Journal of Innovative Image Processing (ISSN: 2582-4252)
www.irojournals.com/iroiip/

Quadratic Luminance Vision Transformer Attention Network for Automated Mitosis Detection in Breast Histopathology Images

Kusuma Sri M.[1], Sathees Kumaran S.[2]

Department of Electronics and Communication Engineering, Anurag University, Hyderabad, India.

E-mail: [1]kusumasriece@anurag.edu.in, [2]satheeskumaranece@anurag.edu.in

Abstract

The primary limitation of CNN-based methods for mitosis detection in breast histopathological image classification is their inability to effectively extract features from potential regions of interest. This study presents a novel Quadratic Stain Luminance Normalized Vision Transformer Attention Network (QSLN-VTAN) designed to enhance feature extraction and classification, resulting in improved accuracy and precision. A feature aggregation sphere utilizing a quadratic discriminant classifier has been developed to integrate these features for the classification of images in mitosis detection. The stain normalization-based preprocessing not only enhances contrast but also preserves background luminance while ensuring robustness, significantly reducing data loss. The performance of QSLN-VTAN is assessed and compared using standard metrics, including precision, recall, accuracy, F1-score, and training time. The QSLN-VTAN demonstrated superior performance compared to other methods when evaluated on the ICPR 2012 dataset, achieving an overall mitosis detection rate of 96%, an F1 score of 93%, and a precision rate of 92%. Additionally, on the MITOSIS-ATYPIA-14 dataset, it exhibited a detection rate of 92%, an F1 score of 91%, and a precision of 96%.

Keywords: Histopathology Images, Deep Learning, Stain Normalization, Vision Transformer, Attention Network, Regions of Interest Segmentation, Feature Extraction and Classification.

DOI: https://doi.org/10.36548/jiip.2025.2.002
Received: 02.04.2025, received in revised form: 30.04.2025, accepted: 14.05.2025, published: 23.05.2025

1. Introduction

Breast tissue analysis in histopathology is a critical diagnostic tool for assessing the severity of breast cancer stages. The identification of mitotic cells demonstrates a strong correlation with cell proliferation, which is an essential metric in cell classification. However, the similarity between abnormal cells and the distinct morphology of mitotic figures can lead to significant inter-classification instability. Despite numerous deep learning techniques proposed for mitosis detection, these methods are often highly susceptible to domain shifts commonly encountered in histological images.

Efficient U-Net (EUNet)-based mitosis detection involves two key stages: candidate segmentation and candidate refinement. The EUNet method effectively segments candidates at lower resolutions, significantly accelerating the detection process of candidate cells. Subsequently, candidate cells are fine-tuned using a deeper classifier network to ensure optimal precision, recall, and F1-score while minimizing processing time.

The Small Mitotic Detector (SMDetector) was introduced in [2], employing deep learning techniques to distinguish between mitotic and non-mitotic regions. To prevent small objects from being overlooked in deeper layers, region proposal networks were utilized to accurately identify small objects through the use of dilated layers in the backbone.

A hybridized approach was proposed in [3], comprising two steps: mitosis detection and classification. Initially, deep learning techniques were employed for detection, followed by the application of fuzzy-based classifiers for precise classification. The mitotic activity index (MAI) is one of the most significant prognostic features considered in disease diagnosis. The advancement of digital pathology has greatly enhanced the efficiency of the diagnostic process. However, it requires a Whole Slide Imaging (WSI) scanner, which remains financially inaccessible for many hospitals due to its high installation costs.

Deep learning techniques were applied in [4] to address challenges faced by current mitosis detection pipelines, achieving higher convergence speed. Additionally, another low-cost histopathological mitosis detection method was presented in [5] with the aim of identifying cancerous mitotic cells. Nevertheless, the lack of publicly available ground-truth data poses a significant barrier to the implementation of contemporary advancements in deep learning. Convolutional neural networks have addressed these challenges, resulting in improved accuracy [6].

An efficient deep learning architecture designed to mitigate overfitting and enhance the understanding of evaluation and validation of breast cancer staging was presented in [7]. A comprehensive review of multi-center evaluations focusing on image quality analysis was conducted in [8]. However, certain factors in the staining reagent process often lead to instability in pathological images. Distillation learning was employed in [9] to facilitate color mapping between source and target images, thereby significantly reducing artifacts. An extensive evaluation of the effectiveness of automated grading techniques was undertaken in [10], aimed at enhancing performance. Nevertheless, the presence of weak labels presents challenges in the mitosis detection process. To address this issue, a CNN-based algorithm was proposed in [11], which segments pixels to generate bounding boxes, enabling accurate and precise mitosis detection.

The issue of data loss in mitosis detection has largely been overlooked in most studies, resulting in a compromise in overall precision and recall rates. To address these concerns, we propose an automated detection method called the Quadratic Stain Luminance Normalized Vision Transformer Attention Network (QSLN-VTAN) for breast cancer histopathology images. Through the implementation of a contrast- and luminance-enhanced stain normalization preprocessing technique, the developed classifier demonstrates superior performance compared to several existing approaches while minimizing data loss. This method employs a contrast- and luminosity-enhanced function that ensures uniform image balancing across multiple samples, effectively reducing the impact of stain-related variations. In this work, we introduce a Quadratic Vision Transformer Attention Network-based classifier for mitosis detection, which accurately and precisely extracts both local and global features.

The subsequent sections of the study are organized as follows. Section 2 outlines the methods relevant to mitosis detection, where machine learning (ML) and deep learning (DL) techniques have been extensively utilized. Section 3 discusses the proposed QSLN-VTAN-based classifier, complete with block diagrams and algorithms. Section 4 details the experimental setup, followed by a performance evaluation and analysis of results obtained using the ICPR 2012 dataset and the MITOSIS-ATYPIA-14 dataset. Finally, Section 5 concludes the paper.

2. Literature Review

A review of several ML and DL techniques employed in mitosis detection by exploring regions of interest and accordingly making overall classifications was presented in [12]. Yet another method focusing on generalization employing deep CNN-based multi-phase detection was proposed in [13]. Here, probing the most probable regions aided in ensuring generalization along with showing good discrimination ability. Nevertheless, the weak label involved in the analysis is not proficient in training mitosis segmentation. To deal with this issue, the single-pixel representation [14] was extended to concentric circles, with the inner circle representing the mitotic region for mitosis analysis. However, the authors did not focus on the computational cost associated with this approach. To address this issue, an Atrous fully convolution-based segmentation method was presented in [15], which, by acquiring mask and bounding box annotations, outperforms prior detection methods. As far as mitosis is concerned, both cell division and duplication are considered. Although the mitotic count plays a pivotal role in cancer diagnosis in clinical practice, it exhibits high inter-rater variability due to the difficulty in distinguishing mitotic from non-mitotic samples.

An ensemble of five CNNs was employed in [16] to improve overall performance, resulting in a higher F1-score and better generalization. In [17], pre-trained large-scale vision-language models were demonstrated by exploiting visual features. A two-stage learning approach was applied in [18] to enhance the detection accuracy in a significant manner. However, pathologists still need to spend significant time examining the frequency of mitotic cells in hotspot regions to determine their predominant occurrence. To address this problem, a two-stage cascaded network was introduced in [19] for efficient mitosis detection.

