Pertanika J. Sci. & Technol. 34 (S1): 111 - 131 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Weighted Bi-directional Feature Pyramid Network and Segment Anything Model (SAM) Specific Knowledge Injection for Fabric Defect Detection

Suresh Kumar* and J. Avanija

Department of Computer Science and Engineering, School of Computing, Mohan Babu University, Sree Sainath Nagar, 517102 Tirupati, Andhra Pradesh, India

ABSTRACT

The quality control in the textile manufacturing industry involves identifying fabric defects. The differences in cloth texture and the limited number of damaged samples are also major problems when it comes to the correct identification of faults in fabrics. In most applications, visual inspection is still a vital component of the quality control process. The ability to check surface defects is an essential part of the quality control methods of fabrics and textiles. A fault can be described as whatever interrupts the pattern of fabric repetition. To solve this problem, this paper proposes a supervised learning approach to the detection of fabric flaws. The main purpose of this study will be to design a rotation-invariant fabric defect detection plan that will attain a high detection rate related to current techniques. The study introduces a novel framework for identifying defects in fabrics by fusing a Weighted Bi-directional Feature Pyramid Network (BiFPN) and the Segment Anything Model (SAM), which have been produced via Specific Knowledge Injection. The Weighted BiFPN enhances the fusion of multi-scale features by assigning dynamic weights to the features at various levels, which makes it possible to detect both minor and large-scale flaws in fabric. At the same time, the Segment Anything Model (SAM) is a sophisticated segmentation model that is fine-tuned with knowledge injection to adapt its generic segmentation abilities to the peculiarities of fabric textures and defects. The combination will guarantee accurate localisation and categorisation of the flaws, even under difficult conditions that may include changes in illumination, texture, and type of defects. Experiments carried out on standard textile datasets prove that the suggested BiFPN-SAM framework is considerably more efficient in terms of accuracy, Intersection over Union (IoU), and performance. The fact that the hybrid

ARTICLE INFO

Article history:
Received: 07 March 2026
Accepted: 20 April 2026
Published: 24 July 2026

DOI: https://doi.org/10.47836/pjst.34.S1.06

E-mail addresses:
ogsureshkumar45@gmail.com (Suresh Kumar)
avans75@yahoo.co.in (J.Avanija)
* Corresponding author

e-ISSN: 2231-8526

method can generalise to different faults also underscores its usefulness in textile quality control systems through automated means.

Keywords: Control, deep learning, fabric defect detection, multi-scale feature fusion, segment anything model, specific knowledge injection, textile quality, weighted BiFPN

INTRODUCTION

The purpose of fabric defect detection is to identify and detect any surface faults on fabrics due to machine malfunctions or operational failures. To modern fabric organisations, it forms one of the essential ways of reducing expenses and enhancing the competitiveness of the products. Human eye investigation is an outdated detection technology with a success rate of 60-75% (Zhou et al.,2024). Nevertheless, this approach does not guarantee the results of the tests to be unbiased or consistent. The limitations can be overcome, the expenses minimised, and a 90 per cent detection success rate is realised with an automatic fabric inspection system. Hence, it is necessary to enhance a quick, precise and non-destructive method of detecting flaws on fabric (Zhang et al., 2022).

Various production imperfections that mostly involve machine breakdown cause the fabric to be thrown away as waste and cause financial losses to Pakistani textile manufacturers (Zhou et al.,2025). They need to be inspected to ascertain whether the fabric manufactured is of the market standards. This process, especially in the developing world, is done manually and involves a great number of qualified labourers. Manual examination is also time-consuming and error prone due to the human factor of fatigue. Thus, it needs an automated system (Ke et al., 2022). Detection methods based on deep learning are popular nowadays due to the development of hardware technology. These procedures are now being applied in a form of real-life scenarios, e.g. detecting defects in fabric (Zhou et al., 2024).

Fabric defect identification has been widely studied, whereby most of the studies have utilised datasets having fabric images placed carefully under a controlled environment. However, datasets pose several difficulties in real-time production applications. Images are often neither well-positioned nor of high quality (Chen et al.,2022; Zhou et al.,2024). Moreover, changes in blurriness, noise, and lighting intensity have an impact on visual perception. There are two primary categories of research in this area: identifying flaws in printed fabrics and identifying flaws in plain fabrics. Additionally, printed textiles can be classified as either regularly printed or irregularly printed. This distinction arises because differentiating between patterns and defects becomes challenging with a high diversity of prints. As a result, distinct approaches are required for each scenario. Models that can simultaneously recognise defects in plain, regularly printed, and irregularly printed textiles have received relatively little attention (Jing et al., 2022; Li et al.,2022).

Research of Our Work

The research in this study that presents an alternative detection method for fabrics by relying on the combination of a WB-FPN and the Segment Anything Model (SAM), which is created with the help of targeted knowledge input. The WB-FPN supports effective acquisition and refinement of multi-scale features, whereas SAM secures effective segmentation of various patterns and textures of fabrics. The structure is better at providing accuracy and flexibility in detecting complex fabric structures by including domain-specific knowledge. This mixed approach is a great advancement in the detection of fabrics with a positive outlook on its application in quality control, automation, and the study of textiles.

Motivation of This Research

This research is driven by the fact that the demand for proper and efficient fabric detection methods is increasing in the context of textile, fashion and manufacturing industries. The current methods usually have limitations where fine-grained pattern recognition, different textures, and different fabric structures may be a problem. Through a BiFPN and by combining the Segment Anything Model (SAM) with injecting domain-specific knowledge, this research is aimed at creating detection accuracy and flexibility. The proposed technique will integrate the enhanced hierarchical processing of features with contextual information, overcoming the current shortcomings and providing a further perspective on more powerful fabric defect detection frameworks in practical scenarios.

