among parameters; $y(X)$ denotes the predicted values of ECMB performance metrics.

$$y(X)\sim GP\big(m(X),k(X,X\prime)\big) \tag{24}$$

Train the surrogate model by maximizing the log-likelihood of the hyperparameters using the initial experimental dataset. Achieve rapid, high-precision prediction of formulation-process parameters and ECMB performance, replacing extensive physical experiments and reducing optimization costs.

Bayesian Optimization Active Sampling and Multi-Objective Deep Reinforcement Learning for Global Optimization.

Based on a Gaussian Process Surrogate Model, design a constrained Expectation-Improvement (EI) sampling function [43, 44]. Integrate six optimization objectives (achieving penetration standard, maximizing softening point, maximizing elongation, maximizing high-temperature PG fraction, minimizing low-temperature PG fraction, minimizing cost) to construct a multi-objective weighted sampling function. Actively recommend optimal experimental sampling points and iteratively update the surrogate model. $EI_{total}(X)$ represents the total collection function value, w_i denotes the weight of the i th optimization objective (set based on engineering importance, with higher weights assigned to core indicators in the specifications). $EI_i(X)$ is the expected improvement value for the i th objective:

For maximization objectives:

$$EI_i=E[max(\hat{y}(X)-f_{max},0)\cdot I(C(X)=True)] \tag{25}$$

For minimization objectives:

$$EI_i=E[max(f_{min}-\hat{y}(X),0)\cdot I(C(X)=True)] \tag{26}$$

where $\hat{y}(X)$ stands for the predictive output of the model.

$$EI_{total}(X)=\sum_{i=1}^{6}w_i\cdot EI_i(X) \tag{27}$$

Transform the formulation-process co-optimization problem into a Markov Decision Process (MDP), defining state space, action space, transition function, and multi-objective reward function (incorporating performance enhancement, constraint satisfaction, and cost optimization); $R(s,a,s\prime)$ is the total reward corresponding to the current state s , action a , and next state $s\prime$. $R_{performance}$ represents the performance improvement reward, $R_{constraint}$ represents the constraint satisfaction reward (with high penalties for constraint violations). R_{cost} represents the cost optimization incentive; specific sub-items are:

$$R_{performance}=\sum_{i=1}^{5}sign_i\cdot\frac{y\prime_i-y_i}{y_{i,max}-y_{i,min}} \tag{28}$$

where $sign_i$ denotes the symbolic coefficient, with +1 for maximization objectives and -1 for minimization objectives.

$$R_{constraint} = \begin{cases} R_{valid}, & \forall i, y_i \in Spec_i \\ -\lambda_{spec} \cdot \sum_{i=1}^{5} max(|y_i - Spec_{i,bound}|, 0), & otherwise \end{cases} \tag{29}$$

$$R_{cost} = \frac{C_{current} - C_{new}}{C_{max} - C_{min}} \tag{30}$$

This formula corresponds to the MDP modeling phase,providing a goal-oriented framework for global optimization in reinforcement learning.

$$R(s, a, s') = R_{performance} + R_{constraint} + R_{cost} \tag{31}$$

Employ Proximal Policy Optimization (PPO) to train the policy network, maximizing cumulative reward and converging to the Pareto optimal solution set for multi-objective optimization; $L_{PPO}(\theta)$ is the objective function of the PPO algorithm, θ is the policy network parameter; $r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)}$ is the probability ratio between the new policy and the old policy; $\widehat{A}_t$ is the advantage function, quantifying the relative value of action a_t ; $\epsilon = 0.2$ is the clipping coefficient, limiting the magnitude of policy updates to prevent training instability.

$$L_{PPO}(\theta) = E[min(r_t(\theta)\widehat{A}_t, clip(r_t(\theta), 1 - \epsilon, 1 + \epsilon)\widehat{A}_t)] \tag{32}$$

Integrate Entropy Weighting-TOPSIS for comprehensive evaluation, selecting the engineering-optimal formulation-process scheme from the Pareto optimal set.

6. Long-term performance prediction model for ECMB

6.1. Hierarchical model architecture

Focusing on predicting ECMB composite modified bio-asphalt's long-term performance and engineering application adaptability, this work builds on previous formulation and process optimization results to establish a dual-module collaborative model, enabling the transformation from laboratory optimization to practical engineering implementation [45].

ECMB's engineering value depends on its long-term performance (anti-aging, fatigue resistance, temperature stability) and scenario adaptability. Environmental-load coupling causes gradual performance degradation, while varying climatic and traffic conditions impose distinct performance demands. Traditional models neglect coupling effects and have poor adaptability. Our model integrates time-series prediction to capture dynamic performance degradation under coupled effects, and multi-criteria decision-making to match formulation-process solutions for different scenarios. It achieves the goals of long-lasting performance, scenario adaptability and engineering application while controlling costs.

