

Received January 19, 2022, accepted March 7, 2022, date of publication March 22, 2022, date of current version May 5, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3161575

Progress of Machine Vision in the Detection of Cancer Cells in Histopathology

WENBIN HE¹, YONGJIE HAN¹, WUYI MING^{1,2}, JINGUANG DU¹, YINXIA LIU³, YUAN YANG⁴, LEIJIE WANG⁵, YONGQIANG WANG⁶, ZHIWEN JIANG¹, CHEN CAO^{1,2}, AND JIE YUAN¹

¹Henan Key Laboratory of Intelligent Manufacturing of Mechanical Equipment, Zhengzhou University of Light Industry, Zhengzhou 450002, China

²Guangdong Provincial Key Laboratory of Digital Manufacturing Equipment, Guangdong HUST Industrial Technology Research Institute, Dongguan 523808, China

³Laboratory Medicine of Dongguan Kanghua Hospital, Dongguan 523808, China

⁴Guangdong Provincial Hospital of Chinese Medicine, Guangzhou 510120, China

⁵School of Mechanical Engineering, Dongguan University of Technology, Dongguan 523808, China

⁶Zhengzhou Coal Mining Machinery Group Company Ltd., Zhengzhou 450016, China

Corresponding authors: Wuyi Ming (mingwuyi@gmail.com) and Yuan Yang (yangyuan@gzucm.edu.cn)

This work was supported in part by the Key Scientific and Technological Projects in Henan Province in 2020 under Grant 202102210395, in part by the Scientific and Technological Innovation Team of Colleges and Universities in Henan Province in 2020 under Grant 20IRTSTHN015, and in part by the Guangdong Provincial Key Laboratory of Manufacturing Equipment Digitization in 2020 under Grant 2020B1212060014.

ABSTRACT In recent years, with the rapid development of artificial intelligence, machine vision technology has been widely used in various fields. Traditional cancer detection methods are time-consuming, labor-intensive, and highly dependent on the experience of pathologists; therefore, these methods cannot meet the needs of modern medical treatment. Machine vision overcomes the disadvantages of traditional detection methods in cancer detection and can help pathologists improve the detection accuracy. According to the requirements of medical detection, this review summarizes the applications of machine vision in the detection of cancer cells in histopathological images and analyzes the advantages and disadvantages of existing methods in image preprocessing, segmentation, feature extraction and recognition. Finally, research on the detection methods of histopathological cancer cells is reviewed and prospected, and future development trends are predicted to provide guidance for follow-up research.

INDEX TERMS Machine vision, histopathological images, cancer cell detection, preprocessing, image segmentation, feature extraction, classification.

I. INTRODUCTION

At present, the rapid increase of cancer has become one of the main causes of death worldwide [1]. According to cancer statistics, in 2018, there were 18.1 million new confirmed cancer cases and 9.6 million cancer deaths. Lung cancer, female breast cancer, prostate cancer and colorectal cancer are the most common cancer diagnoses. The highest death rates were lung cancer, colorectal cancer, gastric cancer and liver cancer [2]. In 2020, breast cancer has replaced lung cancer as the most diagnosed cancer in the world; and breast cancer, lung cancer, liver cancer, gastric cancer, and colon cancer have become the five major causes of cancer deaths [3]. Cancer has become one of the causes of death that cannot be ignored worldwide, regardless of the level

The associate editor coordinating the review of this manuscript and approving it for publication was Sudhakar Babu Thanikanti.

of human development. Figure 1 shows the forecast of the number of new cancers and deaths between 2020 and 2040. If cancer patients can be found early, the cure rate can reach 80%, which is the authoritative conclusion of the World Health Organization [4]. Generally, cancer is a disease caused by cell mutations. Cancer is detected based on cells, so the detection of cells is particularly important [5], [6].

The purpose of cell detection is to determine whether there are specific types of cells in a input image and to identify and locate them [7]. In cell detection, the study of histopathological images is considered the gold standard for cancer diagnosis and grading [8]. One of the greatest advantages of histopathological images is their magnification, which will greatly affect the analysis of tissue images. The higher the magnification is, the better details can be visualized. Figure 2 shows the continuous magnification of prostate histopathological samples, and more cell details can

