Machine Learning Techniques for Automated Nuclear Atypia Detection in Histopathology Images: A Review

Jithy Varghese and J. S. Saleema

Abstract Nuclear atypia identification is an important stage in pathology procedures for breast cancer diagnosis and prognosis. The introduction of image processing techniques to automate nuclear atypia identification has made the very tedious, error-prone, and time-consuming procedure of manually observing stained histopathological slides much easier. In the last decade, several solutions for resolving this problem have emerged in the literature, and they have shown positive incremental advancements in this field of study. The nuclear atypia count is an important measure to consider when assessing breast cancer. This work provides a comprehensive review of automated nuclear atypia scoring process which includes the current advancements and future prospects for this critical undertaking, which will aid humanity in the fight against cancer. In this study, we examine the various techniques applied in detecting nuclear atypia in breast cancer as well as the major hurdles that must be overcome and the use of benchmark datasets in this domain. This work provides a comprehensive review of automated nuclear atypia scoring process which includes the current advancements and prospects for this critical undertaking, which will aid humanity in the fight against cancer.

Keywords Breast cancer · Machine learning · Nuclear pleomorphism · Histopathological image · Nuclear atypia detection · Deep learning

1 Introduction

Cancer is a phrase that refers to diseases that cause abnormal cell proliferation in the body, resulting in tumours [1]. It can spread to other organs of the body from the initial affected site. Every year, cancer claims the lives of huge number of people

J. Varghese (✉) · J. S. Saleema
Department of Computer Science, Christ Deemed to Be University, Bengaluru, Karnataka 560029, India
e-mail: jithylijo@caias.in

J. Varghese
Christ Academy Institute for Advanced Studies, Bengaluru, Karnataka 560083, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

S. Smys et al. (eds.), *Inventive Computation and Information Technologies*, Lecture Notes in Networks and Systems 563, https://doi.org/10.1007/978-981-19-7402-1_51

throughout the world. According to the statistics report of World Health Organization (WHO), the second major reason for mortality in the world is cancer [2].

According to the GLOBOCAN 2018 [3] report, 24.2% of cancer diagnosed amongst women all over the world is breast cancer, and 15% of all cancer-related death is due to breast cancer. Cancer patients' chances of survival can be greatly improved by early detection and treatment [2]. For better early detection of cancer, intensive research investigations are being done worldwide [4]. Many medical multi-imaging modalities are utilized screening and classification of breast cancer, including histopathological images taken through biopsy, digital mammography, sono-mammography, magnetic resonance imaging, and computerized thermography [5].

As biopsy is a clinical procedure of tissue analysis for the screening of cancer [5]. Tissues from vulnerable areas are carefully removed and mounted on microscope slides. These images are called histopathological images. These images can be safely preserved in digital format and transmitted over the Internet for future study [6]. Digital pathology has helped researchers to apply computer-assisted technology in detection of biomarkers in cancer research [2]. The tedious and time-consuming tasks of pathologists can be delegated to computers, and it can be accomplished with high amount of accuracy [7–9]. Whole slide imaging (WSI) technology was introduced to read and store the whole slide at very high magnification [2, 10]. It also aided in the creation of many region of interest (ROI) images that were used to diagnose breast cancer as malignant or non-cancerous [11]. A sample of a high-power field (HPF) image of a breast cancer biopsy is shown in Fig. 1.

The cancer detection method in histopathological images usually includes classifying the image biopsy as malignant or non-cancerous. Pathologists classify the images based on criteria such as form, colour, cytoplasm proportion, and size of the nucleus of the cell. The contrast between normal and malignant cells is depicted in Fig. 2 [12]. Score of nuclear atypia provides an estimate of the prognosis of disease

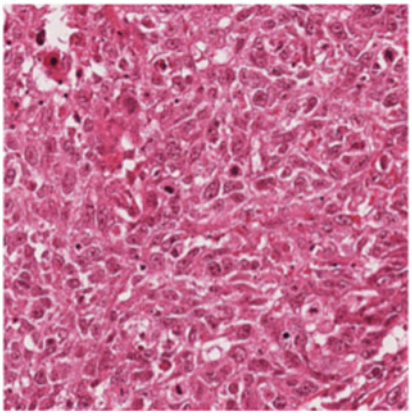

Fig. 1 HPF image of breast cancer histopathological image from MITOSIS-ATYPIA dataset

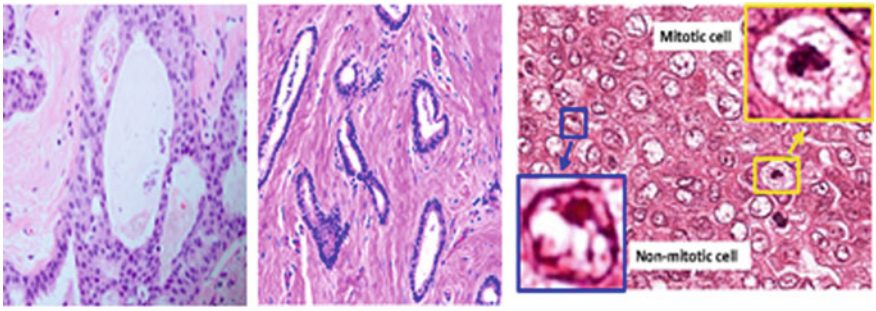

Fig. 2 **a** Nuclear atypia, **b** tubule formation, **c** mitotic cell [13]

in the patient and aids in the development of specific treatment strategies for the patient. According to the number of mitotic cells, tube formation, and nuclear pleomorphism, the cancer is classified as first, second, and third grade breast cancer. Computerized nuclei delineation and classification, which are usually required for cancer diagnosis and grading, are a time-consuming operation made more complex in abnormal images by the majority of nuclei's complex and irregular shape and size.

.

1.1 Grading and Staging

Grading and staging are the two major steps in the cancer treatment plan [14]. The degree to which malignant tumours resemble the surrounding tissue is graded. The main purpose of cancer grading is to identify the disease's aggressiveness. The malignancies which can be differentiated well have a better prognosis than poorly differentiated once because they are less aggressive [2]. The tumour's aggressiveness is a measure of its rate of growth and extension to other organs in the body. Staging in cancer determines up to what extend the growth has advanced from the initial affected organ to other organs of the body. Oncologists use grading and staging to finalize the treatment options and anticipate disease progression [2].

The stage of the breast cancer can be understood using tumour, node, metastasis (TNM) staging [2]. The stage of a tumour is determined by its size, spread, and location. Stage T describes the size and spread of the tumour in the primary site and adjacent organs. Stage N indicates the existence of cancer cells in the lymph node [15]. This is an indication that the cancer is spreading. The lymphatic system is the primary carrier which moves cancer cells to other organs of the body. Stage M specifies that the cancer has extended to an organ other than the one where it was diagnosed first [2].

According to Nottingham grading system (NGS) [2, 16], mitotic count, nuclear pleomorphism (atypia), and tubule development are the three major characteristics utilized for grading breast cancer. Nuclear atypia count is subjective in nature where

Table 1 Breast cancer grading characteristics by NGS [2]

Parameter	Score	Criteria
Mitotic count	1	< 10 mitotic cells
	2	10–19 mitotic cells
	3	≥ 20 mitotic cells
Nuclear pleomorphism (nuclear atypia)	1	Small and similar type nuclei
	2	Moderately variable nuclei shape and size
	3	Many nuclei with clear variation
Tubule formation	1	Tubule forms (> 75%)
	2	Tubule forms (10–75%)
	3	Tubule forms (< 10%)

the grading depends on the proficiency of the pathologist. Table 1 shows breast cancer grading characteristics by NGS [2, 16].