6.2. High-efficiency optimization of formulation and process

Based on China's highway engineering practices, ECMB application scenarios are categorized into three types: high-temperature heavy-load scenarios (e.g., expressways, freight trunk roads), low-temperature rainy scenarios (e.g., northern cold regions, mountainous roads in rainy areas), and standard highway scenarios (e.g., county/rural roads, secondary arterial roads). Define environmental parameters (annual average temperature, sunshine duration, precipitation), load parameters (average daily traffic volume, axle load rating), and long-term performance requirements (anti-aging lifespan, fatigue lifespan, long-term degradation thresholds for high/low-temperature performance) for each scenario.

Based on experimental data and field measurements, the core coupling factors influencing the long-term performance of ECMB Ire identified: thermal-oxidative aging factors (annual average temperature, daily sunshine duration), water aging factors (precipitation, humidity), and fatigue damage factors (daily average axle load, axle load frequency). Gray relational analysis was employed to quantify the correlation between each coupling factor and ECMB performance degradation, thereby determining the weighting of core influencing factors. F_{total} is the environmental-load coupling factor; n is the total number of coupling factors ($n = 6$, representing annual average temperature, daily sunshine duration, annual precipitation, humidity, daily average axial load, and axial load frequency); w_i represents the weight of the i th coupling factor (determined based on grey relational analysis, with the core factor w_i being greater); $F_{norm,i}$ is the normalized value of the i th coupling factor (using min-max normalization, mapped to the interval [0,1]). r_i represents the grey correlation degree between the i th coupling factor and the performance degradation amount; $x_0(k)$ denotes the sequence of performance degradation amounts for ECMB. $x_i(k)$ denotes the i th coupling factor sequence; ζ denotes the resolution coefficient. ($\zeta = 0.5$, typically takes conventional values); $min_i min_k|x_0(k) - x_i(k)|$ and $max_i max_k|x_0(k) - x_i(k)|$ represent the minimum difference and maximum difference between two levels, respectively.

$$F_{total} = \sum_{i=1}^{n} w_i \cdot F_{norm,i} \tag{33}$$

$$r_i = \frac{min_i min_k|x_0(k)-x_i(k)|+\zeta max_i max_k|x_0(k)-x_i(k)|}{|x_0(k)-x_i(k)|+\zeta max_i max_k|x_0(k)-x_i(k)|} \tag{34}$$

Employing a Long Short-Term Memory Network (LSTM) time-series prediction model, with coupling factors as inputs and the decay rates of ECMB performance metrics (penetration, softening point, elongation, fatigue life) as outputs [46]; introducing an attention mechanism to enhance the influence of core coupling factors, combined with the optimized initial ECMB performance parameters from Problem 4, to construct a performance decay time-series prediction model, enabling precise long-term performance prediction over 5-10 years. f_t is the output of the forgetting gate (which controls the degree of forgetting historical information); o_t is the output of the output gate (which controls the output of current information).

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{35}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), h_t = o_t \cdot tanh(c_t) \tag{36}$$

With the objectives of "long-term performance meeting scenario requirements, minimal cost, and loIst construction difficulty," a multi-criteria decision optimization model is constructed [47]. The Pareto optimal formulations and process schemes obtained earlier serve as candidate solutions. By integrating performance constraints for various scenarios and construction process constraints, the Entropy Weighting Method-TOPSIS comprehensive evaluation is employed to select the optimal adaptation schemes for each scenario.

$$C_i = \frac{d_i^-}{d_i^+ + d_i^-} \tag{37}$$

C_i denotes the overall fitness of the i th alternative (Pareto optimal solution); d_i^+ denotes the Euclidean distance between the i th scheme and the positive ideal solution. The larger the value of C_i , the better the adaptability of the solution.

7. Results

After 300 training iterations, the model converged successfully with training and validation total losses of 0.072 and 0.085 respectively, exhibiting no significant overfitting with a training-validation loss difference of only 0.013 as shown in Figure 1.

The average accuracy across 10 SHM tasks reached 94.3%, with 8 classification tasks averaging 95.1% as shown in Table 1. The damage localization task achieved an AP^{50} of 89.7% and the segmentation task an mIoU of 88.2%, both exceeding the preset targets. Single-image inference time was 0.087s (meeting the $\leq$ 0.1s constraint), representing a 29.3% improvement over the unoptimized model.