A novel partially supervised method based on dual parallel deep fully convolutional networks was proposed in [20] using centroid pixel labels. With this type of design, a high F1-score was achieved. An end-to-end multitask learning was applied in [21], by first detecting the mitosis reference region and then classifying the model, ensuring overall detection performance in an extensive manner. A review of deep learning methods was presented in [22] for detecting breast cancer in histopathology images. Yet another precision-recall analysis employing multi-phase deep CNN was presented in [23] for breast cancer histopathology images.

To address the aforementioned research gaps, in this work, a method called Quadratic Stain Luminance A normalized vision transformer attention network is proposed. The Quadratic Vision Transformer Attention Network-based classifier extracts essential features by segmenting both global and local features for accurate classification.

3. Proposed Method

Despite the employment of the Vahadane stain normalization method to preserve the structural properties of stained tissue samples in [1], the edges were not clearly visible, leading to data loss and reduced overall accuracy. A contrastand luminanceenhanced stain normalization preprocessing model is presented in this work. It uses the overall contrast difference between the source sample image and the reference image for analysis, preserving the background luminance of the source sample image and minimizing overall data loss. The speed and time were compromised since improved features were not retrieved, even though the semantic segmentation task was applied using the centroid pixels of the targeted objects rather than all of the nuclei's pixels [2].

To address this issue, a quadratic vision transformer attention network-based classifier is proposed, where both global and local features are extracted as essential for timely and speedy detection of mitosis. Thus, by combining the above two aspects, we propose a method called Quadratic Stain Luminance Normalized Vision Transformer Attention Network (QSLN-VTAN) for mitosis detection.

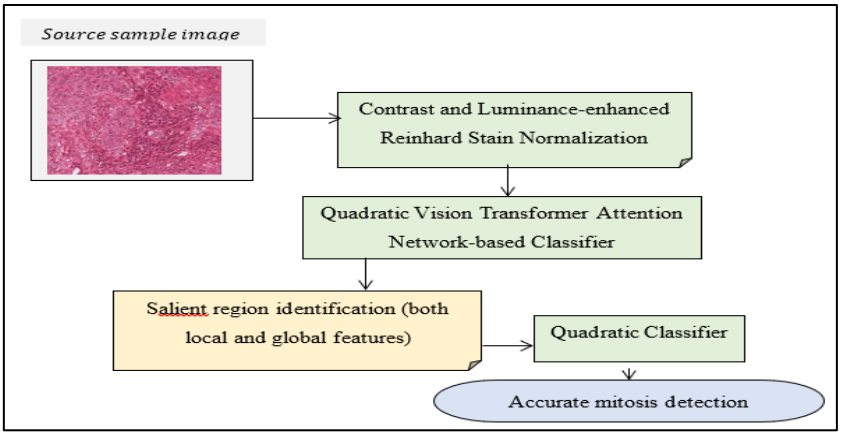

Figure 1. The Overall Architecture of the QSLN-VTAN Method

Figure 1 above illustrates the architecture of the QSLN-VTAN method, with the source sample image provided as input, The former is a noise removal model, and the latter is a

classification model. The noise removal model is applied in our work using contrast- and luminance-enhanced stain normalization-based preprocessing, which not only improves contrast, thereby reducing the computational complexity of the classifier as much as possible, but also preserves background luminance while maintaining robustness, thus extensively reducing data loss, The results of the preprocessed images are then sent to the Quadratic Vision Transformer Attention Network-based classifier to obtain very accurate classification results.

3.1 Contrast and Luminance-Enhanced Stain Normalization – Preprocessing

Stain normalization is a crucial step in tissue analysis, with H&E stains being the most commonly used. While these stains facilitate tissue identification, normalization is necessary to address challenges such as variations in lighting conditions and noise introduced during the staining process.

In this work, a preprocessing model employing contrast- and luminance-enhanced stain normalization is presented. The contrast- and luminance-enhanced stain normalization-based preprocessing model is employed to reduce the influence of differences occurring due to staining and to ensure equilibrium across numerous images. This stain normalization process in turn intends to normalize both color and intensity of staining, thereby constructing the images enable reliable analysis. The proposed model employed in our work initially preserves the background luminance of the given source sample image, minimizes data loss and performs contrast enhancement, thereby aiding in minimizing the computational complexity of the classifier. Figure 2 shows the structure of the contrast- and luminance-enhanced stain normalization-based preprocessing model.

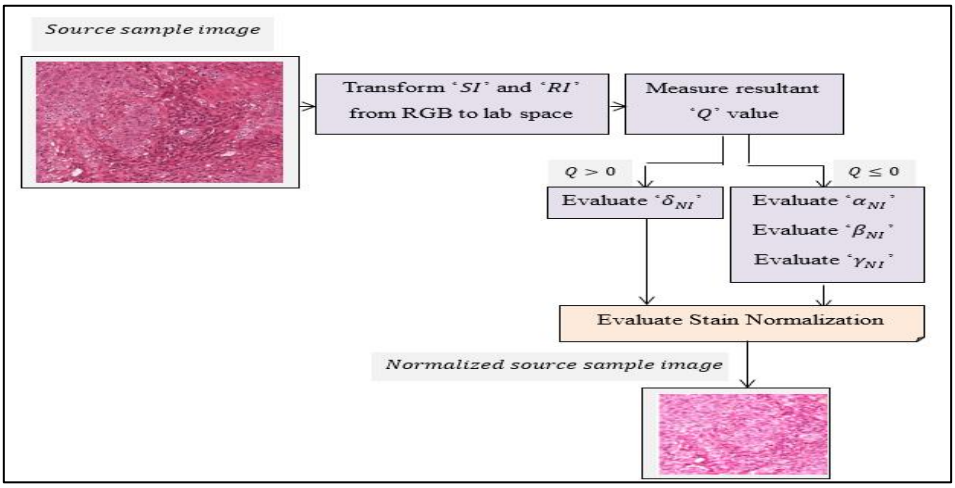

Figure 2. Contrast and Luminance-Enhanced Stain Normalization

As illustrated in the above figure, scanner A and scanner H images are initially read and then converted from RGB to Lab space. Then, it applies the Contrast and Luminance-enhanced function to the conventional Reinhard Stain Normalization to obtain normalized image for further processing. Initially, the source sample image 'SI' and the reference image 'RI' (i.e. the testing image) are transformed from RGB to lab space. This is followed by measuring 'Q' for each source sample image, as given below.

$$Q = \frac{\sigma_{Mean}(RI) - \sigma_{Mean}(SI)}{\sigma_{Mean}(RI)} \quad (1)$$

In the proposed method, 'Q' is a criterion based on the overall contrast difference between the source sample image and the reference (i.e. testing image). Upon identification of the value of 'Q' being negative, it is inferred that the testing image contrast is lesser than the source sample image contrast. This, in turn, ensures that the normalized image contrast greater than that of the source sample image. Therefore, the background luminance of the source sample image is said to be preserved, which aids in reducing data loss significantly. With the resultant value obtained, transformation is performed in lab space as given below (if '$Q > 0$').

$$\delta_{NI} = \mu_{Mean}(\alpha_{SI}) + [\alpha_{SI} - \mu_{Mean}(\alpha_{SI})] * (1 + Q) \quad (2)$$

In addition with the above resultant value obtained transformation is performed in lab space as given below (if '$Q \leq 0$').