1.2 Need for Automated Nuclear Atypia Detection

A pathologist analyzes stained histology slide under a microscope and manually assigns a score to every criteria in NGS. A human grading approach is difficult, tedious, and erroneous due to the inequality in the appearance of the cells and numerous cells per high magnification field (HPF) [2]. These factors contribute to a high amount of diversity in pathological results between observers [17]. Due to lack of qualified pathologists in many nations, early cancer detection is challenging. As a result, cancer death rates in developing countries are high [3, 18]. In this case, automated diagnostic methods can help with a quicker and more effective diagnosis. In this case, automated detection is the best option. In this case, automated diagnostic technologies can help with rapid diagnosis and accurate grading in order to define the optimal treatment strategy, lowering the cancer-related death rate significantly [19].

1.3 Challenges in Nuclear Atypia Detection

In histopathology image analysis, detecting nuclear atypia automatically is a difficult task. Dissimilarity in texture shape and size of the nuclear atypia makes the detection of nuclear atypia detection tough [2]. There are several cellular formations that resemble nuclear atypia, but they are not atypia. Such resemblance increases the false positive rate and thereby decrease the accuracy of classification [2, 20]. In the manual staining of histology slides, there may be variation in the staining due

to the human involvement. Apart from this, the colour intensity of the images varies depending on the scanner used. Aperio and Hamamatsu are the most common type of scanners used in histological images, and they have a significant amount of variation in colour intensity. So if we use colour component in feature extraction, then it might affect the overall accuracy of classification process. To avoid this, various colour normalization techniques are used in nuclear atypia detection approaches [2].

1.4 Dataset in the Public Domain

The researchers started working more in the field of nuclear atypia detection due to the public dataset availability through open challenges. The most common histopathological datasets available in the public domain for nuclear atypia scoring as per the literature are given in the Table 2 [2].

Despite the fact that several datasets for breast cancer histopathological image analysis such as TUPAC 2016, MITOSIS2012, MITOSIS ATYPIA14, AMIDA, BRAEKHIS DATASET10, and BreCaHAD 2019 but none of them contain labelled images for nuclear atypia detection. The MITOS-ATYPIA14 challenge dataset is the best suited dataset for our topic of interest in this regard. As a result, in our comparative study, we employed the MITOSiS-ATYPIA14 dataset.

In this review paper, the major emphasis is given to nuclear atypia characteristics which is used to grade breast cancer. The paper focusses on the various techniques and procedures for grading breast cancer, assesses the challenges that these techniques encounter, and explores the strategy adopted by image analysis methods to overcome these challenges. These tools assist in determining topics that have not been thoroughly investigated in the field of nuclear atypia detection, which is where future study will be focussed. The paper also discusses the various evaluation metrics

Table 2 Public domain datasets for nuclear atypia scoring [2, 21]

Public dataset	Year	Explanation	Image size
MITOS-2012 [22]	2012	50 HPF images at 40 × magnification	A:2084 × 2084 H:2250 × 2252 M:1360 × 1360
AMIDA [23]	2013	311 HPF of 23 subjects at 40× magnification	2000 × 2000
MITOS- ATYPIA14 [24]	2014	1136 HPF images at 40× magnification	A:1539 × 1376 H:1663 × 1485
TUPAC 16 [25]	2016	73 breast cancer of different patients at 40× magnification	2000 × 2000
BREAKHIS DATASET10[26]	2018	9109 WSI of 82 patients	700 × 460
BreCaHAD 2019[27]	2019	162 WSI of 82 patients	1360 × 1024

used in analyzing the detection of nuclear atypia. The following sections make up the document. The first section contains an introduction, breast cancer grading, and the need for nuclear atypia detection, the second section contains a literature review in the field of nuclear atypia detection, the third section summarizes the different evaluation performance measures used for the process of nuclear atypia scoring, the fourth section does an analysis of existing techniques used in the field, and the fifth section gives future prospects in the field of grading of breast cancer [2, 27].

2 Literature Review

For the detection of nuclear atypia, there are two types of histopathological image analysis techniques. They are hand crafted feature-based algorithms that have been constructed and feature-based algorithms. In handmade feature-based approaches, image features must be retrieved precisely. Pre-processing of image, nucleus segmentation, feature extraction, and classification operations is all part of the handcrafted feature-based method.

Pre-processing techniques such as noise smoothing, intensity normalization, colour separation, thresholding, stain normalization, reduction, and augmentation are used to improve the characteristics of histopathological images. After pre-processing, nucleus segmentation is done [21]. For nuclei segmentation, many techniques, such as active contour model, mean shift, and Gaussian model watershed, [28–31] are frequently utilized. One of the most essential requirements for grading breast cancer in histopathology images is extraction of features of biological structures such as lymphocytes and cancer nuclei. The morphological characteristics and shape and size of these structures are frequently utilized as indicators of disease severity. The collected features are given into the classification step, which analyzes them statistically and typically uses machine learning techniques to categorize them into multiple classes. K-means clustering, SVMs, Bayesian classifiers, and artificial neural networks (ANNs) are some of the most often used classifiers. The steps of handcrafted feature-based nuclear atypia detection are shown in Fig. 3.

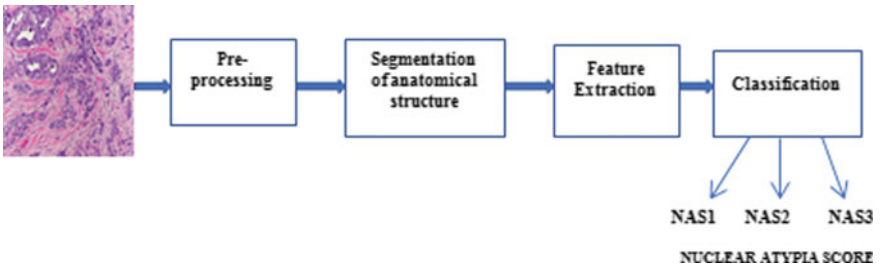

Fig. 3 Steps of handcrafted feature-based nuclear atypia detection

The second set of histopathological image classification methods includes trained feature-based algorithms that extract high-level features without direct feature extraction operation from histopathology images. Since the initiation of convolutional neural networks (CNNs) learned feature-based approaches received a lot of attention of researchers. The majority of newly developed algorithms for breast cancer grading [21] use convolutional neural networks (CNNs), transfer learning (TL), and residual networks (RN) concepts.

2.1 Handcrafted Feature-Based Algorithms

2.1.1 Pre-processing Techniques

Pre-processing removes undesired distortions from images and improves very relevant feature required for grading of histopathological images. The initial step in the pre-processing is elimination of noises and artefacts from the images which has got in at the time of slide preparation [32]. Apart from this stain normalization, colour separation is also done to improve the quality of the image [30, 33]. Prior knowledge on the stain vectors is used for stain normalization. Otsu thresholding, clustering, and trust region optimization techniques are used for stain optimization. After stain normalization, de-convolution technique is used for colour separation [34].

An adaptive technique was proposed in [35] to decrease colour variation. Colour unmixing or colour de-convolution, a special instance of spectral unmixing, is used by Veta et al. [31] and Basavanhally [36] to separate the stains. In paper [30], stain normalization and colour separation are the pre-processing steps used. Wan et al. [37] perform stain normalization which used non-linear mapping technique to adjust image intensity discrepancies due to variations in tissue preparation.

2.1.2 Segmentation of Anatomical Structures

Nuclei identification and segmentation are important processes in cancer diagnosis. Different features of nuclei, such as their morphological structure, size, and mitotic nuclei count, are crucial for identifying the disease and understanding the malignancy and intensity of the disease. Mainly, there are three types of nuclei segmentation algorithms utilized in breast cancer grading. They are threshold-based approaches, region growing-based techniques, and boundary-based techniques.