Our Work Contributions

The Principal Contributions of Our Work Are:

- To design a new fabric defect detection structure, which is a BiFPN integrated with the Segment Anything Model (SAM) and complemented with Specific Knowledge Injection.
- The Weighted BiFPN is developed based on multi-scale feature fusion through transmission-adaptive weights to features at different levels, which allows it to detect the presence of small and large-scale flaws in the fabric structure.
- The experiments conducted with benchmark textile datasets show that the given framework BiFPN-SAM performs significantly better than the current state-of-the-art models.

The rest of the paper is organised in the following way: Section 2 is devoted to the in-depth literature survey, whereas Section 3 presents the description of the proposed technique. The results and the discussion are provided in Section 4, and Section 5 finalises the article by providing the future work outline.

Survey

This survey is a valuable research area that also covers several research fields such as computer vision, materials science and textile engineering. The proposed survey will focus on the developments in automated technologies and methods of identifying, classifying and analysing fabrics. Applications in fabric detection have been applied in both the manufacturing of textile quality control in the design of intelligent fashion and in forensics, as well as in innovative and automated industries. The purpose of this survey was to give a comprehensive overview of the sector in terms of existing trends, challenges, and possible future directions of the fabric detection industry. Dalmini et al. (2022) created a real-time machine imaging system to detect defects in functional textiles, which can be used in the quality control mechanisms of the textile industry. Raw images are first divided into smaller, fitting sizes by way of preparing the data. Spatial filters are commonly used in de-noising. Moreover, there is the YOLOv4 system that is applied to categorise textiles. The described speed of detection is 34 fps, and the estimated time of processing is 21.4 ms per frame. (Dlamini et al., 2022; Garg et al., 2020) suggested a method to identify fabric defects which uses a two-layer neural network method and extracts data of the raw image patches. First, the presence of defects is detected with the help of a convolutional neural network (CNN).

Then, fault localisation is performed with the help of a multi-layer CNN consisting of convolution + ReLU, max pooling, SoftMax layers, and fully connected + ReLU. Li et al. (2020) proposed a way of detecting defects in patternless textiles through the co-occurrence of grey-level operators combined with uniform MDBP operators. A comparison is made using a detection threshold after a multi-directional texture feature matrix has been taken out of the entire fabric image into a feature matrix. Defects are detected with the help of these thresholds. Multiple grayscale images of textiles with different properties may be supported using this approach. Kang et al. (2024) introduced a technique that is based on YOLOv7-tiny and was used to detect flaws in solid-coloured circular knitted textiles. The processing load of the YOLOv7-tiny backbone network is also lower, and its feature extraction functions are improved by using the SPD-Conv module.

This method also integrates spatial and channel attention techniques to create a hybrid awareness module that is then combined into the YOLOv7-tiny system. As a result, the network achieves more accurate spatial information and improved image segmentation accuracy. Mao et al. (2025) presented an enhanced fabric detection method based on the YOLOv8 architecture. The approach integrated an attention mechanism and feature extraction modules to effectively capture multiscale features, which detects small defects accurately. (Ke et al., 2022) The results of defect identification are then obtained by adjusting the feature weights using the Adaptive Fabric Feature Clustering (AFFC) algorithm.

To discover high-resolution saliency targets, Xie et al. (2022) created PGnet. As the backbone, they used parallel connections between ResNet and Swin Transformer to perform feature fusion during the decoding stage. Nevertheless, the SBlock submodule's patch fusion sampling process results in a large loss of information, which leaves inadequate features extracted during decoding. When examining low-resolution fabric defect samples, this restriction results in a higher rate of misdetection.

A two-step technique for detecting fabric defects using the Pyramid Histogram of Edge Orientation Gradients (PHOG) was presented by (Cuifang et al., 2020). Support vector machines (SVMs) were also applied in the classification stage. This technique involves the combination of machine learning and statistical tools to study cloth texture.

Research Gap

Although there have been enormous improvements in the detection of fabrics, the existing technologies now find it hard to resolve the inherent limitations, such as detecting fabrics that look alike, differences in lighting and texture, detecting complex patterns in the real world, etc. Most methods are very dependent on controlled conditions or certain types of fabrics, and as such, they cannot be generalised. Although the introduction of deep learning models has improved accuracy, it may require large volumes of labelled data and a large amount of computing power, which are not feasible at all times. This is in consideration of the fact that more durable, efficient, and versatile fabric detection systems are required that can provide dependable performance in various environments and fabric types (Li et al., 2022).

Problem Identification of Existing System

- The current fabric detection systems are usually experiencing problems in their ability to detect complex patterns, texture and changes in the type of fabrics.
- A lot of systems cannot identify minor flaws, including a small tear, loose thread or differences in the weave, resulting in quality control problems.
- Many of the systems still use manual intervention, and this makes them less efficient and highly prone to human error.
- High implementation costs and complicated installation procedures often characterise advanced systems, and this makes them unavailable to small and medium-sized businesses (Chen et al.,2022).

Proposed System

A new fabric defect detection framework that is presented in this study involves a BiFPN with the Segment Anything Model (SAM) that is supplemented with Specific Knowledge Injection.

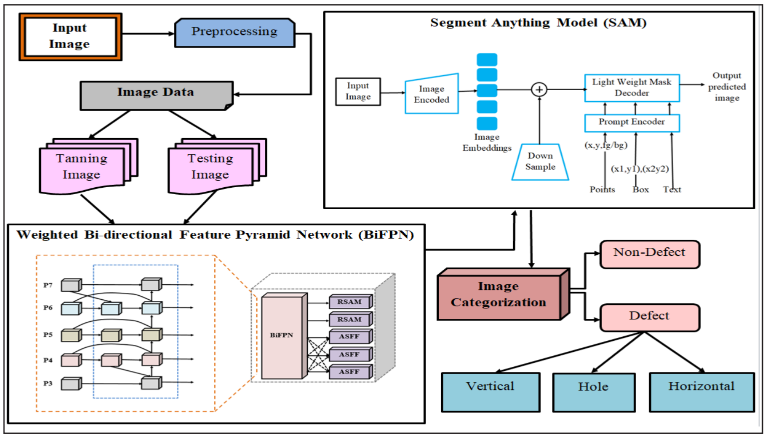

Figure 1. Architecture of BiFPN-SAM

The Weighted BiFPN uses adaptive weights for features of different levels to provide better multi-scale feature fusion to detect small- and large-scale fabric flaws. In the meantime, the Segment Anything Model (SAM) is an advanced segmentation model that is refined by introducing certain knowledge to shape its general segmentation abilities to meet the specifics of the fabric textures and defects. The combination guarantees the accurate localisation and classification of flaws even in unfavourable situations when there is a change in illumination, texture and the type of defect. The architecture of BiFPN-SAM is presented in Figure 1.