$$\alpha_{NI} = \mu_{Mean}(\alpha_{RI}) + [\alpha_{SI} - \mu_{Mean}(\alpha_{SI})] \quad (3)$$

$$\beta_{NI} = \mu_{Mean}(\beta_{RI}) + [\beta_{SI} - \mu_{Mean}(\beta_{SI})] \quad (4)$$

$$\gamma_{NI} = \mu_{Mean}(\gamma_{RI}) + [\gamma_{SI} - \mu_{Mean}(\gamma_{SI})] \quad (5)$$

From the above equations (3), (4) and (5), '$\alpha_{NI}, \beta_{NI}\ \gamma_{NI}$' represents the magnitude of the normalized image in lab space, '$\alpha_{RI}, \beta_{RI}, \gamma_{RI}$' denotes the magnitude of reference image in lab space and '$\alpha_{SI}, \beta_{SI}, \gamma_{SI}$' represents the magnitude of the source sample image, while the overall source sample image mean is denoted by 'μ_{Mean}' . By taking the global variance in equation (2), and substituting the values for '$Q > 0$' and '$Q \leq 0$', the background luminance in the normalized image is preserved by our proposed method, thereby improvingthe overall

contrast in an efficient manner. We outline the algorithmic steps of the Contrast and Luminance-enhanced Stain Normalization in algorithm 1.

Algorithm 1: Contrast and Luminance-Enhanced Reinhard Stain Normalization

1: Initialize Dataset 'DS', Sample Image 'SI', '$N = 100$'

2: Begin

3: For all 'N' do

4: Measure $Q = \frac{\sigma_{Mean}(RI) - \sigma_{Mean}(SI)}{\sigma_{Mean}(RI)}$

5: If '$Q > 0$'

6: Then '$\delta_{NI} = \mu_{Mean}(\alpha_{SI}) + [\alpha_{SI} - \mu_{Mean}(\alpha_{SI})] * (1 + Q)$'

7: End if

8: If '$Q \leq 0$'

9: Evaluate $\alpha_{NI} = \mu_{Mean}(\alpha_{RI}) + [\alpha_{SI} - \mu_{Mean}(\alpha_{SI})]$

10: Evaluate $\beta_{NI} = \mu_{Mean}(\beta_{RI}) + [\beta_{SI} - \mu_{Mean}(\beta_{SI})]$

11: Evaluate $\gamma_{NI} = \mu_{Mean}(\gamma_{RI}) + [\gamma_{SI} - \mu_{Mean}(\gamma_{SI})]$

12: End if

13: Return contrast and luminance enhanced image

14: End for

15: End

3.2 Quadratic Vision Transformer Attention Network-based Classifier for Automated Mitosis Detection

Existing CNN-based methods segment global features from the entire processed image; however, they fail to identify features of potential Regions of Interest (RoI), which may lead to inaccurate mitosis detection. Additionally, a typical histology slide contains both important and irrelevant information, which increases processing time and computational complexity. Segmentation in our work is used for detecting and extracting areas that include more definitive information for processing. Critical shapes, like malignant cells or tumor borders or edges, must to identified accurately and precisely and secluded or isolated by means of segmentation, thereby ensuring more accurate analysis and feature extraction.

An accurate segmentation process aids pathologists in detecting minute differences or deviations, resulting in higher sensitivity for early mitosis detection. Moreover, by secluding or isolating crucial elements, it reduces false positives and provides pathologists with enhanced diagnostic support. This research gap has prompted us to propose a Quadratic Vision Transformer Attention Network-based Classifier model to classify breast histopathological images.

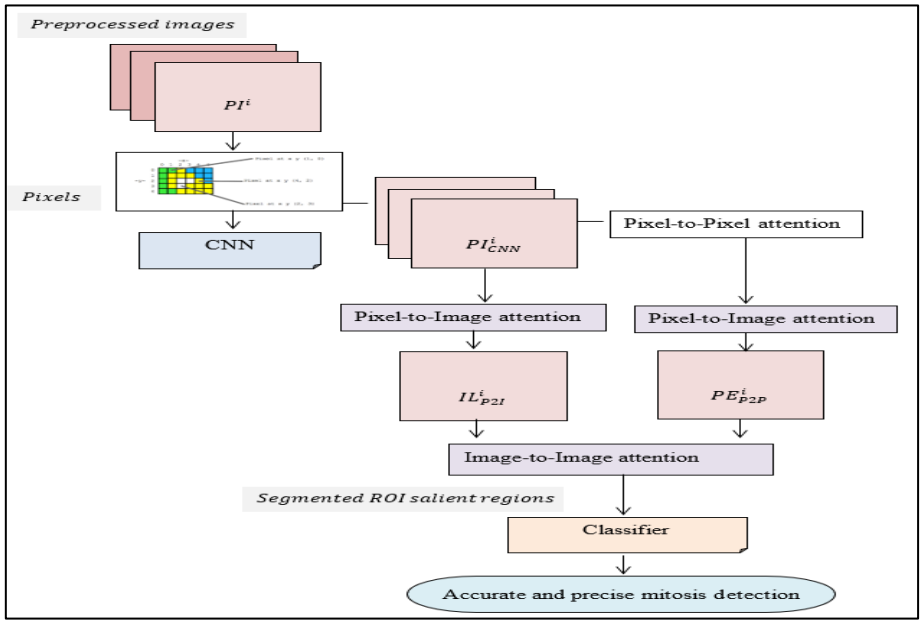

Figure 3. Flow Diagram of Quadratic Vision Transformer Attention Network-based Classifier for automated Mitosis Detection

In the proposed Quadratic Vision Transformer Attention Network-based Classifier model, global and local features are segmented, which are essential for accurate classification, as shown in Figure 3. The Vision Transformer Attention Network interacts with each other (i.e. within pixel limits as well), so that the model automatically identifies salient regions (i.e., detecting edges of segment we want to extract by focusing on both local and global features) on which to focus. The Vision Transformers model employed in our work consists of two sections: a multi-head attention (MHA) mechanism that designs associations between inputs (i.e. preprocessed images) and a feed-forward network (FFN) that learns extensive characterizations. For an input '$PI \in \mathbb{R}^{N*d}$' consisting of '$N\ d-$ dimensional' sample instances (preprocessed images and pixels in our case), the Vision Transformers learn the representations as given below.

$$\text{Res} = \text{Transformer (PI)} = \text{FFN}\left(\text{MHA}\left(\text{PI}_\text{Q} = \text{PI}, \text{PI}_\text{K} = \text{PI}, \text{PI}_\text{V} = \text{PI}\right)\right) \quad (6)$$

From the above equation (6), 'PI_Q', 'PI_K' and 'PI_V' form the input to the query, key, and value spheres in the multi-head attention. Owing to the inclusion of large spatial dimensions of histopathological images, visual representation learning using Vision Transformer is both a laborious and tedious task. Also, with the high computational cost of MHA, applying Vision Transformer to breast histopathology images is computationally intractable. This work extends the vision transformers using the Bag of Visual Words (BoVW) model for learning both local and global representations (i.e., identifying salient regions by detecting edges of the segment we want to extract by focusing on both local and global features) in an end-to-end fashion.

