

Sareenet: saree texture classification via region-based patch generation with an optimized efficient Aquila network

Krishnan Ashok^{1*}, Dharmaraj Karthika Priya², Appathurai Ahilan³, Palanisamy Jayapriya⁴

¹Department of Electronics and Communication Engineering, Coimbatore Institute of Engineering and Technology, Coimbatore, India

²Department of Electronics and Communication Engineering, Thiagarajar College of Engineering, Tamilnadu, India

³Department of Electronics and Communication Engineering, PSN College of Engineering and Technology, Tirunelveli, Tamilnadu, India

⁴Centre for Future Networks and Digital Twin, Department of Computer Science and Engineering, Sri Eshwar Engineering College, Coimbatore, Tamilnadu, India

Article info

Article history:

Received 06 Jul. 2023

Received in revised form 09 Oct. 2023

Accepted 16 Oct. 2023

Available on-line 11 Feb. 2024

Keywords:

Saree texture;
classification;
patch generation;
deep learning;
efficient net;
Aquila optimizer.

Abstract

Online saree shopping has become a popular way for adolescents to shop for fashion. Purchasing from e-commerce is a huge time-saver in this situation. Female apparel has many difficult-to-describe qualities, such as texture, form, colour, print, and length. Research involving online shopping often involves studying consumer behaviour and preferences. Fashion image analysis for product search still faces difficulties in detecting textures based on query images. To solve the above problem, a novel deep learning-based SareeNet is presented to quickly classify the tactile sensation of a saree according to the user's query. The proposed work consists of three phases: i) saree image pre-processing phase, ii) patch generation phase, and iii) texture detection and optimization for efficient classification. The input image is first denoised using a contrast stretching adaptive bilateral (CSAB) filter. The deep learning-based mask region-based convolutional neural network (Mask R-CNN) divides the region of interest into saree patches. A deep learning-based improved EfficientNet-B3 has been introduced which includes an optimized squeeze and excitation block to categorise 25 textures of saree images. The Aquila optimizer is applied within the squeeze and excitation block of the improved EfficientNet to normalise the parameters for improving the accuracy in saree texture classification. The experimental results show that SareeNet is effective in categorising texture in saree images with 98.1% accuracy. From the experimental results, the proposed improved EfficientNet-B3 improves overall accuracy by 2.54%, 0.17%, 2.06%, 1.78%, and 0.63%, for MobileNet, DenseNet201, ResNet152, and InspectionV3, respectively.

1. Introduction

Indian women have worn sarees for centuries and have a significant market-share among their clothes [1]. In India, saree alone accounts for 33% of the women's clothing market. The segment is probable for growing by 5% to 6% by 2023 [2, 3]. Indian women wear the saree, a long strip of unstitched material, as a symbol of their culture and ethnicity. A saree is often draped loosely over the left

shoulder with the other end loosely tucked in at the waist and wrapped around the body with pleats in the middle [4]. One saree can be draped in at least a hundred different ways. It is a very distinguishing feature on both sides of the saree that enhances its beauty. "Anchal" or "Pallav" is the saree end element, which creates much of its decoration [5]. This is the most essential part of the saree. Saree texture recognition involves using artificial intelligence to identify and classify different textures and patterns present in sarees (traditional Indian garments) [6]. This technology can be used for various purposes, such as e-commerce, fashion

*Corresponding author at: ashok_ace01@yahoo.co.in

analysis, and automated inventory management, as well as collecting a diverse and high-resolution dataset of saree images representing various textures, patterns, and colours [7, 8]. Saree texture detection can enhance the user's experience on online shopping platforms, help fashion designers, and facilitate textile industries in inventory management [9]. However, it is essential to approach the development and deployment of such technology ethically and reasonably. Convolutional neural networks (CNNs) are highly effective for image classification tasks. Saree texture detection using deep learning (DL) involves leveraging the power of neural networks, particularly CNNs to automatically learn and recognize patterns and textures present in saree images [10]. One can design a CNN architecture for saree texture detection by possibly using pre-trained models like VGG16 or ResNet and fine-tuning them on the saree dataset.

1.1. Motivation

The purchasing power of women has increased recently because the popular attire has attracted women to purchase online, high-brand consciousness, and a strong sense of fashion due to the images captured of an attractive outfit of another woman. Online saree shopping was a convenient and enjoyable experience; however, it also involved a unique set of difficulties. 1) it is hard to judge the quality of the fabric and embroidery online; 2) the colour of the saree might look different in real life due to variations in screen displays; 3) the consumer cannot physically touch or feel the fabric before buying. To solve this problem, various patch images of a saree must be generated and their textures efficiently categorised using advanced deep learning techniques.

1.2. Contributions

- The key objective of our work is to introduce a novel SareeNet for effective classification of saree textures from the gathered saree dataset.
- The collected saree images are denoised using contrast stretching adaptive bilateral (CSAB) filter for eliminating the noise distortion to retain the colour consistency of the image.
- The denoised images are given as input for generating the patches using the mask region-based convolutional neural network (Mask R-CNN) model rather than physical initialisation, hence eliminating the need for physical interpretation.
- The improved EfficientNet is presented with an optimisation of the deployed squeeze and excitation block for improving the categorisation accuracy.
- The Aquila optimizer is applied to the squeeze and excitation block to normalise the parameters for accurate classification of 25 saree textures.

The rest of this paper was organised as follows. Section 2 includes the prior works related to the cloth segmentation and classification. Section 3 explains the thorough portrayal of the proposed SareeNet. Section 4 contains the experimental findings and discussion and section 5 presents the conclusions and upcoming extension of this work.

2. Literature review

In the present day, numerous approaches have been introduced by researchers, mainly for increasing the classification accuracy in identifying different clothing textures, especially in online shopping. This section covers some analogous research works that have tried for segmenting and classifying the clothing textures.

In 2020, Zhang *et al.* introduced a ClothingOut framework [11] that automatically creates tiled clothing images using generative adversarial networks (GAN). In particular, by studying transformation rules between garments on users and tiled clothes, a unique category-supervised GAN system was created. This strategy entails supplementing a conventional GAN model with category characteristics. The dataset was created with more than 20 000 pairs of cloth images and their matching smooth garment images to train the models. In the cross-scenario of apparel retrieval of images and video advertising, the framework obtained was directly applicable. The developed images, which were separated from the wearers' images, were tested for authenticity and retrieval efficiency.