In an image, threshold can be used for distinguishing foreground objects from the background objects. Choosing a right threshold value helps to convert the image into a binary image which give adequate knowledge about the shape and size of the nucleus [38]. This makes the feature extraction and classification process comparatively easy due to the reduced complexity of the images. Due to its simple concept, many breast cancer classification systems are used thresholding approach for nuclei segmentation [21].

In [39], Petushi et al. use a combination of local morphological methods and optimum adaptive thresholding. Later in [21, 40], Petushi et al. combine edge levelling and morphological filling technique to determine the appropriate threshold value. Weyn et al. [41] combined background smoothing and thresholding based on the median of histogram intensity to get the contour of nuclei. Moncayo et al. [42] use the hematoxylin contribution map and maximally stable extreme regions to segment the nuclear region (MSER) where different thresholds are applied to the image, and the areas which reflect the least change are labelled as maximally stable extreme regions. The nuclei and the surrounds are then distinguished using open and close morphological operations.

Another method for image segmentation is region growth, which segregates neighbourhood pixels based on homogeneity and adds them to a region where the similarity criteria of the class are met. This method is applied for each of the pixels in the neighbourhood region. The technique of region growing begins by choosing a seed location based on certain criteria [21]. These discovered seed sites are then used to expand these regions depending on some conditions. In noisy images with difficulty to detect boundaries or edges, region growing algorithms are the best options.

After segmentation, [29] performs neoplasm localization using a multi-resolution approach. Cell segmentations are performed on high-resolution pictures using Gaussian colour models, whilst [31] uses a marker controlled watershed segmentation approach. To discover the prospective nuclei, pre-processed images with indicated locations with strong radial symmetry and local minima are used. The water shed segmentation algorithm is then utilized to determine the nuclei's contour. In [43], the external margins of nuclei are segmented using a convex grouping technique that is best suited for patchy and open vesicular nuclei found in breast cancer with a high aggressiveness [21]. This irregular nuclei structure is segmented using the K-means clustering algorithm [44].

The edge and intensity discontinuity in image are important features that provide knowledge about object contours and are extensively utilized in image segmentation and item identification. For nuclei segmentation of breast histopathology image, many edge-based segmentation techniques have been used [21].

Cassato et al. [35] apply a difference of Gaussian (DoG) filter to the image. The resultant edge map is then extracted using the Hough transform. With the edge map, the boundary of the nuclei is defined using an active contour model. Faridi et al. [45] used morphological feature extraction procedures and DOG filtering to detect nuclei's centre and the distance regularized level set evolution (DRLSE) technique to recover nuclear border [21]. In paper [46] to find the candidate cell nuclei, the author has employed morphological techniques and a distance transform. The thresholding technique is applied to a gamma-corrected image to generate a binary image before being subjected to morphological procedures like dilatation and erosion. Paper [47] uses thresholding and morphological filtering techniques for pre-processing and later a snake-based method on an image to retrieve nuclear edges [21]. A polar space transformation is conducted, and then, the nuclei boundary is identified with iterative snake technique. For nuclei segmentation, Basavanhally et al. [36] developed a colour

gradient-based geodesic active contour (CGAC) technique. In the gradient-based technique, segmentation is achieved by minimizing an energy function that includes a typical geodesic active contour [21].

For automatic segmentation of nuclei, many algorithms have been developed and experimented. Due to the errors at the time of image capture and the diverse shape and size of the nucleus, finding the region of interest and segmenting is considered as a difficult task. The accuracy of the nuclear atypia detection technique is strongly dependent on the success of the segmentation technique.

2.1.3 Feature Extraction

A crucial phase in the handcrafted feature-based technique, in which the most significant features of an image are retrieved from a big number of features, is called feature extraction. Excess cell proliferation in cancer tissues is caused by disruptions in the cell life cycle, resulting in impaired cellular function [21]. The various grades of cancer have different cellular characteristics. The majority of these characteristics are graph-based topographical, morphological, and textural features. Various feature extraction methods available in the literature for breast cancer grading are covered in this section [21]. When compared to normal cells or nuclei, malignant cells differ in size and shape, and pathologists utilize this difference to grade malignancy. The form and size of a cell are frequently revealed by morphological trait. Smoothness, symmetry, length of axes, roundness, and concavity are used to describe their shape, whereas the radius, perimeter, and area of segmented nuclei are commonly used to describe their size. For cancer identification and grading, morphological aspects of the nucleus structures have been extracted and used by various algorithms.

For each of the images, Petushi et al. [39] calculated standard deviation, area, intensity mean, the minimum intensity value, major and minor axes of the nuclei that have been segmented. The feature vector generated is given for clustering along with a binary decision tree. The features like mean value of intensity, area, and count of nuclei are extracted from the source image in paper [45]. For cancer grading and prediction, Veta et al. [31] considered the standard deviation and mean of the area of the segmented nuclei for feature extraction.

Textural characteristics provide crucial information on the intensity variation of pixel over a surface with respect to smoothness and regularity. Statistical, spectral, and structural approaches are commonly used to extract textural information. This section discusses about the papers which use various texture features for cancer grading and classification.

In [40], texture features which help to calculate the density of cell nuclei are used in supervised classification approaches such neural networks, decision trees, linear classifiers, and quadratic classifiers. For nuclear grading of breast tumours, Khan et al. [48] developed a textural-based characteristic called the geodesic mean of region covariance descriptors (gmRC). In the next phase, K-nearest neighbour (KNN) classifier is applied on gmRC matrix. Ojansivu et al. [49] used descriptors such as local phase quantization (LPQ) and local binary patterns (LBPs) to create

histograms that describe the image's statistical textural qualities. Later, support vector machine classifier is applied on the pre-processed image data for categorizing the nuclear atypia score. A bag of features (BoF) with multi-scale descriptors is used to represent the discovered nuclei in [42], and later, the K-means clustering algorithm is utilized to split the extracted descriptors, which are then employed as the dictionary's atoms. Rezaeilouyeh in [50] used textural characteristics like shearlets for cancer tissue classification using convolutional neural network (CNN) classifier. For nuclear grading, Lu et al. [30] retrieved a set of morphological and textural parameters like size of nuclei, mean, standard deviation, sum, and entropy of image features. The histogram of each of these attributes is used for grading using an SVM classifier. [43] uses features to reflect the size differences between mitotic nuclei and normal nuclei. The feature set is made up of the mean and standard deviation of ten parameters like contrast, skewness, solidity, grey value, eccentricity, entropy, diameter, area, smoothness, and symmetry retrieved from the segmented nuclei. Gandomkar et al. [22] used texture-based feature extraction method to extract attributes from histopathology slides which can identify the grade of the cancer. These features are later analyzed and coupled with the ensemble of trees to evaluate and to calculate the atypia count.

Topological features in a tumour tissue provide information on the structure and spatial arrangement of nuclei. The spatial interdependence of the cells is represented in this way using various forms of graphs, from which the required information for classification is derived. For cancer grading, graph-based criteria are frequently combined with morphological or textural features.

