

Regularised Layerwise Weight Norm Based Skin Lesion Features Extraction and Classification

S. Gopikha* and M. Balamurugan

School of Computer Science of Engineering, Bharathidasan University, Tiruchirapalli, 620023, India

*Corresponding Author: S. Gopikha. Email: gopikha.in@gmail.com

Received: 13 February 2022; Accepted: 07 April 2022

Abstract: Melanoma is the most lethal malignant tumour, and its prevalence is increasing. Early detection and diagnosis of skin cancer can alert patients to manage precautions and dramatically improve the lives of people. Recently, deep learning has grown increasingly popular in the extraction and categorization of skin cancer features for effective prediction. A deep learning model learns and co-adapts representations and features from training data to the point where it fails to perform well on test data. As a result, overfitting and poor performance occur. To deal with this issue, we proposed a novel Consecutive Layerwise weight Constraint MaxNorm model (CLCM-net) for constraining the norm of the weight vector that is scaled each time and bounding to a limit. This method uses deep convolutional neural networks and also custom layer-wise weight constraints that are set to the whole weight matrix directly to learn features efficiently. In this research, a detailed analysis of these weight norms is performed on two distinct datasets, International Skin Imaging Collaboration (ISIC) of 2018 and 2019, which are challenging for convolutional networks to handle. According to the findings of this work, CLCM-net did a better job of raising the model's performance by learning the features efficiently within the size limit of weights with appropriate weight constraint settings. The results proved that the proposed techniques achieved 94.42% accuracy on ISIC 2018, 91.73% accuracy on ISIC 2019 datasets and 93% of accuracy on combined dataset.

Keywords: Norm; overfitting; regularization; melanoma; weight constraints

1 Introduction

In the United States, skin cancer is a deadly and shared disease [1]. More than half of all new cases diagnosed between 2008 and 2018 were malignant, according to recent reports. Deaths from this disease are predicted to increase over the next decade. If diagnosed at a later stage, the survival rate drops to less than 14%. It is possible, however, to save nearly 97% of people with early-stage skin cancer. Malignant lesions can be diagnosed using a dermoscopic visual examination by dermatologists [2]. Detecting skin cancer using dermoscopy can be difficult because of the variety of textures and wounds on the surface of the skin. A manual examination of dermoscopic images, however, makes it hard to accurately diagnose skin cancer. The dermatologist's experience affects the precision with which a lesion can be correctly



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

diagnosed. Biopsy and macroscopic are the only other methods available for the detection of skin cancer. Skin lesions necessitate more time and attention because of their complexity [3].

Observation with the naked eye, followed by dermoscopy (microscopically magnifying lesions) and a biopsy, are the most common steps taken by a skilled dermatologist. While the patient would have to wait, this could result in the patient moving on to more advanced stages. As a result, an accurate diagnosis relies on the clinician's skill. Dermatologists have an accuracy of less than 80% in properly identifying skin cancer, according to a recent study. To make matters worse, the world's public health systems have a dearth of dermatologists with the necessary training [4]. The development of computer image analysis algorithms has been extensively researched in order to diagnose skin cancer quickly and at the earliest stage and solve some of the difficulties mentioned above. These algorithms were mostly parametric, requiring normally distributed data to work [5]. These methods would not be able to precisely diagnose the disease because the nature of the data cannot be controlled. It's not necessary for a non-parametric solution to assume normal distributions, but it can still be useful.

Computer Aided Design (CAD) techniques have been introduced by a number of medical imaging researchers. A four-step CAD process includes pre-processing images, identifying infected parts, extracting features, and classifying them. Dermatologists can use a computerised method as a second opinion to confirm the results of a manual diagnosis [6]. Convolutional Neural Network (CNNs) are a type of deep learning that can be used to automatically extract features from a dataset. Using a computer vision technique known as a CNN, an image can be automatically recognized and distinguished from other images [7]. Convolutional layer, rectified linear activation unit (ReLU) layer, normalisation layer, pooling layer, fully connected layer, and Softmax layer [8] are typical layers in a simple CNN. For classification tasks, researchers used pre-trained deep learning models in numerous techniques. AlexNet, Visual Geometry Group (VGG), GoogleNet, InceptionV3, and residual neural network (ResNet) are a few publicly available pre-trained deep learning models. These models were used by researchers through transfer learning. In several cases, feature selection and fusion were utilised to enhance recognition accuracy. Dermatologists and physicians can benefit from computer-aided diagnostic tools, which can help them make judgments and minimise testing expenses. In order to develop an automated skin lesion identification mechanism, there are numerous difficulties to overcome, including changing appearance and imbalanced datasets.

1.1 Problem Statement

Skin lesion classification accuracy is pretentious by some factors. The following matters are addressed in this study:

- Skin lesion classification accuracy is affected by a number of factors. The following issues are addressed in this study.
- The overlaps of skin lesions may leads to less classification accuracy, hence it requires appropriate segmentation techniques for high accuracy on skin cancer.
- Types of skin lesions have comparable shapes, colours and textures and so extract similar characteristics. At this point, the features are categorised into the wrong skin type.
- As a result of its inability to perform on test data, feature representations learned from training data are prone to overfitting.

1.2 Major Contribution

The novelty of this research work is explained as follows:

- To classify the correct skin lesion class, suitable pre-processing technique is required. Here, De Trop Noise Exclusion with in-painting is introduced for hair removal process.

- To improve the classification accuracy, segmentation techniques includes patch detection, image conversion, edge preservation and k-means clustering technique are used, before extracting the important features.
- For feature extraction and classification, *CLCM-net* is designed, where overfitting issue is solved by regularising the weights of the proposed model. Its performance is enhanced by training the network with small weights, i.e., weight constraints in the network are scaled each time and bound to a limit.

1.3 Structure of the Research Paper

The basic introduction about skin cancer, deep learning models and its pre-trained models along with problem statement is given in Section 1. The in-depth study of existing techniques for skin cancer classification is provided in Section 2. The brief explanation of proposed model for pre-processing, segmentation and classification is depicted in Section 3. The experiment is conducted on two publicly available datasets for analysing the performance of proposed model with existing models, which is shown in Section 4. Finally, the scientific contribution of the research work with future development is signified in Section 5.