FashionOn network [12] was designed by Hsieh to combine images of individuals wearing various outfits in random postures to provide full information about the appropriate outfits. In particular, FashionOn synthesises the trail images through three key phases: posture-guided parsing translation, segmentation region colouring, and salient region refining. These three steps give a raw image, an in-store apparel image, and a goal posture. Numerous trials show that FashionOn preserves clothing information and addresses the body occlusion problem, resulting in the most cutting-edge virtual trial efficiency in terms of both quality and quantity.

In 2021, Leithardt presented Fashion-MNIST dataset [13] to demonstrate four distinct CNNs. Fashion-MNIST was developed for classifying the type of products that fall into this category, such as clothing, and comparing various classification algorithms to determine which is the most effective. The major objective of the study was to improve the technique of classification comparisons for subsequent studies. The suggested CNNs were used to address this issue, and the classification outcomes are contrasted with the initial ones. This technique accuracy was increased by 89.7%

Condition-CNN [14] was a revolutionary hierarchical image classification technique that reports some of the limitations of branching CNN based on training time and classification accuracy. This algorithm employs Teacher Forcing training, which uses actual class labels rather than projected ones when training lower-level branches. As a result of Condition-CNN learning about the links between various levels of categories and the image data for each level of classification, class predictions are estimated during scoring. The testing results from the gathered dataset show that Level1, Level2, and Level3 attain accuracy of 99.8%, 98.1%, and 91.0%, respectively.

In 2020, Zhang and Liu [15] suggested a colour-texture-based unsupervised image segmentation technique for clothing-related images. To extract more precise colour attributes, a typical three-dimensional colour space was first split into a six-dimensional colour space using block truncation encoding. The apparel image was then described by colour characteristics using a texture feature in terms of

an improved local binary pattern (LBP) technique. A bisection method was suggested for the segmentation operation based on the numerical presence rule of the item and backdrop features in the cloth images. Bisection image segmentation will be completed more quickly because the image was separated into numerous sub-image blocks.

FashionSegNet [16] is a model constructed with encoder–decoder for a high-precision semantic segmentation of clothing images. By reducing the impact of variability in image shots and similarity between cloth classes and boundary complexity on clothing image segmentation accuracy, the accuracy of cloth image segmentation was improved. An upgraded spatial pyramid pooling component together with a global feature extraction division of a large convolutional filter was built to accomplish fusion of multiscale feature in various images and enhance the capability of the technique to recognize fashion and its border characteristics. The experimental results of the FashionSegNet attained the mIoU of 74.5% and bIoU of 57.5%, respectively.

A variety of factors affect the outer view of saree surfaces, including shape, reflectance, illumination, and viewing direction, which makes identifying fabric materials from photos challenging [17, 18]. Statistical learning approaches [19–21] have been applied to databases containing items shot in various lighting conditions and viewing angles. Despite very high classification rates [22], the data gathered show that the problem is complex. A sense of touch can help distinguish between materials with minute variances. In contrast to other datasets [23, 24], this fine-tuned classification challenge is significantly more challenging when fabrics include material classes or have intra- or inter-class variance. There have been a few datasets introduced recently, such as FashionMNIST and DeepFashion, that are used to categorise various types of clothing [25, 26].

Research gaps regarding the proposed research question were identified after a thorough review of the literature.

Shirts, jeans, jackets, and shorts are among the most common clothing items in the FashionMNIST and DeepFashion databases. Traditional Indian clothing, including sarees, kurtas, dhotis, and other types of clothing, contrasts sharply with Western culture. In this study, the authors classify 25 different fabrics preferred by Indian women, including silk and cotton [27]. A real-time saree classification system is difficult to develop, especially for images taken on mobile devices, due to two factors. The first challenge is the representation of fabrics and the segmentation of saree fields. Secondly, different types of sarees have properties that are difficult for humans to understand and require extensive calculations. Furthermore, traditional craft features cannot extract the essential features of clothing, so they are not appropriate for reliably categorising fabrics images. The authors deployed the DL approach to increase the accuracy in the identification and classification of saree textures.

3. Proposed SareeNet

This section presents a novel SareeNet which was proposed for the efficient saree texture classification using DL models. The inclusive workflow of the proposed framework is shown in Fig. 1.

3.1. Dataset acquisition

In this work, the Indian saree dataset (ISDA) [8] is created including 40 different saree class images taken at various scale, illumination, resolution, and angles. For a variety of research challenges relating to clothes, such as classification of fabric type and texture, DeepFashion and FashionMNIST are two textile databases that are frequently used. There is no publicly accessible database available yet for Indian sarees. Hence, a new database dedicated to Indian sarees is created by Ref. 8 under the conditions like change in camera resolution, lighting, and point of view.

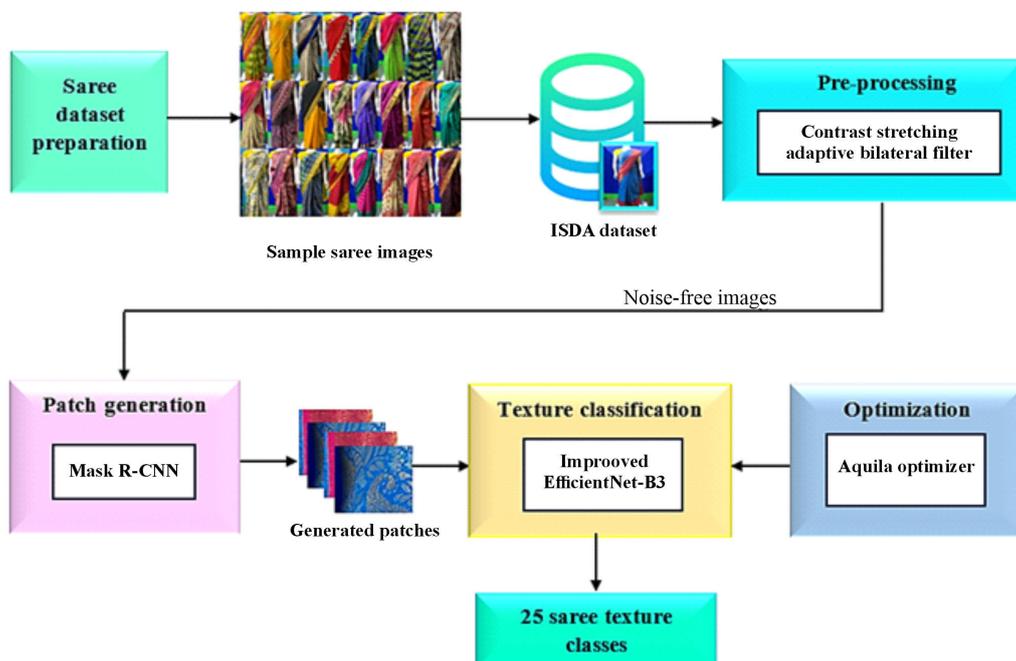


Fig.1. Schematic portrayal of the proposed saree texture classification model.