Naik et al. [51], Das et al. [21] extracts many morphological and graphical representations of features using Voronoi diagram (VD), Delaunay triangulation (DT), and minimal spanning tree (MST) for automated breast cancer grading. Apart from that morphological features as well as boundary features derived from nuclear structure are subjected to principal component analysis (PCA) for feature reduction, and classification is carried out later using an SVM classifier. Textural features and graphical representation of features using MST, VD, and DT are used to represent the architecture of nuclei and grade cancer in Doyle et al. [21, 52]. The textural features like standard deviation, minimum to maximum ratio, and average are included in the study along with graph-based features, and these two feature vectors are subjected to spectral clustering (SC) techniques to reduce the number of features, which is then supplied to the SVM classifier. In paper [53], Wan et al. apply Gabor filters, kirsch filters to access pixel level features, object-based features are extracted through MST, DT, VDs, and convolutional neural networks (CNNs) helps to extract semantic level features, and this helps to identify heterogeneity of cancerous tissue. Later, graph embedding technique is used to reduce the dimensionality of the image, and the resultant image is given to SVM classifier [21]. [48] presents a novel image level descriptor based on area covariances for assessing breast full slide images. The image is divided into non-overlapping area and region covariance (RC) which is calculated for each area. To extract features of this area, maximum response 8(MR8) filter banks are used [21]. The extracted RC descriptors form points on Riemannian manifold.

This is integrated to get single image descriptor called geodesic geometric mean of region covariance, and later, geodesic K-nearest neighbour classifier (GKNN) is used for grading the breast cancer [21].

2.1.4 Image Classification

The features retrieved from tumour tissue are required for cancer classification and grading. Classifiers are divided into learning and testing phase. The characteristics retrieved from digital annotated slides are used to train the classifier during the learning phase [21]. These classifiers are then put to the test with previously unseen data. Algorithms such as Gaussian mixture models, K-nearest neighbourhood (KNN) algorithm, decision tree, random forest classifier (RF), Bayes classifiers, and supervised learning techniques are used for nuclear atypia detection widely. Table 3 gives details of the few most frequently used machine learning techniques in the domain of nuclear atypia scoring. The most widely used deep learning (DL) techniques can avoid explicit feature extraction process [21] and organize discriminative information of the data as its in built functionality.

The cancer features derived from histopathology images are used in an automated breast cancer grading system. This could necessitate precise identification and segmentation of biological structures, which is a difficult operation due to the complex structure of histologic data. As a result, computerized intelligent systems for nuclear atypia scoring are in high demand.

2.2 Learned Feature-Based Algorithms

Deep neural networks (DNNs) have been widely used to solve a variety of medical image analysis problems like genetic disorder identification, speech recognition, Alzheimer's disease classification, etc. DL methods have outperformed traditional methods in several areas. This does not necessitate the feature extraction phase. The layers are capable of learning an implicit representation of the raw input by themselves. Several deep learning algorithms for nuclear atypia scoring have recently appeared in the literature. In this section, we will go through the existing deep learning techniques used in the domain of nuclear atypia detection. Figure 4 shows the learned feature-based nuclear atypia detection process.

Rezaeilouyeh et al. [50] employed the magnitude and phase shearlet coefficients as secondary information for the neural network to train CNN and classify cancer. For the classification of breast cancer images, Rakhlin et al. [55] employ a gradient boosting method pre-trained CNN on Image Net. Bardou et al. [56] used CNN with a fully connected classifier layer to classify cancer subtypes using handcrafted features [21].

To extract details from nuclei in different scale in paper [57], Araujo et al. applied CNN architecture. In [58], author proposed a deep convolutional activation feature

Table 3 Nuclear atypia scoring using handcrafted feature-based machine learning technique

Paper	Pre-processing technique	Segmentation technique	Feature extraction technique	Classification technique
Cosatto et al. [35]	Adaptive thresholding technique	Active contour using Hough transform	Texture and shape features	SVM
Naik et al. [51]	–	Bayesian classifier	Morphological and graph-based features	SVM
Dalle et al. [29]	–	Region growing	Morphological features	–
Dalle et al. [46]	–	Boundary	Size, shape, and texture features	Gaussian mixture model
Huang et al. [47]	–	Snake-based algorithm	Size and textural features	Bayesian classifier
Veta et al. [21, 31]	Colour normalization	Watershed segmentation	Morphological features	–
Basavanhally et al. [21, 36]	Orthonormal transformation of RGB vectors	Snake-based segmentation	Textural and graph-based features	Random forest
Lu et al. [30]	Stain normalization and colour de-convolution	Region growing	Morphological features	SVM
Maqlin et al. [21, 43]	–	K-means clustering	Mean and standard deviation of textural and morphological features	Artificial neural network
Moncayo et al. [21, 42]	Colour de-convolution	Threshold	Textural features	SVM
Faridi et al. [45]	Unmixing of colour channels	Boundary	Morphological features	SVM
Wan et al. [37]	Stain normalization	Snake-based segmentation	Textural, graph-based, and CNN derived features	SVM-based cascaded ensemble classifiers
Gandomkar et al. [21, 54]	Stain normalization and colour de-convolution	Threshold	Textural features	Multiple regression trees
Khan et al. [21, 48]	–	–	Geodesic mean of region covariance descriptors	Geodesic K-nearest neighbour classifier(GKNN)

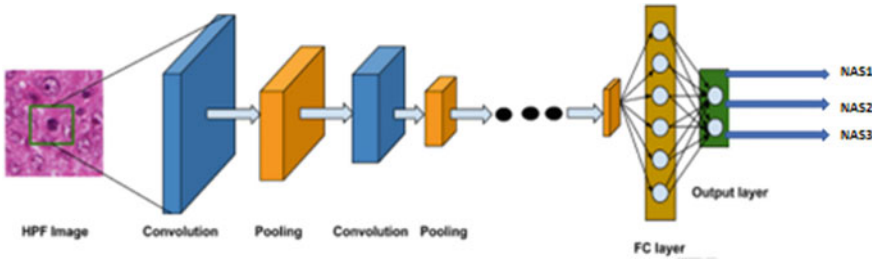

Fig. 4 Learned feature-based nuclear atypia detection process

(DCAF) for classification of histopathological images. In this, a pre-trained CNN for feature extraction is used, and later, it is fed into a new classifier that has been trained on problem specific data. Nahid et al. [59] categorize histopathological images using long short-term memory (LSTM) and an integrated CNN + LSTM model, and a softmax and SVM layers employ the extracted global and local information to determine the class. Guo et al. [60] used a CNN architecture that combines patch level voting, merging module and reduces generalization of error using hierarchy voting approach and bagging technique [21].

Golatkar et al. [61] used Inception-V3 CNN technique on patches extracted based on nuclei density for histological breast cancer classification as the initial step and in the later stage of the process, majority voting technique was used for the final classification [21]. For grading breast cancer, Jannesari et al. [62] employed pre-trained ResNet, ImageNet, and inception model networks. For the multiclass histopathological image classification, Jiang et al. [63] integrated the squeeze and excitation patch and the ResNet model, which reduced the number of parameters and thereby avoiding the problem of model over fitting [21]. In [33] in order to adaptively identify the sample, the author has proposed batch mode active learning on the Riemannian manifold for breast cancer nuclear atypia scoring [21, 64]. In [65], the author proposed sparse coding and dictionary learning on symmetric positive definite (SPD) matrices. The sparse coding problem on the SPD manifold is a convex problem in the higher dimensional reproducing kernel Hilbert space (RKHS), which allows for greater use of the reparability of histopathological images [21, 65].

Recently, generative adversarial networks (GANs) are recently widely employed in biomedical area of research [26] such as medical image reconstruction, image generation, segmentation, classification, and histopathological image analysis since its beginnings [66]. GAN can be used for artificially staining digital pathology images which helps to avoid the problems created through manual staining and thereby reduce the effort and expense of staining process [66]. GAN models are also used to eliminate unknowing and intentional discolorations in stained histopathology pictures that could impair analysis using deep learning algorithm. On breast histology, Xu et al. [67] employed a multi-resolution convolutional network with plurality voting approach for nuclear atypia detection [66].

