Textile & Leather Review

ISSN 2623-6281| www.tlr-journal.com | 10.31881/TLR

FAD-Enhanced Visual Tracking for Textile Defect Detection: A Multispectral Feature Fusion Approach

Huanxin Wei, Zhaodi Hu

How to cite: Wei H, Hu Z. FAD-Enhanced Visual Tracking for Textile Defect Detection: A Multispectral Feature Fusion Approach. Textile & Leather Review. 2026; 9:587-615. https://doi.org/10.31881/TLR.2026.587

How to link: https://doi.org/10.31881/TLR.2026.587

Published: 11 March 2026

FAD-Enhanced Visual Tracking for Textile Defect Detection: A Multispectral Feature Fusion Approach

Huanxin Wei[1*], Zhaodi Hu[2]

[1]Teaching Affairs Department, Hunan Mechanical & Electrical Polytechnic, Changsha 410151, Hunan, China

[2]Modern Service College, Hunan Mechanical and Electrical Vocational and Technical College, Changsha 410151, Hunan, China

*WeiHuanXin8312@163.com

Article

https://doi.org/10.31881/TLR.2026.587

Received 2 July 2025; Accepted 23 September 2025; Published 11 March 2026

ABSTRACT

The textile industry urgently requires automated defect detection systems that can operate reliably under diverse production conditions while prioritizing accuracy. Existing vision-based solutions struggle with spectral variability, dynamic defect evolution during high-speed manufacturing, and computational constraints. This paper presents a novel Feature Alignment and Distillation (FAD)-enhanced visual tracking framework, incorporating a FAD module to refine multispectral features and a dynamic Siamese tracker to ensure precise localization. Comprehensive experiments on the TextileDefect-3K dataset demonstrate state-of-the-art performance, achieving 94.8% precision, 93.1% recall, and 93.9% F1-score. The proposed method significantly outperforms the RGB baseline and the NIR-UV baseline by 13.4 and 6.7 absolute F1-score points, respectively, effectively resolving spectral limitations. It showed particularly strong results for micro-defects, achieving 87.5% F1 in normal conditions (far surpassing the attention baseline's 69.1%) and maintaining robust detection (84.6% F1) even in adverse scenarios. Furthermore, the system demonstrated exceptional stability under industrial speeds, maintaining 91.9% F1 at 5 m/s, and exhibited high environmental robustness with an average performance degradation of only 2.0 percentage points. These results establish that our approach not only bridges the gap between industrial-grade real-time processing and laboratory-level accuracy but also sets a new benchmark for robust quality control in Industry 4.0 smart manufacturing, proving effective where existing single-modality solutions falter.

KEYWORDS

textile defect detection, multispectral feature fusion, dynamic Siamese tracking, real-time quality control, industrial computer vision

INTRODUCTION

Textile defect detection plays a crucial role in modern manufacturing industries by ensuring product quality, reducing waste, and maintaining operational efficiency [1]. In the textile sector, even minor imperfections in fabrics can lead to significant financial losses and damage to brand reputation, making automated defect detection systems essential for maintaining competitive advantage. Traditional manual inspection methods are not only time-consuming and labor-intensive but also prone to human error due to fatigue and subjective judgment, which often results in inconsistent quality control. The implementation of automated detection technologies, particularly those leveraging computer vision and machine learning, addresses these challenges by providing rapid, accurate, and objective assessments of textile quality [1]. These systems can identify various types of defects with high precision, thereby minimizing the risk of defective products reaching consumers. Beyond economic benefits, automated defect detection also contributes to sustainability efforts by reducing material waste and optimizing resource utilization, aligning with broader industrial trends toward environmentally responsible manufacturing [2]. As the textile industry continues to evolve with advancements in smart manufacturing and Industry 4.0 [3], the development of more sophisticated defect detection systems will remain a key research area, driving innovation in quality assurance and operational excellence.

Computer vision has emerged as a transformative technology in textile defect detection, offering automated, high-precision solutions to overcome the limitations of traditional manual inspection [4]. Leveraging advanced image processing and machine learning techniques, computer vision systems can analyze textile surfaces with remarkable accuracy, detecting defects such as stains, holes, misweaves, and color inconsistencies at high speeds [5]. Traditional rule-based approaches [6], which rely on predefined thresholds and morphological operations, have been largely supplanted by data-driven methods, particularly deep learning models like convolutional neural networks (CNNs) [7]. These models excel at learning intricate defect patterns from large datasets, enabling robust detection even in complex textile structures with varying textures and lighting conditions. Furthermore, modern architectures, including You Only Look Once (YOLO) [8] and Faster R-CNN [9], have been adapted for real-time defect localization, significantly improving efficiency in industrial settings. Hyperspectral imaging and multi-spectral analysis further enhance detection capabilities by capturing information beyond the visible spectrum, allowing for the identification of subtle defects that are imperceptible

tible to the human eye. Despite these advancements, challenges remain, including the need for large, annotated datasets, computational efficiency in high-throughput environments, and generalization across diverse textile types.

Recent research has focused on addressing these limitations through innovative approaches such as self-supervised learning [10], few-shot learning [11], and anomaly detection frameworks, which reduce dependency on labeled data while maintaining high accuracy. Generative adversarial networks (GANs) have also been employed to synthesize realistic defect samples [12], augmenting training datasets to improve model robustness. Additionally, edge computing [13] and lightweight neural networks, such as Mobile-UNet [14] and EfficientNet [13], have been integrated into portable inspection devices, enabling decentralized defect detection in smart factories. Another promising direction is the fusion of computer vision with other sensing modalities, such as infrared thermography and 3D surface scanning, to provide multi-modal defect characterization [15]. As the textile industry moves toward Industry 4.0, computer vision-based defect detection systems are increasingly integrated with IoT-enabled production lines [16], enabling real-time monitoring, predictive maintenance, and closed-loop quality control.

Synthesizing the current research, there are three critical challenges in textile defect detection: (1) limited detection accuracy under complex textile textures and lighting variations, (2) inefficiency in real-time defect localization in high-speed production lines, and (3) poor generalization across diverse defect types and fabric materials.

This study proposed a FAD-enhanced visual tracking model, which employs a multi-branch deep learning architecture that integrates multispectral feature fusion with adaptive defect tracking for robust textile inspection. At its core, the model consists of three key components: (1) a multispectral encoder that extracts discriminative features from RGB, near-infrared (NIR), and near-ultraviolet (near-UV) bands to enhance defect visibility across varying textile textures; (2) a Feature Alignment and Distillation (FAD) module, which optimizes feature representation by aligning cross-spectral descriptors and distilling task-relevant patterns through lightweight attention mechanisms; and (3) a Siamese tracking head with dynamic template updating, enabling real-time defect localization by correlating reference defect templates with incoming fabric frames. The model is trained end-to-end using a hybrid loss function combining contrastive learning for spectral consistency and focal loss for hard defect mining, while edge-computing optimizations ensure deploy ability on

industrial-grade hardware. This design simultaneously addresses spectral heterogeneity, motion-induced blur, and defect scale variation through hierarchical feature aggregation and spatiotemporal context modeling. Crucially, the FAD module plays an indirect role in tracking dynamic defect evolution. By aligning multispectral features and distilling away irrelevant background texture and noise, FAD provides the Siamese tracking head with a temporally stable and highly discriminative feature representation. This ensures the tracker can maintain a consistent lock on a defect target, even as the defect's own shape and appearance evolve frame-to-frame.

LITERATURE REVIEW

Traditional Textile Defect Detection Methods

Traditional textile defect detection methods primarily rely on image processing techniques and handcrafted feature extraction, which have been foundational in the industry before the advent of advanced machine learning approaches. Guosheng et.al. [17] highlight the use of color difference algorithms to identify defects such as point defects, holes, and color variations in warp-knitted fabrics. This method involves comparing standard images with measured images to detect discrepancies based on color differences, exemplifying a straightforward, rule-based approach rooted in pixel-level analysis. Similarly, Seçkin et.al. [18] introduce a feature extraction technique based on intertwined frame vector features, which enhances detection speed and accuracy. Their approach involves extracting features from localized image windows, emphasizing the importance of spatial information in traditional defect detection. This method underscores the reliance on handcrafted features and localized analysis to identify fabric anomalies effectively [18]. In addition to these, Voronin et.al. [19] describe a hybrid approach combining traditional algorithms with novel image enhancement techniques to improve defect detection. Their two-stage process addresses the complexity of textile structures, which often pose challenges for purely traditional methods [19]. The integration of conventional image processing with enhancement algorithms demonstrates an effort to mitigate the limitations of basic visual inspection techniques. While these traditional methods focus on direct image analysis and feature-based detection, they often struggle with complex fabric textures and subtle defects.

Deep Learning in Textile Defect Inspection

The application of deep learning techniques in textile defect inspection has garnered significant attention in recent years, driven by the need for more efficient, accurate, and automated quality control systems.