The data collected (images of sarees) are segregated as grade 1 and grade 2 based on their texture quality where grade 1 has more quality compared to grade 2. In this dataset 12 500 images are taken in which each class consists of 500 images with a dimension of 181×317 and the cropped images are reduced into 121×215 with the size of 148 kB. Table 1 illustrates the differentiation of 25 varieties of sarees that are used in the experiment.

3.2. Data pre-processing

The saree images are pre-processed into using a CSAB filter to increase the image quality without losing the important features. By linearly mixing pixel values from nearby input pixels, linear filters produce an output pixel value. CSAB filter can be an efficient way of enhancing and improving images, when dealing with images that have poor contrast, noise, and fluctuating local characteristics. CSAB filtering processes pixels in a local area surrounding each image pixel. The proposed method uses the entire image as a neighbourhood, whereas adaptive bilateral filtering uses a window of a defined size surrounding each pixel.

The CSAB filter computes weights to assess the nearby pixels effect on the processed pixel. A weighted average of pixel values in neighbouring pixels is computed using the weights. The proposed methods update the pixel value of the processed pixel after computing the weights and considering nearby pixel values. The transformation function used for contrast stretches each pixel in the image as follows:

$$R = \begin{cases} s * t & 0 \leq t < c \\ n * (t - c) + u & c \leq t < d \\ m * (t - d) + v & d \leq t < s - 1 \end{cases}, \quad (1)$$

where $s, t, u,$ and v are the range of pixel intensity values to cover the images full dynamic range. Now, assume that U is the input saree image and the input intensity range is calculated as a range

$$\text{Range} = |U_{max} - U_{min}|, \quad (2)$$

where U_{max} and U_{min} are the input image maximum and minimum values for the new intensity. The width of the kernel is determined by creating the function $P(y)$ using (3)

$$P(y) = \gamma(y)^{-1} \sum_{yew_l} \omega(Y - L)K_0(y)(f(y), f(l))f(l). \quad (3)$$

By calculating the weighted average of the pixel values in the immediate proximity, the filtered pixel value at the

position (y, l) is determined. The noise-free image with the augmented and annotated dataset of saree is shown in Fig. 2. The pre-processed saree images are fed into the Mask R-CNN for patch generation.

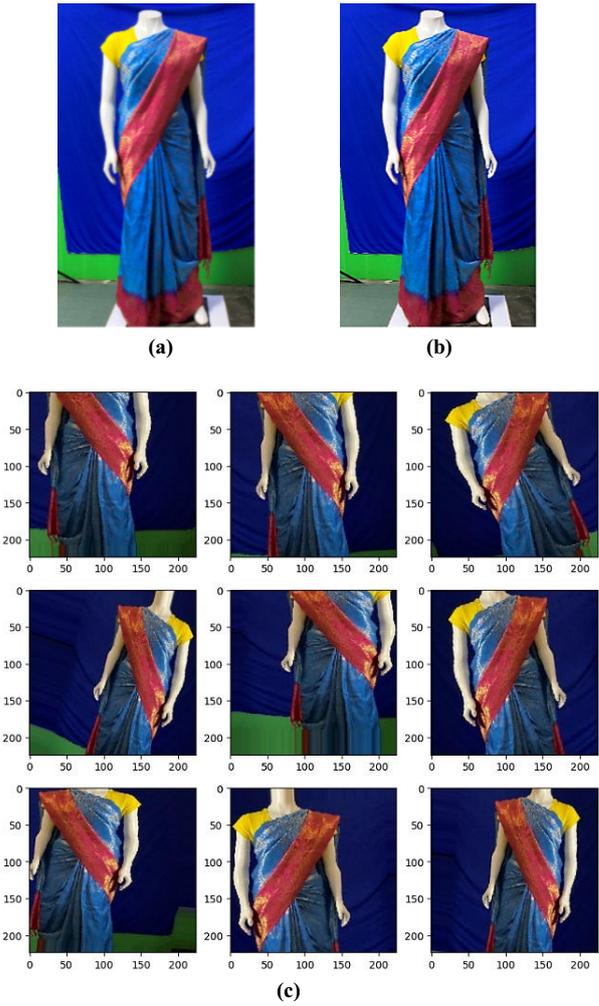


Fig. 2. Sample image of input (a), noise-free image (b), augmented image (c).

3.3. Patch generation

In this phase, Mask-RCNN is used for generating saree patches to classify different textures of saree images. Mask R-CNN segments saree images and generates patches based on the trained data. The Mask R-CNN network consists of three layers: the backbone, the region proposal network (RPN), and the region of interest (RoI) as shown in Fig. 3.

This includes four layers of CNNs and six batches of normalisation layers, and the inclusion of a prediction branch to produce feature maps. A convolution operation is

Table 1. List of saree fabric types.