Conditional GANs [68] are a type of GAN conditioned on labelled images whilst attempting to match the target data directly. This necessitates precisely labelled and registered dataset for training, which might be costly and laborious in histopathological dataset [66]. A coupled GAN (CoGAN) [69] is a GAN version that share weights of layers and combines two GANs to learn a joint distribution from only the residuals [66]. Disadvantage of this GAN model is mode collapse and the gradient growth and instabilities during the process of training. GAN variant [53] recommends that the Wasserstein distance can used to improve the model. The convergence of these methodologies is still uncertain, its implications are still unknown, and the level of computational complexity is likewise quite high [66].

In [66], the author proposed NAS-SGAN to overcome the shortage of labelled images. It is a stacked feature matching semi-supervised generative adversarial network (GAN) technique used for breast cancer grading. The author used the ability of semi-supervised GAN (SGAN) for representing the distribution of data by utilizing the unlabelled and labelled samples [66]. A combined training of the discriminator with the use of feature matching and layered architecture also makes the model much more stable than other GAN models. Deep learning methods for nuclear atypia scoring are summarized in Table 4

2.3 Evaluation Metrics

The evaluation metrics are used to assess the trained classifier model's generalization ability as well as a model evaluation assessor. They are used to evaluate and summarize the efficiency of the classifier model when it is put to the test with data that has not been seen before. For quantitatively summarizing the outcomes, the majority of the approaches we looked at used confusion metrics as the validation metrics.

Performance and quality of classification model are interpreted with confusion matrix table. The true negative (TN), true positive (TP), false positive (FP), and false negative values of the confusion metrics are used to calculate the following metrics like accuracy (Acc), precision (Pr), specificity (TNR), sensitivity or recall (Re), and F1 score (FS), Matthew's correlation coefficient (MCC), false negative rate (FNR), true negative rate (TNR), true positive rate (TPR), false positive rate (FPR), and negative predictive value (NPV) to evaluate the performance of the model classifier.

Receiver operating characteristics (ROC) graphs are an excellent way for analyzing classifiers and visualizing their performance. The relative trade-off between the true positive rate (TPR) and the false positive rate (FPR) at various thresholds is depicted by a ROC curve. The area under the curve (AUC) can be measured to see how well the extracted features can distinguish between different categories of breast cancer. The system is better if the area beneath the curves is larger. Table 5 shows the most commonly used validation metrics in the nuclear atypia detection papers which discussed in the literature review.

Table 4 Nuclear atypia scoring using deep learning techniques

Paper	Classification technique used	Pros of the technique	Cons of the technique
Araujo et al. [57]	CNN	Extract details from nuclei in different scale	Imbalances in the class is not considered
Rakhlin et al. [55]]	CNN	Transfer learning using ImageNet	No pre-processing
Bardou et al. [56]	CNN	Combination of handcrafted and learned features	Multi-class classification is not efficient
Spanhol et al. [58]	DL	Pre-trained deep convolutional activation feature	Complicated structures in the image will pull down the accuracy
Nahid et al. [59]	DL	Integrated CNN and long short-term memory (LSTM) model guided by structural and statistical features	Small dataset
Guo et al. [60]	DL	CNN combined with hierarchy voting and bagging technique which shows good performance	Sensitivity value is low
Golatkar et al. [26]	CNN variants	Pre-trained CNN Inception-v3 is used	Grading accuracy is low
Rezaeilouyeh et al. [50]	DL	Improved accuracy due to DNN fed with shearlet coefficients	Very complex process
Xu et al. [67]	DL	Multi-resolution convolutional network combined with plurality voting	Multiple levels of training required
Jannesari et al. [62]	CNN variants	High sensitivity achieved through combination of residual network and inception network	Requires fine tuning of parameters
Jiang et al. [63]	CNN variants	ReSNet with reduced set of parameters	Very complex process

(continued)

Table 4 (continued)

Paper	Classification technique used	Pros of the technique	Cons of the technique
Asha Das et al. [64]	Batch mode active learning(BMAL)	Active learning on Riemannian manifold helped to reduce manual effort in labelling the images	Very complex process
Asha Das et al. [65]	Kernel-based sparse coding and dictionary learning(KSCDL)	A sparse coding and dictionary learning on SPD matrices provide a better discrimination	Very complex process
Gulrajani et al. [53]	WGAN-GP	GAN model which use Wasserstein distance	Very complex computation process
Liu et al. [69]	CoGAN	Coupled GAN which share weights and learn with join distribution	Gradient expansion, vanishing, and mode collapse aresues whist training
Mirza et al. [68]	CGAN	GAN model which is conditioned on labelled images	Needs perfectly annotated data
Asha Das et al. [66]	CNN variant NAS-SGAN model	A semi-supervised generative adversarial training, which improves accuracy and robustness with limited annotated data	Not using clustering properties for improved classification accuracy

2.4 Performance Analysis of Nuclear Atypia Scoring Algorithms

The histopathological image classification continues to be a difficult task due to the huge storage requirement of the image data. Furthermore, the processing required for the image analysis and classification is huge and can take several hours. As a result, computer analysis of histopathological images frequently necessitates the employment of highly efficient computing technologies such as multi-core processors and graphics processing units (GPUs) that are required to speed up the processing. A comparison of the many strategies detailed in the literature is time-consuming because each method uses its own unique dataset, and the findings are given using different assessment measures. This section discusses a comparative examination of some papers which had worked on breast cancer nuclear atypia scoring [24].

The papers included in the comparative analysis are geodesic-distance-based K-nearest neighbour (GKNN) classifier [48, 66], support vector machine [30], deep neural network shearlet transform [50], multi-resolution convolutional network with plurality voting (MR-CN-PV) model [66, 67], batch mode active learning (BMAL)

Table 5 Validation metrics in the nuclear atypia detection

Metrics	Formula	Description
TP		Count of nuclear atypia cells classified correctly
TN		Count of non-nuclear atypia cells classified correctly
FP		Count of nuclear atypia classified incorrectly
FN		Count of non-nuclear atypia classified incorrectly
Accuracy (Ac)	$\frac{TP+TN}{TP+TN+FN+FP}$	It calculates the proportion of right classifications to the total number of test data analyzed
Precision (Pr)	$\frac{TP}{TP+FP}$	Count of correctly identified positives
F1-score (Fs)	$\frac{2*Precision*Recall}{Precision+Recall}$	Harmonic mean of sensitivity gives F1-score
Specificity	$\frac{TN}{TN+FP}$	Count of correctly identified negatives
Sensitivity	$\frac{TP}{TP+FN}$	Calculates the count of correctly identified positives
Error rate	$\frac{FP+FN}{TP+TN+FP+FN}$	The ratio of inaccurate classifications to the total number of test data analyzed
MCC	$\frac{(TP*TN-FP*FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$	The observed classifications' correlation coefficient with the projected classifications
NPV	$\frac{TN}{TN+FN}$	The percentage of negatives properly detected out of the total samples in the negative class
FPR	$\frac{FP}{FP+TN}$	The proportion of negative samples that were incorrectly classified as positive to the count of negative samples
FNR	$\frac{FN}{FN+TP}$	The count of positive samples that was incorrectly forecasted as negative to the count of negative samples

[64, 66], kernel-based sparse coding and dictionary learning (KSCDL) [65, 66], Wasserstein GAN-gradient penalty (WGAN-GP) [53], coupled GAN (CoGAN) [69], conditional GAN (CGAN) [66, 68], and nuclear atypia scoring semi-supervised generative adversarial network (NAS-SGAN) [66]. Tables 6 and 7 show analysis of performance of nuclear atypia scoring on the Aperio and Hamamatsu dataset.