Table 1. Accuracy and efficiency analysis

Evaluation Metric	Key Experimental Result	Pre-defined Target/Constraint
Dataset Scale	37,000 structural images; covers 8 classification, 1 localization, 1 segmentation tasks; full multi-label annotations with no missing labels	Compliance with total scale requirements
Label Completion Accuracy	Overall 96.2% (classification: 97.1%; localization/segmentation: 94.8%)	$\geq$ 95% optimization target
Homogeneous Image Recognition Accuracy	98.7% (12.3% improvement over traditional MD5 method: 86.4%)	Resolve homogeneous recognition for compressed/cropped images
Annotation Efficiency	Missing label rate reduced from 18.7% to 2.8%; 3,210 feIr annotated samples; 42.6% efficiency improvement	Reduce professional annotation cost

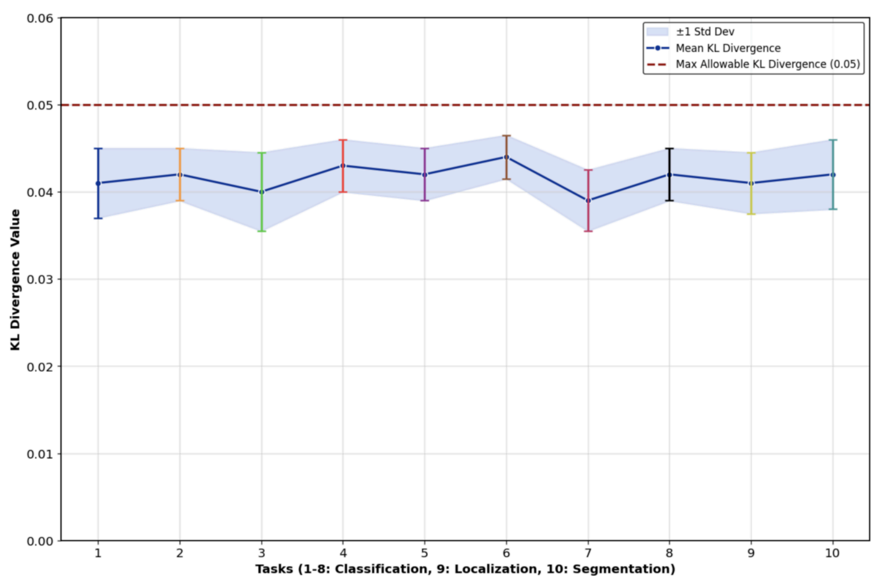

Figure 1. KL divergence value

Compared to MT-YOLOv7 and MultiTask-Transformer, MAMT2 improved average precision by 8.5% and 6.7% while reducing inference time by 32.1% and 27.8% respectively. By integrating dynamic task routing, causal attention and Transformers, MAMT2 balances accuracy and real-time performance for visual SHM, demonstrating significant engineering and academic value.

The target domain achieved expected performance across all six tasks in the test set, as showed below:

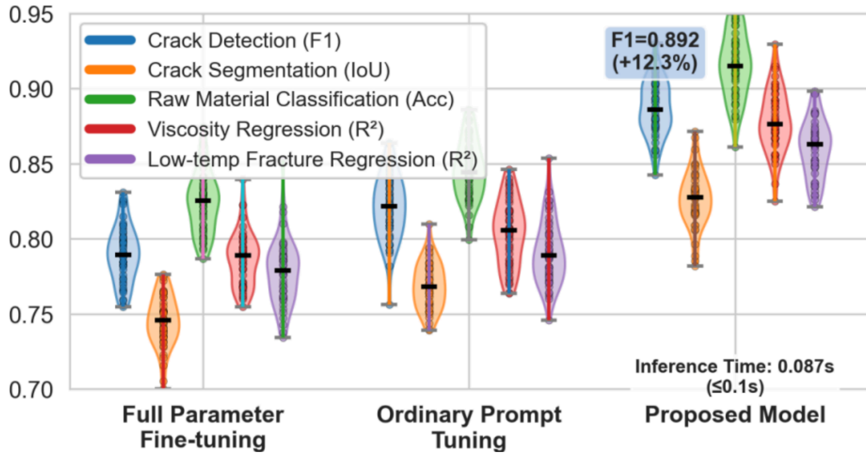

Figure 2. The performance of tasks

As shown in Figure 2, the proposed method effectively alleviates domain shift by extracting damage-invariant features, achieving a 12.3% improvement in crack detection F1-score and 9.7% improin track detection F1-score and 9.7% improvement in segmentation mloU compared to full-parameter fine-tuning. By extracting invariant damage features via cross-domain contrastive learning and introducing domain feature distribution regularization, it suppresses discrepancies caused by material attributes and imaging differences between concrete and bioasphalt, and realizes reliable cross-material model migration.