Dataset

The three types of defective photos in our fabric defect database, vertical, horizontal, and hole, along with the corresponding mask images for each defective sample, are displayed in Figure 2. The raw photos are in the 'caught' folder (Khodier et al., 2022). The image size is 640×360.

Figure 2. Sample image of fabric defect dataset

Preprocessing

Initially, the samples were reduced in size, and the RGB images were further transformed into greyscale for processing and model training. Contrast-Limited Adaptive Histogram Equalisation (CLAHE) was utilised to increase the contrast of the images (Roy et al., 2021). Before computing the cumulative distribution function (CDF), CLAHE clips the histogram at a predetermined value, preventing noise amplification in contrast to traditional Adaptive Histogram Equalisation (AHE). Two important factors that can be used to modify the quality of image enhancement are Block Size (BS) and Clip Limit (CL). $M \times N$ Tiles are created by first segmenting the image into contextual areas that do not overlap. Each contextual region's contrast-limited histogram is then calculated using the CL value as mentioned in Equation 1:

$$N_{avg} = \left(N_{px} \times N_{py}\right) / N_{gray} \qquad [1]$$

Where N_{avg} is the average number of pixels, N_{px} and N_{py} are the region's number of pixels in the X and Y directions, respectively, and N_{gray} is the grey level number of the contextual region. The actual clip limit can be estimated as specified in Equation 2:

$$N_{CL} = N_{clip} \times N_{avg} \qquad [2]$$

Wherever, N_{clip} is the standardised clip limit in the range of [0, 1]. If the number of pixels is greater than the N_{CL}, the pixels are clipped. Distributing the fraction that surpasses the clip limit throughout all histogram bins is the preferred method. The procedure is repeated if the redistribution results in some values exceeding the clip limit once more until the excess is minimal (Dlamini et al., 2022).

Weighted Bi-directional Feature Pyramid Network (BiFPN)

The BiFPN is an advanced version of the Feature Pyramid Network (FPN) designed to enhance feature fusion by assigning learnable weights to different features during the upsampling and downsampling processes. The approach effectively combines low-level fine-grained details with high-level semantic information (Huang et al., 2024; Zhou et al., 2022). This bi-directional approach ensures better propagation of contextual and spatial information, making it particularly useful for tasks requiring fine-grained detection, such as fabric defect detection. Figure 3 illustrates the architecture of the BiFPN model.

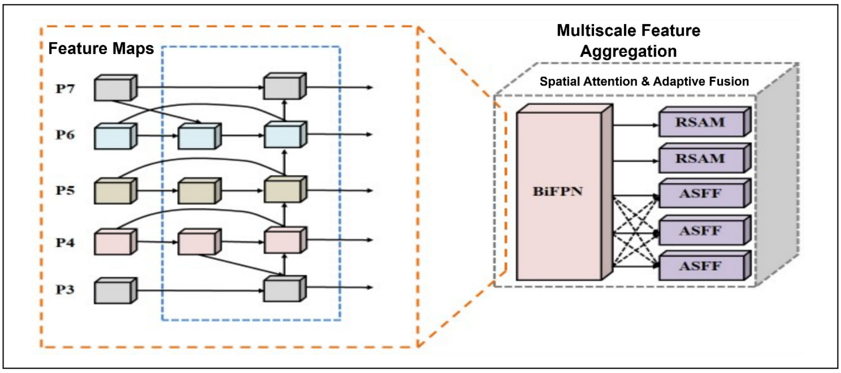

Figure 3. Architecture of the BiFPN model

Specific Knowledge Injection involves embedding domain-specific knowledge, such as patterns, textures, or structural properties unique to fabric materials, into the network. This is accomplished using pre-trained models or auxiliary inputs that guide the network in learning fabric-specific representations. When combined with BiFPN, this approach improves the network's capability to detect even subtle fabric defects.

Bi-directional Feature Fusion

BiFPN uses learnable weights for each input feature. For a given set of feature maps $F_1, F_2, \ldots, F_n$, the fusion equation can be expressed as in Equation 3:

$$F_{output} = \sum_{i=1}^{n} w_i \cdot F_i \quad [3]$$

Where F_i are input feature maps from various pyramid levels, w_i is a learnable weight for each input feature map, constrained such that $\sum w_i = 1$ (softmax normalisation).

Feature Scaling and Normalisation

To align spatial dimensions before fusion, Equation 4 is applied:

$$F_i^{'} = Resize(F_i, target\_size) \quad [4]$$

Bi-directional Pathway

Top-down pathway for upsampling is mentioned in Equation 5:

$$F_{td}^{l} = Upsample(F^{l+1}) + w^{td} \cdot F^{l} \quad [5]$$

Bottom-up pathway for downsampling is specified in Equation 6:

$$F_{bu}^{l} = Downsample\left(F^{l-1}\right) + w^{bu} \cdot F^{l} \quad [6]$$

Specific Knowledge Injection

Domain knowledge K is integrated as an auxiliary input or additional loss function as in Equation 7:

$$F_{enhanced} = F_{Output} + \alpha \cdot \phi(K) \quad [7]$$

Where $\phi(K)$ are knowledge embedding functions, α is a weight that controls the impact of specific knowledge.