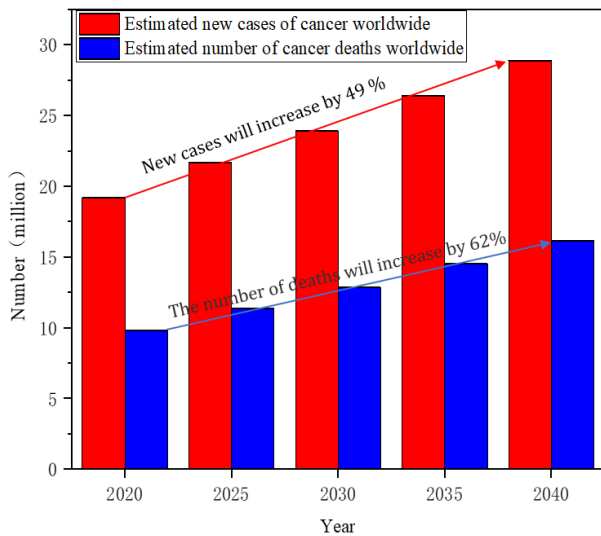


FIGURE 1. Estimation of new and death cases of all cancer types worldwide between 2020 and 2040.

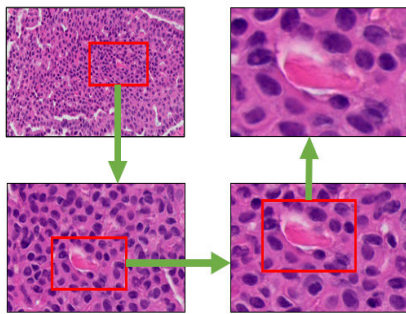


FIGURE 2. Enlarged stained image of prostate tissue.

be seen from left to right [9]. Hematoxylin and Eosin (H&E) staining is the most commonly used light microscope staining method in histopathological laboratories [10]. In traditional histopathological testing, the pathologist uses his eyes to directly observe whether there are cancer cells in a tissue sample. The traditional method is time-consuming and labor-intensive. In addition, traditional detection methods are subjective and lack quantitative feature parameters as a reference, resulting in different recognition accuracies. Researchers in the field of image analysis and pathology have recognized the importance of the quantitative analysis of pathological images. Therefore, efficient, stable and accurate quantitative analysis, detection and identification of cell images have become the focus of research work [11], [12]. In recent years, detection methods based on machine vision have largely overcome the shortcomings of traditional pathological detection methods.

Machine vision is a method of extracting useful information from images of real objects through optical components, visual sensors and digital technology. In simple terms, machine vision uses machines instead of human eyes for

measurement and judgment [13]. Figure 3 shows a machine vision detection system, which mainly includes image acquisition, image preprocessing, image segmentation, feature extraction and classification and recognition modules [11]. Image acquisition is the process of obtaining digital images through optical system mapping and then obtaining human observable data through the subsequent processing of the collected images by a computer. The image preprocessing module mainly includes image denoising, image enhancement, color normalization and other steps. In image preprocessing, due to the influence of the field environment, CCD image photoelectric conversion, transmission circuit and electronic components, images will contain noise, which reduces the quality of images and has adverse effects on image processing and analysis [14]. In addition, factors such as chemical substances and environmental conditions cause changes in the staining information of pathological images, causing unnecessary problems [15]. After preprocessing, images become clearer, which enhances the recognition effect of images and is conducive to further target segmentation. The purpose of image segmentation is to extract the region of interest from an image. Commonly used methods are the threshold method, active contour method, clustering method, watershed method and neural network method. The main difference in cell structure between normal cells and abnormal cells is as follows. Normal cells have precise shapes and sizes, and abnormal cells do not have any precise boundaries like normal cells [16]. Therefore, histopathological image segmentation mainly segments the contours of the cell body, nucleus and cytoplasm. Image feature extraction extracts expressions that can describe the characteristics of a target from an image and map the high-dimensional feature space to the low-dimensional feature space, which is conducive to subsequent image recognition [17]. The commonly used methods of histopathological image feature extraction include texture features, shape features, color features and others. Due to the differences in tissue structures and cell shapes, texture features and shape features are the most widely used methods. In the feature extraction process, a single feature cannot achieve a good recognition effect. Many researchers use the method of combining multiple features for feature extraction to achieve better recognition results. As the last step of image recognition, classification is very important. The accuracy and stability of classification depends not only on the performance of the classifier itself but also on the selection of appropriate feature extraction methods [18].