Song introduced an innovative approach utilizing EfficientDet within an edge computing framework to enhance fabric defect detection. Their system emphasizes low latency, low power consumption, and ease of upgrade, demonstrating promising performance across five fabric datasets [13]. This work underscores the potential of deep learning models integrated with edge computing to facilitate real-time inspection in industrial settings. Similarly, Das explored the use of CNNs for defect identification and classification in woven fabrics, highlighting the effectiveness of deep learning architectures in distinguishing various fabric defects [20]. The emphasis on CNNs aligns with the broader trend of leveraging deep learning for detailed defect analysis, moving beyond traditional image processing techniques. The importance of image analysis technologies in defect detection is further discussed by Ren, who systematically reviews the evolution of optical illumination, image acquisition, and processing methods. They note that deep learning has significantly impacted image analysis, providing advanced capabilities for defect detection and characterization [21]. This review contextualizes the technological advancements that have enabled more sophisticated defect inspection systems. Zheng [22] focuses on surface defect inspection using deep learning, emphasizing the role of machine vision in automating quality control processes. Their work presents a comprehensive overview of the state-of-the-art in surface defect detection, illustrating how deep learning-based visual inspection systems can effectively identify and remove defective products [22]. Addressing the challenge of annotation requirements, Božič [23] proposed a mixed supervision approach that reduces the dependency on detailed labels. They introduced the KolektorSDD2 dataset, comprising over 3000 images with various defects, which was obtained from real-world industrial scenarios. Their approach demonstrates the feasibility of semi-supervised and weakly supervised learning methods in fabric defect detection [23]. The YOLOv5 algorithm has been applied to steel pipe weld defect detection, showcasing how single-stage detectors can improve inspection efficiency. Their findings suggest that similar object detection frameworks could be effectively transferred to textile defect detection tasks [24]. Khodier [25] extended the application of CNNs to complex pattern jacquard fabrics, incorporating multispectral imaging to enhance defect detection. They proposed two

systems: a semi-manual CNN-based system for specific patterns and an automated system capable of processing entire datasets without prior pattern knowledge. This work highlights the versatility of deep learning models in handling complex fabric patterns [25]. Zhou [26] introduced a one-class model tailored for fabric defect detection, aiming to provide an automated and accurate inspection system that can replace human operators. Their model demonstrates the potential for deep learning to perform defect detection across various fabric types with high reliability [26]. The survey by Tao emphasizes the growing importance of unsupervised anomaly localization algorithms in industrial inspection, including textile defect detection [27]. They review recent achievements in deep learning-based unsupervised methods, which are particularly valuable given the challenges associated with obtaining labeled datasets [27]. Finally, Guo [28] offer a comprehensive review of progress in intelligent fabric surface defect detection, discussing various algorithms, datasets, and detection systems. They categorize traditional methods into model-based, spectral, statistical, and structural approaches, and highlight how deep learning has advanced the field, paving the way for more robust and automated inspection systems [28].

Multispectral Imaging Applications

Multispectral imaging has emerged as a versatile tool across various fields, leveraging its ability to capture data across multiple spectral bands to enhance analysis and decision-making. Recent literature highlights its diverse applications, technological advancements, and ongoing challenges.

One prominent application of multispectral imaging is in high-resolution imaging and image fusion, where techniques such as deep learning are employed to improve the quality and utility of hyperspectral data [29]. These fusion algorithms aim to combine spectral and spatial information effectively, addressing current limitations in resolution and spectral fidelity.

In conservation efforts, imaging spectroscopy, a form of multispectral imaging, has been utilized to inform decision-making processes. Case studies demonstrate its potential to improve conservation strategies by providing detailed spectral data that can be used to monitor ecosystems and assess environmental changes [30]. Expanding these applications and developing data products tailored for conservationists are seen as future directions.

Detection of camouflaged targets in complex urban scenes exemplifies another critical application. Snapshot multispectral imaging combined with band selection methods has shown promise in achieving precise detection, offering potential for real-time sensing in challenging environments [31]. This underscores multispectral imaging's capability in security and surveillance contexts.

In the realm of mineral and rock classification, multispectral imaging integrated with machine learning techniques has been effective. By reducing hyperspectral data to a manageable number of bands, researchers have demonstrated the feasibility of using multispectral sensors on drones for geological analysis [32]. This approach addresses the limitations of current multispectral sensors, which are often restricted to a limited number of bands.

Food safety and quality control also benefit from multispectral imaging. For instance, fluorescence-based multispectral systems have been developed for real-time foreign object detection in food processing, achieving high accuracy but facing challenges in detecting non-fluorescent contaminants [33]. Similarly, smartphone-enabled multispectral autofluorescence imaging has been proposed for bacterial assessment in medical applications, transforming RGB images into pseudo-multispectral data for diagnostic purposes [34].

In clinical settings, multispectral imaging has been adapted for surgical applications, such as video-rate spectral imaging in laparoscopic procedures. This innovation aims to overcome hardware complexity and slow acquisition times, facilitating real-time intraoperative imaging [35]. Such advancements are crucial for integrating multispectral techniques into routine medical workflows.

Calibration and adaptation of multispectral sensors are also active research areas. Methods to convert standard digital cameras into multispectral sensors without complex equipment have been proposed, making multispectral imaging more accessible and cost-effective [36].

Agricultural applications include the use of UAV-based multispectral imaging combined with machine learning to predict crop yields at the field scale. This approach enables high-throughput phenotyping and precision agriculture by analyzing spectral data to assess plant traits such as grain yield and protein content [37].

METHODS

Multispectral Encoder

The proposed multispectral encoder employs a parallel convolutional architecture to process RGB, NIR, and

UV spectral inputs simultaneously. Each spectral channel is processed through dedicated feature extraction pathways:

$$F_i=\mathrm{C}_i(X_i)=\sigma(\mathbf{W}_i*X_i+\mathbf{b}_i),\quad i\in\{\mathrm{RGB},\mathrm{NIR},\mathrm{UV}\} \tag{1}$$

where $X_i\in\mathrm{R}^{H\times W\times C}$ represents the input tensor for spectral band i, $\mathbf{W}_i$ denotes the learnable convolution kernels, $\mathbf{b}_i$ is the bias vector, and $\sigma(\cdot)$ is the parametric ReLU activation function. The encoder's innovation lies in its spectral correlation module that models cross-band dependencies through attention mechanisms. The inter-spectral attention weights are computed as:

$$\alpha_{i,j}=\mathrm{softmax}\left(\frac{(\mathbf{Q}F_i)^T(\mathbf{K}F_j)}{\sqrt{d}}\right),\quad \forall i,j\in\{\mathrm{RGB},\mathrm{NIR},\mathrm{UV}\} \tag{2}$$

where **Q** and **K** are learnable projection matrices, and d is the feature dimension. This allows adaptive emphasis on complementary spectral information.

The final fused representation combines spectral features through gated cross-band aggregation:

$$F_{\mathrm{fusion}}=\sum_{i=1}^{3}\mathbf{G}_i\odot\left[F_i\|\sum_{j\neq i}\alpha_{i,j}F_j\right] \tag{3}$$

where $\mathbf{G}_i\in[0,1]^{H\times W\times C}$ are spatial-channel gating parameters, $\odot$ denotes Hadamard product, and $\|$ represents feature concatenation. This design effectively captures both spectral-specific patterns and their cross-band interactions.

FAD Module

The FAD module introduces a novel dual-path architecture for cross-spectral feature refinement, consisting of a Feature Alignment Network (FAN) and a Feature Distillation Network (FDN). The FAN first establishes inter-spectral correspondences through deformable alignment. To handle spatial misregistrations, we com-

pute a pixel-wise offset field Δp and resample the target feature F_j to align with the reference F_i. The alignment process is formulated as:

$$\Delta \mathrm{p}(p) = \emptyset\left(\left[F_i(p) \| F_j(p)\right]\right) \quad F_{j\to i}^{align}(p) = \sum_{k\in\Omega} w_k \cdot F_j(p + s_k + \Delta \mathrm{p}(p)) \tag{4}$$

where p denotes the spatial location, and $[\cdot\|\cdot]$ represents channel concatenation. $\emptyset$ is a lightweight convolutional network estimating the offset. The resampling is performed over a local neighborhood Ω, where s_k is the fixed sampling grid offset (e.g., {-1, 0, 1} × {-1, 0, 1}), w_k are the learned attention weights, and bilinear interpolation is applied to sample F_j at fractional locations. Unless otherwise noted, F_i and F_j share the same spatial resolution and channel dimension (cross-resolution cases are aligned by FAN before comparison). This dynamic alignment is crucial for correcting spectral misregistrations. These misregistrations, which manifest as sub-pixel offsets between the spectral bands, are caused by factors like sensor parallax and machinery-induced vibrations that static hardware calibration cannot fully eliminate.