3.2.1 Pixel-to-Pixel Attention Model

The Pixel-to-Pixel attention model consists of a vision transformer unit with Multi-Head Attention (MHA) and Feed Forward Network (FFN), permitting us to design the interactions between pixels and identify important or salient pixels in the whole slide or preprocessing image. The preprocessed image '$\text{PI} \in \mathbb{R}^{\text{w}*\text{h}}$' with width 'w' and height 'h' is initially split into 'n' non-overlapping preprocessed images, '$\text{PI} = (\text{PI}^1, \text{PI}^2, \dots, \text{PI}^\text{n}) \in \mathbb{R}^{\frac{\text{w}}{\text{n}}*\frac{\text{h}}{\text{n}}}$', where '$\text{PI}^\text{i}$' denotes the '$\text{i}-\text{th}$' image. Each preprocessed image 'PI^i', is then split into 'm' non overlapping pixels '$\text{PI}^\text{i} = \left(\text{P}_1^\text{i}, \text{P}_2^\text{i}, \dots, \text{P}_\text{m}^\text{i}\right) \in \mathbb{R}^{\frac{\text{w}}{\text{mn}}*\frac{\text{h}}{\text{mn}}}$', with '$\text{P}_\text{j}^\text{i}$' denoting the '$\text{j}-\text{th}$'pixel in the '$\text{i}-\text{th}$' preprocessed image. Followed by which the pixels 'P_1^i' inside each preprocessed image 'PI^1' are fed to CNN with the intent of generating pixel level characterizations for each images as given below.

$$\text{PI}_\text{CNN}^\text{i} = \left(\text{P}_1^\text{i}, \text{P}_2^\text{i}, \dots \text{P}_\text{m}^\text{i}\right) \quad (7)$$

Inter-pixel closeness in each preprocessed image '$\text{PI}_\text{CNN}^\text{i}$' is encoded using the Vision Transformer to generate contextualized pixel embedding '$\text{PE}_\text{P2P}^\text{i}$' results as given below.

$$\text{PE}_\text{P2P}^\text{i} = \text{FFN}\left(\text{MHA}\left(\text{PI}_\text{Q} = \text{PI}_\text{CNN}^\text{i}, \text{PI}_\text{K} = \text{PI}_\text{CNN}^\text{i}, \text{PI}_\text{V} = \text{PI}_\text{CNN}^\text{i}\right)\right) \quad (8)$$

From the above equation, the MHA employed on one hand ensures encoding of inter-pixel closeness, and on the other hand, the FFN permits learning learn extensive image characterization.

3.2.2 Pixel-to-Image Attention Model

The Pixel-to-Pixel attention model generates contextualized pixel embedding results for each image. These pixel-level characterizations are combined to generate image-level characterizations by linearly aggregating the pixels inside each image, 'PI^{i}_{P2I}'. In other words, each pixel in 'PE^{i}_{P2P}' the image is mapped using a projection function 'ψ'. This is followed by a linear transformation 'LT_{P2I}' and Exponential Tilting functions, which are then applied to generate results as given below.

$$IL^{i}_{P2I} = ET\left(\psi\left(PE^{i}_{P2P}\right)LT_{P2I}\right)PE^{i}_{P2P} \tag{9}$$

From the above equation (10), by employing the linear transformation 'LT_{P2I}' the pixels embedded results 'PE^{i}_{P2P}' are linearly combined to generate image-level 'IL^{i}_{P2I}' characterization, therefore extracting global features. In a similar manner, the pixel-level characterization acquired from the CNN for each image is also aggregated using projection function 'ψ' as given below.

$$PL^{i}_{P2I} = ET\left(\psi\left(PE^{i}_{CNN}\right)LT_{P2I}\right)PE^{i}_{CNN} \tag{10}$$

From the above equation (), results pixel-level characterization results 'PL^{i}_{P2I}' are obtained by employing the projection function 'ψ', linear transformation, 'PE^{i}_{CNN}' respectively; therefore, local features are extracted.

3.2.3 Image-to-Image Attention Model

Finally, to encode inter-image closeness, the Image-to-Image attention model is employed in this section. By using model both local and global information are segmented efficiently. Initially, multi-head attention is applied to segment inter image characterization as given below.

$$IRes_{I2I} = MHA\left(PI_{Q} = PL^{i}_{P2I}, PI_{K} = PL^{i}_{P2I}, PI_{V} = PL^{i}_{P2I}\right) \tag{11}$$

Finally with the above multi head attention intermediate results '$IRes_{I2I}$' the image-to-image characterization is mathematically stated as given below to obtain overall segmented results, therefore segmenting both global and local features extracted essential for accurate classification.

$$\mathrm{Res_{I2I}} = \mathrm{FFN}\left(\mathrm{MHA}\left(\mathrm{PI_Q} = \mathrm{IRes_{I2I}}, \mathrm{PI_K} = \mathrm{IL^{i}_{P2I}}, \mathrm{PI_V} = \mathrm{IL^{i}_{P2I}}\right)\right) \quad (12)$$

3.2.4 Quadratic Discriminant Classifier

In the classification process, feature vectors extracted are employed as input to the classification algorithm performed via three phases: training, testing, and validation. The Quadratic Discriminant Classifier is employed in the proposed method to classify '$\mathrm{Res_{I2I}}$' into 'C' diagnosis classes as given below.

$$\mathrm{PO} = \mathrm{LR}(\mathrm{Res_{I2I}}, \delta_{\mathrm{C}}) \quad (13)$$

$$\mathrm{LR} = \frac{\sqrt{2\pi(\Sigma_1)}^{-1} \exp\left(-\frac{1}{2}(\mathrm{PO}-\mu_1)^{\mathrm{T}} \sum_1^{-1}(\mathrm{PO}-\mu_1)\right)}{\sqrt{2\pi(\Sigma_0)}^{-1} \exp\left(-\frac{1}{2}(\mathrm{PO}-\mu_0)^{\mathrm{T}} \sum_1^{-1}(\mathrm{PO}-\mu_0)\right)} < \mathrm{Th} \quad (14)$$

From the above equations (13) and (14) during evaluation, the pointer that has the highest confidence outcomes in 'PO' is selected as the predicted class label. We outline the algorithmic steps of the Quadratic Vision Transformer Attention Network-based Classifier in algorithm 2.

Algorithm 2: Quadratic Vision Transformer Attention Network-based Classifier

1: Initialize Dataset 'DS', Sample Image 'SI', '$\mathrm{N} = 100$', preprocessed image 'PI'

2: Begin

3: For all 'N' do

4: Measure global pixel information $\mathrm{IL^{i}_{P2I}} = \mathrm{ET}\left(\psi\left(\mathrm{PE^{i}_{P2P}}\right)\mathrm{LT_{P2I}}\right)\mathrm{PE^{i}_{P2P}}$

5: Measure local pixel information characterization$\mathrm{PL^{i}_{P2I}} = \mathrm{ET}\left(\psi\left(\mathrm{PE^{i}_{CNN}}\right)\mathrm{LT_{P2I}}\right)\mathrm{PE^{i}_{CNN}}$

6: Segment inter image characterization $\mathrm{IRes_{I2I}} = \mathrm{MHA}\left(\mathrm{PI_Q} = \mathrm{PL^{i}_{P2I}}, \mathrm{PI_K} = \mathrm{PL^{i}_{P2I}}, \mathrm{PI_V} = \mathrm{PL^{i}_{P2I}}\right)$

7: Measure overall segmented results $\mathrm{Res_{I2I}} = \mathrm{FFN}\left(\mathrm{MHA}\left(\mathrm{PI_Q} = \mathrm{IRes_{I2I}}, \mathrm{PI_K} = \mathrm{IL^{i}_{P2I}}, \mathrm{PI_V} = \mathrm{IL^{i}_{P2I}}\right)\right)$

8: Return segmented results $\mathrm{Res_{I2I}}$

9: Evaluate predicted output $\mathrm{PO} = \mathrm{LR}(\mathrm{Res_{I2I}}, \delta_{\mathrm{C}})$

10: Return predicted output results PO

11: End

4. Results And Discussion

In this section, we evaluate the performance of our Quadratic Stain Luminance Normalized Vision Transformer Attention Network (QSLN-VTAN) using two different datasets, namely, the ICPR 2012 dataset and the MITOSIS-ATYPIA-14 dataset.