Machine vision is a rapidly growing field that provides advantages such as remote diagnosis and image analysis to improve the efficiency of decision-making processes. The application of cancer cell detection methods based on machine vision in disease prediction and risk assessment is increasing, and it has become a research hotspot [19], [20]. There are many studies on machine vision in cancer cell detection [21]–[23], which promote the development of subsequent cancer cell detection. The mitotic activity of cancer cells is one of the most decisive signs of cancer

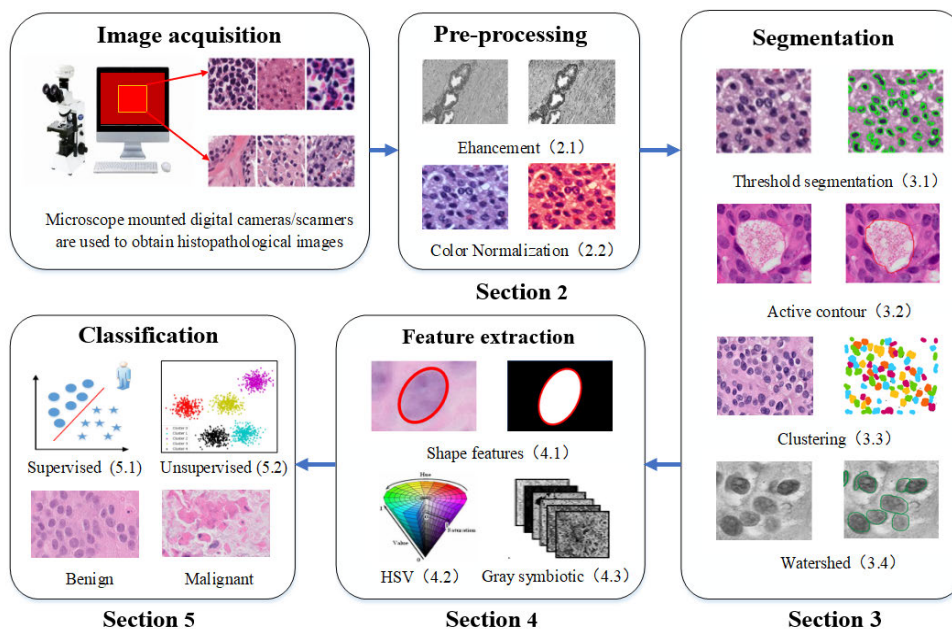


FIGURE 3. Machine vision detection system for the detection of cancer cells in histopathology.

development [24]. Many studies have detected pathological images of mitosis [25], [26]. There are several noteworthy challenges in the detection and localization of cancer cells: (1) Huge changes in cell shape. (2) Histopathological images may contain very complex cellular structures in a disordered manner that make the detection of cancer cells difficult. (3) Cell images have high variability, such as overlapping cells, dust, impurities, and uneven lighting. Hence, reliably segmenting nuclear images is still a challenging task [27].

Histopathological image detection based on machine vision has the advantages of high work efficiency and low work intensity. Therefore, this study conducted a comprehensive review of the machine vision methods used to detect histopathological images. Image detection is divided into the following processes: image preprocessing, image segmentation, image feature extraction and image recognition. These analysis processes are usually applicable to all imaging modes. The first section of this paper introduces the application and progress of machine vision in cell detection. In 2 section, the histopathological image preprocessing, color normalization and image enhancement steps are summarized, analyzed and discussed. In 3 section, we discussed the latest progress in histopathological image segmentation and analyzed the advantages and disadvantages of various segmentation methods. The fourth section specifically discusses the main methods of histopathological image feature extraction. In 5 section, we summarize the advantages and disadvantages of the classifiers and their applications in cancer cell detection. In Section 6, the problems of cancer cell detection algorithm in histopathological images are discussed. Finally, some potential problems and future development directions

of machine vision in histopathological image detection are proposed.

II. IMAGE PREPROCESSING

During the image acquisition process, due to the interference generated by an image itself or introduced from the outside, the useful information in the image will be overwhelmed or lost in the subsequent processing, thereby affecting the subsequent processing. Therefore, image preprocessing is a necessary process that can improve the image quality, highlight and strengthen the part of interest, and suppress useless information. Histopathological image preprocessing mainly includes image enhancement and color normalization.

A. IMAGE ENHANCEMENT

Image enhancement enhances the overall or local features of an image, and its purpose is to improve image quality problems caused by objective factors such as light illumination, signal transmission, and the collection environment. Image enhancement algorithms include grayscale transformation, histogram equalization and filtering algorithms. Most filtering algorithms are processed in the spatial and frequency domains. Interventionary studies involving animals or humans, and other studies that require ethical approval, must list the authority that provided approval and the corresponding ethical approval code.

Mouelhi *et al.* [28] used gray histogram equalization to enhanced the contrast and eliminated the background of breast tissue image. Aswathy *et al.* [16] adopted a contrast-limited adaptive histogram equalization method, which improved the contrast of the input image by changing the hue saturation value of the breast tissue image.

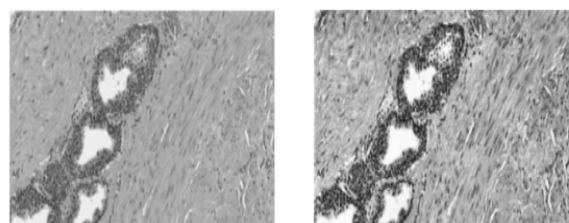
TABLE 1. Histopathological image enhancement.