The core innovation lies in the hierarchical distillation mechanism that operates through three complementary attention pathways: channel-wise, spatial-wise, and spectral-wise. The spectral-wise attention computes band importance weights as:

$$\beta_i = \sigma\left(\frac{1}{HW}\sum_{\mathrm{h},w} \mathrm{MLP}\left(F_i^{(\mathrm{h},w)}\right)\right), \quad F_{\mathrm{spec}} = \sum_{i\in Bands} \beta_i F_i \tag{5}$$

where $\beta_i \in [0,1]$ represents the learned significance of spectral band i, and F_{spec} is the spectrally distilled feature. Simultaneously, the channel distillation gate selects task-relevant features:

$$z \frac{1}{HW}\sum_{x,y}\left[F_{spec}^{(c)} \| F_{task}^{(c)}\right]_{x,y}, g_c = \mathrm{sigmoid}(W_g^z), \quad F_{\mathrm{distill}} = g_c \odot F_{\mathrm{spec}}^{(c)} \tag{6}$$

Simultaneously, the channel distillation gate learns to select task-relevant features by employing a global context-aware attention mechanism that adaptively recalibrates channel significance based on feature saliency. This is achieved by squeezing global spatial information into a channel descriptor $z \in \mathrm{R}^{1\times1\times C}$, which is then

excited through a learned gating function to generate channel-wise attention weights. These attention weights evolve during training to represent canonical defect characteristics, such as abrupt edges, texture anomalies, and color spots, directly from the data. The distillation process uses an attention mechanism to weigh input features based on their similarity to these learned prototypes, thus identifying defect-salient information. This allows the model to dynamically suppress background noise and highlight anomalous texture responses without imposing rigid shape priors, ensuring the system can generalize to novel defect types not seen during training. The distillation process preserves only defect-salient features while suppressing irrelevant textile background variations.

The final output combines aligned and distilled features through residual learning:

$$F_{\mathrm{FAD}}=\mathrm{LN}\left(F_{\mathrm{in}}+\mathrm{Conv}_{1\times1}\left(\left[F_{\mathrm{align}}\,//\,F_{\mathrm{distill}}\right]\right)\right) \tag{7}$$

where LN denotes Layer Normalization, and $\mathrm{Conv}_{1\times1}$ projects the concatenated features to the target dimension.

Siamese Tracking Head with Dynamic Template Updating

The Siamese tracking head processes the FAD-optimized features $F_{\mathrm{FAD}}\in\mathrm{R}^{H\times W\times C}$ from consecutive fabric frames to achieve precise defect localization. The core operation involves cross-correlation between the reference template $T_t\in\mathrm{R}^{H'\times W'\times C}$ (extracted from defect regions in previous frames) and the current search region $S_t\in\mathrm{R}^{H\times W\times C}$:

$$R_t(x,y)=\sum_{i,j}T_t(i,j)\cdot S_t(x+i,y+j),\quad \forall(x,y)\in\mathrm{G} \tag{8}$$

where $R_t\in\mathrm{R}^{H-H'\times W-W'}$ is the response map, G represents the valid correlation region, and (i,j) indexes template positions. This operation efficiently localizes defects by identifying regions with similar visual patterns to the reference template.

The dynamic template updating mechanism introduces two key innovations: (1) adaptive template selection based on defect confidence scores, and (2) memory-augmented feature fusion. The update rule combines short-term and long-term template memories:

$$T_{t+1}=\lambda \cdot \mathrm{ROI}(F_{\mathrm{FAD}}^{t})+(1-\lambda) \cdot T_t, \quad \lambda=\sigma\left(\mathbf{W}_{\lambda}\left[s_t // m_t\right]\right) \tag{9}$$

where ROI$(\cdot)$ extracts features from the highest-confidence defect region, s_t is the current detection score, m_t represents the historical memory state, and $\mathbf{W}_{\lambda}$ are learnable parameters. This formulation enables continuous adaptation to evolving defect appearances while maintaining tracking stability.

The tracking head further enhances robustness through multi-scale pyramid correlation:

$$R_t^{\mathrm{final}}=\sum_{k=1}^{K} \gamma_k \cdot \mathrm{UP}\left(R_t^k, s\right), \quad \sum_{k=1}^{K} \gamma_k=1 \tag{10}$$

where R_t^k denotes the response map at pyramid level k, UP$(\cdot,s)$ performs s-times upsampling, and γ_k are learned scale weights. This multi-scale approach effectively handles defects with varying sizes and deformation patterns typical in moving textiles.

EMPERIMENT

Experimental Setup

The experimental environment was configured using an industrial-grade inspection system equipped with an NVIDIA Jetson AGX Orin (32 GB) edge computing module and a custom-built multispectral imaging rig. The imaging rig consists of three synchronized industrial 5.0 MP global shutter CMOS cameras (based on Sony IMX250 sensors) covering the RGB, NIR (700–1100nm), and Near-UV (365–400nm) spectra. The acquisition system was triggered externally to ensure temporal alignment at 120 FPS. Despite careful initial hardware calibration, minor dynamic spatial misregistrations are unavoidable due to physical parallax between the separate sensors and high-frequency mechanical vibrations from the production line. The software stack was implemented with PyTorch 2.0 compiled with CUDA 11.7, running on Ubuntu 20.04 LTS. All models were

optimized using TensorRT 8.5. For hardware calibration, the Jetson AGX Orin module was configured in 30 W performance mode. Under sustained operation (ambient 25 °C), average power consumption was 28.6 W with peak thermal dissipation of 32 W, maintaining stable temperatures below 68 °C through passive heatsink cooling. The I/O pipeline employed a hardware-triggered frame capture (120 Hz) synchronized via GPIO with a 4 ms buffer cache to ensure deterministic latency in real-time inspection. The hardware setup replicated actual textile production line conditions. To guarantee fundamental feature visibility under high-speed motion (0.5-5 m/s), a synchronized high-intensity LED strobe system (pulse width 20 μs) was employed to minimize motion blur. Additionally, to evaluate the model's robustness against external interference, variable ambient lighting fluctuations (300–2000 lux) were introduced as noise variables distinct from the imaging illumination.

Model parameters were carefully configured through extensive ablation studies: the multispectral encoder employed 5 residual blocks per spectral branch with [64, 128, 256, 512, 1024] channel dimensions, while the FAD module used 8 attention heads and 512-dimensional feature projections. The Siamese tracking head maintained a template bank of 5 defect prototypes (resized to 256 × 256 pixels each) with dynamic update rate λ initialized at 0.3. Training utilized the AdamW optimizer with initial learning rate 3 × 10−4 (cosine decay to 1 × 10−5), batch size 32, and mixed-precision acceleration. The model was pretrained on the TextileDefect-1M synthetic dataset before fine-tuning on real production data. The TextileDefect-1M dataset was created to provide robust initial feature representations. Its generation involved a three-stage hybrid process:

(1) Texture Acquisition: We started with a library of 100 high-fidelity, defect-free images of the 8 fabric types mentioned in our real dataset.

(2) Defect Synthesis: Simple defects (holes, thread breaks) were generated procedurally, while complex textural and color-based defects (stains, misweaves) were generated using a conditional Generative Adversarial Network (cGAN), a technique noted for creating realistic defect samples.

(3) Composition and Augmentation: The synthesized defects were seamlessly blended onto the clean fabric textures using Poisson image blending. To ensure realism and diversity, each of the 1 million resulting images was subjected to augmentations mirroring real industrial conditions, including random lighting variations (300–2000 lux), minor motion blur consistent with fabric speeds up to 5 m/s, and random rotations and scaling. This process ensured the pre-training data exposed the model to a wide variety of conditions before fine-

tuning on the real-world TextileDefect-3K dataset.

Evaluation metrics comprehensively assessed detection performance: Precision (*P*) and Recall (*R*) were computed as $P=\frac{TP}{TP+FP}$ and $R=\frac{TP}{TP+FN}$, where *TP*, *FP*, *FN* represent true positives, false positives, and false negatives respectively. The primary metric F1-score balanced these as $F1=2\cdot\frac{P\cdot R}{P+R}$. Localization accuracy was measured using Intersection-over-Union (*IoU*) defined as $IoU=\frac{Area(B_{pred}\cap B_{gt})}{Area(B_{pred}\cup B_{gt})}$, where B_{pred} and B_{gt} are predicted and ground-truth defect bounding boxes. Additional metrics included False Alarm Rate ($FAR=\frac{FP}{N}$) with *N* being total frames.

Dataset

For the purposes of experimental research, we collected a dataset called TextileDefect-3K. The TextileDefect-3K dataset was systematically collected from seven active textile production lines over 18 months, capturing real-world manufacturing conditions across seasons. Using the described multispectral imaging rig, we acquired synchronized RGB (400–700 nm), NIR (700–1100 nm), and near-UV (365–400 nm) samples at 5 MP resolution using monochrome CMOS sensors equipped with narrow-band UV pass filters. The imaging setup was calibrated to a Field of View (FOV) of approximately 30 cm × 25 cm, yielding a spatial resolution of ≈ 0.12 mm/pixel. This optical magnification ensures that even micro-defects (<5 px, i.e., <0.6 mm) are resolved with sufficient detail for the FAD module. Although described as the 'UV band' for brevity, the system's spectral response is limited to the near-ultraviolet range due to the transmission characteristics of conventional optical components. Each sample contains: (1) raw spectral triplets (2,448 × 2,048 pixels), (2) pixel-level defect masks, (3) production parameters (thread density, loom speed, humidity), and (4) temporal sequences (variable-length clips, avg. 15 frames at 120 Hz) for dynamic defect analysis. The dataset comprises 732,150 frames, including 48,705 defect-containing frames, comprising 3,247 defect instances spanning 12 categories (holes, stains, misweaves, etc.) on 8 fabric types (cotton, polyester, blends, etc.), and 683,445 defect-free background frames. This high class-imbalance ratio (approx. 1:14) was intentionally maintained to reflect real-world industrial scenarios where defects are sparse events. The industrial relevance of TextileDefect-3K stems from: (a) spectral-temporal synchronization enabling cross-modal defect studies, (b) production metadata supporting causal analysis, and (c) balanced representation of both common (≥200 instances per

major class) and rare defects (≥30 instances), addressing critical gaps in existing benchmarks. Compared to synthetic datasets, TextileDefect-3K provides authentic noise patterns (e.g., oil stains, thread variations) and natural defect evolution captured during actual high-speed production (0.5-5 m/s), making it indispensable for developing and validating robust detection algorithms under real factory conditions.