2 Related Works

With a deep convolutional neural network, a 2018 study by Haenssle et al. [9] found 86.6 percent accuracy and specificity in the classification of dermatoscopy melanocytic pictures. Chaturvedi et al. [10] approach for multiclass skin cancer categorization was given. Classification models (pretrained deep learning), fine-tuning, feature extraction and performance evaluation were all part of the strategy given here. The individual model (ResNet-101) attained a maximum accuracy of 93.20 percent during the evaluation procedure, but the ensemble model (InceptionResNetV2 + ResNet-101) reached a complete precision of 92.83 percent. Ultimately, they concluded that the best hyper-parameter settings for deep learning models might outperform even ensemble models.

Huang et al. [11] provided an algorithm for the detection of skin lesions that is both quick and easy to use. More dependable, practicable, and user-friendly than the previous method. The Human Against Machine with 10000 training images (HAM10000) dataset was used for the experiment, and the accuracy rate was 85.8%. In addition, this approach was evaluated on a five-class Kaohsiung Chang Gung Memorial Hospital (KCGMH) dataset and attained an accuracy of 89.5 percent.

To compensate for the HAM10000 dataset's intrinsic class imbalance, Le et al. [12] developed an ensemble of ResNet50 networks using class-weighting and a focused loss function. Lesions were segregated during the pre-processing stage. Although this method yielded a decrease in accuracy, it suggests that the skin surrounding the lesion is critical to the network's learning of distinguishing traits.

For skin cancer classification, Hekler et al. [13] studied the impact of label noise on CNNs. There is a risk of systematic error because numerous skin cancer classification studies employ non-biopsy-verified training images. This association was found amid models trained with dermatologists' diagnoses and good quality results on a dermatologist-created test set. Dermatologists were able to recognise the qualities that CNNs were able to identify, but CNNs were also able to identify the sources of dermatological errors. A CNN's accuracy declined from 75.03 percent to 64.24 percent when tested against a biopsy-verified ground truth using a CNN trained with majority judgments, according to the study's authors. Nevertheless, this research has a few flaws, including the following: All lesions had been biopsied, which are difficult to categorise by nature and hence comprise edge cases, with the authors noting that the addition of simpler samples would likely increase network accuracy. (ImageNet pre-trained ResNet50 was utilised.)

Classification and detection testing for the 2014 Inception ImageNet Large Scale Visual Recognition Challenge (ILSVRC) used GoogleNet [14]. AlexNet, one of the first classification networks, has a total of 22 layers, whereas this network has a total of 24 layers. More neurons were added to each level of the network, further increasing its complexity. (3) A collection of convolution layers and a max pooling with ReLU activation are included in the Inception module. The input is a 224 by 224 Red Green Blue (RGB) image, and the output is a 1000-class probability vector.

Based on the change of the connections between layers, DenseNet201 [15] is a deep network. When compared to typical nets, DenseNet connects one layer to all of the following layers, i.e., all of the preceding feature maps are used as input for the following layers. For a total of 201 layers, dense blocks are connected by convolutional, max pooling, and activation layers, which are used as transitions between each block and the next. High performance can be achieved with minimal computing by reducing the number of parameters in the network, as proved by tests on ImageNet, Canadian Institute For Advanced Research (CIFAR-10 and 100), and Street View House Numbers (SVHN) dataset. This strategy also has the benefit of reducing the vanishing-gradient problem by performing feature reuse and propagation on the network.

It is a hybrid of two well-known deep networks, Inception-ResNetV2 [16], which seeks to gain from ResNet's residual connections in order to speed up Inception's training time. In practice, the summations are applied to simpler Inception blocks, followed by filter-expansion layers to increase the dimensionality of the Inception block, and then the summation is applied. This 164-layer convolutional neural network was obtainable at the 2015 ILSVRC challenge with the goal of enhancing the ILSVRC 2012 classification performance.

Research Gap:

1. All the pre-trained models such as ResNet, AlexNet, GoogleNet, DenseNet and InceptionNet are tested with different datasets such as ILSVRC, CIFAR, SVHN3, but none of these techniques are tested with ISIC challenging datasets.
2. Mostly all of the pre-trained models considered the input directly from the dataset, no segmentation and patch detection are considered for accurate skin cancer classification, which leads only 85% of accuracy.
3. Overfitting issues are occurred during the learning of feature representations from training data.

These issues are addressed by developing the proposed model CLCM-net along with patch detection technique and evaluated with ISIC challenging datasets.

3 Proposed System

In this research work, two datasets from ISIC are considered as an input images for pre-processing techniques. To get an enhanced images, four process along with DeTrop Noise Exclusion are carried out during pre-processing. Then, patch detection using You Only Look Once-Version 5 (Yolo-V5) network, image conversion, edge preservation using 2D-Otsu thresholding and active contour model along with k-means clustering are used for segmenting the skin lesions from pre-processed images. Then, pre-trained model of *CLCM-net* is introduced to solve the issues of overfitting. Finally, the performance validation is carried out by using training and testing dataset. Fig. 1 displays the working flow of the proposed model.

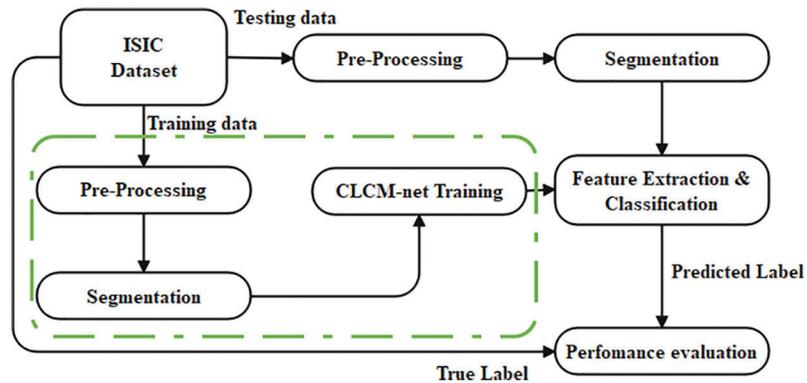


Figure 1: The proposed system model

3.1 Dataset Description

The dataset is acquired from ISIC (International skin imaging collaboration) 2018 challenge dataset archive which contains benign and malignant skin lesions dermoscopy images. There are 2650 images of training images and 712 testing image [17]. Sample Figures from the dataset is provided in Fig. 2.