Furthermore, hierarchical prompt tuning combined with MAML meta-learning restrains overfitting under sparse target-domain annotations, achieving accurate microscopic damage detection with limited bioasphalt datasets and lowering experimental labeling costs.

Meanwhile, the model supports simultaneous execution of six tasks including crack identification, segmentation and performance regression, realizing integrated multi-task structural health monitoring for bioasphalt.

A total of 12 Pareto optimal solutions were acquired. The recommended optimal engineering parameters are: swelling duration 40 min, shear rate 4,500 r/min, shear temperature 170 ℃, shear duration 50 min.

Optimal Solution Performance and Cost shown below as shown in Figure 3.

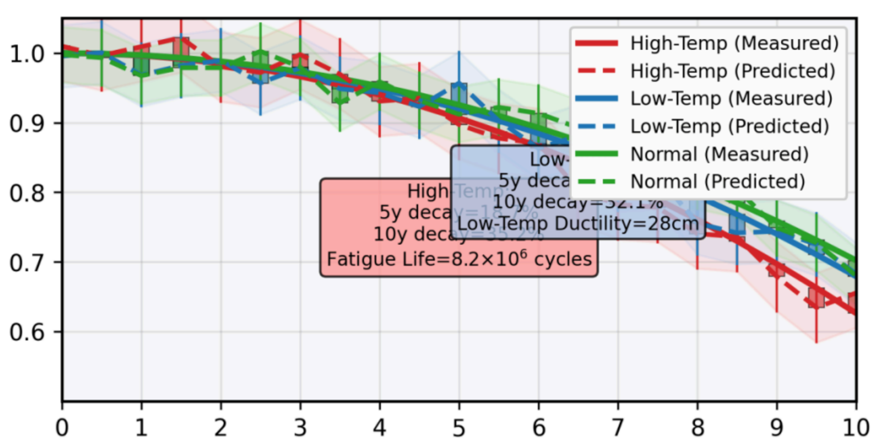

Figure 3. Tendency of service time as temp changed

Fully meeting all Chinese highway engineering specification constrains,with a 10-year fatigue life of 8.2×106 cycles, low-temperature ductility of 28 cm, and material cost reduced by 12% compared to conventional SBS modified asphalt.

LSTM Prediction Model demonstrates excellent prediction accuracy, precisely capturing the performance degradation patterns of ECMB.

Optimal engineering adaptation solutions as shown in Figure 4.

① High-temperature heavy-load scenarios: WCO 34%, RB 31%, Resin 30%, LDPE 5%, shear temperature 175°C, shear rate 4,800 r/min (emphasizing high-temperature stability and fatigue resistance);

② Low-temperature, high-precipitation scenarios: WCO 36%, RB 29%, Resin 30%, LDPE 5%, shear temperature 165°C, shear rate 4,200 rpm (emphasizing low-temperature crack resistance and water aging resistance);

③ Standard highway conditions: WCO 35%, RB 30%, Resin 31%, LDPE 4%, shear temperature 170°C, shear rate 4,500 rpm (balancing performance and cost).

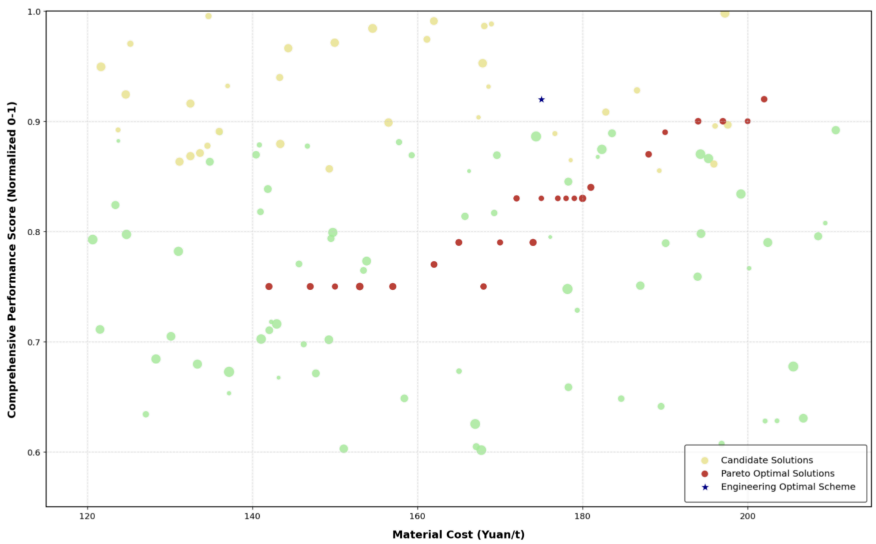

Figure 4. Pareto optimal solution set

We quantified environment-load coupling factors and developed an attention-based LSTM model to precisely capture ECMB performance degradation temporal patterns. This overcomes traditional models' limitations of low accuracy and neglected coupling effects, enabling accurate 5-10 year long-term performance forecasting and providing a scientific basis for pavement service life evaluation as shown in Figure 5.