Saree classes which are used in the experiment				
Benares	Bengal cotton	Binnysilk	Chettinad cotton	Crepe
Georgette-grade 1	Jute cotton	Kanchipuram silk grade 1	Kerala cotton	Kora muslin - grade 1
Kovai cotton	Linen grade 1	Net cotton grade 1	Poly cotton	Poonam
Samuthrika silk	Sana silk	Silk cotton grade 1	Tissue linen	Vasthrakala silk grade 1
Brasso	Crushed satin and georgette	Faux georgette	Georgette and net fabric	Kotta cotton

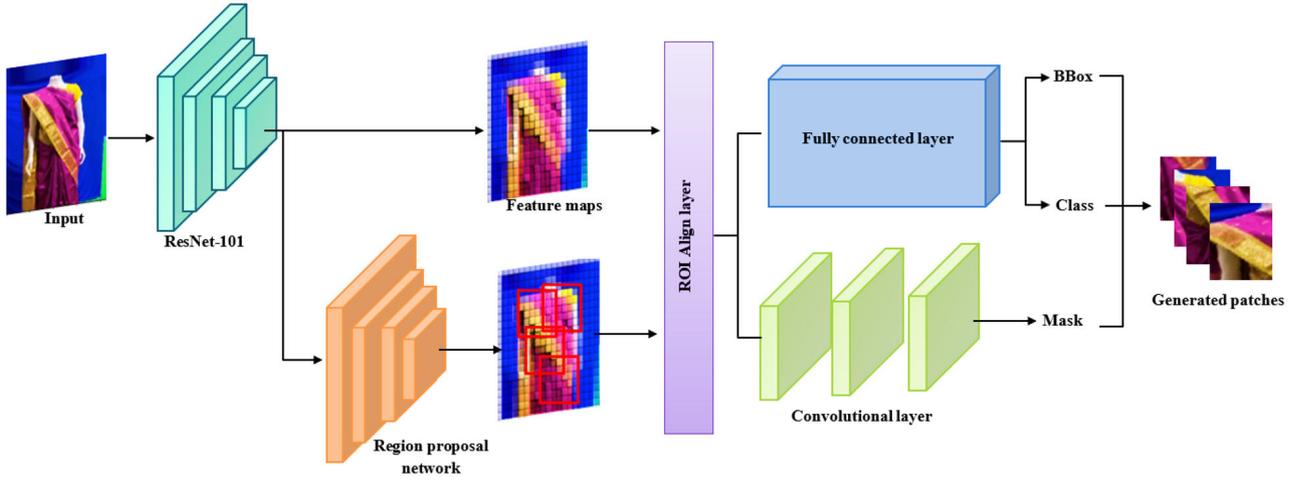


Fig. 3. Mask R-CNN architecture for automatic patch generation.

performed on the input images to extract the conceptual features. In the first step, these convolution feature maps provide the basis for the RPN, which offers proposals that may include foreground, and for the RoI network, which refines the proposal results and obtains the segmentation in the second stage. ResNet-101 is used as the backbone for this framework. Three types of anchors with various length-to-width ratios are available at the RPN stage. The loss function of the RPN is calculated as

$$RPN_{loss} = C_{loss} + R_{loss} = \sum_{i=1} C_{loss}(z_i) + R_{loss}(z_i), \quad (4)$$

where z_i signifies the features from the i^{th} scale. Boundary boxes are created using intersection over union (IoU) for more precise things. RoI Align was used to obtain the improved feature maps after the RoI network stage had rectified the boundary boxes, which were then fed into the mask layer for object segmentation.

Figure 4 shows the generated patches from the pre-processed saree images in the preceding phase. In this phase, a convolutional layer is applied as an alternative of a fully connected layer for classification and regression. The loss function of the RoI layer is calculated as

$$RoI_{loss} = C_{loss} + R_{loss} + G_{loss}. \quad (5)$$

The RoI layer loss function entails with the classification loss C_{loss} , regression loss R_{loss} and patch generation loss G_{loss} . The RoI Align operation will use a bilinear

interpolation to convert a feature map of multiple scales into a single scale. In addition, the segmentation layer is responsible for dividing the RoI. A sparse prediction differs from the previous classification layer that was made with a convolutional layer and a loss function for R-CNN as follows:

$$S_{loss} = \alpha CE_{loss} + \beta B_{loss} + \gamma PG_{loss}, \quad (6)$$

where CE_{loss} signifies the cross-entropy loss, B_{loss} signifies the boundary loss, PG_{loss} signifies the cut loss while generating patches of saree images, and $\alpha = 1$, $\beta = 1.5$, and $\gamma = 1.5$ are the weights for above-mentioned losses. The generated results of Mask R-CNN model of each image is partitioned into small multiple patches and the results of patch is given for classification.

3.4. Texture classification

In this classification, an Aquila optimized EfficientNet is presented with an addition of a bio-inspired optimizer and a squeeze and excitation block for effectual categorisation of saree textures. The architecture of improved EfficientNet is constructed with the reversed bottleneck MB Conv as shown in Fig. 5.

The MBConvblock is composed of convolutional layers that boost and then reduce the channels in order to extract the texture features as t_f . The relations between bottleneck layers are used which have significantly less channels than extension layers. The compound coefficient μ is applied for the compound scaling with certain rules in (7)

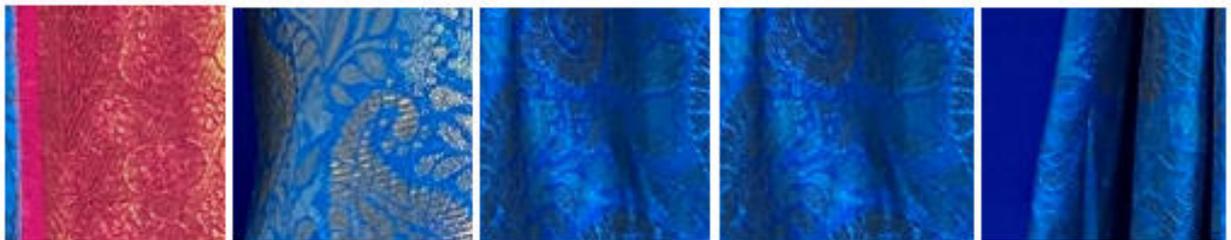


Fig. 4. Generated sample patches of the pre-processed saree image.