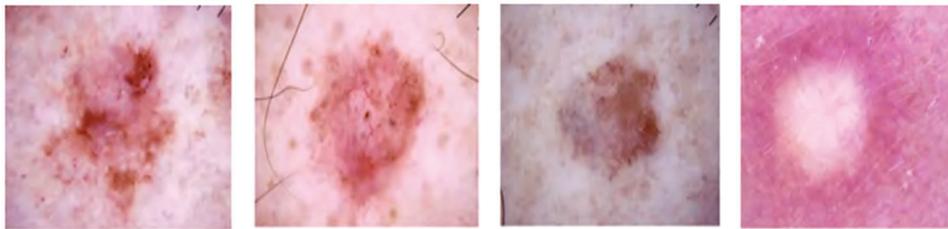


Figure 2: Samples of the seven different classes of images

The dataset is acquired from ISIC (International skin imaging collaboration) 2019 challenge dataset archive, which contains benign and malignant skin lesions dermoscopy images. There are 2637 images of training images and 660 testing image [18].

3.2 Pre-processing

Initially, the hair is removed from the input images, five process are presented in the proposed research that includes vintage boosting, gray contrast stretching, filtering, mask creation and DeTrop Noise Exclusion using inpainting. Digital inpainting allows for the rebuilding of minor damaged areas of an image. Digital inpainting can be used to remove text and logos from still photos or films, recreate scans of damaged images by eliminating scratches or stains, or create aesthetic effects. The missing area should be inserted to estimate the gray value and gradient from that area when it is inserted. Navier-Stokes is used to perform inpainting in the proposed method.

Inpainting is an incompressible fluid according to Navier-Stokes. The picture intensity function serves as the stream function, with isophote lines defining the flow's streamlines. When the user selects the regions to be repaired, inpainting zone is filled in automatically by the algorithm. The fill-in is done so that isophote lines that approach the region's borders are completed inside. When a new piece of information is discovered, it doesn't matter where it came from. This is done automatically (and quickly), allowing for the simultaneous filling in of a large number of unique zones, each with their own distinct structures and surrounding backdrops. Furthermore, there are no restrictions on the region's topology when painting it. There is simply one human contact required by the programme: designating the areas that need to be

painted. Example output image of pre-processing technique is shown in Fig. 3, which illustrates the DeTrop Noise elimination procedure described in Algorithm 1.

Algorithm 1: DeTrop Noise Exclusion using In-painting

Input: $Mask(BTH)(NX_{I=1, \dots, n})$

Output: $De(NX_{I=1, \dots, n})$

Compute

$De(NX_{I=1, \dots, n}) \rightarrow ((NX_{I=1, \dots, n}), Mask(BTH)(NX_{I=1, \dots, n}), inpainting - rad, Inpaint - NS)$

Return $De(NX_{I=1, \dots, n})$

Where,

NS Navier strokes

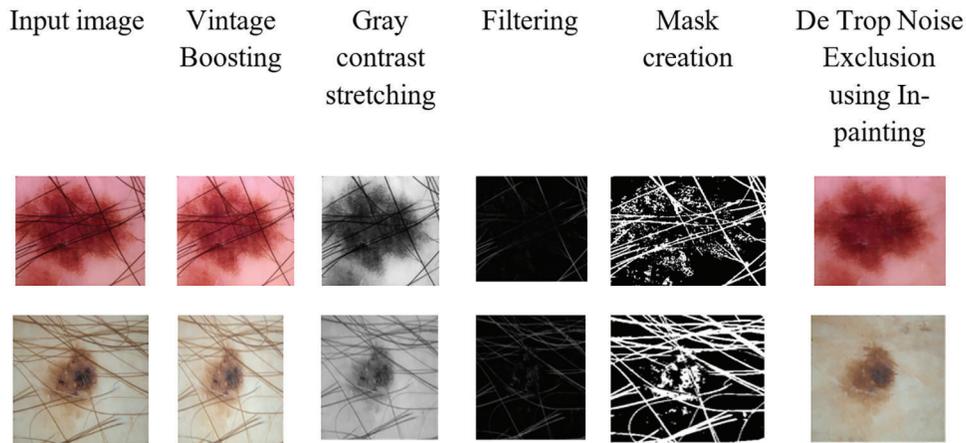


Figure 3: Sample output for hair removal

3.3 Segmentation

In image annotation, the python script was used to build rectangle bounding boxes over the affected area of skin patches and lesions, and the object was annotated as Lesion. For patch detection, there is also an FPN-based feature pyramid structure in YOLOv5. For example, the FPN layer carries high semantic characteristics, while the feature pyramid carries strong placement characteristics. Improve the network's ability to detect targets at various scales by using feature aggregation from several feature layers. Based on the YOLOv5 network, development of network for segmentation (Dseg-net) is introduced and a patch is detected.

After the patch detection, the color image is rehabilitated into gray scale images by using gray conversion technique called 2D-Otsu algorithm and active contour model is used to preserve the edges of the images. Finally, k-means clustering is used, in which the sum of the distances between the centroid of each cluster and the objects in all clusters is minimised iteratively. In order to allocate the points in a cluster that have the shortest Euclidean distance between them, each cluster's centroid is recalculated once the grouping process has been completed. Finally, in Fig. 4, the segmented output is shown.