By integrating previous formulation-process optimization results with a multi-criteria decision model, this study identified optimal adaptation schemes for diverse scenarios, resolving ECMB application challenges of insufficient scenario adaptability and cost-performance imbalance. Model validation ensures result reliability, facilitating ECMB's transition from laboratory optimization to practical engineering and providing comprehensive technical support for its large-scale promotion.

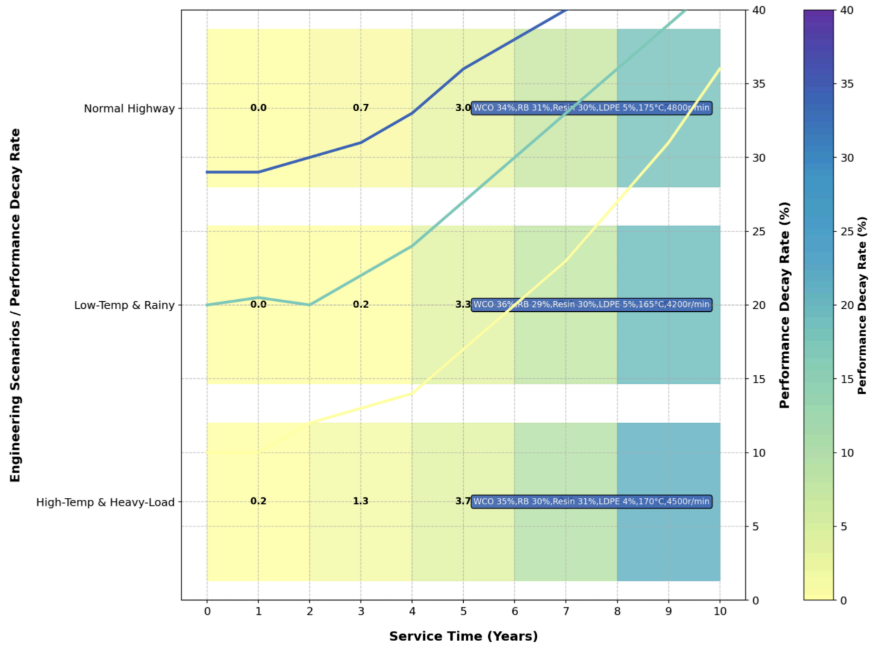

Figure 5. 10 years performance decay trend

8. Conclusions and extensions

Research Significance (Theoretical and Applied): The proposed model theoretically expands the application of multi-attribute, multi-task learning and cross-domain transfer methods, establishing a new paradigm for structural health monitoring and road material optimization. In practice, the research findings provide intelligent tools for transportation engineering, enabling efficient simultaneous detection of 10 SHM tasks with 0.087s single-image inference time,meeting the real-time monitoring requirement of $\leq$ 0.1s and multi-objective optimization of composite modified bio-asphalt formulations. For example, the proposed method can provide a basis for selecting and formulating eco-asphalt materials under various operating conditions on expressways, which is of great significance for improving road performance and reducing environmental impact.

Existing work has limitations in design and scope. Model training relies on the limited φ-NeXt dataset, and domain transfer validation is limited to only two materials—concrete and bio-asphalt—without considering the diversity of additional material types or real-world scenarios; the performance metrics (e.g., durability, cost) and environmental variables selected for formulation optimization are relatively limited; numerous factors (such as temperature and load variations) are assumed to be fixed or linear, and the universality and generalizability of the approach remain to be verified; furthermore, the hierarchical Prompt and meta-learning methods employed involve numerous parameters, which may lead to a risk of overfitting. This study did not conduct large-scale field tests and lacks an evaluation of the model's real-time performance and robustness (such as verifying damage detection accuracy in real traffic environments). Due to space constraints, a comparative analysis of different optimization algorithms and computational efficiency was not conducted in depth; these are areas that require further exploration in the future.