Loss Function for Defect Detection

A combined loss function is often employed to balance detection accuracy and domain adaptation, as specified in Equation 8.

$$L = Ldomain_{det} \quad [8]$$

Where, L_{det} is the loss for fabric defect detection, L_{domain} is the loss for knowledge consistency or domain alignment, β is balancing the weights of the two loss terms (Jing et al., 2022).

These components together enhance the BiFPN's performance for fabric defect detection by leveraging both structural network improvements and domain-specific knowledge.

Segment Anything Model (SAM)

The SAM represented in Figure 4 is a state-of-the-art deep learning framework designed to perform image segmentation tasks with high generalisability. It uses a combination of Vision Transformers (ViTs) and a promptable interface to segment any object in an image with minimal input. However, for specialised applications such as fabric defect detection or segmentation, SAM's performance can be enhanced by incorporating domain-specific knowledge. Figure 4 illustrates the architecture of the SAM model.

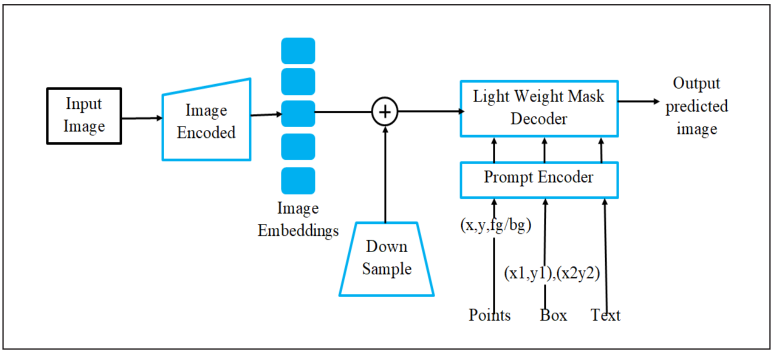

Figure 4. Architecture of Segment Anything Model (SAM)

SAM Segmentation Framework

SAM uses a vision encoder E to extract image features F, and a prompt encoder P to process inputs like points, bounding boxes, or masks as in Equations 9 and 10:

$$F = E(I) \quad [9]$$

$$M = D\left(F, P(p)\right) \quad [10]$$

Wherever I is the input image, D is the decoder, $P(p)$ is the encoded prompt, and M is the segmented mask.

Specific Knowledge Injection

Injecting specific knowledge can be achieved by modifying the encoder or the prompt input using Equation 11.

$$F^{'} = E^{'}(I, K)$$

Here, K is the domain-specific knowledge (e.g., defect patterns, texture features). This can involve pretraining $E^{'}$ on a fabric dataset or appending K to feature embeddings.

Loss Function for Domain Adaptation

To adapt SAM to fabric detection, a customised loss function L can combine segmentation loss with domain-specific penalties as specified in Equation 12:

$$L = L_{seg} + \lambda L_{domain} \quad [12]$$

- L_{seg}: Standard segmentation loss (e.g., Dice Loss, Cross-Entropy Loss)
- L_{domain}: Loss capturing fabric-specific properties (e.g., texture continuity, defect localisation)
- λ: Weight balancing the two terms.

Prompt Augmentation for Fabric Detection

Specific prompts P_{fabric} can encode domain-relevant information using Equation 13:

$$\mathbf{p}_{\mathbf{fabric}} = \text{encode(fabric-particular patterns user comments cloth)} \quad [13]$$

Fusion of Features and Prior Knowledge

In case of feature fusion, we integrate the features of SAM of fabric-describing nature using Equation 14.

$$F^{''} = \alpha F + (1 - \alpha) F_K \quad [14]$$

- F_K: Features extracted from fabric-specific knowledge
- α: Weight controls the contribution of each feature.

The model incorporates the generalisation capabilities of SAM and induced domain knowledge to arrive at the model, resulting in improved performance on complex textures and subtle defects.

Advantages of Proposed Method

- Uses bi-directional feature aggregation that is weighted to improve feature representation. for fabric textures.
- Utilises SAM's general segmentation capabilities, specifically tailored for fabric-related applications.
- Optimised for high performance with minimal overhead compared to traditional models.
- Ensures improved feature propagation and context understanding across scales (Khodier et al., 2022).

RESULT AND DISCUSSIONS

Experimental Setup

The following is the configuration of the experiment's environment: Microsoft Windows 10 as the operating system, CUDA 11.5 and CUDNN 11.1 as GPU-accelerated libraries,

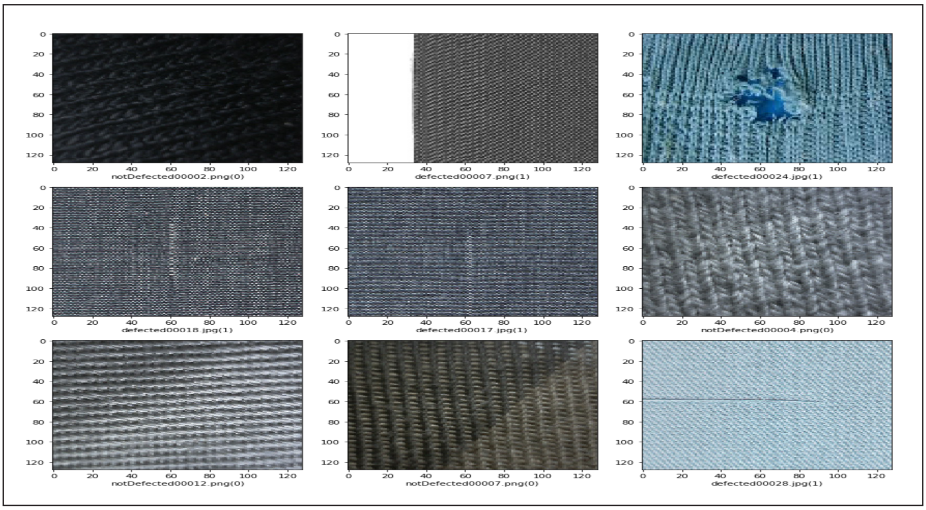

Figure 5. Final prediction images for the dataset

PyCharm as the integrated development environment, Python 3.6 as the programming language, TensorFlow-GPU 1.14.0 and Keras 2.2.5 as the deep learning frameworks, and 16 GB of RAM, AMD R7-4800H CPU, and Nvidia GeForce RTX 2060 6 GB GPU. Figure 5 shows the final prediction images for the dataset.