The KSCDL technique was chosen for analysis because they investigate histological breast images in a non-Euclidean framework, and its performance was good [65]. In addition, the SVM, DNN-shearlet, GKNN, and MR-CN-PV were also taken into consideration for comparison because of their outstanding results in textural, morphological, deep learning, and transfer learning-based algorithms. Furthermore, batch

Table 6 Performance analysis of nuclear atypia scoring algorithms on Aperio dataset [21]

Paper	Algorithm	Acc	Re	TNR	Pr	FS	MCC
Khan et al. [48]	GKNN	0.797	0.562	0.782	0.631	0.565	0.393
Lu et al. [30]	SVM	0.780	0.455	0.771	0.445	0.447	0.262
Rezaeilouyeh et al. [50]	DNN-shearlet	0.720	0.333	0.666	0.250	0.285	0.163
Xu et al. [67]	MR-CN-PV	0.800	0.325	0.672	0.330	0.317	0.030
Asha et al. [64]	BMAL	0.853	0.785	0.874	0.762	0.790	0.665
Asha et al. [65]	KSCDL	0.826	0.762	0.855	0.721	0.741	0.621
Gulrajani et al. [53]	WGAN-GP	0.875	0.813	0.902	0.786	0.813	0.701
Liu et al. [69]	CoGAN	0.873	0.811	0.900	0.785	0.812	0.698
Mirza et al. [68]	CGAN	0.857	0.792	0.880	0.767	0.790	0.667
Asha et al. [66]	NAS-SGAN	0.982	0.964	0.974	0.965	0.963	0.940

Table 7 Performance analysis of nuclear atypia scoring algorithms on Hamamatsu dataset [21]

Paper	Algorithm	Acc	Re	TNR	Pr	FS	MCC
Khan et al. [48]	GKNN	0.812	0.571	0.779	0.627	0.569	0.395
Lu et al. [30]	SVM	0.756	0.406	0.726	0.423	0.402	0.181
Rezaeilouyeh et al. [50]	DNN-shearlet	0.747	0.333	0.666	0.249	0.285	0.173
Xu et al. [67]	MR-CN-PV	0.702	0.476	0.726	0.520	0.492	0.224
Asha et al. [64]	BMAL	0.864	0.797	0.880	0.769	0.781	0.653
Asha et al. [65]	KSCDL	0.829	0.761	0.844	0.722	0.742	0.621
Gulrajani et al. [53]	WGAN-GP	0.885	0.815	0.904	0.795	0.814	0.677
Liu et al. [69]	CoGAN	0.868	0.812	0.901	0.793	0.813	0.674
Mirza et al. [68]	CGAN	0.868	0.799	0.881	0.770	0.782	0.654
Asha et al. [66]	NAS-SGAN	0.984	0.975	0.982	0.975	0.974	0.956

mode active learning on Riemannian manifold (BMALR) approach was considered for the comparison due to the active learning used in it [64, 66]. The method was used to reduce the requirement for a labelled dataset. The problem of getting precisely labelled and registered data which is often costly and laborious in the case of histopathology images is resolved to an extend by the use of by the generative adversarial network [66]. Three papers which use GAN concept were chosen for comparative analysis due to exceptional accuracy level when compared to their techniques. Out of the three techniques, complexity and issue of convergence were the disadvantages of CoGAN and CGAN when compared to NAS-SGAN. Amongst all the three techniques which use GAN concept, the NAS-SGAN which use semi-supervised GAN concept was able to resolve the stability issue to an extend and SGANs using its stacked feature matching it was able to collect the data distribution more effectively using both labelled and unlabelled data [66]. Figures 5 and 6

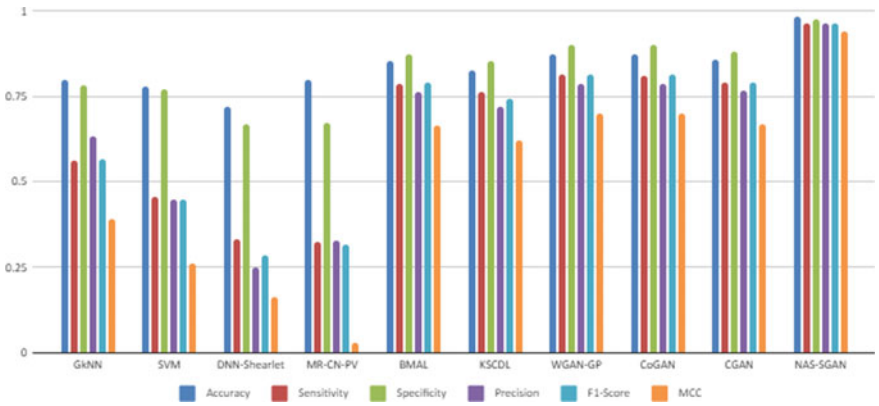

Fig. 5 Performance analysis of nuclear atypia scoring algorithms on Aperio dataset

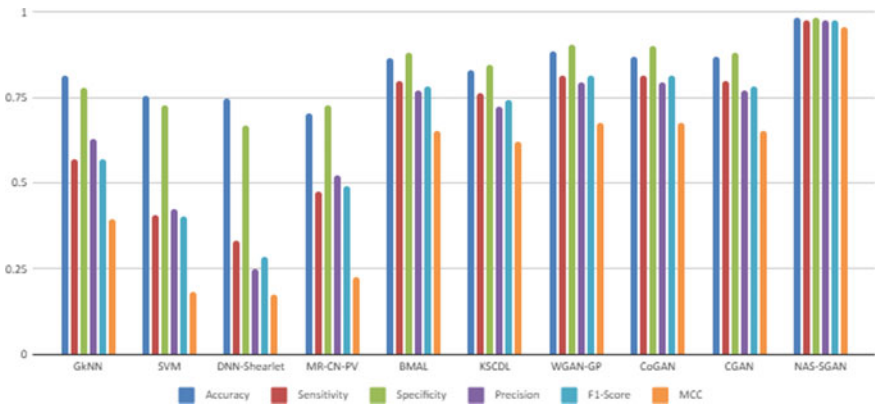

Fig. 6 Performance analysis of nuclear atypia scoring algorithms on Hamamatsu dataset

show the performance analysis of nuclear atypia scoring algorithms on Aperio and Hamamatsu dataset [21]

2.5 Conclusion

The significant challenges of nuclear atypia detection in histopathological breast tissue image analysis are distortions of image caused by inaccurate slide preparation, complicated structure of the underlying biological pattern, lack of stability as well as inter- and intra-observer variation, which can affect diagnostic accuracy and therapy planning. These challenges can be overcome with the help of computer aided analytical techniques. This study reviews the methodologies used for nuclear

atypia detection process in breast cancer grading, obstacles in existing computerized automated grading process, prospective techniques of computer-assisted grading for the disease diagnosis. This research is an attempt to synthesize information to learn about the most recent changes in analyzing characteristics for breast cancer grading and provide a summary of the precision and reliability of various methods in the domain of nuclear atypia detection. The findings shows that deep learning has exploded in popularity in the field of nuclear atypia detection as a result of its superior results. Approaches based on deep learning offer higher prospects in the future. In deep learning technique, generative adversarial network (GAN) outperforms non-adversarial and deep learning networks because adversarial training uses the discriminator network to train the generator network, resulting in improved model performance. GAN also provides provision for automatic staining, and it performs well even without perfectly annotated dataset. We also found that public datasets and open challenges have a good impact on research on histopathological image analysis, and most of the studies in the literature used public dataset. In the future variations, of GAN technique can be developed with acceptable performance levels for clinical applications. This study could serve as inspiration for future research.