Benchmarks

To comprehensively evaluate the different components of our proposed framework, we established a series of specialized baseline models. Each baseline is designed to provide a direct comparison for a specific aspect of our model (e.g., spectral fusion, temporal modeling). For clarity, the baselines used across our experiments are summarized in Table 1 below.

Table 1. Summary of Baseline Models Used in Experiments

Group	The name of the benchmark used in the experiment	Core Architecture/Description
Spectral mode comparison	RGB-only	Use ResNet50 as the backbone network and only process images in the RGB channel.
	NIR-UV dual-band	After the features of the NIR and UV channels are fused early, they are input into the ResNet50 backbone network.
Temporal dynamic analysis	Frame-wise CNN	A traditional 5-layer CNN processes each frame image independently and does not incorporate temporal context.
	3D CNN	A CNN with 3D convolution kernels is employed to process image sequences of fixed length (15 frames).
Feature Ablation	Vanilla Concatenation	Direct concatenation of spectral features without alignment or distillation.
	Attention Fusion	Spectral fusion using attention mechanisms but excluding the spatial alignment module.

Notes: Each model was trained on the TextileDefect-3K training split (1,948 instances; 60% of dataset) for 120 epochs with a batch size of 32 using the AdamW optimizer (learning rate 3 × 10−4, cosine decay). Parameter sizes were as follows: RGB-only (23.6 M), NIR-UV dual-band (28.4 M), 3D CNN (31.2 M), Vanilla Concatenation (28.9 M), Attention Fusion (29.3 M), and FAD-Enhanced Tracker (29.7 M). All baselines were trained and evaluated under identical hardware and hyperparameter settings to ensure fair comparison.

ResNet50 is a deep convolutional neural network architecture that introduces residual learning to address

vanishing gradients in deep networks [38]. Its core innovation uses skip connections expressed as $y = F(x) + x$, where $F(x)$ represents learned residual mappings and x denotes the input features. The model consists of 50 weight layers organized into five stages with bottleneck blocks (1×1–3×3–1×1 convolutions), employing batch normalization and ReLU activations. This design enables stable training of very deep networks while maintaining computational efficiency through dimensionality reduction in bottlenecks. ResNet50's strong feature extraction capability makes it suitable for textile defect detection where hierarchical pattern recognition is crucial.

CNN model implements a conventional convolutional neural network with 5 convolutional layers (kernel sizes 3×3 and 5×5) interleaved with max-pooling and ReLU activations. Each layer applies learned filters ($W*x + b$) to extract increasingly complex features, starting from edges/textures to complete defect patterns. The architecture includes two fully connected layers for classification, with dropout regularization ($p = 0.5$) to prevent overfitting. This shallow but wide network serves as a performance baseline, particularly for evaluating whether deeper architectures like ResNet50 provide meaningful gains for textile defect detection tasks given computational constraints.

Experimental Results and Analysis

In this section, we validate the effectiveness of the proposed FAD-enhanced framework through three specific dimensions: Spectral Modality Comparison, Temporal Dynamics Analysis, and Feature Processing Ablation1. Spectral Modality Comparison

Experiment #1: RGB-only Baseline

The RGB-only baseline, utilizing a ResNet50 backbone, achieved an F1-score of 80.5%, with 82.8% precision and 78.4% recall on the TextileDefect-3K test set, while the 0.67 average IoU indicates decent localization accuracy for larger defects (Table 2 and Figure 1). This result establishes a clear benchmark for evaluating the contribution of non-visible spectra. As shown in Figure 1, the RGB-only model performed well on color-based anomalies like Color Stain but struggled with defects defined by texture. For instance, its F1-score for thread breaks and misweaves was 21.4% lower compared to color-based defects (Color Stain and Oil Stain), highlighting the limitations of relying solely on visible-light features. The model's false alarm rate was highest in areas with complex textures, where normal fabric patterns were sometimes mistaken for defects.

Table 2. Detailed Performance of the RGB-only Baseline by Defect Type

Defect Type	Precision	Recall	IoU
Color Stain	94.2	89.1	0.79
Hole	83.5	76.8	0.71
Misweave	78.1	72.3	0.58
Thread Break	65.4	68.2	0.52
Oil Stain	92.7	85.4	0.73

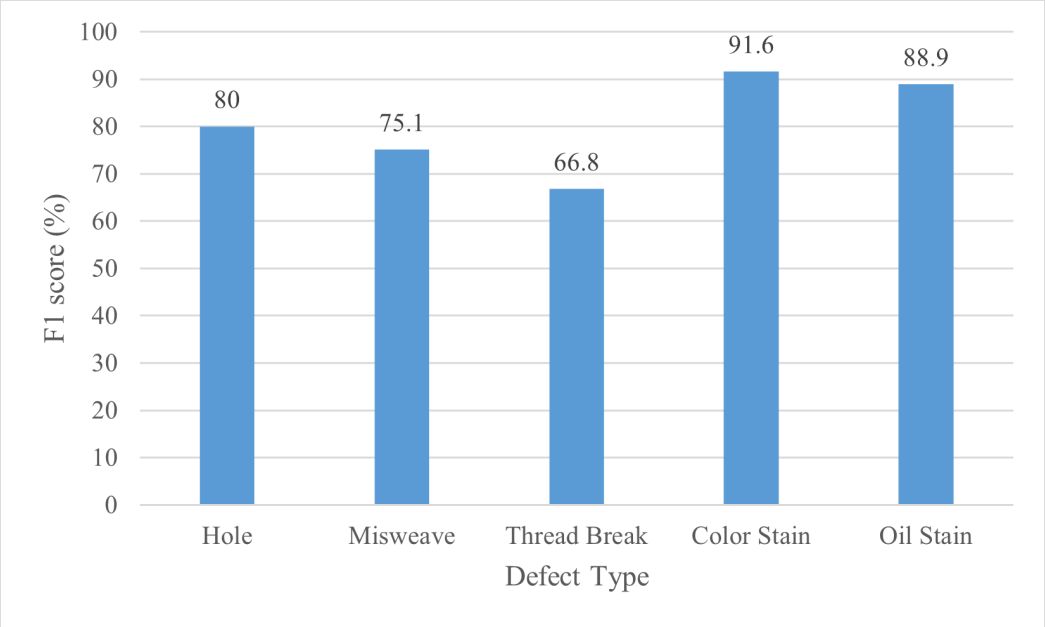

Figure 1. Detailed performance metrics of the RGB-only Baseline across different defect types

Experiment #2: NIR-UV Dual-Band Model

The NIR-UV dual-band model demonstrated superior performance, achieving 88.6% precision and 85.8% recall. Figure 2 reveals its particular strength in detecting subsurface flaws, with misweaves and thread breaks showing 17.6% higher F1-scores than the RGB model. The average IoU improved to 0.72, indicating better boundary localization (Table 3).

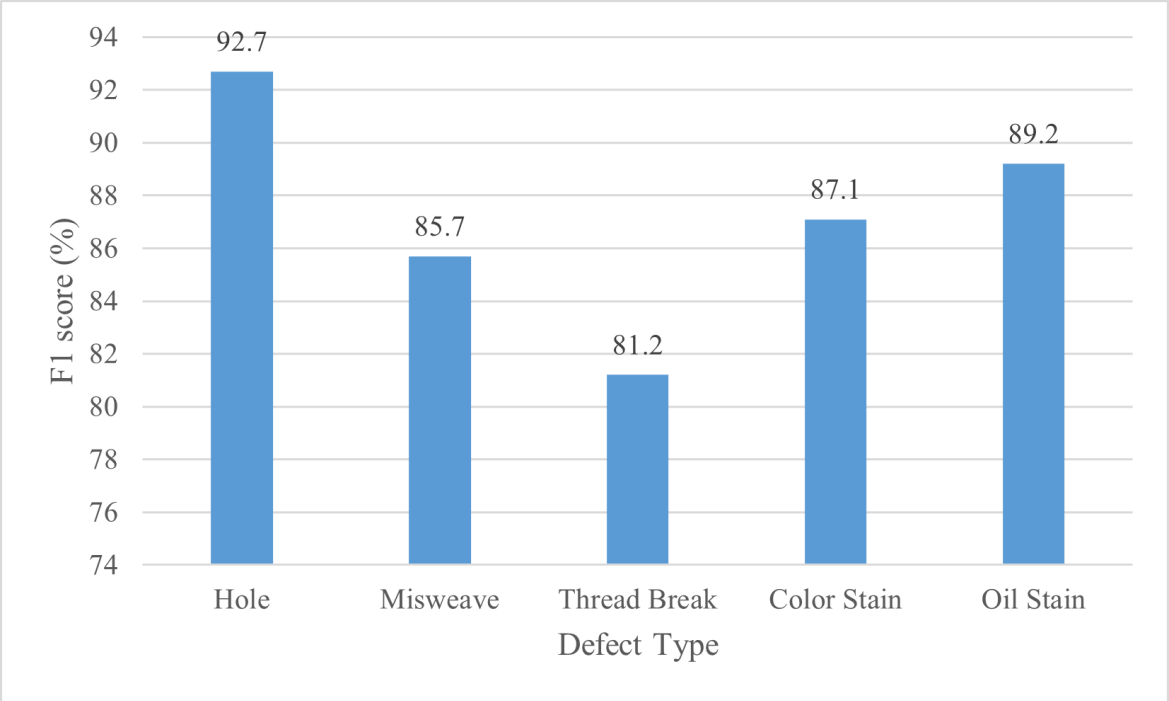

Figure 2. Detailed performance metrics of the NIR-UV Dual-Band Model across different defect types