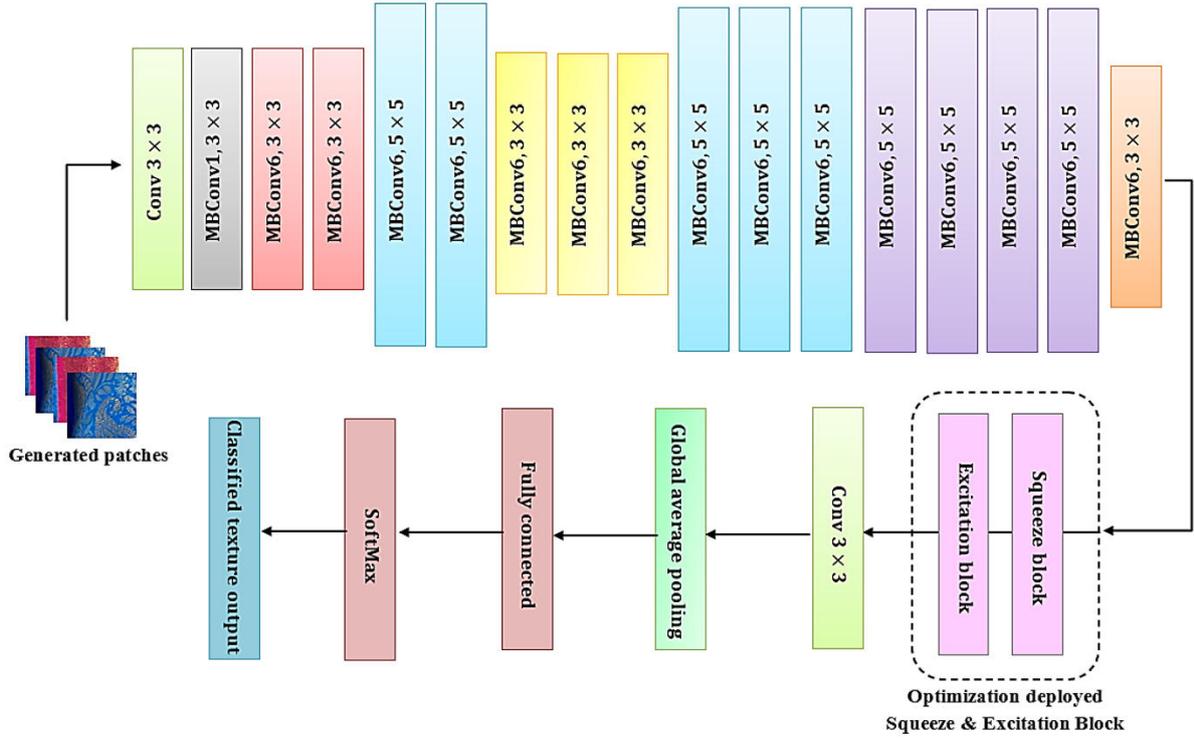


Fig. 5. Architecture of an improved EfficientNet-B3.

$$\begin{aligned} \text{depth: } \mathbb{D} &= \alpha^\mu \\ \text{width: } \mathbb{W} &= \beta^\mu \\ \text{resolution: } \mathbb{R} &= \gamma^\mu, \end{aligned} \quad (7)$$

where $\alpha, \beta, \gamma \geq 1$ are the constants for estimating by searching the grid with the compound coefficient μ . In the standard convolution block, the flops are correlated to $\mathbb{D}, \mathbb{W}^2, \mathbb{R}^2$. The improved block is integrated in the DL-based EfficientNet for efficient extraction of texture features. In this module, an optimization block using squeeze and excitation is used for collecting the extracted features and selecting the relevant features. The extracted texture features t_f from the saree images are applied to the next stage which incorporates an optimizer block for efficient and rapid classification of saree types.

3.5. Aquila optimization

Aquila optimization is used to normalise the parameters of the squeeze and excitation block within the EfficientNet for better classification of saree textures. The parameters of the EfficientNet-B3 are boosted by the addition of the optimizer and this block mainly concentrates on learning texture features based on the Aquila mechanism. Equation (8) is applied to generate the initial population Z that is collected of N solutions.

$$Z_i = lb + R \text{ and } (1, V) * (ub - lb), \quad (8)$$

where R and $(1, V)$ denote a random vector with V values. ub and lb are the search space limits. Then, Z_i is changed to binary by (9)

$$BZ_{ij} = \begin{cases} 1 & \text{if } Z_{ij} > 0.5 \\ 0 & \text{otherwise} \end{cases}. \quad (9)$$

Thus, by ignoring irrelevant features with zero values in BZ_i , equation (9) reduces the number of features selected and the fitness is calculated as

$$\text{fitness} = \beta * \gamma_i + (1 - \beta) * \left(\frac{|BZ_i|}{D} \right), \quad (10)$$

where $\beta \in [0, 1]$ means the weights used to balance the relation between important features $\frac{|BZ_i|}{D}$ and the classification error γ_i . The testing set has been condensed to match the binary of Z_b . Several efficiency metrics have been established based on the reduced data to assess the effectiveness of the classification process for saree texture detection.

The improved EfficientNet was validated and its effectiveness assessed by partitioning the samples into training and test sets. In a single training cycle, multiple trials are conducted to tune hyperparameters. The training application is fully executed with values for the selected hyperparameters set of the network. The number of epochs, neurons, activation function, learning rate, and batch size are the hyperparameters as shown in Table 2 dedicated to EfficientNet. The relative weights of each connection are typically the parameters of a neural network. During the training phase, these parameters are learned. The model hyperparameters are adjusted while the ideal weights are retained. The learning rate of the EfficientNet is 0.1 and the weights are changed by altering the assigned error amount. The proposed EfficientNet has ReLU as input activation, while Sigmoid provides output activation with 100 epochs.

Table 2.
Hyperparameter setting of EfficientNet.

Parameter	Operations	Output map size
–	Input	256 × 256
Stem	Conv3 × 3, BN, Swish	128 × 128
Block 1	MB Conv1, k3 × 3	128 × 128
Block 2	MB Conv6, k3 × 3	64 × 64
Block 3	MB Conv6, k5 × 5	32 × 32
Block 4	MB Conv6, k3 × 3	16 × 16
Block 5	MB Conv6, k5 × 5	8 × 8
Block 6	MB Conv6, k5 × 5	7 × 7
Block 7	MB Conv6, k3 × 3	7 × 7
–	Conv1 × 1, BN, Swish	7 × 7
–	Global average pooling, ReLU	1 × 1
Optimization	Aquila	–

4. Results and discussion

The experimental preparation of this study has been executed with MATLAB 2019b, a DL toolbox. The images from the ISDA dataset [8] were used to identify the saree textures as discussed in this results analysis. The dataset was divided into two parts: 75% as a training set and 25% as a testing set, respectively. The performance of the proposed model and the comparison of machine learning (ML) and DL algorithms are also provided in this section.