References

1. WHO—cancer key facts; 2018. https://www.who.int/news-room/fact-sheets/detail/cancer. Accessed 20 Nov 2020
2. Mathew T et al (2020) Computational methods for automated mitosis detection in histopathology images: a review. Biocybern Biomed Eng. https://doi.org/10.1016/j.bbe.2020.11.005
3. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. Cancer J Clin 68(6):394–424
4. Wardle J, Robb K, Vernon S, Waller J (2015) Screening for prevention and early diagnosis of cancer. Am Psychol 70(2):119
5. Mahmood T, Li J, Pei Y, Akhtar F, Imran A, Rehman KU (2020) A brief survey on breast cancer diagnostic with deep learning schemes using multi-image modalities. IEEE Access 8:165779–165809. https://doi.org/10.1109/ACCESS.2020.3021343
6. Bejnordi BE, Zuidhof G, Balkenhol M, Hermsen M, Bult P, van Ginneken B, Karssemeijer N, Litjens G, van der Laak J (2017) Context-aware stacked convolutional neural networks for classification of breast carcinomas in whole-slide histopathology images. J Med Imag 4(4):1–8
7. Al-Janabi S, Huisman A, Van Diest PJ (2012) Digital pathology: current status and future perspectives. Histopathology 61(1):1–9
8. Higgins C (2015) Applications and challenges of digital pathology and whole slide imaging. Biotechn Histochem 90(5):341–347
9. Manoharan JS (2021) Study of variants of extreme learning machine (ELM) brands and its performance measure on classification algorithm. J Soft Comput Paradigm (JSCP) 3(02):83–95
10. Al-Janabi S, Huisman A, Vink A, Leguit RJ, Offerhaus GJA, Ten Kate FJ et al (2012) Whole slide images for primary diagnostics of gastrointestinal tract pathology: a feasibility study. Human Pathol 43(5):702–707

11. Shibusawa M, Nakayama R, Okanami Y, Kashikura Y, Imai N, Nakamura T, Kimura H, Yamashita M, Hanamura N, Ogawa T (2016) The usefulness of a computer-aided diagnosis scheme for improving the performance of clinicians to diagnose non-mass lesions on breast ultrasonographic images. J Med Ultrason 43(3):387–394
12. Kumar R, Srivastava R, Srivastava S (2015) Detection and classification of cancer from microscopic biopsy images using clinically significant and biologically interpretable features. J Med Eng 2015(2015):1–14. https://doi.org/10.1155/2015/457906
13. Wan T et al (2014) Wavelet-based statistical features for distinguishing mitotic and non-mitotic cells in breast cancer histopathology. In: 2014 IEEE international conference on image processing (ICIP) 2290–2294
14. Singletary SE, Allred C, Ashley P, Bassett LW, Berry D, Bland KI et al (2002) Revision of the American joint committee on cancer staging system for breast cancer. J Clin Oncol 20(17):3628–3636
15. O'Sullivan B, Brierley J, Byrd D, Bosman F, Kehoe S, Kossary C, Piñeros M, Van Eycken E, Weir HK, Gospodarowicz M (2017) The TNM classification of malignant tumours-towards common understanding and reasonable expectations. Lancet Oncol 18(7):849–851. https://doi.org/10.1016/S1470-2045(17)30438-2. PMID: 28677562; PMCID: PMC5851445.
16. Elston CW, Ellis IO (1991) Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology 41(3a). 151. Elston CW, Ellis IO (2002) Author commentary. Histopathology 19:403–410
17. Fuchs TJ, Buhmann JM (2011) Computational pathology: challenges and promises for tissue analysis. Comput Med Imaging Graph 35(7–8):515–530
18. Hamdan YB (2021) Construction of statistical SVM based recognition model for handwritten character recognition. J Inf Technol 3(02):92–107
19. Paul A, Mukherjee DP (2015) Mitosis detection for invasive breast cancer grading in histopathological images. IEEE Trans Image Process 24(11):4041–4054
20. Raj JS, Joe MCV (2021) Wi-Fi network profiling and QoS assessment for real time video streaming. IRO J Sustain Wirel Syst 3(1):21–30
21. Das A, Nair MS, Peter SD (2020) Computer-aided histopathological image analysis techniques for automated nuclear atypia scoring of breast cancer: a review. J Digit Imaging 33(5):1091–1121. https://doi.org/10.1007/s10278-019-00295-z.PMID:31989390;PMCID:PMC7573034
22. MITOS – dataset; 2012. http://ludo17.free.fr/mitos_2012/dataset.html. Accessed 20 Nov 2020
23. Veta M, Van Diest PJ, Willems SM, Wang H, Madabhushi A, Cruz-Roa A et al (2015) Assessment of algorithms for mitosis detection in breast cancer histopathology images. Med Image Anal 20(1):237–248
24. MITOS-ATYPIA—dataset; 2014. https://mitos-atypia-14.grand-challenge.org/Dataset/. Accessed 20 Nov 2020
25. TUPAC – dataset; 2016. http://tupac.tue-image.nl/node/3. Accessed 20 Nov 2020
26. Spanhol FA, Oliveira LS, Petitjean C, Heutte L (2016) A dataset for breast cancer histopathological image classification. IEEE Trans Biomed Eng 63(7):1455–1462. https://doi.org/10.1109/TBME.2015.2496264
27. Aksac A, Demetrick DJ, Ozyer T et al (2019) BreCaHAD: a dataset for breast cancer histopathological annotation and diagnosis. BMC Res Notes 12:82. https://doi.org/10.1186/s13104-019-4121-7
28. Beevi KS, Nair MS, Bindu GR (2016) Automatic segmentation of cell nuclei using Krill Herd optimization based multithresholding and localized active contour model. Biocybernetics Biomed Eng 36(4):584–596. https://doi.org/10.1016/j.bbe.2016.06.005
29. Dalle JR, Leow WK, Racoceanu D, Tutac AE, Putti TC (2008) Automatic breast cancer grading of histopathological images. In: 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp 3052–3055. https://doi.org/10.1109/IEMBS.2008.4649847