Performance Metrics

DSR (Detection Success Rate) is a performance metric that measures a model's accuracy in identifying the correct class. It is determined by dividing the total number of cases by the proportion of correctly detected occurrences, as mentioned in Equation 15.

$$DSR = \frac{TruePositives}{TruePositives\ + FalseNegatives} \times 100 \quad [15]$$

A metric called PPV (Positive Predictive Value) gauges how well a model forecasts the future, as in Equation 16. It shows the proportion of actual positive results out of all the positive forecasts.

$$PPV = \frac{TP}{TP+FP} \quad [16]$$

NPV (Negative Predictive Value) is a metric utilised to calculate the percentage of true negative results among a model's negative predictions, as in Equation 17. It helps assess the model's accuracy in recognising negative cases.

$$NPV = \frac{TN}{TN+FN} \quad [17]$$

The F-measure is a statistic that provides a fair assessment of a model's performance by combining recall and precision, as in Equation 18. When there is an imbalance in the distribution of classes, it is very helpful. The following formula is utilised to determine the F-measure:

$$F - Measure = 2 \times \frac{Pr\ ecision \times Re\ call}{Pr\ ecision + Re\ call} \qquad [18]$$

The accuracy of a model is expressed as the number of correct forecasts divided by the number of predictions, as mentioned in Equation 19. It can be calculated with the following Equation 19 (Roy et al., 2021):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad [19]$$

Comparative Methods

MTF-Net (Multitask Feature Fusion Network)

The collaboration among a pixel-level branch network, an object-level branch network, and a reconstruction branch network enhances the model's detection capability (Zhang et al., 2023).

CCN (Colour Conversion Network

The suggested CCN container should be implemented as an add-on module that seamlessly integrates with existing strategies to succeed universal enhancement (Fang et al.,2022).

CNN (Convolutional Neural Network)

The proposed approach is likely more reliable and applicable than traditional visual techniques. Data were developed as fabric datasets. A robotic arm fitted with touch sensors was used to automate and standardise the collection process.

Table 1
Performance metrics comparisons of the various models to detect the fabric type

Models	Fabric Types	DSR (%)	PPV (%)	NPV (%)	F-Measure (%)	Accuracy (%)
MTF-Net	Hole	87.54	70.33	66.13	61.19	52.19
	Vertical	65.98	81.22	60.44	84.98	76.98
	Horizontal	60.22	65.54	76.56	81.22	70.55
CCN	Hole	76.98	89.33	87.22	80.76	87.14
	Vertical	89.76	56.98	81.22	73.36	91.88
	Horizontal	70.44	50.33	59.91	55.14	93.77

Table 1 (*continued*)

Models	Fabric Types	DSR (%)	PPV (%)	NPV (%)	F-Measure (%)	Accuracy (%)
CNN	Hole	87.21	75.45	81.11	82.18	90.32
	Vertical	66.33	88.87	91.98	89.91	89.22
	Horizontal	90.34	81.33	88.23	91.88	81.77
Yolov8	Hole	90.12	88.45	89.76	89.10	92.84
	Vertical	91.67	90.21	92.33	91.05	91.05
	Horizontal	92.88	91.76	93.45	92.10	92.10
Proposed	Hole	93.76	92.65	94.56	95.18	96.18
	Vertical	93.91	94.44	94.91	95.65	97.77
	Horizontal	95.55	96.11	97.78	96.11	98.89

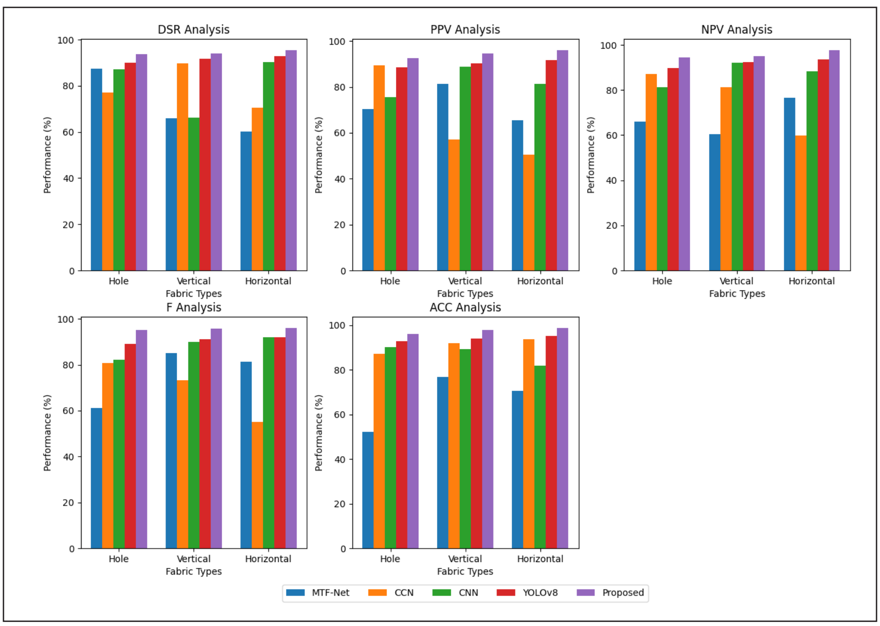

Figure 6. Comparison of performance metrics of different models for fabric type detection

Table 1 and Figure 6 indicate significant differences in detection metrics, including DSR (Detection Sensitivity Rate), PPV, NPV, F-Measure, and Accuracy. The MTF-Net model produces different scores, with a maximum accuracy of 76.98 per cent in identifying vertical fabric types, and horizontal and hole fabrics have poor performance. Instead, the CCN model attains a high level of detection of hole textiles (87.14) and vertical fabrics (91.88). The CNN model and YOLOv8 deliver middle performance with good accuracy of 90.32 and 92.84 per cent on hole textiles but marginally reduced precision in horizontal garments (81.77%). The proposed model is always better than the others in entirely different fabrics, and reaches the utmost standards in every criterion, among others, an astonishing.