To address the above limitations, future research can be expanded in multiple directions: Expand datasets and application scenarios—collect more multimodal SHM data (such as damage images and field effects under different materials and complex environments), and validate model performance in actual road projects; In-depth mechanism modeling—incorporating more nonlinear factors (such as temperature-viscosity coupling and high-temperature aging) into material models to improve long-term prediction accuracy; Algorithm optimization—exploring more efficient few-shot learning and transfer learning strategies (such as self-supervised learning and graph neural networks) to reduce reliance on labeled data; Lightweight and Real-time Solutions — Develop lightweight model architectures and online learning algorithms to meet the requirements of embedded monitoring devices and real-time feedback; Multi-objective and Multi-scale Optimization — Incorporate additional objectives (e.g., service life, maintainability) and multi-stage processes (production-construction-operation and maintenance) into the optimization framework, and employ deep reinforcement learning or game theory methods to find global optimal solutions; Collaborative Research and Standardization — Establish unified standards and data-sharing platforms for SHM and material performance evaluation, and strengthen interdisciplinary collaboration. These directions will effectively address the shortcomings of current research and drive the field toward higher levels of development.

Authorship

Weiran Duan and Juanyong Qiu contributed equally to this work and should be considered co-first authors.

Weiran Duan: Conceptualization, Methodology, Software, Data curation, Writing - Original Draft, Visualization. Specifically, he was responsible for the construction of the φ-NeXt standardized dataset, the design of the MAMT2 hierarchical Transformer model, and the development of the cross-domain transfer adaptation strategy.

Juanyong Qiu: Validation, Formal analysis, Investigation, Resources, Writing - Review & Editing, Supervision. Specifically, he was responsible for the ECMB formulation multi-objective optimization framework, the long-term performance prediction model based on attention-LSTM, and the experimental verification of all proposed methods.

References

[1] Lv, G., Dong, F., Li, P., Wang, Y., Sheng, Y., Cheng, F., & Wang, B. (2026). Highway engineering construction safety risk evaluation based on decision tree algorithm. *International Journal of Computational Intelligence Systems, 19*(1), Article 166.

[2] Duan, M., & Jin, H. (2026). Characteristics of mixed traffic flow in highway tunnels as affected by changes in illumination. *Journal of Transportation Engineering, Part A: Systems, 152*(6), Article 04026026. https: //doi.org/10.1061/JTEPBS.TEENG-9275

[3] Hu, Q., Zou, Y., Wu, S., Zhang, S., Zhang, Y., & Wu, L. (2025). Exploring the impact of climate change on road fatalities: A macroscopic panel data analysis. *Journal of Advanced Transportation, 2025*(1), Article 4693354.

[4] Usanga, I. N., Inyang, E. O., & Ikeagwuani, C. C. (2026). Understanding the stress-strain performance of asphalt mixtures modified with palm kernel shell ash. *International Journal of Pavement Research and Technology.* Advance online publication. https: //doi.org/10.1007/s42947-026-00767-w

[5] Shen, P. (2026). Refined mixing strategy for Buton rock asphalt mixtures: Optimization of heating temperature and mixing sequence. *International Journal of Pavement Research and Technology*. Advance online publication. https: //doi.org/10.1007/s42947-026-00751-4

[6] You, Z., Mills-Beale, J., Fini, E., Goh, S. W., & Colbert, B. (2011). Evaluation of low-temperature binder properties of warm-mix asphalt, extracted and recovered RAP and RAS, and bioasphalt. *Journal of Materials in Civil Engineering, 23*(11), 1569–1574.

[7] Fini, E. H., Al-Qadi, I. L., You, Z., Zada, B., & Mills-Beale, J. (2012). Partial replacement of asphalt binder with bio-binder: Characterisation and modification. *International Journal of Pavement Engineering, 13*(6), 515–522.

[8] Hajj, E. Y., Souliman, M. I., Alavi, M. Z., & Salazar, L. G. L. (2013). Influence of Hydrogreen bioasphalt on viscoelastic properties of reclaimed asphalt mixtures. *Transportation Research Record: Journal of the Transportation Research Board, 2371*(1), 13–22.

[9] Yang, X., You, Z., Dai, Q., & Mills-Beale, J. (2014). Mechanical performance of asphalt mixtures modified by bio-oils derived from waste wood resources. *Construction and Building Materials, 51*, 424–431.

[10] Sunil, S., Holla, B. R. K., Pal, A., & Siddiqui, S. A. (2026). The changing landscape of concrete bridge infrastructure health monitoring and informed maintenance strategies: A decade of developments and trends. R*esults in Engineering, 30*, Article 110264.

[11] More, F., Fabbrocino, G., & Sandoli, A. (2026). Structural health monitoring of timber structures: Methods, sensors and real-world applications. *Procedia Structural Integrity, 78*, 944–951.

[12] Liu, Z. (2026). Smart sensors for structural health monitoring and nondestructive evaluation: 2nd edition. *Sensors, 26*(5), Article 1453.