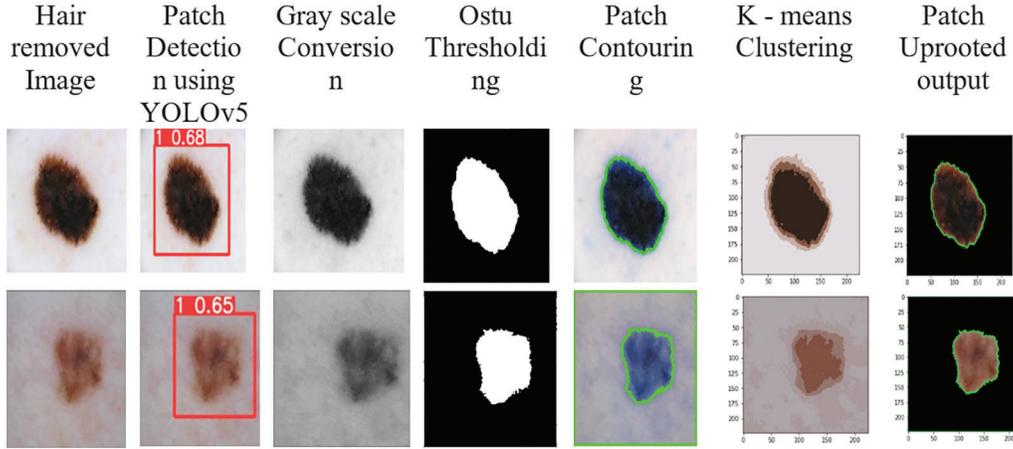


Figure 4: Sample output for segmentation

3.4 Feature Extraction and Classification

Nowadays, the researchers used CNN technique in various applications namely face recognition, image classification, action recognition, re-identification of vehicles from similar ones and scene labelling, because it is important deep learning architectures. The various fixed structures of input signals are obtained by dividing the spectrograms and these fixed structures are given as input to small neural networks. The feature map is represented by extracting every features and formed as a set of arrays. In the proposed network, the primary element is defined as the convolutional layers, where the local information about shape, color, texture of image and other features are extracted by using the convolution kernel. The noise interferences are reduced by improving the input features using convolutional operations. The mathematical Eq. (1) shows the mapping process that are occurred in convolution process.

$$x_j^l = f_c \left(\sum_{i \in M_j} x_i^{l-1} \times k_{i,j}^l + \theta_j^l \right) \quad (1)$$

where, the convolution layer is depicted as x_j^l or l and the mapping set as j . The i^{th} feature set is represented as x_i^{l-1} in the convolution layer of $(l-1)$. In the layer l , the convolution kernel between j^{th} mapping set and i^{th} feature set is illustrated as $k_{i,j}^l$. The activation function is described as f_c and bias is defined as θ_j^l . During training, the pooling process is the next step in the network, where the possibility of overfitting is minimized by the pooling. The mathematical Eq. (2) explains the pooling process.

$$x_j^l = f_p \{ \beta_j^l \text{down}(x_i^{l-1}) + \theta_j^l \} \quad (2)$$

where, the down sampling method is described as $\text{down}(\cdot)$ from $(l-1)$ layer to l^{th} layer. In general, the average pooling and maximum pooling are presented in the CNN. Therefore, the additive bias is illustrated as θ_j^l and multiplicative bias is represented as β_j^l . The activation function in the pooling layer is denoted as $f_p(\cdot)$. A rasterization layer is formed by sequentially arranging the final pooling layer's matrix features in a vector form and that corresponds to the fully connected layer. Eq. (3) describes the output value of any node j as follows:

$$h_j = f_h \left(\sum_{i=0}^{n-1} W_{i,j} x_i - \theta_j \right) \quad (3)$$

where, the connection weight of the input vector x_i is described as W_{ij} . The activation function is denoted as $f_h(\cdot)$ and the node threshold is represented as θ_j for the node j . The class labels and score are developed for classification process by using these fully connected layers. In every convolutional layer, it is observed that the training effect and training speed is highly influenced by the total number of filters. The performance of CNN model is minimized, when there are high total of filters and smaller number of samples present in the network.

When a neural network is modelled, the model learns features based on weights and predictions depending on the parameters learned. The values that control the system's behaviour are called weights. Typically, vectors have weight values that are large in magnitude (for example, 538.1234 or -1098.5921) rather than small (for example, 3.8392 or -2.0944). We can control whether a model is more likely to overfit or underfit. Models with high weights can overfit by memorising properties of the training set that do not serve them well on the test set. In order to resolve the overfitting issue in the neural network model, a deep learning model of CLCM-net is proposed. It consists of 15 layers composed of four sets of consecutive layerwise maxnorm conv2D constraint layers, max pooling followed by global average pooling, flattening, Maxnorm Dense layer and finally the classification layer. The patched uprooted skin lesion dataset is used for training and testing. The patch uprooted RGB image is fed as an input with a size of (224, 224, 3) to the proposed model. For each maxnorm weight constraint layer, the vector norm of the incoming weight at each hidden unit is bound to a limit m . The magnitude of the vector is calculated using maxnorm. By customising the weights in each max norm conv2d layer, kernel constraints and bias constraints values are fixed. This tends to constrain the whole weight matrix directly. If the maxNorm of your weights exceeds m , scale your whole weight matrix by a factor that reduces the norm to m . Here, a simple and effective CNN implementation with proposed *CLCM-net and its algorithm* are explained. Fig. 5 displays the architecture of proposed model.

Algorithm of the Proposed CLCM-net Architecture

```

Input: PUI (i = 1,2,..n)
Output: CLI (i = 1,2,..n)
for each PUI (i)
  Read PUI (i) ( $w(\vec{I_x})$ )
  for each PUI (i) ( $w(\vec{I_z})$ )  $\rightarrow$  calculate ( $\|w(\vec{I_x})\|_{maxnorm} < m$ )
  for each layers
    ( $L(\|w(\vec{L_x})\|_{maxnorm} < m)$ ), ( $\|k_c\|_{maxnorm} < z$ ), ( $\|B_c\|_{maxnorm} < t$ )
  For Dense layer
    LD( $\|w(\vec{I_x})\|_{maxnorm} < m$ ), ( $\|k_c\|_{maxnorm} < z$ ),
    CLI (i = 1,2,..n)

```

where, PUI (i = 1,2,..n) is patched uprooted skin lesion image, CLI (i = 1,2,..n) is classified output image, ($\|w(\vec{I_x})\|_{maxnorm}$) is maxnorm regularization for weight vector, m , z and t are constants, $L(\|w(\vec{L_x})\|_{maxnorm})$ is layer wise maxnorm weight constrain, $\|k_c\|_{maxnorm}$ is maxnorm kernel constraints, $\|B_c\|_{maxnorm}$ is maxnorm bias constraints and LD ($\|w(\vec{I_x})\|_{maxnorm}$) is maxnorm Dens layers.