Horizontal fabrics had an accuracy of 98.89%. This brings out the high efficacy of the proposed approach. in fabric type detection.

A comparison of the Average Percentage Error (APE) and Mean Absolute Error (MAE) for fabric type detection across several models is presented in Table 2 and Figure 7. The results indicate that the proposed model always has the smallest error values, with APE having a range between 5.87% and 10.44% and MAE from 15.65% to 19.45%. On the contrary, the MTF-Net model displays higher error rates, especially when dealing with the Vertical and Horizontal fabric types, with the APE values going up to 32.67% and MAE up to 51.98%. There are moderate errors in the CNN and CCN models. CNN doing better than CCN, and YOLOv8 is better than the other two. Overall, the proposed model is superior to any other models, proving to be more accurate and with a lower error rate.

Table 2
APE and MAE comparison of various models used in the fabric type detection

Models	Fabric Types	APE (%)	MAE (%)
MTF-Net	Hole	22.76	45.76
	Vertical	27.19	51.98
	Horizontal	32.67	48.18
CCN	Hole	19.45	34.19
	Vertical	23.87	39.66
	Horizontal	20.91	41.11
CNN	Hole	12.17	27.19
	Vertical	18.19	34.33
	Horizontal	15.76	31.91
Yolov8	Hole	8.92	21.36
	Vertical	11.48	25.72
	Horizontal	9.76	23.18
Proposed	**Hole**	**5.87**	**16.44**
	Vertical	**7.87**	**19.45**
	Horizontal	**10.44**	**15.65**

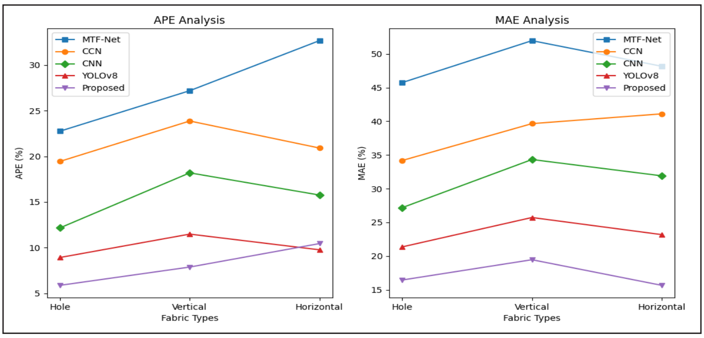

Figure 7. APE and MAE comparison between the APE and MAE of various models in fabric type detection

Table 3 and Figure 8 provide a comparison of the detection time in milliseconds (ms) for different models in fabric type detection. The suggested model has the lowest detection times, and its values tend to be between 1.675 ms and 5.876 ms in all kinds of fabrics. Comparatively, the CNN model has a slightly higher detection time with a range of 8.445 ms to 10.887 ms. There is also competitive performance between the CNN model and YOLOv8, and the detection times are in the range of 12.786 ms up to 15.154 ms. The MTF-Net model exhibits the highest detection times, especially when it comes to the Horizontal type of fabric, with a maximum value of around 18.876 ms. In general, the suggested model presents the shortest detection times, which points to its effectiveness in fabric type recognition.

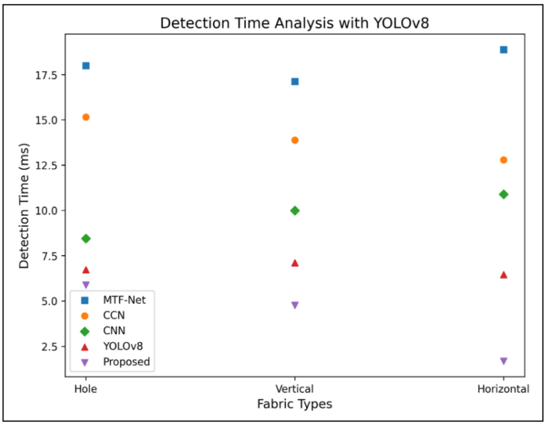

Figure 8. Comparison of detection time analysis of different models for fabric type detection

Table 3
Detection time analysis of the various models to estimate fabric type detection

Models	Fabric Types	Detection Time (ms)
MTF-Net	Hole	17.987
	Vertical	17.113
	Horizontal	18.876
CCN	Hole	15.154
	Vertical	13.876
	Horizontal	12.786
CNN	Hole	8.445
	Vertical	9.987
	Horizontal	10.887
YOLOv8	Hole	6.732
	Vertical	7.115
	Horizontal	6.458
Proposed	Hole	5.876
	Vertical	4.765
	Horizontal	1.675

Loss and Accuracy of Training

The estimation of fabric defect detection relying on training and validation loss, as illustrated in Figure 9, portrays a complete picture of how the model will be presented, besides having training and validation accuracy. Training loss is monitored so that the model is learning from the dataset, but the validation loss measures whether the model can generalise to new data (Huang et al., 2024). On the same note, high accuracy in training signifies effective learning throughout the training, with high accuracy in validation signifying the stability and viability of the model to be used in the real world. It is important to balance the metrics since significant differences can indicate overfitting or underfitting. All these metrics confirm the reliability and efficiency of the suggested fabric defect detection structure.