[13] Huang, L., Fan, G., & Li, J. (2026). Bridge surface damage detection with structural context via boundary-aware component segmentation. *Engineering Structures, 358*, Article 122633.

[14] Gao, Y., Yang, J., Qian, H., & Mosalam, K. M. (2023). Multiattribute multitask transformer framework for vision-based structural health monitoring. *Computer-Aided Civil and Infrastructure Engineering, 38*(17), 2358–2377.

[15] Gao, J., Guo, G., Wang, H., Xu, N., & Jin, D. (2025). Multiscale characterization of bio-based polyurethane modified asphalt: Macro-micro experiments and molecular dynamics simulations. *Construction and Building Materials, 491*, Article 142699.

[16] Pan, W., Guan, H., Deng, J., & Yang, H. (2026). 'Rigid-Flexible' interpenetrating network: Performance regulation & multi-scale mechanism of lignin – Polyurethane composite asphalt. *Case Studies in Construction Materials, 24*, Article e05772.

[17] Sun, C., Li, Q., Liang, C., Zhu, X., Guo, S., Du, P., & Zhang, H. (2026). Improving compatibility of devulcanized rubber/bio-oil composite rejuvenated SBS modified asphalt using maleic anhydride-grafted polymer toward sustainable waste utilization. *Case Studies in Construction Materials, 24*, Article e05968.

[18] Shu, L., Cao, X., Shan, B., Deng, L., Jiang, T., Yuan, Y., Yang, X., & Tang, B. (2026). Dynamic bond-driven bio-based polyurethane-modified asphalt: Preparation, performance evaluation, and self-healing behavior. *Construction and Building Materials, 513*, Article 145416.

[19] Sagodi, A., Schniertshauer, J., & van Giffen, B. (2022). Engineering AI-enabled computer vision systems: Lessons from manufacturing. *IEEE Software, 39*(6), 51–57. https: //doi.org/10.1109/MS.2022.3189904

[20] Rivera-Castillo, J., Flores-Fuentes, W., Rivas-Lopez, M., Oleg Sergiyenko, Gonzalez-Navarro, F. F., Rodriguez-Quinonez, J. C., Hernandez-Balbuena, D., Lindner, L. H., & Basaca-Preciado, L. C. (2017). Experimental image and range scanner datasets fusion in SHM for displacement detection. *Structural Control & Health Monitoring, 24*(10), Article e1967.

[21] Toulaoui, A., Khalfi, H., & Hafidi, I. (2026). Efficient vision transformers via patch selective soft-masked attention and knowledge distillation. *Applied Soft Computing, 196*, Article 115151.

[22] Wei, J., Qiao, K., Zhu, D., Feng, Z., Wang, X., & Liu, Z. (2025). Dynamic reliability assessment method based on Gaussian process for engineering structures. *Structures, 80*, Article 109643.

[23] Kurnyta, A., Wrąbel, K., Baran, M., & Leski, A. (2026). Printed crack detection sensors for SHM based on direct ink write additive manufacturing. *Materials, 19*(5), Article 870.

[24] Elwalwal, H. M., Mahzan, S. B. H., & Abdalla, A. N. (2017). Crack inspection using guided waves (GWs)/structural health monitoring (SHM): Review. *Journal of Applied Sciences, 17*(8), 415–428.

[25] Rajender, A., Samanta, A. K., & Paral, A. (2024). Comparative study of corrosion-based service life prediction of reinforced concrete structures using traditional and machine learning approach. *International Journal of Structural Integrity, 16*(3), 591-621.

[26] Zhang, G., Li, H., Xiao, C., & Sobhani, B. (2022). Multi-aspect analysis and multi-objective optimization of a novel biomass-driven heat and power cogeneration system; utilization of grey wolf optimizer. *Journal of Cleaner Production, 355*, Article 131442.

[27] Yang, Y., Mei, H., & Fei, T. (2026). Multi-objective optimization of energy consumption and thermal-humidity environment in ice sports buildings: An integrated XGBoost-MLP framework. *Building and Environment, 295*, Article 114485.

[28] Alaeimoghadam, S., Karimi, M., & Mansourian, A. (2026). Optimizing a multi-scale cellular structure for urban growth modeling using NSGA-II. *Transactions in GIS, 30*(1), Article e70182.

[29] Guo, Y., Zhu, Q., & Mou, J. (2024). The path tracking control of unmanned surface vehicles based on an improved non-dominated sorting genetic algorithm II-based multi-objective nonlinear model predictive control method. *Journal of Marine Science and Engineering, 12*(12), Article 2188.