4.1 Implementation

The whole QSLN-VTAN method is implemented using PyCharm (Python 3.9) on the MS Windows platform on a computer with an i5-2350 processor and 64 GB RAM. We performed contrastand luminance-enhanced Reinhard stain normalization-based preprocessing and the quadratic vision transformer attention network-based classifier using the Python high-level general-purpose programming language.

4.2 Description of the Datasets

The samples obtained from the ICPR 2012 [24] dataset consist of five breast cancer biopsy slides. These slides were scanned using two different devices: the Aperio XT scanner (Scanner A) and the Hamamatsu NanoZoomer scanner (Scanner H). Scanner A has a resolution of 0.2456 μm per pixel (2084 × 2084 pixels), while Scanner H has a resolution of 0.22753 μm per pixel (2252 × 2250 pixels). The dataset includes a total of 226 mitotic cell samples for training and 101 mitotic cell samples for testing.

The second MITOSIS-ATYPIA dataset employed in our work is found to be comparatively larger in size than the ICPR 2012 dataset. The samples acquired from the MITOSIS-ATYPIA dataset are scanned using both the A-type scanner and the H-type scanner. The individual pixel size is found to be ‘0.2456 μm’ for the A-type scanner, with the size of HPFs being ‘1539 × 1376 pixels’ and ‘0.2273 μm’ for the H-type scanners with size of HPFs being ‘1663 × 1485 pixels’ respectively. The MITOSIS-ATYPIA-14 dataset overall includes 1696 HPFs for a single type of scanner, with 1,200 having 749 samples labeled as mitotic cells and 496 samples being unlabeled, respectively.

4.3 Pre-processing Results

With the aid of Algorithm 1, enhanced images are obtained using two different datasets: the ICPR 2012 dataset and the MITOSIS-ATYPIA-14 dataset. The enhanced images from Scanner A are shown in images (a) and (b), while those from Scanner H are shown in images

(c) and (d), respectively. Their corresponding preprocessed results are shown in figure 4(e)–(h).

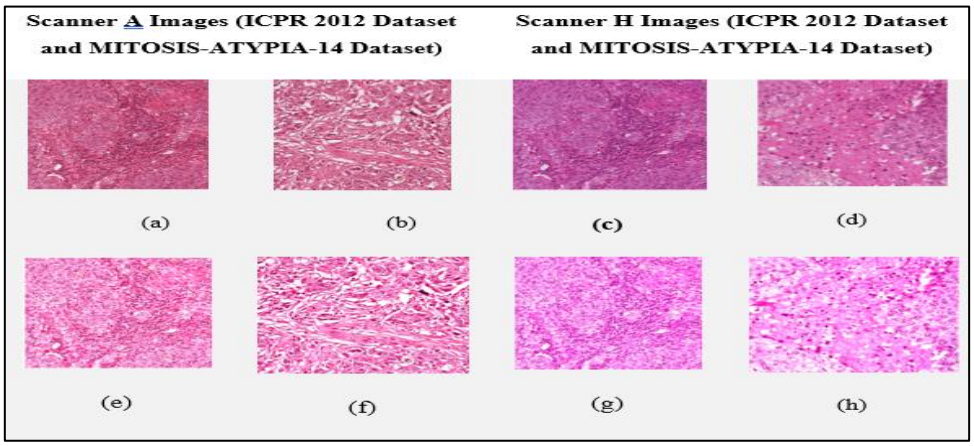

Figure 4. Source Sample Images (A–D) (from Dataset 1 and Dataset 2) and their Stained Normalized Results (E–H) (from Dataset 1 and Dataset 2) after Applying our Contrast and Luminance-Enhanced Reinhard Stain Normalization

As illustrated in the above figure, we show how our contrast and luminance-enhanced Reinhard stain normalization minimizes stain volatility across images from different datasets. As source sample images, we have two datasets of scanner A images and two datasets of scanner H images; similarly, the source sample images are used for the rest of the paper. To generate these stain-normalized images as output, we have used the contrast and luminance-enhanced Reinhard stain normalization algorithm.

4.4 Detection and Classification Results

With the aid of the algorithm 2, the segmented and classified results are obtained using the ICPR 2012 dataset and the MITOSIS-ATYPIA-14 dataset via scanner A. An image is illustrated in figure 5.

Figure 5 shows the segmentation and classification process employing two datasets, ICPR 2012 and MITOSIS-ATYPIA-14, respectively. In 5(a) and 5(b), the blue square represents the local portions, whereas the green square represents the global portion. Employing these two portions, the classification process is performed and shown as results in 5(b) and 5(d).

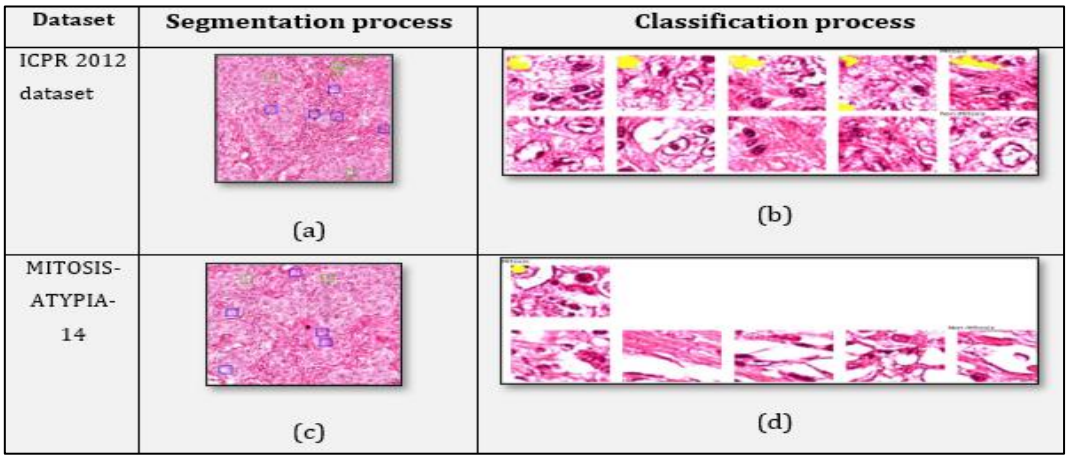

Figure 5. Segmentation and Classification Results(a) Segmentation Process using ICPR 2012 Dataset (b) Classified Output using ICPR 2012 Dataset (c) Segmentation Process using MITOSIS-ATYPIA-14 Dataset (d) Classified Output using MITOSIS-ATYPIA-14 Dataset