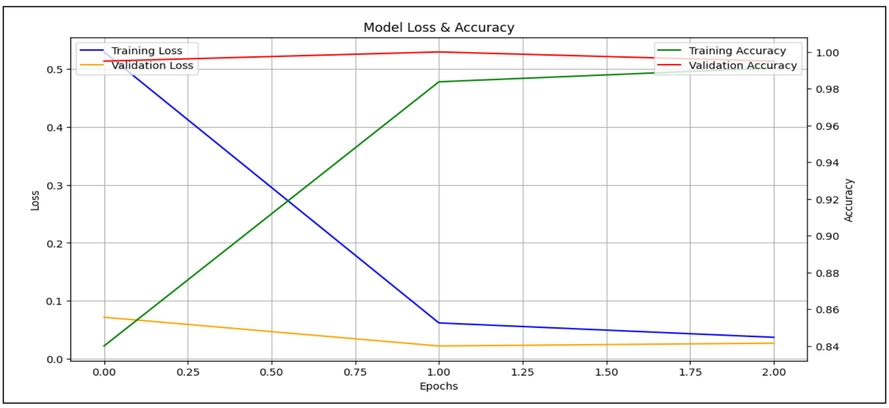

Figure 9. Training validation loss and accuracy

RESULT AND DISCUSSIONS

The fact that the evaluation of performance indicators of various models to identify types of fabric provides an opportunity to conclude that our proposed model has more benefits in comparison with the previous studies (Ding et al., 2024; Fang et al., 2022; Zhang et al., 2023). In the major measures, our model is always ahead of others, primarily in accuracy, with a success rate of up to 98.89% in the case of the Horizontal fabric type. This is much better than the best accuracy indicated by Zhang et al. (2023) of 93.77%. Also, our model has significantly lower error rates, which are indicated by the APE and MAE values. As an example, it records an APE of 5.87% and an MAE of 16.44% on the Hole fabric type, whereas the most successful models, including Fang et al. (2022), record an APE of approximately 12.17% and an MAE of approximately 27.19%. Our model is very efficient in terms of detection time, as the Horizontal fabric type is processed in the shortest time of 1.675 ms, which is much less than the other models, which require between 8.445 ms and 18.876 ms.

These results highlight the better results of our model in accuracy and computational efficiency, as it is a formidable competitor in real-time applications of fabric type diagnosis.

Ablation Study

Figure 10 and Table 4 show the ablation study, which shows the performance improvements attained through the various elements of the suggested model of fabric defect diagnosis. The foundation is the baseline model that an accuracy of 87.34 was achieved. Utilising only the BiFPN led to a decrease in the accuracy to 77.33, and this indicates that more features are required. On the other hand, the use of the sole SAM provided a commendable accuracy of 93.44%. Strong segmentation strengths. The complete combination of BiFPN and SAM in the BiFPNSAM structure produced the maximum accuracy of 98.89, which highlighted the synergistic effect of integrating the multi-scale fusion of features and fine-tuning segmentation to achieve the best defect detection.

Table 4
Ablation study for accuracy analysis

Models	Accuracy
Baseline Model	87.34
BiFPN Only	77.33
SAM Only	93.44
BiFPN-SAM	**98.89**

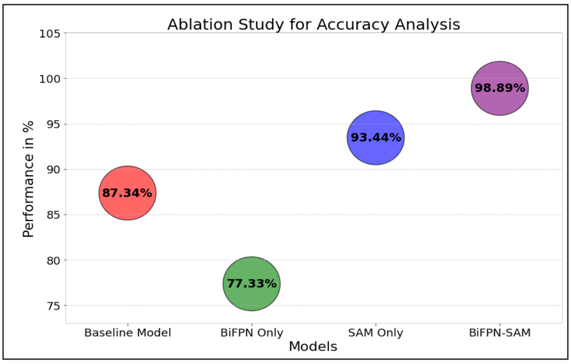

Figure 10. Ablation study for accuracy analysis

Limitations and Challenges

Although the BiFPN and the Segment Anything Model (SAM) with Specific Knowledge Injection exhibit stunning performance in cloth flaw identification, some shortcomings and obstacles persist. The training and fine-tuning of these models require the use of large-scale computational resources; the model containers are a hindrance to small-scale industries. Regardless of its flexibility, SAM can have difficulties with identifying very subtle or overlapping defects in the absence of refinement. Additionally, domain expertise and careful parameter injection are necessary to incorporate knowledge injection. tuning which may complicate deployment. Both real-time performance and high performance must be ensured. Also, accuracy is an issue in cases of varying fabric textures and dynamic production environments. These problems need to be resolved to achieve wider applicability in real-world settings.

CONCLUSION

Finally, the BiFPN and the Segment Anything Model (SAM) were combined and used. Specific Knowledge Injection has defined the exemplary presentation in fabric defect detection. The multi-scale feature fusion approach based on the BiFPN and the SAM, based on the accurate segmentation of the defects, successfully deals with the issues of texture difference and reduced defect samples. It is further enhanced by the fact that the system incorporates knowledge injection ability to adjust to the peculiarities of the fabric defects, achieving state-of-the-art accuracy and computational effectiveness in experimental tests. This is not only a way of guaranteeing accuracy, localisation, and categorisation of fabric defects but also underlines its possible application in real-world uses in automated applications of quality control of textiles.

Future Work

The structure may be expanded in future work to have the capability to detect defects in real-time, meeting the demands of fast manufacturing environments. As to its flexibility for other applications and in industrial settings, including the inspection of metal surfaces or composite materials, it could also be extended to show that it is versatile. In addition, incorporation of advanced self-supervised learning techniques and domain adaptation methods would make it more resilient to various datasets and manufacturing situations, preparing for a more universal and scalable quality control solution.

ACKNOWLEDGEMENT

Author Contributions

All authors contributed equally, and all authors read and approved the final version of the paper.