4.1. Performance analysis

The proposed SareeNet was assessed based on criteria such as accuracy, precision, recall, and F1-score.

$$\text{Precision} = \frac{T_{pos}}{T_{pos} + F_{pos}} \tag{11}$$

$$\text{Recall} = \frac{T_{pos}}{T_{pos} + F_{neg}} \tag{12}$$

$$\text{Accuracy} = \frac{T_{pos} + T_{neg}}{\text{Total no. of samples}} \tag{13}$$

$$\text{F1 score} = 2 \left(\frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \right). \tag{14}$$

The efficiency of the proposed network for the classification of different types of texture from class 1–25 is illustrated in Table 3 together with the evaluation metrics of the improved EfficientNet-B3, and the EfficientNet also categorises the textures of saree images into 25 categories.

When evaluating 25 various types of saree textures, the proposed improved EfficientNet-B3 attains an accuracy of 98.16%. All scenarios of hyperparameter selection and regularisation techniques are incorporated in the performance measurement of the improved EfficientNet-B3 model. Performance evaluation of the proposed SareeNet is shown in Fig. 6.

As the number of epochs increases, the accuracy of the improved EfficientNet-B3 progresses as depicted in Fig. 7.

Table 3.
Performance metrics of improved EfficientNet-B3.

Class	Accuracy	Precision	Recall	F1-score
1	98.12	99.24	98.07	97.25
2	97.25	98.05	97.25	98.15
3	97.04	98.14	97.21	97.04
4	96.17	96.45	96.08	98.75
5	98.42	97.26	97.28	98.82
6	97.24	97.08	96.21	98.05
7	97.54	97.23	97.22	98.14
8	98.15	96.24	96.23	97.05
9	97.24	98.15	96.25	97.27
10	96.28	97.05	97.18	97.36
11	97.08	98.82	97.07	97.15
12	98.75	98.05	98.05	98.75
13	98.07	97.08	98.12	97.15
14	96.21	98.14	98.75	97.25
15	97.28	97.25	96.21	98.14
16	96.07	97.08	98.14	98.07
17	97.05	97.27	97.36	98.17
18	97.28	98.14	98.17	98.14
19	98.14	97.25	98.07	97.08
20	96.07	96.15	97.05	97.05
21	98.17	97.36	98.42	98.14
22	97.05	98.03	97.05	98.82
23	98.07	98.25	97.25	98.42
24	97.05	97.04	98.17	96.15
25	98.16	98.14	97.15	98.17

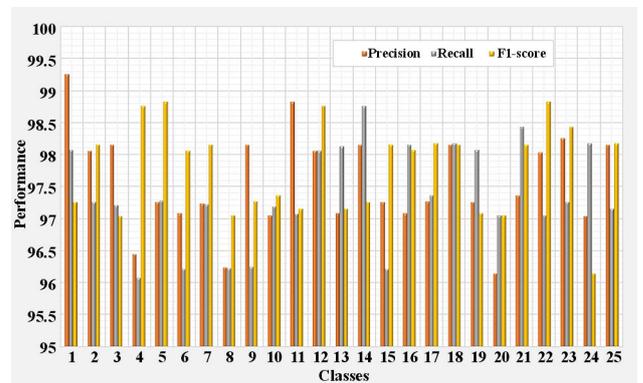


Fig. 6. Performance evaluation of the proposed SareeNet.

Figure 8 displays the epochs and loss that reveal that when the epochs are increased, the loss of the improved EfficientNet-B3 declines. The proposed improved EfficientNet-B3 achieves high-accuracy range in identifying different classes. Several training epochs was first calculated in this research in order to achieve the highest level of testing accuracy of 98.16%.

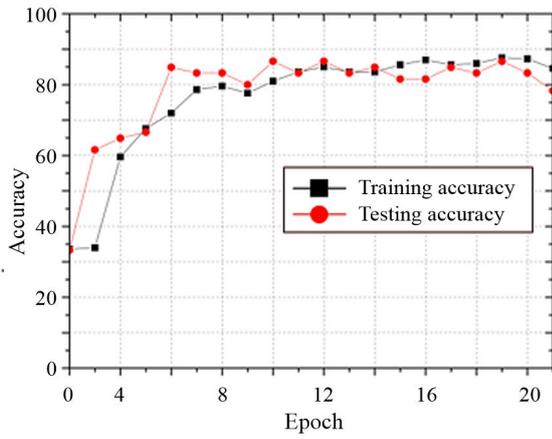


Fig. 7. Accuracy curve of the proposed improved EfficientNet-B3.

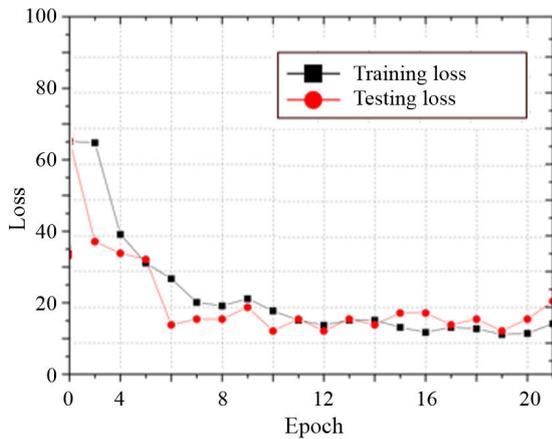


Fig. 8. Loss curve of the proposed improved EfficientNet-B3.

4.2. Comparative analysis

The competence of each DL network was evaluated to assess the results of the performance of the proposed model, which achieves higher accuracy. In the proposed comparison model, five DL classifiers were included: MobileNet, DenseNet-201, ResNet152-v2, InspectionV3, and EfficientNet-B3. Several measures of DL efficiency were used: precision, specificity, recall, accuracy, and the F1-score. The accuracy obtained by the proposed improved EfficientNet-B3 was higher than that of conventional DL networks.

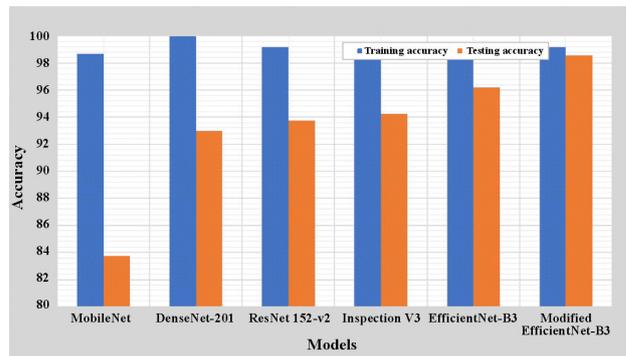
Table 4 shows the classification accuracy for those models with a size of 76 MB. The proposed improved EfficientNet-B3 classification accuracy of the model is 98.16%. The resulting accuracy for the variability of neural network is depicted in Table 3. The improved EfficientNet-B3 was shown to be effective in categorising saree textures with an accuracy of 98.16%. Figure 9 illustrates the comparison of training and testing graphs of the existing and Modified EfficientNet-B3. The accuracy of the proposed Modified EfficientNet-B3 is better than the existing conventional network as shown in Fig. 9(a). The loss curve is determined in the Fig. 9(b) which is comparatively lower than the existing networks.