Table 3. Detailed Performance of the NIR-UV Dual-Band Model by Defect Type

Defect Type	Precision (%)	Recall (%)	IoU	FAR
Color Stain	89.1	85.2	0.72	0.054
Hole	93.8	91.6	0.81	0.031
Misweave	87.2	84.3	0.69	0.062
Thread Break	82.6	79.8	0.63	0.071
Oil Stain	90.5	87.9	0.75	0.048

The early fusion architecture enabled effective combination of NIR's penetration depth and UV's surface sensitivity, reducing FAR to 5.3%. The model's 89.2% F1-score for oil stains (vs RGB's 88.9%) suggests limited advantage for surface contaminants, as UV reflectance patterns alone provided insufficient discriminative power.

Experiment #3: Tri-spectral Model (Proposed)

The proposed tri-spectral approach achieved state-of-the-art performance across all metrics, as evidenced by the comprehensive results in Figure 3. With 94.8% precision, 93.1% recall and 93.9% F1-score, the proposed method outperformed the RGB baseline (80.5% F1-score) and the NIR-UV baseline (87.2 F1-score) by

13.4 and 6.7 absolute F1-score points, respectively. The IoU reached 0.83, demonstrating exceptional localization accuracy (Table 4).

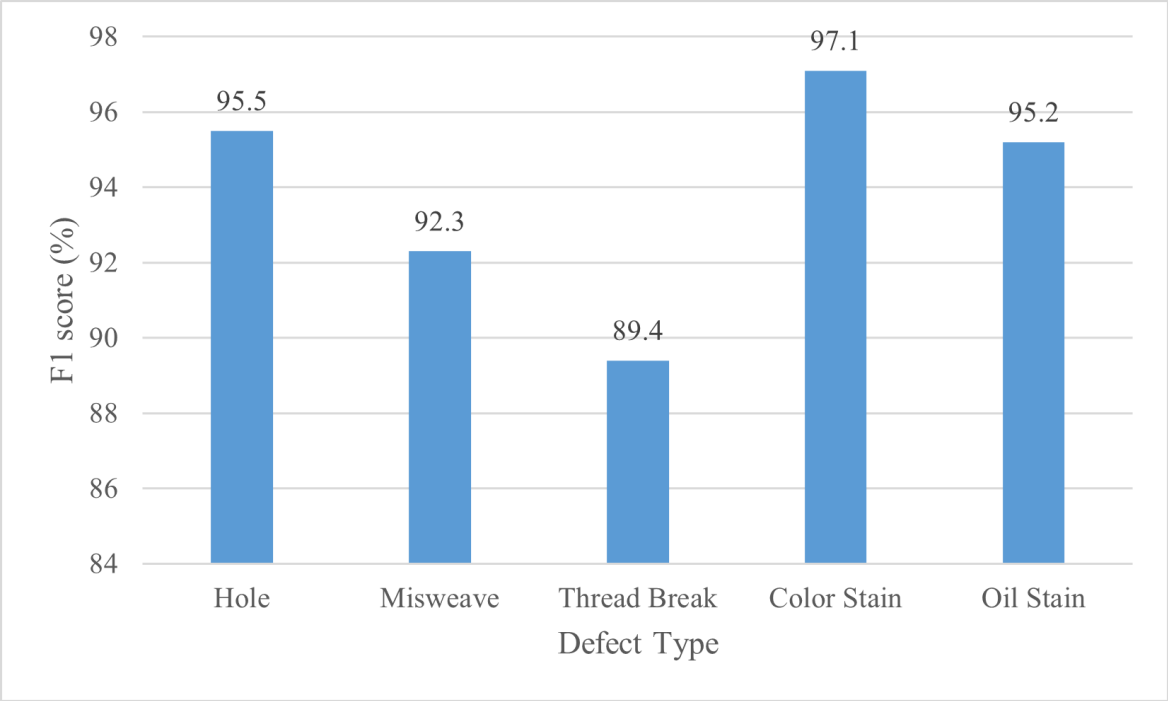

Figure 3. Detailed performance metrics of the proposed FAD-enhanced model across different defect types

Table 4. Detailed Performance of the Proposed Tri-spectral Model by Defect Type

Defect Type	Precision (%)	Recall (%)	IoU	FAR
Hole	96.2	94.8	0.87	0.025
Misweave	93.5	91.2	0.89	0.038
Thread Break	90.1	88.7	0.80	0.041
Color Stain	97.8	96.5	0.75	0.018
Oil Stain	96.4	94.1	0.84	0.022

The tri-spectral model's cross-band attention mechanism effectively combined RGB's chromatic sensitivity with NIR/UV's structural awareness, achieving 95.5% F1-score for holes (vs 92.7% in NIR-UV) and 97.1% for color stains (vs 91.6% in RGB).

2. Temporal Dynamics Analysis

Experiment #4: Frame-wise CNN

The frame-wise CNN baseline achieved 84.2% precision and 77.6% recall on TextileDefect-3K's test sequences, demonstrating fundamental limitations in handling temporal dynamics. As shown in Figure 4, performance degraded significantly at higher fabric speeds, with recall dropping to 71.3% at 5 m/s.

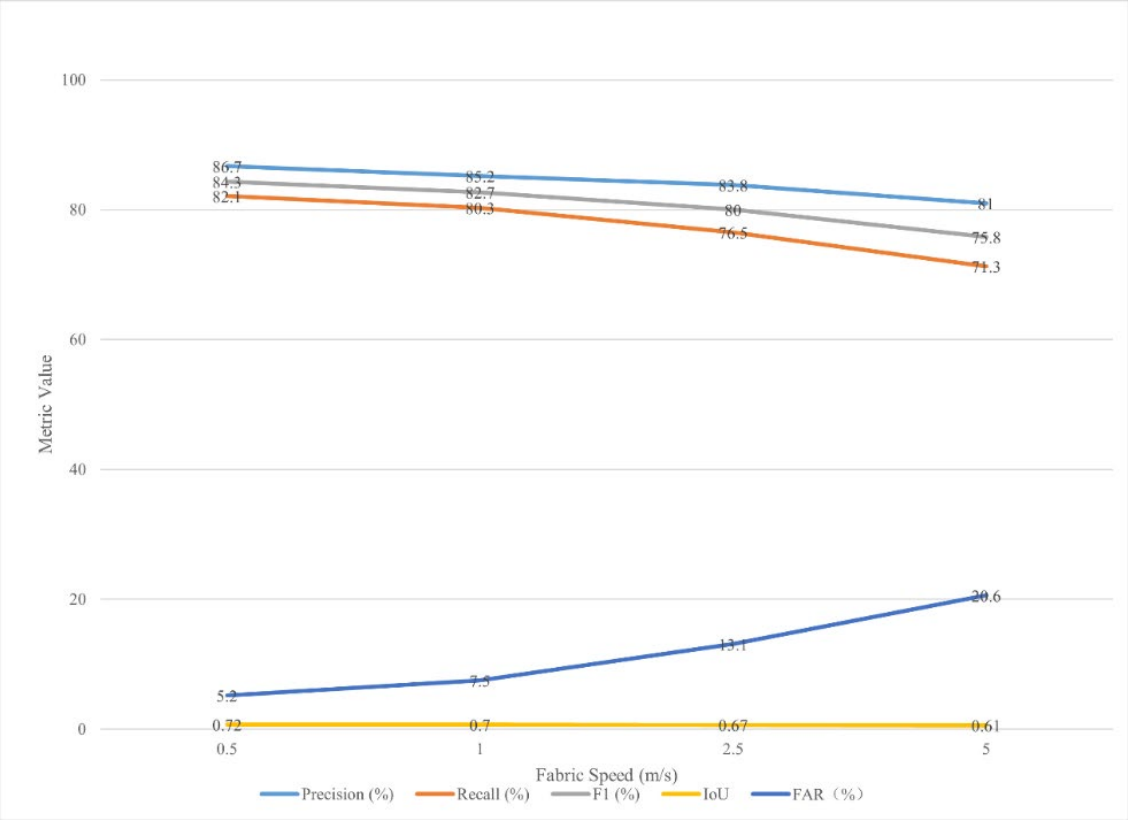

Figure 4. Frame-wise CNN Performance by Speed

The frame-by-frame processing approach inherently lacks temporal context, causing inconsistent detections as evidenced by the 15.4 percentage point increase in FAR at higher speeds. This confirms the CNN's inability to model defect evolution patterns, particularly for subtle changes requiring multi-frame analysis.