4.5 Performance Evaluation

The performance of the proposed QSLN-VTAN is assessed based on its ability to correctly identify mitotic cells. The training and validation sets are created by using the two datasets that we employed in our work. Two types of training datasets, the ICPR 2012 dataset and the MITOSIS-ATYPIA-14 dataset, comprising 100 samples and 304 samples, respectively, are employed. First, the samples in our work are randomly divided into 80% for training, 10% for validation, and 10% for testing. The pathologists might lose mitotic images if the precision is not high for the mitosis detection model. The Precision rate is mathematically represented as given below.

$$\text{Pre} = \frac{\text{TP}}{\text{TP+FP}} \tag{15}$$

From the above equation (15) the precision rate 'Pre' is analyzed using the true positive rate 'TP' and false positive rate 'FP' respectively. Recall on the other hand calculates how many of the actual positives our model obtained by labeling them as positive (i.e. true positive). For example, if a mitotic region is incorrectly predicted as non-mitotic, the results may lead to a wrong diagnosis.

$$\text{Rec} = \frac{\text{TP}}{\text{TP+FN}} \tag{16}$$

From the above equation (16) the recall rate ‘Rec’ is measured employing the true positive rate ‘TP’ and false negative rate ‘FN’. Next, F1-socre is employed with the intent of seeking a balance between precision and recall. This is represented as given below.

$$\text{F1 score} = 2 * \frac{\text{Pre} * \text{Rec}}{\text{Pre} + \text{Rec}} \tag{17}$$

4.6 Confusion Matrix Evaluation

The confusion matrix is considered a table that is frequently utilized in estimating the classification performance on a set of test data (i.e., two datasets) for which the true values are known [25]. The matrix portrays the number of sample instances generated by the method on the test data. First, a true positive is when the method correctly predicted a positive outcome and the actual outcome was positive. The second factor is the true negative, where the method correctly predicted a negative outcome and the actual outcome was negative. The third factor is the false positive, where the method incorrectly predicted a positive outcome and the actual outcome was negative, and the fourth factor is the false negative, where the method incorrectly predicted a negative outcome and the actual outcome was positive. Table 1 given below analyzes the confusion matrix for the MITOSIS-ATYPIA-14 dataset.

As illustrated in Figure 6, the confusion matrix is essential for analyzing and validating the performance of classification methods in mitosis detection, particularly in terms of accuracy, precision, recall, and F1 score.

Table 1. Confusion Matrix using QSLN-VTAN

	Predicted [MITOSIS-ATYPIA-14] using QSLN-VTAN			
		Mitosis	Non-mitosis	Total
Actual	Mitosis	$\text{TP} = 24$	$\text{FN} = 3$	27
	Non-mitosis	$\text{FP} = 4$	$\text{TN} = 9$	13
	Total	28	12	40

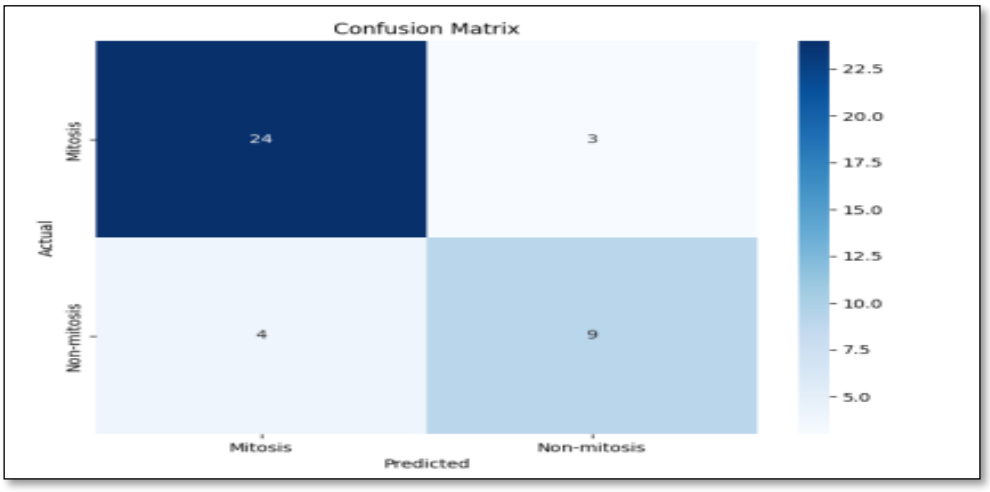

Figure 6. Confusion Matrix for QSLN-VTAN

4.6.1 Performance Measure of Precision, Recall and F1-Score on ICPR 2012 and MITOSIS-ATYPIA-14 Dataset

In this section, the performance metrics such as precision, recall, and F1 score are calculated and compared with existing approaches. Table 2 below presents a comparative analysis of the proposed QSLN VTAN method against two existing methods: EUNet [1] and SMDetector [2].

Table 2. Performance Comparison of Proposed QSLN-VTAN with EUNet and SMDetector

Methods	**ICPR 2012 Dataset**			**MITOSIS-ATYPIA-14 Dataset**		
	Precision	Recall	F1-score	Precision	Recall	F1-score
EUNet	0.75	0.83	0.78	0.85	0.89	0.86
SMDetector	0.60	0.75	0.66	0.78	0.83	0.80
QSLN-VTAN	0.92	0.96	0.93	0.96	0.92	0.91

Table 2 lists the calculated performance measures using the QSLN-VTAN method, along with two state-of-the-art methods: EUNet [1] and SMDetector [2]. With a single mitosis image obtained as input from the ICPR 2012 validation dataset and 30 regions identified as

mitosis, the true positive rate, true negative rate, false positive rate, and false negative rate using the QSLN-VTAN method were observed to be 17, 3, 8, and 2, respectively. By substituting these measured value into equations (15), (16), and (17), the precision rate was found to be 0.85, recall 0.89, and the F1-score 0.86. In the case of [1], the true positive rate, true negative rate, false positive rate, and false negative rate were observed to be 15, 5, 7, and 3, and the precision, recall, and F1-score were found to be 0.92, 0.96, and 0.93. Finally, by employing [2], with 12, 8, 6, and 4 being the true positive, true negative, false positive, and false negative values, the overall precision, recall, and F1-score were observed to be 0.60, 0.75, and 0.66, respectively. Thus, the three performance metrics when applied with the QSLN-VTAN method were found to be comparatively better than those of [1] and [2].

In a similar manner, with a single mitosis image obtained as input from the MITOSIS-ATYPIA-14 validation dataset and 50 regions identified as mitosis, the true positive, false positive, true negative, and false negative of the QSLN-VTAN method were 24, 4, 9, and 3; the true positive, false positive, true negative, and false negative for [1] were 22, 6, 8, and 4; and finally, the true positive, false positive, true negative, and false negative of [2] were 20, 8, 6, and 6. The overall precision, recall, and F1-score using the QSLN-VTAN method were found to be 0.91; using [1], they were found to be 0.85, 0.89, and 0.86; and applying [2], they were identified to be 0.78, 0.83, and 0.80, respectively. With these simulation results, the precision, recall, and F1-score of the QSLN-VTAN method were found to be comparatively better than those of [1] and [2].