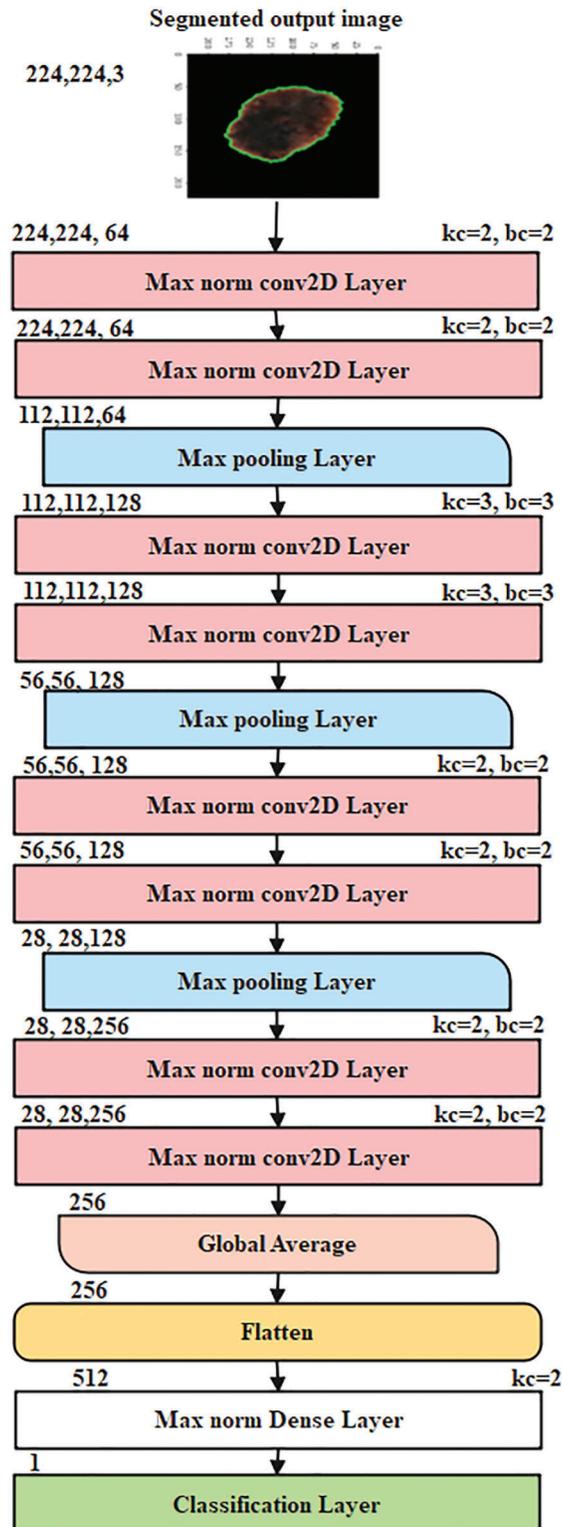


Figure 5: The proposed CNN CLCM-net architecture

4 Results and Discussion

There were 64GB of DDR4 2400 MHz Dual Channel RAM and a GTX1080 Ti 11GB GPU used to train the networks. Tensorflow GPU 2.4.1 and Keras 2.3.1 were used in the software configuration on Windows 10. For baseline experiments, we trained 5 of the most widely used deep learning architectures: DenseNet, GoogleLeNet, ResNetV5, InceptionV3, Mobile Network version 2 (MobileNetV2). For training data, we trained and validated our model on an 80:20 split of photos from the ISIC 2018–2019 datasets, utilising a balanced dataset. The goal of this work is to offer baseline data without the use of extra methods, hence transfer learning was not employed in any of the trials.

Stochastic Gradient Descent (SGD) was used to train the 19 networks with 32 batches for 50 epochs and an initial learning rate of 0.01 for each network. Early stopping was implemented until each network converged, which was defined by a patience of 10 epochs. Several data augmentation techniques were used to make up for the small size of the training set, such as random rotations, random zooms, random shifts in width and height, and horizontal and vertical fusing.

4.1 Performance Evaluation Metric

The accuracy, precision, recall, and F-measure of each model and each deep network were compared using the conventional classification methods. Accuracy measures the percentage of correct detections, whereas precision measures the percentage of relevant examples retrieved, and recall measures the percentage of relevant instances retrieved over the total number of relevant instances. The F-measure is a useful tool for evaluating a method's overall performance, taking into account both precision and recall.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

$$F - measure = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (7)$$

4.2 Performance of Proposed Model Evaluation

Here, the performance of proposed model is tested with various existing pre-trained models using two datasets along with combined dataset in terms of different parameters. Initially, [Tab. 1](#) shows the validation analysis of different pre-trained models on ISIC 2018 dataset.

Table 1: ISIC 2018 dataset performance evaluation

Model	Accuracy	Recall	Precision	F-Measure
DenseNet	0.8834	0.8009	0.8175	0.8091
GoogleLeNet	0.8862	0.7492	0.8657	0.8032
ResNetV5	0.8605	0.6661	0.8035	0.7284
InceptionV3	0.8390	0.6523	0.7667	0.7049
MobileNetV2	0.8532	0.7045	0.8134	0.6843
<i>CLCM-net</i>	0.9442	0.9050	0.9203	0.9126

In the accuracy analysis, the DenseNet and GoogleNet achieved nearly 88%, ResNetV5 achieved 86%, InceptionV3 and MobileNet achieved nearly 84%, but the proposed model achieved 94.42%. The reason for better performance of proposed model is that the before giving the input to the proposed model, five different pre-processing techniques are implemented in this research work. While comparing with all pre-trained models, Inception V3 achieved poor performance and this is due to the requirement of more computational resources for training that leads to overfitting issues and the speed is less. In the analysis of recall, ResNetV5 and InceptionV3 achieved nearly 66%, DenseNet achieved 80% and proposed model achieved 90.50%, where DenseNet decreases the computation and parameter efficiency that leads to less performance, when compared with proposed model. The *CLCM-net* achieved 92.03% of precision and 91.26% of F-measure, where other pre-trained models achieved nearly 76% to 85% of precision and F-measure. The next experiments is carried out on ISIC 2019 dataset for validating its performance and it is shown in [Tab. 2](#).