In conclusion, the proposed model is capable of achieving good results with higher accuracy in classifying ISDA dataset and is further evaluated with a test set of

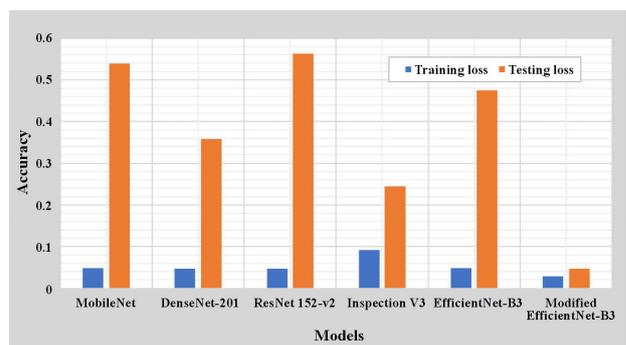
state-of-art models to assess its detection performance. Several traditional DL networks are compared in Table 3, along with their classification accuracy. In comparison to the upgraded Efficient-B3, the standard DL networks did not yield better results. In comparison to MobileNet [24], DenseNet-201 [25], ResNet152-v2 [26], InspectionV3[27], and EfficientNet-B3 [28], the proposed model improves accuracy by 2.54%, 0.17%, 2.06%, 1.78%, and 0.63%, respectively. However, the suggested network outperformed the current networks. The results of our experiments clearly demonstrate that our technology is superior to alternatives. Due to this fact, the results of the SareeNet in determining saree texture seem quite reliable.

Table 4. Classification accuracy of improved EfficientNet-B3.

Model	Training accuracy	Testing accuracy	Training loss	Testing loss	Time (s/epoch)	Step time (ms/step)
MobileNet [24]	95.66	83.75	0.048	0.5396	65 s	522 ms
DenseNet-201 [25]	97.99	93	0.0473	0.359	58 s	461 ms
ResNet152-v2 [26]	96.19	93.75	0.0465	0.5629	38 s	614 ms
InspectionV3 [27]	96.41	94.25	0.0924	0.2442	15 s	243 ms
EfficientNet-B3 [28]	97.54	96.20	0.048	0.4746	47 s	373 ms
Improved EfficientNet-B3	98.16	97.55	0.0296	0.0476	28 s	222 ms



(a)



(b)

Fig. 9. Training and testing accuracy graph (a), loss graph (b).

5. Conclusions

This work introduces a DL-based SareeNet to recognize saree textures. Segmenting the saree field area was performed using an optimized RCNN. After the improved EfficientNet-B3, squeeze and excitation attention blocks were applied to extract high-level texture characteristics. To better categorise saree textures, these blocks were added to EfficientNet-B3. In addition, the improved EfficientNet was optimized with the Aquila optimizer to normalise the parameters for efficient classification. The proposed SareeNet was evaluated for accuracy, precision, recall, and F1-score. The outcomes were also studied in comparison to other deep models and baseline methods. The improved EfficientNet outperforms existing methods, achieving a training accuracy of 98.16%. Experiments have shown that the SareeNet outperforms existing techniques in terms of accuracy level. The proposed enhanced EfficientNet improves the overall accuracy range by 2.54%, 0.17%, 2.06%, 1.78%, and 0.63% better than MobileNet, DenseNet-201, ResNet152-v2, InspectionV3, and EfficientNet-B3, respectively. This approach uses hyperparameters during its training phase, which increases its computing efficiency and makes it more suitable for e-commerce applications. In the future, the authors will focus on using cutting-edge DL or optimization approaches to categorise the 40 different varieties of saree fabric.

Authors' statement

The authors confirm contribution to the paper as follows: study conception and design: K.A., D.K.P.; data collection: A.A.; analysis and interpretation of results: P.J.; draft manuscript preparation: K.A., P.J. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgements

The authors would like to thank the reviewers for all of their careful, constructive, and insightful comments on this work.