The reason was that by applying the contrast- and luminance-enhanced stain normalization-based preprocessing algorithm, the background luminance of a given source sample image was preserved, which in turn aided in reducing the false positives and false negatives significantly. Moreover, by applying the contrast- and luminance-enhanced function to the conventional Reinhard Stain Normalization, overall contrast differences between the source sample image and the reference image were initially obtained. Following this, by taking global variance to obtain transformation in lab space, the true negatives and true positives were improved considerably. This, in turn, improved the precision, recall, and F1-score of the QSLN-VTAN method by 10% and 25% compared to [1] and [2], 6% and 14% compared to [1] and [2], and 8% and 20% compared to [1] and [2], respectively, for the ICPR 2012 dataset, and 7% and 14% compared to [1] and [2], 8% and 12% compared to [1] and [2], and 10% and 11% compared to [1] and [2], respectively, for the MITOSIS-ATYPIA 2014 dataset. The figure

below shows the graphical representation of the F1 score of three different methods using two datasets.

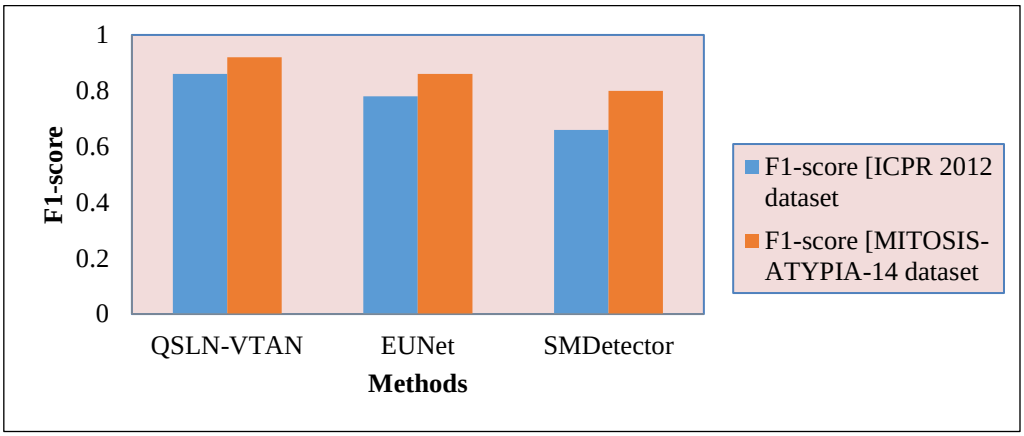

Figure 7. Graphical Representations of F1-Score Values

From Figure 7, we have also analyzed the influence of fusing the probability scores of the local regions and global regions during segmentation tasks. The figure shows the change in the performance of the QSLN-VTAN method on the validation set. We can see that the QSLN-VTAN method achieves excellent performance compared to [1] and [2].

4.6.2 Performance Comparison of Different Mitosis Detection Methods in Preprocessing and Classification Stages

In this section, we explore the influence of different methods on mitosis detection with respect to preprocessing and classification on two different datasets, ICPR 2012 and MITOSIS-ATYPIA-14. Because the backbone mitosis detection network also contributes to the detection performance, several experiments were conducted using different backbone network models, QSLN-VTAN, EUNet [1], and SMDetector [2] in preprocessing and classification stages. Table 3 shows the effect of the performance of different backbone mitosis detection methods on accuracy.

Table 3. Performance Comparison of Different Backbone Mitosis Detection Methods (Preprocessing and Classification Stage)

Methods	**Preprocessing Stage**						**Classification Stage**					
	ICPR 2012 Dataset			MITOSIS-ATYPIA-14 Dataset			ICPR 2012 Dataset			MITOSIS-ATYPIA-14 Dataset		
	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score
EUNet	0.79	0.87	0.82	0.88	0.92	0.89	0.82	0.9	0.85	0.89	0.84	0.81
SMDetector	0.64	0.79	0.7	0.81	0.86	0.83	0.67	0.82	0.73	0.82	0.8	0.8
QSLN-TAN	0.89	0.93	0.9	0.95	0.95	0.95	0.92	0.96	0.93	0.96	0.92	0.91

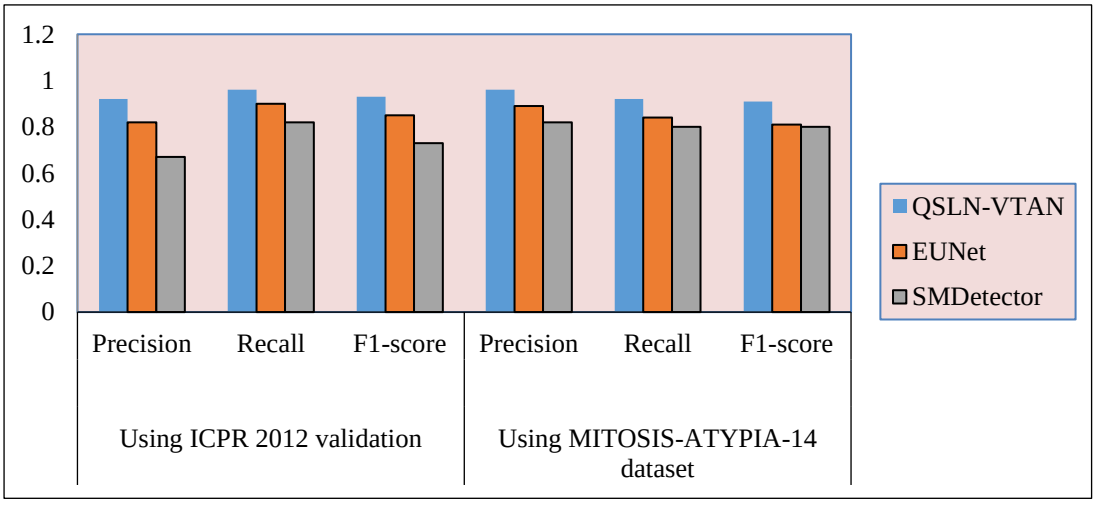

Figure 8. Performance Comparisons of Different Backbone Mitosis Detection Methods (Classifier Stage)

Figure 8 above shows that our proposed method detected the largest number of true positives compared with other methods and, hence, reached the highest recall value. Compared with some state-of-the-art mitosis detection methods that are extensively utilized, our improved proposed QSLN-VTAN method can identify the most positive samples, or, to be more specific, the true positive value.

5. Conclusion

A novel Quadratic Vision Transformer Attention Network (QSLN-VTAN) is proposed to enhance the accuracy of mitosis detection and improve the assessment of breast cancer severity. Initially, raw sample source images obtained from the ICPR 2012 dataset and the MITOSIS-ATYPIA 14 dataset underwent preprocessing through contrast- and luminance-enhanced stain normalization. This process ensured that the luminance contrast was improved in the normalized images for effective classification. Subsequently, the preprocessed images were analyzed using a Quadratic Vision Transformer Attention Network-based classifier. By segmenting both global and local features, which are critical for accuracy, we established a foundation for minimizing false positive and false negative rates. Additionally, comparative experiments conducted on the ICPR 2012 and MITOSIS-ATYPIA-14 datasets demonstrated a significant enhancement in mitosis detection performance, as indicated by improved precision and recall metrics. The results underscore the potential of the QSLN-VTAN method to substantially increase the detection accuracy of mitotic sample images, achieving 92% precision, 96% recall, and a 93% F1 score. Compared to the EUNet model, the QSLN-VTAN method improved precision, recall, and F1 score by 10%, 6%, and 8%, respectively.