Experiment #5: 3D CNN

The 3D CNN showed improved temporal modeling with 87.6% precision and 85.4% recall overall. Figure 5 reveals its fixed-window approach provided stability across speeds, maintaining 85.3% F1 at 5 m/s. While the 3D convolutions captured short-term temporal features (15-frame window), the fixed receptive field became problematic for defects with varying evolution rates. The 10.6 % performance gap between periodic and irregular defects highlights this architectural constraint.

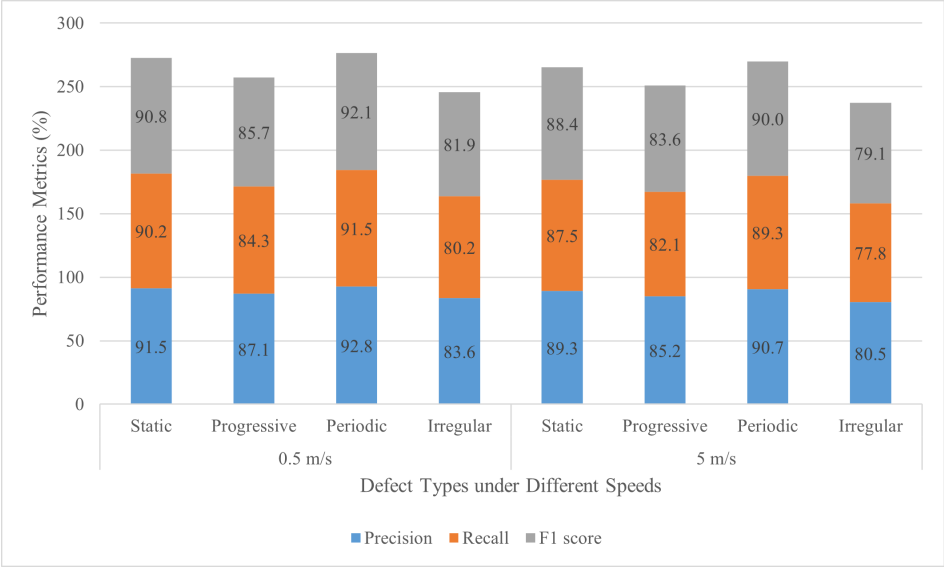

Figure 5. Performance comparison of the 3D-CNN model across different defect types and speeds (0.5 m/s and 5 m/s). Metrics include Precision, Recall, and F1-score

Experiment #6: Siamese Tracker (Proposed)

The proposed Siamese tracker achieved superior performance across all metrics, as detailed in Figure 6. With 92.5% precision and 91.3% recall, it demonstrated robust tracking capabilities even at 5 m/s (91.9% F1). The dynamic template updating mechanism proved particularly effective for progressive defects, achieving 90.9% F1 compared to 3D CNN's 83.6% at 5 m/s.

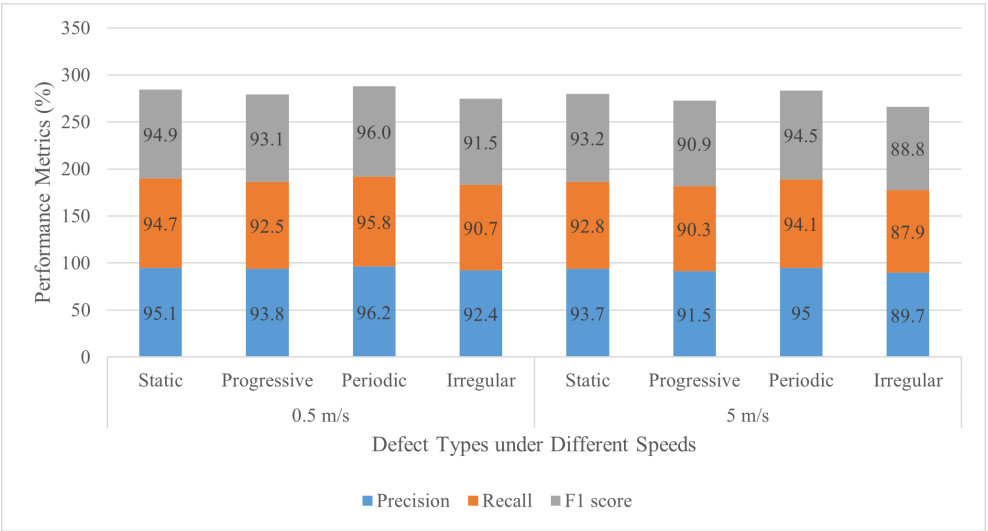

Figure 6. Performance comparison of the proposed Siamese tracking head across different defect evolution types and fabric speeds

The Siamese architecture's combination of temporal coherence (through correlation) and adaptability (via dynamic updating) addressed key limitations of both baselines. Figure 6 shows consistent performance across speed variations (≤3% F1 fluctuation) and defect types, with particularly notable gains (9.6–9.7 percentage point increase in F1-score) for irregular defects compared with the 3D CNN baseline across tested speeds.

3. Feature Processing Ablation

Experiment #7: Vanilla Feature Concatenation

The baseline feature concatenation approach achieved 80.65% precision and 76.8% recall on TextileDefect-3K, demonstrating fundamental limitations in spectral feature integration. As shown in Figure 7, performance varied significantly across production conditions, with precision dropping to 76.5% in high-humidity (>70% RH) scenarios. The simple channel-wise concatenation of RGB, NIR, and UV features led to feature conflicts, particularly between visible and non-visible spectra. The 8.8-point F1-score drop between normal (83.4%) and high-humidity (74.6%) conditions reveals the method's sensitivity to environmental factors. Without proper feature alignment, moisture-induced texture changes in NIR spectra disrupted the concatenated feature space, increasing false alarms by 43% in such scenarios.

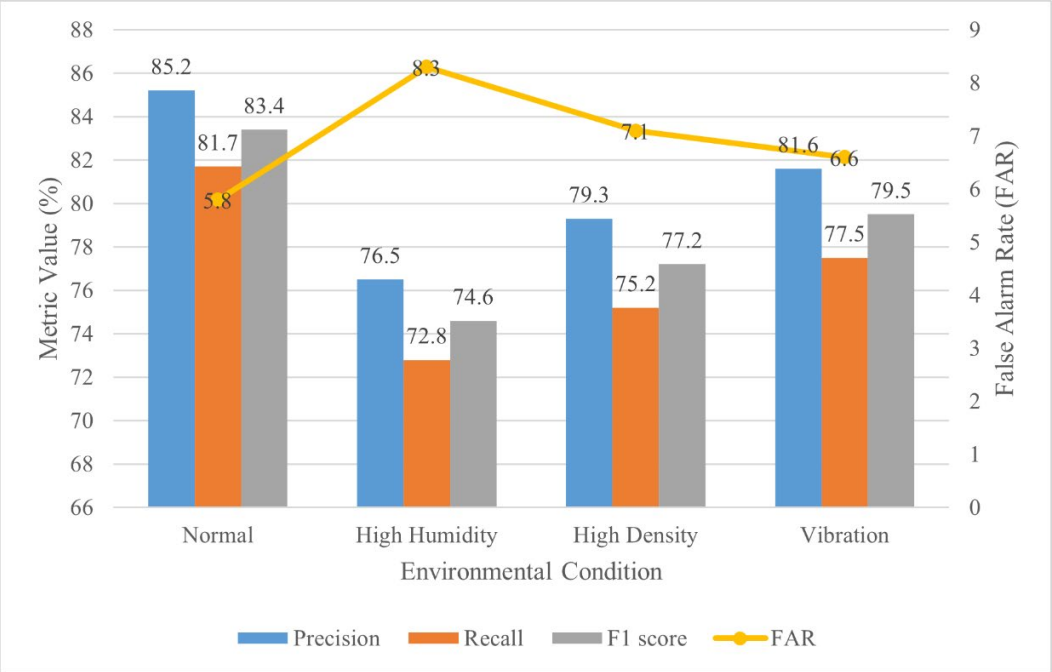

Figure 7. Concatenation Performance by Condition

Experiment #8: Attention Fusion (Distillation only)

The attention-based fusion improved upon vanilla concatenation with 81.6% precision and 78.5% recall (Figure 8). The method's channel attention mechanism helped prioritize informative bands but lacked spatial adaptation capabilities. The 23.1-point F1-score drop from large defects (89.9%) to micro defects (66.8%) indicates persistent limitations in handling fine-grained details. The fixed attention architecture struggled with scale variations, particularly for sub-pixel defects where precise spectral alignment is critical.