[30] Wang, L., Wang, L., & Chen, Y. (2026). Optimized model predictive controller using multi-objective whale optimization algorithm for urban rail train tracking control. *Biomimetics, 11*(1), Article 60.

[31] Song, L., Wang, Y., Shi, L., Yu, J., Li, F., & Xiang, S. (2025). Transformer with token attention and attribute prediction for image captioning. *Pattern Recognition Letters, 188*, 74–80.

[32] Chen, Y., Zhu, J., Long, J., Lin, H., Liu, H., & Xiang, B. (2026). Application of the dual-model interpretability framework of XGBoost-SHAP and GAT-GNNExplainer to investigate the impact of the built environment on traffic accidents at intersections. *Accident Analysis & Prevention, 224*, Article 108305.

[33] Wang, K., Jiang, J., He, H., & Chen, J. (2026). Spatial–temporal prediction of traffic accident risks using a fusion of GNN-LSTM. *Neural Processing Letters, 58*(3), Article 34.

[34] Gao, Y., Kong, B., & Mosalam, K. M. (2019). Deep leaf-bootstrapping generative adversarial network for structural image data augmentation. *Computer-Aided Civil and Infrastructure Engineering, 34*(9), 755–773.

[35] Huang, J.-X., Li, Q.-S., & Han, X.-L. (2026). Estimation of missing stress data for structural health monitoring based on multitask learning and multisource information. *Journal of Structural Engineering, 152*(3), Article 04026004. https: //doi.org/10.1061/JSENDH.STENG-14959

[36] Zhang, Q., Liu, Y., Zhang, Y., Zong, M., & Zhu, J. (2023). Improved YOLOv3 integrating SENet and optimized GIoU loss for occluded pedestrian detection. *Sensors, 23*(22), 9089.

[37] Tian, X., Li, B., & Pei, Z. (2025). The joint entity–relation extraction model based on hierarchical architecture for Chinese programming knowledge with complex semantics. International Journal of Web Information Systems, 22(21), 121-140.

[38] Jiang, R., Zhang, W., Zhou, X., Yuan, J., Zhou, X., Zuo, J., & Ran, M. (2025). Effects of biochar on the microstructure and micromorphology of bio-asphalt based on molecular dynamics and microscopic characterization. *International Journal of Pavement Engineering, 26*(1), 2587168

[39] Min, C., Ge, J., Li, X., Yang, Z., & Wen, G. (2026). Interpretable rate-of-penetration prediction using a Bayesian-optimized MAML-LSTM-Transformer under small-sample conditions. *Energy and AI, 24*, Article 100734.

[40] Nath, S. K., Zhang, Y., & Li, J. V. (2025). Earnings call scripts generation with large language models using few-shot learning prompt engineering and fine-tuning methods. Applied AI Letters, 6(1), Article e110.

[41] J, S. S., Chauhan, A., H, S., & Quadir Md, A. (2026). Hybrid facial deepfake detection framework using Swin V2 vision transformers and multi-scale textural descriptors. *Signal, Image and Video Processing, 20*(4), Article 233.

[42] Li, J., Wan, W., Feng, Y., & Chen, J. (2025). Meta-task interpolation-based data augmentation for imbalanced health status recognition of complex equipment. *Computers in Industry, 165*, Article 104226.

[43] Maghaireh, Z. N., AlAkash, H. M., Hamdiya, A. A., & Ababnah, M. (2026). A Gaussian process regression–based framework for optimizing green wall plant selection and indoor air quality performance in sustainable buildings. *Asian Journal of Civil Engineering, 27*, 3225–3240. https: //doi.org/10.1007/s42107-026-01665-z

[44] Xue, Y., Guo, Y., Zhang, H., & Chen, Y. (2026). Gaussian belief propagation with information-domain measurement refinement for multi-sensor multi-target tracking. *Digital Signal Processing, 177*, Article 106095.

[45] Toufiq, D. M., Sagheer, A. M., & Hadi Veisi. (2022). Brain tumor segmentation from magnetic resonance image using optimized thresholded difference algorithm and rough set. *TEM Journal, 16*(3), 631–638.

[46] Bae, M., Kang, J., & Kim, S. (2025). Enhanced long-term thermal performance prediction of glass wool vacuum insulation panels: A comprehensive accelerated testing approach. *Case Studies in Thermal Engineering, 75*, Article 107260.

[47] Klanatsky, P., Veynandt, F., Heschl, C., Stelzer, R., Zogas, P., Siokas, G., & Balomenos, A. (2025). Real long-term performance evaluation of an improved office building operation involving a data-driven model predictive control. *Energy and Buildings, 338*, Article 115590.