References

- [1] Vijayaraj, A. *et al.* Deep learning image classification for fashion design. *Wirel. Commun. Mob. Comput.* **2022**, 7549397 (2022). <https://doi.org/10.1155/2022/7549397>
- [2] Rajput, P. S. & Aneja, S. IndoFashion: Apparel Classification for Indian Ethnic Clothes. in *IEEE/CVF Conference on Computer Vision And Pattern Recognition Workshops* 3930–3934 (IEEE, 2021). <https://doi.org/10.1109/cvprw53098.2021.00440>
- [3] Revathy, G. & Kalaivani, R. Fabric defect detection and classification via deep learning-based improved Mask RCNN. *Signal Image Video Process.* **17**, 1–11 (2023). <https://doi.org/10.1007/s11760-023-02884-6>
- [4] Usmani, U. A., Haponen, A. & Watada, J. Enhanced Deep Learning Framework for Fine-Grained Segmentation of Fashion and Apparel. in *Intelligent Computing Proc. of the 2022 Computing Conference* (ed. Arai, K.) **2**, 29–44 (Springer International Publishing, 2022). https://doi.org/10.1007/978-3-031-10464-0_3
- [5] Almeida, T., Moutinho, F. & Matos-Carvalho, J. P. Fabric defect detection with deep learning and false negative reduction. *IEEE Access* **9**, 81936–81945 (2021). <https://doi.org/10.1109/ACCESS.2021.3086028>
- [6] Sonawane, C., Singh, D. P., Sharma, R., Nigam, A. & Bhavsar, A. Fabric Classification and Matching Using CNN and Siamese Network for E-commerce. in *Computer Analysis of Images and Patterns*. (eds. Vento, M. & Percannella, G.) vol. 11679 (Springer, 2019). https://doi.org/10.1007/978-3-030-29891-3_18
- [7] Padilha, R. & da Silva Vasconcelos, R. C. Prototype for Recognition and Classification of Textile Weaves Using Machine Learning. in *High Performance Computing and Networking*. (eds. Satyanarayana, C., Samanta, D., Gao, X. Z. & Kapoor, R. K.) vol. 853 (Springer, 2022). https://doi.org/10.1007/978-981-16-9885-9_17
- [8] Priya, D. K., Bama, B. S., Ramkumar, M. P. & Roomi, S. M. M. STD-net: saree texture detection via deep learning framework for E-commerce applications. *Signal Image Video Process.* **27**, 1–9 (2023). <https://doi.org/10.1007/s11760-023-02757-y>
- [9] Chaitra, H. V., Shadaab, M., Dubey, K., Subudhi, R. & Maurya, S. Image Segmentation and Classification for Fashion Apparel. in *2022 IEEE North Karnataka Subsection Flagship International Conference (NKCon)* 1–6 (IEEE, 2022). <https://doi.org/10.1109/nkcon56289.2022.10126905>
- [10] Dinesh Jackson, S. R. *et al.* Real time violence detection framework for football stadium comprising of big data analysis and deep learning through bidirectional LSTM. *Comput. Netw.* **151**, 191–200 (2019). <https://doi.org/10.1016/j.comnet.2019.01.028>
- [11] Zhang, H. *et al.* ClothingOut: a category-supervised GAN model for clothing segmentation and retrieval. *Neural Comput. Appl.* **32**, 4519–4530 (2020). <https://doi.org/10.1007/s00521-018-3691-y>
- [12] Hsieh, C. W. *et al.* FashionOn: Semantic-Guided Image-Based Virtual Try-On with Detailed Human and Clothing Information. in *Proc. of the 27th ACM International Conference on Multimedia* 275–283 (Association for Computing Machinery, 2019). <https://doi.org/10.1145/3343031.3351075>
- [13] Steffens Henrique, A. *et al.* Classifying garments from fashion-MNIST dataset through CNNs. *Adv. Sci. Technol. Eng. Syst. J.* **6**, 989–994 (2021). https://comum.rcaap.pt/bitstream/10400.26/37237/1/ASTESJ_0601109.pdf
- [14] Kolisnik, B., Hogan, I. & Zulkernine, F. Condition-CNN: A hierarchical multi-label fashion image classification model. *Expert Syst. Appl.* **182**, 115195 (2021). <https://doi.org/10.1016/j.eswa.2021.115195>
- [15] Zhang, J. & Liu, C. A study of a clothing image segmentation method in complex conditions using a features fusion model. *Automatika* **61**, 150–157 (2020). <https://doi.org/10.1080/00051144.2019.1691835>
- [16] Xiang, Z., Zhu, C., Qian, M., Shen, Y. & Shao, Y. FashionSegNet: a model for high-precision semantic segmentation of clothing images. *Vis. Comput.* **39**, 1–17 (2023). <https://doi.org/10.1007/s00371-023-02881-3>
- [17] Varma, M. & Zisserman, A. A statistical approach to texture classification from single images. *Int. J. Comput. Vis.* **62**, 61–81 (2005). <https://doi.org/10.1007/s11263-005-4635-4>
- [18] Varma, M. & Zisserman, A. A statistical approach to material classification using image patch exemplars. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**, 2032–2047 (2009). <https://doi.org/10.1109/TPAMI.2008.182>
- [19] Dana, K., Van-Ginneken, B., Nayar, S. & Koenderink, J. Reflectance and texture of real-world surfaces. *ACM Trans. Graph.* **18**, 1–34 (1999). <https://doi.org/10.1145/300776.300778>
- [20] Padilha, R. & da Silva Vasconcelos, R. C. Prototype for Recognition and Classification of Textile Weaves Using Machine Learning. in *High Performance Computing and Networking: Select Proceedings of CHSN* (eds. Satyanarayana, C., Samanta, D., Gao, X. Z. & Kapoor, R. K.) 205–213 (Springer, 2022). https://doi.org/10.1007/978-981-16-9885-9_17
- [21] Deng, J. *et al.* Imagenet: A Large-Scale Hierarchical Image Database. in *IEEE Conf. on Computer Vision and Pattern Recognition. (CVPR)* 248–255 (IEEE, 2009). <https://doi.org/10.1109/CVPR.2009.5206848>
- [22] Sharan, L., Rosenholtz, R. & Adelson, E. Material perception: What can you see in a brief glance? *J. Vis.* **9**, 784 (2009). <https://doi.org/10.1167/9.8.784>
- [23] Nocentini, O., Kim, J., Bashir, M. Z. & Cavallo, F. Image classification using multiple convolutional neural networks on the Fashion-MNIST Dataset. *Sensors* **22**, 9544 (2022). <https://doi.org/10.3390/s22239544>

- [24] Wang, W. *et al.* A novel image classification approach via dense-MobileNet models. *Mob. Inf. Syst.* **2020**, 7602384 (2020). <https://doi.org/10.1155/2020/7602384>
- [25] Sanghvi, H. A. A deep learning approach for classification of COVID and pneumonia using DenseNet-201. *Int. J. Imaging Syst. Technol.* **33**, 18–38 (2023). <https://doi.org/10.1002/ima.22812>
- [26] Uddin, A. H., Mahamud, S. S. & Arif, A. S. M. A Novel Leaf Fragment Dataset and Resnet for Small-Scale Image Analysis. in *Intelligent Sustainable Systems: Proceedings of ICISS 2021* (eds. Raj., J. S., Palanisamy, R., Perikos, I. & Shi, Y.) 25–40 (Springer, 2021). https://doi.org/10.1007/978-981-16-2422-3_3
- [27] Xia, X., Xu, C. & Nan, B. Inception-v3 for Flower Classification. in *2017 2nd International Conference on Image, Vision and Computing (ICIVC)* 783–787 (IEEE, 2017). <https://doi.org/10.1109/icivc.2017.7984661>
- [28] Alquzi, S., Alhichri, H. & Bazi, Y. Detection of COVID-19 Using EfficientNet-B3 CNN and Chest Computed Tomography Images. in *International Conference on Innovative Computing and Communications Proc. of ICICC 2021* (eds. Khanna, A., Gupta, D., Bhattacharyya, S., Hassanien, A. E., Anand, S. & Jaiswal, A.) **1**, 365–373 (Springer, 2022). https://doi.org/10.1007/978-981-16-2594-7_30