Table 2: ISIC 2019 dataset performance evaluation

Model	Accuracy	Recall	Precision	F-Measure
DenseNet	0.8604	0.6845	0.7818	0.7299
GoogleLeNet	0.8234	0.6234	0.7323	0.6735
ResNetV5	0.8280	0.7086	0.6855	0.6969
InceptionV3	0.8277	0.7992	0.7522	0.6002
MobileNetV2	0.8467	0.7186	0.8090	0.7045
<i>CLCM-net</i>	0.9173	0.8480	0.8530	0.8505

When comparing with ISIC 2018 dataset, the performance of the all pre-trained models achieved less performance in ISIC 2019 dataset. For instance, the *CLCM-net* achieved only 91.73% of accuracy and nearly 85% of precision, recall and F-measure, where the existing techniques achieved nearly 68% to 73% of recall, 82% to 86% of accuracy, 73% to 80% of precision and 66% to 72% of F-measure. This shows that the data plays a major role in the classification accuracy for skin cancer detection. Even though, MobileNet minimized the parameters, size and performs faster, it achieved less accuracy than DenseNet and proposed model. [Fig. 6](#) shows the graphical comparison of various models in terms of accuracy on two datasets.

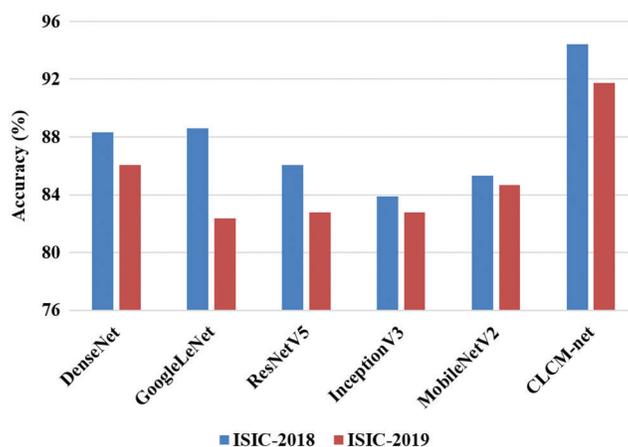


Figure 6: Comparative analysis of proposed *CLCM-net* in terms of accuracy on two different datasets

4.3 Performance of Proposed Model for Combined Dataset

While combining both datasets, the performance of the proposed model is validated and tabulated in Tab. 3 and Fig. 7.

Table 3: ISIC 2018 & ISIC 2019 combined dataset performance evaluation

Model	Accuracy	Recall	Precision	F-Measure
DenseNet	0.8718	0.7427	0.7996	0.7695
GoogleLeNet	0.8548	0.6863	0.7990	0.7383
ResNetV5	0.8442	0.6873	0.7445	0.7126
InceptionV3	0.8311	0.5755	0.7594	0.6525
MobileNetV2	0.8499	0.7115	0.8112	0.6944
CLCM-net	0.9307	0.8765	0.8866	0.8965

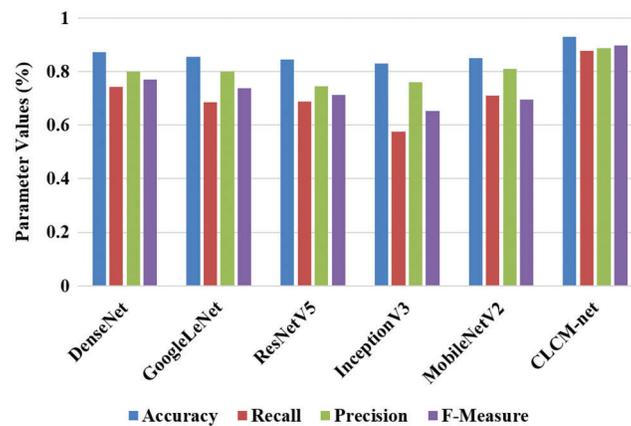


Figure 7: Graphical representation of all pre-trained models on combined dataset

The performance of CLCM-net achieved better performance than ISIC-2019 dataset, but it achieved high performance than ISIC-2018 dataset. For instance, CLCM-net achieved 93% of accuracy, 87.65% of recall, 88.66% of precision and 89.65% of F-measure, where the DenseNet achieved better performance than existing pre-trained models i.e., 87.18% of accuracy, 74% of recall, 79% of precision and 76.95% of F-measure. The other models achieved nearly 84% of accuracy, 69% of recall, 74% of precision and 71% of F-measure. In the proposed model, patch detection is carried out for improving the classification accuracy, where other models didn't considered such detection that leads poor performance. The accuracy for training, testing and validation for all pre-trained models on combined dataset is experimented and tabulated in Tab. 4.

While training the all the models with pre-processing and segmentation techniques on combined dataset, accuracy is calculated. For instance, ResNetV5 and InceptionV3 achieved 89% of training accuracy, MobileNet and GoogleNet achieved 86% of training accuracy and proposed model achieved 92.56% of training accuracy. While testing the combined dataset, the proposed model achieved only 89.31% of testing dataset, where the other pre-trained model achieved nearly 81% to 84% of testing accuracy. The reason is there is a differences between the kind of data that is used for training as well as testing dataset. Finally, the CLCM-net achieved 90.47% of validation accuracy, where DenseNet and GoogleNet

achieved nearly 81% of validation accuracy and other models achieved nearly 84% of validation accuracy. The Central processing unit (CPU) time for all models on combined dataset is experimented and tabulated in [Tab. 5](#).