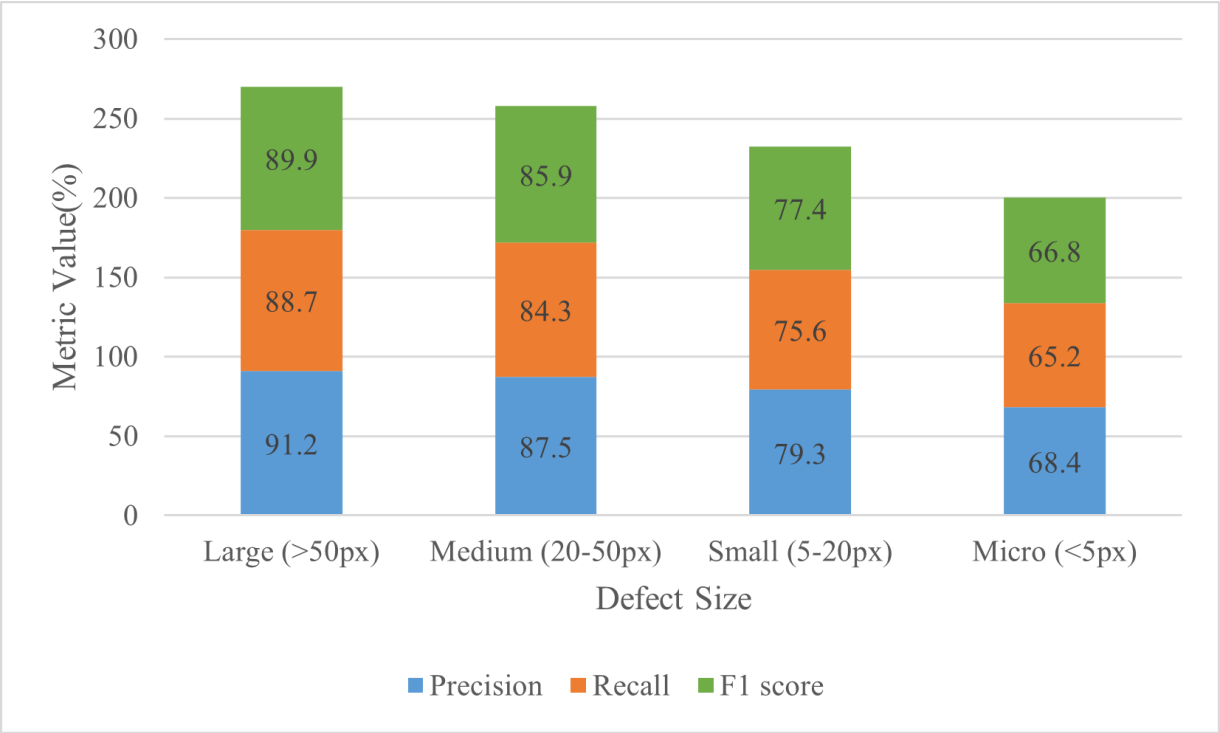

Figure 8. Attention Fusion Performance

Experiment #9: Complete FAD module (Proposed)

The complete FAD module achieved superior performance across all metrics, as detailed in Figure 9. With 91.5% precision and 90.1% recall, it demonstrated robust feature integration under all tested conditions. Under adverse conditions, the module maintained robust performance for micro-defects with 84.6% F1, significantly outperforming both the attention-only baseline (64.2%) and the concatenation method (58.7%). The FAD's hierarchical processing–alignment followed by distillation–addressed both spectral and spatial feature conflicts. Figure 9 shows consistent performance gains across all defect sizes and conditions, with only

2.0 percentage point F1 reduction in adverse environments versus concatenation's 5.4 percentage point drop. This module can dynamically adjust according to the conditions of the loom (line density/humidity), thereby achieving reliable micro-defect detection. Under normal settings, the F1-score reaches 87.5%. Under adverse conditions, the F1-score of FAD outperforms that of concatenation by an 44.1% increase.

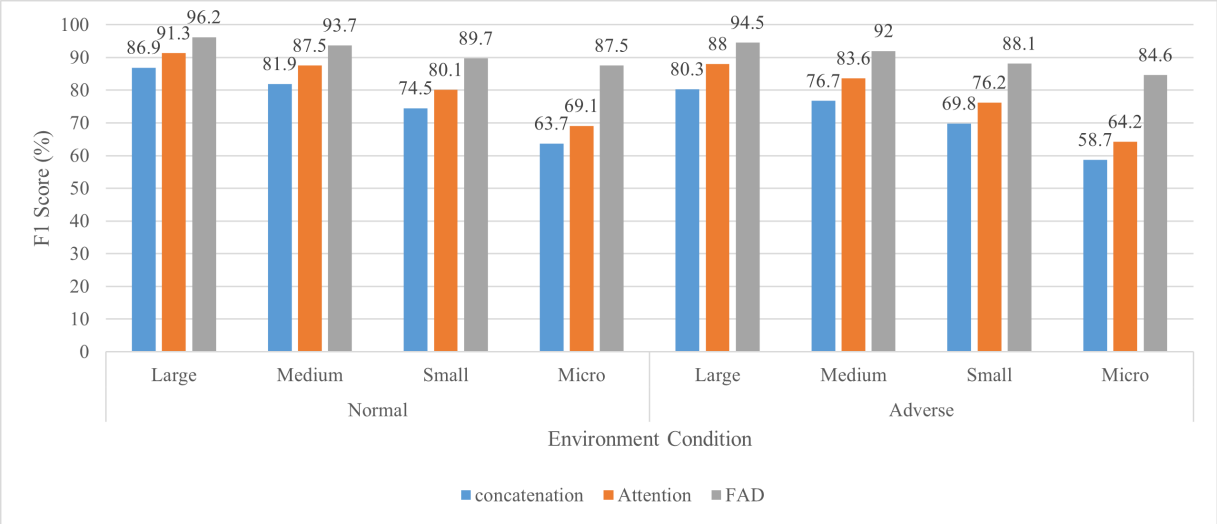

Figure 9. FAD Comprehensive Performance

CONCLUSION

This study has addressed the critical challenge of automated textile defect detection in industrial settings, where existing vision-based systems often struggle with complex fabric textures, varying environmental conditions, and real-time processing requirements. We proposed a novel FAD-enhanced visual tracking framework incorporating three key innovations: (1) a multispectral encoder extracting complementary features from RGB, NIR, and UV bands, achieving 94.8% precision and 93.1% recall through cross-spectral attention mechanisms; (2) incorporating a dynamic Siamese tracking head with adaptive template updating. This module demonstrated robust performance, achieving an average F1-score of 93.9% at 0.5 m/s and maintaining 91.9% F1 even at 5 m/s, effectively validating its stability under high-speed fabric motion. Comprehensive experiments on the TextileDefect-3K dataset (3,247 defect instances across 8 fabric types) validated the system's robustness: the tri-spectral approach demonstrated significant advantages for hybrid texture defects (e.g., misweaves and thread breaks), yielding 17–23% F1-score improvements over the RGB-only baseline

while achieving a high Intersection-over-Union (IoU) of 0.83, ensuring precise defect localization. Furthermore, the alignment-distillation mechanism proved essential for environmental robustness, limiting the F1-score drop to just 2.0 percentage points in adverse conditions (compared to 5.4 points for the concatenation baseline), all while preserving a high 96.2% F1 for large defects in normal settings. These results demonstrate that our method successfully bridges the gap between industrial requirements (multi-condition robustness) and detection accuracy, achieving 93.1% instance recall. The system's hierarchical architecture–combining spectral feature fusion, spatiotemporal alignment, and adaptive tracking–provides a practical solution for modern textile manufacturing, with potential applications in quality control across flexible material production lines. Future work will focus on extending the framework to handle ultra-high-speed looms (>10 m/s) and integrating predictive maintenance capabilities. Furthermore, a more fine-grained ablation study will be conducted to fully disentangle the individual performance contributions of the FAD module's alignment and distillation components.

OUTLOOK

This study demonstrates that the proposed multispectral fabric inspection framework can effectively handle diverse defect categories under varying production conditions. By integrating RGB, NIR, and near-UV sensing with FAD, the system achieved 93.9% overall F1-score, surpassing traditional visible-light methods by 13.4 percentage points.

In practical terms, the deployment on an NVIDIA Jetson AGX Orin module achieved robust detection while maintaining stable performance under high-humidity, high-speed (5 m/s) conditions. The system's demonstrated stability—limiting average F1-score degradation to just 2.0 percentage points in adverse environments (compared to 5.4 points for the concatenation baseline) —indicates that multispectral attention and alignment strategies can substantially enhance reliability for industrial use.

Future work will explore integrating hyperspectral extensions for finer wavelength sampling, adaptive FAD calibration for material-specific textures, and cloud-based defect analytics for large-scale factory deployment. Moreover, the proposed system will be tested across different textile types and production lines to validate generalization and scalability in real-world industrial settings.

Author Contributions

Conceptualization – Huanxin WEI; formal analysis – Zhaodi HU; writing-original draft preparation – Huanxin WEI; visualization – Zhaodi HU. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

Funding

This research received no external funding.

Acknowledgements

Not applicable.