References

[1] Jahanifar, Mostafa, Adam Shephard, Neda Zamanitajeddin, Simon Graham, Shan E. Ahmed Raza, Fayyaz Minhas, and Nasir Rajpoot. "Mitosis detection, fast and slow: robust and efficient detection of mitotic figures." Medical Image Analysis 94 (2024): 103132.

[2] Khan, Hameed Ullah, Basit Raza, Munawar Hussain Shah, Syed Muhammad Usama, Prayag Tiwari, and Shahab S. Band. "SMDetector: Small mitotic detector in histopathology images using faster R-CNN with dilated convolutions in backbone model." Biomedical Signal Processing and Control 81 (2023): 104414.

[3] Nemati, Nooshin, Refik Samet, Emrah Hancer, Zeynep Yildirim, and Eyyup Ensar Akkas. "A hybridized Deep learning methodology for mitosis detection and classification from histopathology images." Journal of Machine Intelligence and Data Science (JMIDS) 4, no. 1 (2023): 35-43.

[4] Farooq, Hasan, Saira Saleem, Iffat Aleem, Ayesha Iftikhar, Umer Nisar Sheikh, and Hammad Naveed. "Toward interpretable and generalized mitosis detection in digital pathology using deep learning." Digital Health 10 (2024): 20552076241255471.

[5] Shabbir, Bilal, Saira Saleem, Iffat Aleem, Nida Babar, Hammad Farooq, Asif Loya, and Hammad Naveed. "Low-Cost Histopathological Mitosis Detection for Microscope-acquired Images." AMIA Summits on Translational Science Proceedings 2024 (2024): 409.

[6] Albarqouni, Shadi, Christoph Baur, Felix Achilles, Vasileios Belagiannis, Stefanie Demirci, and Nassir Navab. "Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images." IEEE transactions on medical imaging 35, no. 5 (2016): 1313-1321.

[7] Saha, Monjoy, Chandan Chakraborty, and Daniel Racoceanu. "Efficient deep learning model for mitosis detection using breast histopathology images." Computerized Medical Imaging and Graphics 64 (2018): 29-40.

[8] Michielli, Nicola, Alessandro Caputo, Manuela Scotto, Alessandro Mogetta, Orazio Antonino Maria Pennisi, Filippo Molinari, Davide Balmativola et al. "Stain normalization in digital pathology: Clinical multi-center evaluation of image quality." Journal of pathology informatics 13 (2022): 100145.

[9] Kang, Hongtao, Die Luo, Weihua Feng, Shaoqun Zeng, Tingwei Quan, Junbo Hu, and Xiuli Liu. "Stainnet: a fast and robust stain normalization network." Frontiers in Medicine 8 (2021): 746307.

[10] Voon, Wingates, Yan Chai Hum, Yee Kai Tee, Wun-She Yap, Humaira Nisar, Hamam Mokayed, Neha Gupta, and Khin Wee Lai. "Evaluating the effectiveness of stain normalization techniques in automated grading of invasive ductal carcinoma histopathological images." Scientific Reports 13, no. 1 (2023): 20518.

[11] Lu, Wenjing. "A Two-Phase Mitosis Detection Approach Based on U-Shaped Network." BioMed Research International 2021, no. 1 (2021): 1722652.

[12] Rashmi, R., Keerthana Prasad, and Chethana Babu K. Udupa. "Breast histopathological image analysis using image processing techniques for diagnostic purposes: A methodological review." Journal of Medical Systems 46, no. 1 (2022): 7.

[13] Sohail, Anabia, Asifullah Khan, Noorul Wahab, Aneela Zameer, and Saranjam Khan. "A multi-phase deep CNN based mitosis detection framework for breast cancer histopathological images." Scientific Reports 11, no. 1 (2021): 6215.

[14] Li, Chao, Xinggang Wang, Wenyu Liu, Longin Jan Latecki, Bo Wang, and Junzhou Huang. "Weakly supervised mitosis detection in breast histopathology images using concentric loss." Medical image analysis 53 (2019): 165-178.

[15] Kausar, Tasleem, Mingjiang Wang, M. Adnan Ashraf, and Adeeba Kausar. "SmallMitosis: Small size mitotic cells detection in breast histopathology images." IEEE Access 9 (2020): 905-922.

[16] Sohail, Anabia, Asifullah Khan, Humaira Nisar, Sobia Tabassum, and Aneela Zameer. "Mitotic nuclei analysis in breast cancer histopathology images using deep ensemble classifier." Medical image analysis 72 (2021): 102121.

[17] Ding, Ruiwen, James Hall, Neil Tenenholtz, and Kristen Severson. "Improving mitosis detection on histopathology images using large vision-language models." In 2024 IEEE International Symposium on Biomedical Imaging (ISBI), IEEE, 2024, 1-5.

[18] Çayır, Sercan, Gizem Solmaz, Huseyin Kusetogullari, Fatma Tokat, Engin Bozaba, Sencer Karakaya, Leonardo Obinna Iheme et al. "MITNET: a novel dataset and a two-stage deep learning approach for mitosis recognition in whole slide images of breast cancer tissue." Neural Computing and Applications 34, no. 20 (2022): 17837-17851.

[19] Wang, Huadeng, Hao Xu, Bingbing Li, Xipeng Pan, Lingqi Zeng, Rushi Lan, and Xiaonan Luo. "A novel dataset and a two-stage mitosis nuclei detection method based on hybrid anchor branch." Biomedical Signal Processing and Control 87 (2024): 105374.

[20] Sebai, Meriem, Tianjiang Wang, and Saad Ali Al-Fadhli. "PartMitosis: a partially supervised deep learning framework for mitosis detection in breast cancer histopathology images." IEEE Access 8 (2020): 45133-45147.

[21] Alom, Md Zahangir, Theus Aspiras, Tarek M. Taha, T. J. Bowen, and Vijayan K. Asari. "MitosisNet: end-to-end mitotic cell detection by multi-task learning." IEEE Access 8 (2020): 68695-68710.

[22] Priya CV, Lakshmi, Biju VG, Vinod BR, and Sivakumar Ramachandran. "Deep learning approaches for breast cancer detection in histopathology images: A review." Cancer Biomarkers 40, no. 1 (2024): 1-25.

[23] Sohail, Anabia, Asifullah Khan, Noorul Wahab, Aneela Zameer, and Saranjam Khan. "A multi-phase deep CNN based mitosis detection framework for breast cancer histopathological images." Scientific Reports 11, no. 1 (2021): 6215.

[24] Roux, Ludovic, Daniel Racoceanu, Nicolas Loménie, Maria Kulikova, Humayun Irshad, Jacques Klossa, Frédérique Capron, Catherine Genestie, Gilles Le Naour, and Metin Gurcan. "Mitosis detection in breast cancer histological images." ICPR International Contest (2014).

[25] Li, Chao, Xinggang Wang, Wenyu Liu, Longin Jan Latecki, Bo Wang, and Junzhou Huang. "Weakly supervised mitosis detection in breast histopathology images using concentric loss." Medical image analysis 53 (2019): 165-178.

[26] Vivekanandan, S., S. Mounika, P. Monisha, and M. Balaganesh. "Robust Breast Cancer Prognosis Prediction: Adaptive Outlier Removal using SVM and K-Means Clustering." Journal of Soft Computing Paradigm 6, no. 1 (2024): 85-99.