30. Lu C, Ji M, Ma Z, Mandal M (2015) Automated image analysis of nuclear atypia in high-power field histopathological image. J Microsc 258(3):233–240. https://doi.org/10.1111/jmi.12237
31. Veta M, Kornegoor R, Huisman A, Verschuur-Maes AHJ, Viergever MA, Pluim JPW, van Diest PJ (2012) Prognostic value of automatically extracted nuclear morphometric features in whole slide images of male breast cancer. Mod Pathol 25(12):1559–1565. https://doi.org/10.1038/modpathol.2012.126
32. Man R, Yang P, Xu B (2020) Classification of breast cancer histopathological images using discriminative patches screened by generative adversarial networks. IEEE Access 8:155362–155377. https://doi.org/10.1109/ACCESS.2020.3019327
33. Niethammer M, Borland D, Marron J, Woosley JT, Thomas NE (2010) Appearance normalization of histology slides. In: MLMI, Springer, Berlin, pp 58–66
34. Macenko M, Niethammer M, Marron JS, Borland D, Woosley JT, Guan X, Schmitt C, Thomas NE (2009) A method for normalizing histology slides for quantitative analysis. In: Proceedings—2009 IEEE international symposium on biomedical imaging: from nano to macro. ISBI 2009, pp 1107–1110. https://doi.org/10.1109/ISBI.2009.5193250
35. Cosatto E, Miller M, Graf HP, Meyer JS (2008) Grading nuclear pleomorphism on histological micrographs. 2008 ICPR 2008 19th international conference on (August 2016) Pattern Recognition, pp 1–4. https://doi.org/10.1109/ICPR.2008.4761112
36. Basavanhally A, Ganesan S, Feldman M, Shih N, Mies C, Tomaszewski J, Madabhushi A (2013) Multi-field-of-view framework for distinguishing tumor grade in ER+ breast cancer from entire histopathology slides. IEEE Trans Biomed Eng 60(8):2089–2099. https://doi.org/10.1109/TBME.2013.2245129
37. Wan T, Cao J, Chen J, Qin Z (2017) Automated grading of breast cancer histopathology using cascaded ensemble with combination of multi-level image features. Neurocomputing 229:34–44. https://doi.org/10.1016/j.neucom.2016.05.084
38. Xing F, Yang L (2016) Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: a Comprehensive review. IEEE Rev Biomed Eng 9:234–263. https://doi.org/10.1109/RBME.2016.2515127
39. Petushi S, Katsinis C, Coward C, Garcia F, Tozeren A (2004) Automated identification of microstructures on histology slides. In: 2004 IEEE international symposium on biomedical imaging: nano to macro, IEEE, pp 424–427
40. Petushi S, Garcia FU, Haber MM, Katsinis C, Tozeren A (2006) Large-scale computations on histology images reveal gradedifferentiating parameters for breast cancer. BMC Med Imaging 6(1):14. https://doi.org/10.1186/1471.2342.6.14
41. Weyn B, Van De Wouwer G, Van Daele A, Scheunders P, Van Dyck D, Van Marck E, Jacob W (1998) Automated breast tumor diagnosis and grading based on wavelet chromatin texture description. Cytometry 33(1):32–40. https://doi.org/10.1002/(SICI)1097-0320(19980901)33:1.32:AID-CYTO4.3.0.CO;2-D
42. Moncayo R, Romo-Bucheli D, Romero E (2015) A grading strategy for nuclear pleomorphism in histopathological breast cancer images using a bag of features (bof). Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics) 9423:75–82. https://doi.org/10.1007/978.3.319.25751.8.10
43. Maqlin P, Thamburaj R, Mammen JJ, Manipadam MT (2015) Automated nuclear pleomorphism scoring in breast cancer histopathology images using deep neural networks. Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics) 9468:269–276. https://doi.org/10.1007/978.3.319.26832.3.26
44. Xue M, Shivakumara P, Zhang C et al (2019) Curved text detection in blurred/non-blurred video/scene images. Multimed Tools Appl 78:25629–25653. https://doi.org/10.1007/s11042-019-7721-2
45. Faridi P, Danyali H, Helfroush MS, Jahromi MA (2016) Cancerous nuclei detection and scoring in breast cancer histopathological images. arXiv:161201237
46. Dalle Jr, Racoceanu D, Putti TC (2009) Nuclear pleomorphism scoring by selective cell nuclei detection. In: IEEE workshop on applications of computer vision: 7–8

47. Huang CH, Veillard A, Roux L, Lomenie N, Racoceanu D (2011) Time-efficient sparse analysis of histopathological whole slide images. Comput Med Imaging Graph 35(7–8):579–591. https://doi.org/10.1016/j.compmedimag.2010.11.009
48. Khan AM, Sirinukunwattana K, Rajpoot N (2015) A global covariance descriptor for nuclear atypia scoring in breast histopathology images. IEEE J Biomed Health Inform 19(5):1637–1647. https://doi.org/10.1109/JBHI.2015.2447008
49. Ojansivu V, Linder N, Rahtu E, Pietikainen M, Lundin M, Joensuu H, Lundin J (2013) Automated classification of breast cancer morphology in histopathological images. Diagn Pathol 8(1):S29
50. Rezaeilouyeh H, Mollahosseini A, Mohammad MH (2016) Microscopic medical image classification framework via deep learning and shearlet transform. J Med Imaging 3(4):044,501. https://doi.org/10.1117/1.JMI.3.4.044501
51. Naik S, Doyle S, Agner S, Madabhushi A, Feldman M, Tomaszewski J (2008) Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology. In: 2008 5th IEEE international symposium on biomedical imaging: from nano to macro, proceedings, ISBI, pp 284–287. https://doi.org/10.1109/ISBI.2008.4540988
52. Doyle S, Agner S, Madabhushi A, Feldman M, Tomaszewski J (2008) Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features. In: 2008 5Th IEEE international symposium on biomedical imaging: from nano to macro, proceedings, ISBI, pp 496–499. https://doi.org/10.1109/ISBI.2008.4541041
53. Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville A (2017) Improved training of Wasserstein GAN's. arXiv preprint arXiv:1704.00028, 2017
54. Gandomkar Z, Brennan PC, Mello-Thoms C (2019) Computer-assisted nuclear atypia scoring of breast cancer: a preliminary study. J Digit Imaging. https://doi.org/10.1007/s10278-019-00181-8
55. Rakhlin A, Shvets A, Iglovikov V, Kalinin AA (2018) Deep Convolutional Neural Networks for Breast Cancer Histology Image Analysis. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol 10882 LNCS, pp 737–744. https://doi.org/10.1007/978-3-319-93000-8-83, 1802.00752.
56. Bardou D, Zhang K, Ahmad SM (2018) Classification of breast cancer based on histology images using convolutional neural networks. IEEE Access 6:24,680–24,693
57. Ara´ujo T, Aresta G, Castro E, Rouco J, Aguiar P, Eloy C, Polonia A, Campilho A Classification of breast cancer histology ´ images using convolutional neural networks. PloS one 12(6), 2017.
58. Spanhol FA, Oliveira LS, Cavalin PR, Petitjean C, Heutte L (2017) Deep features for breast cancer histopathological image classification. In: 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, pp 1868–1873
59. Nahid AA, Mehrabi MA, Kong Y (2018) Histopathological breast cancer image classification by deep neural network techniques guided by local clustering. BioMed research international 2018.
60. Guo Y, Dong H, Song F, Zhu C, Liu J (2018) Breast Cancer Histology Image Classification Based on Deep Neural Networks. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol 10882 LNCS, pp 827–836. https://doi.org/10.1007/978-3-319-93000-8-94, 1803.04054
61. Golatkar A, Anand D, Sethi A (2018) Classification of breast cancer histology using deep learning. In: Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), vol 10882 LNCS, pp 837–844. https://doi.org/10.1007/978-3-319-93000-8 95
62. Jannesari M, Habibzadeh M, Aboulkheyr H, Khosravi P, Elemento O, Totonchi M, Hajirasouliha I (2019) Breast cancer histopathological image classification: a deep learning approach. In: Proceedings - 2018 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2018, pp 2405–2412. https://doi.org/10.1109/BIBM.2018.8621307.
63. Jiang Y, Chen L, Zhang H, Xiao X (2019) Breast cancer histopathological image classification using convolutional neural networks with small SE-resnet module. PLoS ONE 14(3). https://doi.org/10.1371/journal.pone.0214587.

64. Das A, Nair MS, Peter DS (2020) Batch mode active learning on the Riemannian manifold for automated scoring of nuclear pleomorphism in breast cancer. Artif Intell Med 103:101805. ISSN 0933-3657. https://doi.org/10.1016/j.artmed.2020.101805
65. Das A, Nair MS, Peter SD (2019) Sparse representation over learned dictionaries on the Riemannian manifold for automated grading of nuclear pleomorphism in breast cancer. IEEE Trans Image Process 28(3):1248–1260. https://doi.org/10.1109/TIP.2018.2877337
66. Das A, Devarampati VK, Nair MS (2021 Nov 26) NAS-SGAN: a semi-supervised generative adversarial network model for atypia scoring of breast cancer histopathological images. IEEE J Biomed Health Inform. https://doi.org/10.1109/JBHI.2021.3131103. Epub ahead of print. PMID: 34826299
67. On breast histology, Xu et al. [76] describe a deep learning–based approach for computerized Nuclear Atypia Grading (NAS). The computerized scoring is done via a Multi-Resolution Convolutional Network with Plurality Voting
68. Mirza M, Osindero S (2014) Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784
69. Liu M, Tuzel O (2016) Coupled generative adversarial networks. CoRR abs/1606.07536. [Online]. Available http://arxiv.org/abs/1606.07536