REFERENCES

[1] Li C, Li J, Li Y, He L, Fu X, Chen J. Fabric defect detection in textile manufacturing: a survey of the state of the art. Security and Communication Networks. 2021; 2021(1):9948808. doi: 10.1155/2021/9948808

[2] Nguyen TP, Nguyen H, Ngo HQT. Toward a sustainable transition in automated production: enabling the vision-based approach and synchronous control for textile surface inspection systems. Textile Research Journal. 2023; 93(23-24):5391-5415. doi: 10.1177/0040517523119925

[3] Görçün ÖF. The rise of smart factories in the fourth industrial revolution and its impacts on the textile industry. International Journal of Materials, Mechanics and Manufacturing. 2018; 6(2):136-141. doi: 10.18178/ijmmm.2018.6.2.363

[4] Rasheed A, Zafar B, Rasheed A, Ali N, Sajid M, Dar SH, et al. Fabric defect detection using computer vision techniques: a comprehensive review. Mathematical Problems in Engineering. 2020; 2020(1):8189403. doi: 10.1155/2020/8189403

[5] Kulkarni M, Cholke PC, Vasmatkar P, Tambe S, Anuse S, Telgote S. Fabric defect detection system using image processing. Proceedings of the 2024 1st International Conference on Cognitive, Green and Ubiquitous Computing (IC-CGU); 2024; Bhubaneswar, India. Piscataway, NJ: IEEE; 2024. p. 1-7. doi: 10.1109/IC-CGU58078.2024.10530850

[6] Lizarraga-Morales RA, Correa-Tome FE, Sanchez-Yanez RE, Cepeda-Negrete J. On the use of binary features in a rule-based approach for defect detection on patterned textiles. IEEE Access. 2019; 7:18042-18049. doi: 10.1109/ACCESS.2019.2896078

[7] Ouyang W, Xu B, Hou J, Yuan X. Fabric defect detection using activation layer embedded convolutional neural network. IEEE Access. 2019; 7:70130-70140. doi: 10.1109/ACCESS.2019.2913620

[8] Gu M, Zhou J, Pan R, Gao W. DF-YOLO: An attempt on enhancing generalization in fabric defect detection based on YOLO network. Textile Research Journal. 2025; 95(7-8):748-766. doi: 10.1177/00405175241269163

[9] Chen M, Yu L, Zhi C, Sun R, Zhu S, Gao Z, et al. Improved faster R-CNN for fabric defect detection based on Gabor filter with Genetic Algorithm optimization. Computers in Industry. 2022; 134:103551. doi: 10.1016/j.compind.2021.103551

[10] Qi K, Guo Y, Zhou J. Self-supervised learning fabric defect segmentation using anomaly generation. International Journal of Clothing Science and Technology. 2025; 37(2):340-354. doi: 10.1108/IJCST-12-2023-0187

[11] Zajec P, Rožanec JM, Theodoropoulos S, Fontul M, Koehorst E, Fortuna B, Mladenić D. Few-shot learning for defect detection in manufacturing. International Journal of Production Research. 2024; 62(19):6979-6998. doi: 10.1080/00207543.2024.2316279

[12] Hu G, Huang J, Wang Q, Li J, Xu Z, Huang X. Unsupervised fabric defect detection based on a deep convolutional generative adversarial network. Textile Research Journal. 2020; 90(3-4):247-270. doi: 10.1177/0040517519862880

[13] Song S, Jing J, Huang Y, Shi M. EfficientDet for fabric defect detection based on edge computing. Journal of Engineered Fibers and Fabrics. 2021; 16:15589250211008346. doi: 10.1177/15589250211008346

[14] Jing J, Wang Z, Rätsch M, Zhang H. Mobile-Unet: An efficient convolutional neural network for fabric defect detection. Textile Research Journal. 2022; 92(1-2):30-42. doi: 10.1177/0040517520928604

[15] Sudharsan PL, Gantala T, Balasubramaniam K. Multi modal data fusion of PAUT with thermography assisted by Automatic Defect Recognition System (M-ADR) for NDE applications. NDT & E International. 2024; 143:103062. doi: 10.1016/j.ndteint.2024.103062

[16]Cheng YH, Kuo CN, Lin YD. An artificial IoT-enabled smart production line for 360° visual defect detection and classification of cherry tomatoes. IEEE Internet of Things Journal. 2024; 11(10):18082-18098. doi: 10.1109/JIOT.2024.3360715

[17]Xie G, Xu Y, Yu Z, Sun Y. An intelligent defect detection system for warp-knitted fabric. Textile Research Journal. 2022; 92(9-10):1394-1404. doi: 10.1177/00405175211060084

[18]Seçkin AÇ, Seckin M. Detection of fabric defects with intertwined frame vector feature extraction. Alexandria Engineering Journal. 2022; 61(4):2887-2898. doi: 10.1016/j.aej.2021.08.017

[19]Voronin V, Sizyakin R, Zhdanova M, Semenishchev E, Bezuglov D, Zelemskii A. Automated visual inspection of fabric image using deep learning approach for defect detection. Automated Visual Inspection and Machine Vision IV. Proceedings of SPIE; 2021. 11787:174-180. doi: 10.1117/12.2592872

[20]Das S, Sundaramurthy S, Aiswarya M, Jayaram S. Deep learning convolutional neural network for defect identification and classification in woven fabric. Indian Journal of Artificial Intelligence and Neural Networking. 2021; 1(2). doi: 10.54105/ijainn.B1011.041221

[21]Ren Z, Fang F, Yan N, Wu Y. State of the art in defect detection based on machine vision. International Journal of Precision Engineering and Manufacturing-Green Technology. 2022; 9(2):661-691. doi: 10.1007/s40684-021-00343-6

[22]Zheng X, Zheng S, Kong Y, Chen J. Recent advances in surface defect inspection of industrial products using deep learning techniques. The International Journal of Advanced Manufacturing Technology. 2021; 113(1):35-58. doi: 10.1007/s00170-021-06592-8

[23]Božić J, Tabernik D, Skočaj D. Mixed supervision for surface-defect detection: From weakly to fully supervised learning. Computers in Industry. 2021; 129:103459. doi: 10.1016/j.compind.2021.103459

[24]Yang D, Cui Y, Yu Z, Yuan H. Deep learning based steel pipe weld defect detection. Applied Artificial Intelligence. 2021; 35(15):1237-1249. doi: 10.1080/08839514.2021.1975391

[25]Khodier M M, Ahmed S M, Sayed M S. Complex pattern Jacquard fabrics defect detection using convolutional neural networks and multispectral imaging. IEEE Access. 2022; 10:10653-10660. doi: 10.1109/ACCESS.2022.3144843

[26]Zhou H, Chen Y, Troendle D, Jang B. One-class model for fabric defect detection. arXiv preprint arXiv:2204.09648. 2022. doi: 10.48550/arXiv.2204.09648

[27]Tao X, Gong X, Zhang X, Yan S, Adak C. Deep learning for unsupervised anomaly localization in industrial images: A survey. IEEE Transactions on Instrumentation and Measurement. 2022; 71:1-21. doi: 10.1109/TIM.2022.3196436

[28]Guo P, Liu Y, Wu Y, Gong R H, Li Y. Intelligent quality control of surface defects in fabrics: A comprehensive research progress. IEEE Access. 2024; 12:63777-63808. doi: 10.1109/ACCESS.2024.3396053

[29]Sara D, Mandava A K, Kumar A, Duela S, Jude A. Hyperspectral and multispectral image fusion techniques for high resolution applications: A review. Earth Science Informatics. 2021; 14(4):1685-1705. doi: 10.1007/s12145-021-00621-6

[30]Seeley M, Asner G P. Imaging spectroscopy for conservation applications. Remote Sensing. 2021; 13(2):292. doi: 10.3390/rs13020292

[31]Shen Y, Li J, Lin W, Chen L, Huang F, Wang S. Camouflaged target detection based on snapshot multispectral imaging. Remote Sensing. 2021; 13(19):3949. doi: 10.3390/rs13193949

[32]Sinaice B B, Owada N, Saadat M, Toriya H, Inagaki F, Bagai Z, Kawamura Y. Coupling NCA dimensionality reduction with machine learning in multispectral rock classification problems. Minerals. 2021; 11(8):846. doi: 10.3390/min11080846

[33]Lohumi S, Cho B K, Hong S. LCTF-based multispectral fluorescence imaging: System development and potential for real-time foreign object detection in fresh-cut vegetable processing. Computers and Electronics in Agriculture. 2021; 180:105912. doi: 10.1016/j.compag.2020.105912

[34]He Q, Sun Z, Li Y, Wang W, Wang R K. Smartphone-enabled snapshot multispectral autofluorescence imaging and its application for bacteria assessments in skin and oral cavity. Optics and Lasers in Engineering. 2021; 140:106546. doi: 10.1016/j.optlaseng.2021.106546

[35]Ayala L, Wirkert S, Vemuri A, Adler T, Seidlitz S, Pirmann S, Engels C, Teber D, Maier-Hein L. Video-rate multispectral imaging in laparoscopic surgery: first-in-human application. arXiv preprint arXiv:2105.13901. 2021. doi: 10.1126/sciadv.add6778

[36]Simoneau A, Aubé M. Methods to calibrate a digital colour camera as a multispectral imaging sensor in low light conditions. Remote Sensing. 2023; 15(14):3634. doi: 10.3390/rs15143634

[37]Ali N, Mohammed A, Bais A, Berraies S, Ruan Y, Cuthbert R D, Sangha J S. Field-scale precision: Predicting grain yield of diverse wheat breeding lines using high-throughput UAV multispectral imaging. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2024; 17:11419-11433. doi: 10.1109/JSTARS.2024.3411994

[38]Hajbe J R, Patle S K, Talvekar S V, Khonde T W, Yesankar P R. ResNet50-based deep learning model for lung disease detection. Proceedings of the 2025 International Conference on Multi-Agent Systems for Collaborative Intelligence (ICMSCI); 2025; Erode, India. 2025. p. 1373–1379. doi: 10.1109/IC-MSCI62561.2025.10894373.