Table 4: Accuracy results of the combined dataset with pre-process and segmentation classifier

Model	Accuracy		
	Training	Validation	Testing
DenseNet	0.8465	0.8093	0.8013
GoogleLeNet	0.8669	0.8347	0.8322
ResNetV5	0.8955	0.8487	0.8412
InceptionV3	0.8928	0.8457	0.8431
MobileNetV2	0.8725	0.8566	0.8439
CLCM-net	0.9256	0.9047	0.8931

Table 5: CPU times comparison for the combined dataset models

Model	Training (s)	Testing (s)
DenseNet	7.311 ± 0.076	0.025 ± 7.839 e-4
GoogleLeNet	3.223 ± 0.008	0.019 ± 9.284 e-4
ResNetV5	7.519 ± 0.092	0.019 ± 2.876 e-4
InceptionV3	4.733 ± 0.004	0.018 ± 6.642 e-4
MobileNetV2	3.513 ± 0.021	0.020 ± 3.435 e-4
CLCM-net	5.950 ± 0.082	0.017 ± 2.760e-4

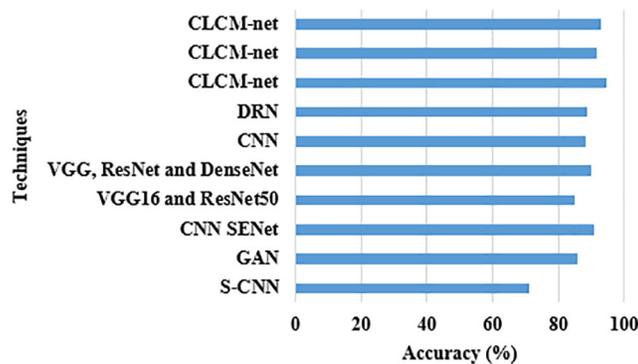
When training the combined dataset, the CLCM-net has slight high training time (i.e., 5.950 s) than GoogleNet (3.223 s), Inception (4.733 s) and MobileNet (3.513 s). This is because, the CLCM-net carried out five process on pre-processing for hair removal as well segmenting the affected lesions by five different process, while other pre-trained models directly classifies the skin cancer. However, while testing, the proposed model provides less testing time than other models i.e., 0.017 s, where all other models achieved 0.018 s to 0.025 s. This shows that the proposed model achieved better performance even on combined dataset.

4.4 Performance of Proposed Model with Existing Techniques

In this section, the comparison is considered with existing techniques for skin cancer classification on ISIC dataset that is designed by different authors are undertaken. [Tab. 6](#) and [Fig. 8](#) shows the comparative analysis, which provides the information about techniques, author names and dataset name.

Table 6: Comparative analysis of proposed model with existing techniques on ISIC dataset

Author	Technique	Dataset	Accuracy (%)
Sedigh et al. [19]	S-CNN	ISIC	71
Rashid et al. [20]	Generative Adversarial Network (GAN)	ISIC 2018	86
Pacheco et al. [21]	CNN for Segmentation (CNN-SENet)	ISIC 2019	91
Li et al. [22]	VGG16 and ResNet50	ISIC 2018	85
Rahi et al. [23]	VGG, ResNet and DenseNet	HAM10000 & ISIC	90
Daghrir et al. [24]	CNN	ISIC	88.4
Vinay et al. [25]	Deep Residual Network (DRN)	ISIC	88.7
Proposed	CLCM-net	ISIC 2018	94.42
Proposed	CLCM-net	ISIC 2019	91.73
Proposed	CLCM-net	ISIC 2018 & ISIC 2019	93.07

**Figure 8:** Graphical representation of comparative models on ISIC with combined dataset

The existing GAN, VGG16 and ResNet50 achieved nearly 86% of accuracy and proposed model achieved nearly 94.42% of accuracy on ISIC-2018 dataset, where CNN SENet achieved 91% of accuracy and the proposed model achieved 91.73% of accuracy on ISIC-2019 dataset. The other techniques such as DRN, S-CNN and CNN achieved 88% of accuracy on ISIC dataset, where proposed model achieved 93.07% of accuracy on combined dataset. The reason for better performance of CLCM-net is effectively removed the hair and segment the skin lesions from the input images by using DeTrop Noise Exclusion Technique and Yolov5 network with k-means clustering models. In recent years, Yolov3 model with CNN [26,27] is used to develop a real-time small object detection (RSOD) algorithm to improve the accuracy of object detection in Unmanned Aerial Vehicles (UAVs), but in our proposed, Yolov5 network is used for patch detection, which is very effective in skin cancer classification.

5 Conclusion

The proposed CLCM-net view emphasises the resilient performance of LayerWise weight constraints that are implemented consecutively to limit the weight restrictions. Consecutive layerwise maxnorm weight restrictions in the CLCM model regularisation solve the overfitting problem by keeping weights small in CNN, and model performance improves in the testing phase. Furthermore, several procedures are