LIST OF ABBREVIATIONS

WB-FPN	:	Weighted Bi-directional feature pyramid network
SAM	:	Segment Anything Model
BiFPN	:	Weighted Bi-directional feature pyramid network
IoU	:	Intersection over Union
CLAHE	:	Contrast-limited adaptive histogram equalisation
CDF	:	Cumulative distribution function
AHE	:	Adaptive histogram equalisation
BS	:	Block size
CL	:	Clip limit
DSR	:	Detection success rate

PPV : Positive predicted value
NPV : Negative predicted value
TP : True Positive
TN : True Negative
FP : False Positive
FN : False Negative
MTF-Net : Multitask feature fusion network
CCN : Colour conversion network
CNN : Convolutional neural network
APE : Average percentage error
MAE : Mean absolute error

REFERENCES

Chen, Z., Tian, S., Shi, X., & Lu, H. (2023). Multiscale shared learning for fault diagnosis of rotating machinery in transportation infrastructures. *IEEE Transactions on Industrial Informatics, 19*(1), 447-458. https://doi.org/10.1109/TII.2022.3148289

Cuifang, Z., Yu, C., & Jiacheng, M. (2020). Fabric defect detection algorithm based on PHOG and SVM. *Indian Journal of Fibre & Textile Research, 45*(1), 123-126.

Ding, M., Pan, L., & Chi, M. (2024). A multitask feature fusion network for woven fabric density analysis. *IEEE Access, 12*, 36229-36238. https://doi.org/10.1109/ACCESS.2024.3371175

Dlamini, S., Kao, C.-Y., Su, S.-L., & Kuo, C.-F. J. (2022). Development of a real-time machine vision system for functional textile fabric defect detection using a deep YOLOv4 model. *Textile Research Journal, 92*(5-6), 675-690. https://doi.org/10.1177/00405175211034241

Fang, B., Long, X., Sun, F., Liu, H., Zhang, S., & Fang, C. (2022). Tactile-based fabric defect detection using convolutional neural network with attention mechanism. *IEEE Transactions on Instrumentation and Measurement, 71*, Article 2503309. https://doi.org/10.1109/TIM.2022.3165254

Garg, M., & Dhiman, G. (2020). Deep convolutional neural network approach for defect inspection of texture surfaces. *Journal of the Institute of Electronics and Computer, 2*(1), 28-38. https://doi.org/10.33969/JIEC.2020.21003

Huang, Z., Jing, H., Liu, Y., Yang, X., Wang, Z., Liu, X., Gao, K., & Luo, H. (2024). Segment Anything model combined with multi-scale segmentation for extracting complex cultivated land parcels in high-resolution remote sensing images. *Remote Sensing, 16*(18), Article 3489. https://doi.org/10.3390/rs16183489

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

Kang, X., & Li, J. (2024). AYOLOv7-tiny: Towards efficient defect detection in solid colour circular weft fabric. *Textile Research Journal, 94*(1-2), 225-245. https://doi.org/10.1177/00405175231205898

Ke, Y. Y., & Tsubono, T. (2022). Recursive contour-saliency blending network for accurate salient object detection. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)* (pp. 2940-2950). IEEE. https://doi.org/10.1109/WACV51458.2022.00143

Khodier, M. M., Ahmed, S. M., & Sayed, M. S. (2022). Complex pattern Jacquard fabrics defect detection using convolutional neural networks and multispectral imaging. *IEEE Access, 10*, 10653-10660. https://doi.org/10.1109/ACCESS.2022.3144843

Li, F., Yuan, L., Zhang, K., & Li, W. (2020). A defect detection method for unpatterned fabric based on multidirectional binary patterns and the grey-level co-occurrence matrix. *Textile Research Journal, 90*(7-8), 776-796. https://doi.org/10.1177/0040517519879904

Li, L., Wang, Y. Q., Gao, H., Qi, J., & Zhou, T. (2022). Automatic recognition method for the three-elementary woven structures and defects of carbon fabric prepregs. *Composite Structures, 291*, Article 115527. https://doi.org/10.1016/j.compstruct.2022.115527

Mao, Y., Wang, G., Ma, Y., & Gui, X. (2025). Enhancing fabric defect detection with attention mechanisms and optimised YOLOv8 framework. *IEEE Access, 13*, 96767-96781. https://doi.org/10.1109/ACCESS.2025.3570455

Roy, K., Hasan, M., Rupty, L., Hossain, M. S., Sengupta, S., Taus, S. N., & Mohammed, N. (2021). Bi-FPNFAS: Bi-directional feature pyramid network for pixel-wise face anti-spoofing by leveraging Fourier spectra. *Sensors, 21*(8), Article 2799. https://doi.org/10.3390/s21082799

Xie, C., Xia, C., Ma, M., Zhao, Z., Chen, X., & Li, J. (2022). Pyramid grafting network for one-stage high-resolution saliency detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 11717-11726). IEEE. https://doi.org/10.1109/CVPR52688.2022.01142

Zhang, C., Qi, Y., & Wang, Y. (2024). Learning the colour space for effective fabric defect detection. *IEEE Transactions on Emerging Topics in Computational Intelligence, 8*(1), 981-991. https://doi.org/10.1109/TETCI.2023.3320565

Zhang, H., Chen, X., Lu, S., Yao, L., & Chen, X. (2023). A contrastive learning-based attention generative adversarial network for defect detection in colour-patterned fabric. *Colouration Technology, 139*(3), 248-264. https://doi.org/10.1111/cote.12642

Zhou, Q., Shi, H., Xiang, W., Kang, B., & Latecki, L. J. (2025). DPNet: Dual-path network for real-time object detection with lightweight attention. *IEEE Transactions on Neural Networks and Learning Systems, 36*(3), 4504-4518. https://doi.org/10.1109/TNNLS.2024.3376563

Zhou, Q., Wang, L., Gao, G., Kang, B., Ou, W., & Lu, H. (2024). Boundary-guided lightweight semantic segmentation with multi-scale semantic context. *IEEE Transactions on Multimedia, 26*, 7887-7900. https://doi.org/10.1109/TMM.2024.3372835

Zhou, Q., Wu, X., Zhang, S., Kang, B., Ge, Z., & Latecki, L. J. (2022). Contextual ensemble network for semantic segmentation. *Pattern Recognition, 122*, Article 108290. https://doi.org/10.1016/j.patcog.2021.108290