carried out in order to effectively learn the features, such as denoising the hair in dermoscopic images, lesion detection, and patch uprooting of skin lesions. The CLCM-net model uses patch-uprooted images as input and learns the foreground artifacts in the skin lesion images. Skin lesion feature extraction and classification are achieved by adequately learning the features. The experiments are carried out on ISIC-2018 and 2019 challenging datasets in terms of various parameters for testing the performance of the proposed model. From the results, the CLCM-net achieved 94.42% accuracy and 92% of precision on ISIC-2018, 91.73% of accuracy and 85% of precision on ISIC-2019 and 93.07% of accuracy and 88% of precision on the combined dataset, where the existing pre-trained models achieved nearly 83% to 88% of accuracy and nearly 75% to 80% of precision on all three datasets. In addition, the proposed model achieved 90.47% of validation accuracy and all existing pre-trained models achieved nearly 82% to 85% of validation accuracy on the combined dataset. However, the results are not satisfactory on a combined dataset, which needs modification in deep learning techniques. Weights in deep learning, for instance, can be optimised in the future by employing a feature selection procedure, and weight constraints can be applied row-wise or column-wise vectors to learn the features exactly.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] A. Masood and A. Ali Al-Jumaily, "Computer aided diagnostic support system for skin cancer: A review of techniques and algorithms," *International Journal of Biomedical Imaging*, vol. 2013, pp. 1–22, 2013.
- [2] A. K. Verma, S. Pal and S. Kumar, "Comparison of skin disease prediction by feature selection using ensemble data mining techniques," *Informatics in Medicine Unlocked*, vol. 16, pp. 100202, 2019.
- [3] T. Akram, H. M. J. Lodhi, S. R. Naqvi, S. Naeem, M. Alhaisoni *et al.*, "A multilevel features selection framework for skin lesion classification." *Human-Centric Computing and Information Sciences*, vol. 10, no. 1, pp. 12, 2020.
- [4] B. Erkol, R. H. Moss, R. J. Stanley, W. V. Stoecker and E. Hvatum, "Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes," *Skin Research and Technology*, vol. 11, no. 1, pp. 17–26, 2005.
- [5] M. Kumar, M. Alshehri, R. AlGhamdi, P. Sharma and V. Deep, "A DE-ANN inspired skin cancer detection approach using fuzzy c-means clustering," *Mobile Networks and Applications*, vol. 25, no. 4, pp. 1319–1329, 2020.
- [6] M. A. Khan, T. Akram, Y. D. Zhang and M. Sharif, "Attributes based skin lesion detection and recognition: A mask RCNN and transfer learning-based deep learning framework," *Pattern Recognition Letters*, vol. 143, pp. 58–66, 2021.
- [7] U. Nazar, M. A. Khan, I. U. Lali, H. Lin, H. Ali *et al.*, "Review of automated computerized methods for brain tumor segmentation and classification," *Current Medical Imaging Formerly Current Medical Imaging Reviews*, vol. 16, no. 7, pp. 823–834, 2020.
- [8] S. Sharma, S. Sharma and A. Athaiya, "Activation functions in neural networks," *International Journal of Engineering Applied Sciences and Technology*, vol. 4, no. 12, pp. 310–316, 2020.
- [9] H. A. Haenssle, C. Fink, R. Schneiderbauer, F. Toberer, T. Buhl *et al.*, "Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists," *Annals of Oncology*, vol. 29, no. 8, pp. 1836–1842, 2018.
- [10] S. S. Chaturvedi, J. V. Tembhurne and T. Diwan, "A multi-class skin cancer classification using deep convolutional neural networks," *Multimedia Tools and Applications*, vol. 79, no. 39–40, pp. 28477–28498, 2020.
- [11] H. Huang, B. W. Hsu, C. Lee and V. S. Tseng, "Development of a light-weight deep learning model for cloud applications and remote diagnosis of skin cancers," *The Journal of Dermatology*, vol. 48, no. 3, pp. 310–316, 2021.

- [12] D. N. T. Le, H. X. Le, L. T. Ngo and H. T. Ngo *et al.*, “Transfer learning with class-weighted and focal loss function for automatic skin cancer classification.” ArXiv: 2009.05977 [Cs], 2020, <http://arxiv.org/abs/2009.05977>.
- [13] A. Hekler, J. N. Kather, E. K. Henning, J. S. Utikal, F. Meier *et al.*, “Effects of label noise on deep learning-based skin cancer classification,” *Frontiers in Medicine*, vol. 7, pp. 177, 2020.
- [14] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed *et al.*, “Going deeper with convolutions,” in *2015 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, pp. 1–9, 2015.
- [15] G. Huang, Z. Liu, L. V. D. Maaten and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, pp. 4700–4708, 2017.
- [16] C. Szegedy, S. Ioffe, V. Vanhoucke and A. A. Alemi, “Inception-v4, inception-resnet and the impact of residual connections on learning,” in *Thirty-First AAAI Conf. on Artificial Intelligence*, San Francisco, California USA, pp. 4278–4284, 2017.
- [17] N. Codella, V. Rotemberg, P. Tschandl, M. E. Celebi, S. Dusza *et al.*, “Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (ISIC),” ArXiv: 1902.03368 [Cs], 2019, <http://arxiv.org/abs/1902.03368>.
- [18] P. Tschandl, C. Rosendahl and Kittler, H., “The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions,” *Scientific data*, 5(1), pp. 1–9, 2018.
- [19] P. Sedigh, R. Sadeghian and M. T. Masouleh, “Generating synthetic medical images by using gan to improve cnn performance in skin cancer classification,” in *2019 7th Int. Conf. on Robotics and Mechatronics (ICRoM)*, Tehran, Iran, pp. 497–502, 2019.
- [20] H. Rashid, M. A. Tanveer and H. A. Khan, “Skin lesion classification using GAN based data augmentation,” in *2019 41st Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, the Berlin Institute of Health (BIH), pp. 916–919, 2019.
- [21] A. G. C. Pacheco, A. R. Ali and T. Trappenberg, “Skin cancer detection based on deep learning and entropy to detect outlier samples,” ArXiv: 1909.04525 [Cs, Stat], 2020, <http://arxiv.org/abs/1909.04525>.
- [22] X. Li, J. Wu, E. Z. Chen and H. Jiang, “From deep learning towards finding skin lesion biomarkers,” in *2019 41st Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, the Berlin Institute of Health (BIH), pp. 2797–2800, 2019.
- [23] M. M. I. Rahi, F. T. Khan, M. T. Mahtab, A. A. Ullah, M. G. R. Alam *et al.*, “Detection of skin cancer using deep neural networks,” in *2019 IEEE Asia-Pacific Conf. on Computer Science and Data Engineering (CSDE)*, Melbourne, VIC, Australia, pp. 1–7, 2019.
- [24] J. Daghri, L. Tlig, M. Bouchouicha and M. Sayadi, “Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach,” in *2020 5th Int. Conf. on Advanced Technologies for Signal and Image Processing (ATSIP)*, Sousse, Tunisia, pp. 1–5, 2020.
- [25] B. Vinay, P. J. Shah, V. Shekar and H. Vanamala, “Detection of melanoma using deep learning techniques,” in *2020 Int. Conf. on Computation, Automation and Knowledge Management (ICCAKM)*, Dubai, United Arab Emirates, pp. 391–394, 2020.
- [26] W. Sun, G. Z. Dai, X. R. Zhang, X. Z. He and X. Chen, “TBE-Net: A three-branch embedding network with part-aware ability and feature complementary learning for vehicle re-identification,” *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–13, 2021. <https://ieeexplore.ieee.org/abstract/document/9635721>.
- [27] W. Sun, L. Dai, X. R. Zhang, P. S. Chang and X. Z. He, “RSOD: Real-time small object detection algorithm in UAV-based traffic monitoring,” *Applied Intelligence*, pp. 1–16, 2021. <https://link.springer.com/article/10.1007/s10489-021-02893-3>.