Methods	Authors, year	Objective	Comments
Median filter (nonlinear)	Logombal <i>et al.</i> (2015) [4]	Breast	Simple method, easy to cause discontinuity of the image.
Gaussian filter (linear)	Kleczek <i>et al.</i> (2020) [10], Kowal <i>et al.</i> (2018) [29]	Skin, Breast	Suitable for eliminating Gaussian noise.
Anisotropic diffusion (nonlinear)	Sertel <i>et al.</i> (2009) [31]	Neuroblastoma	Smooth uneven areas and preserve edges.
Wiener filter (linear)	Öztürk <i>et al.</i> (2018) [30]	Tissue	Wide adaptability, but not suitable for non-stationary noise.
Matched filter (linear)	Kamel <i>et al.</i> (2010) [32]	Follicle	It has the ability to smooth noise and preserve image contours.
Morphology	Wang <i>et al.</i> (2019) [27], Vahadane <i>et al.</i> (2014) [33], Mohammed <i>et al.</i> (2013) [35]	Cervix, Prostate, lymph	Logic-based filtering, more diversified calculation methods, and enhanced image contrast.
Dynamic threshold	Dundar <i>et al.</i> (2011) [36]	Breast	Eliminate low-brightness pixels.

Logombal *et al.* [4] used median filter to remove the noise generated in the process of breast tissue image dyeing, and preserved the edge while removing the noise. Kleczek and Kowal *et al.* [10], [29] applied Gaussian filter to smooth the input RGB image slightly, which could reduce the influence of CCD noise. Öztürk *et al.* [30] aimed at the problems of low contrast, uneven color and noise in histopathological images, converted the image from RGB color space to HSV color space, and used Wiener filter to remove noise from the vertical component to increase the clarity of cells. Sertel *et al.* [31] adopted anisotropic diffusion to smooth relatively uneven regions while preserving important edge information.

In addition, Kamel *et al.* [32] proposed matched filters to flatten the background of regional follicles (inside the follicular area). Compared with the traditional Gaussian flat filter and anisotropic diffusion, the novelty of the matched filter lied in its ability to smooth noise and preserve image contours. Vahadane *et al.* [33] used Gaussian smoothing to remove high-frequency noise, and further used morphological filtering (open operation) for image enhancement. It enhances the segmentation and poor dyeing of different nuclei, and further smooths high-intensity background and low-intensity foreground, while retaining structure and edge information. Wang and Gurcan *et al.* [27], [34] adopted a top-down hat transformation to enhance the contrast between the nucleus and other areas, as shown in Fig. 4. Mohammed *et al.* [35] applied morphological methods to remove the noise generated by the lymphocyte image threshold and Canny edge detection step. Dundar *et al.* [36] adopted dynamic threshold segmentation to eliminate blue-purple pixels with low brightness in histopathological images.

In summary, the above filtering methods are briefly summarized in Table 1. The spatial filtering algorithm is relatively simple, the running time period is shorter, and the sharpening effect is obvious. The frequency domain filtering algorithm is complex, the calculation is slow, there is a ringing effect, and the image effect is displayed relatively smoothly. Spatial filtering is also divided into linear filtering and nonlinear filtering. Linear filtering (Gaussian filtering, Wiener filtering,

**FIGURE 4. Morphological enhancement of histopathological images.**

and matched filtering) not only eliminates the mutation information in an image but also reduces the edge definition of the image. Nonlinear filtering (median filtering and anisotropic diffusion) can highlight more details, but it will take more time.

B. COLOR NORMALIZATION

During the preparation of tissue sections, the appearance of histological staining usually has great variability, which is affected by the ratio of the staining agent, staining platform and imaging platform. There is a great difference in color, which greatly affects the image processing accuracy and quantitative analysis ability. Color normalization is a necessary step to remove unwanted color changes in histopathological images [37], which can reduce staining variability and increase the stability of image processing.

Reinhard *et al.* [38] proposed the histogram normalization method to remove undesirable color bias in the image, such as converting a daylight image into a night scene. Furthermore, Sertel *et al.* [31] adopted histogram normalization method to normalize the color of neuroblastoma pathological images, and balance the color distribution of slides under different staining conditions. Rabinovich *et al.* [39] proposed a non-negative matrix factorization (NMF) method that can decompose tissue samples stained with multiple tissue dyes. Macenko *et al.* [40] applied a singular value decomposition (SVD) method, which improves the quantitative analysis ability of tissue images. However, the results

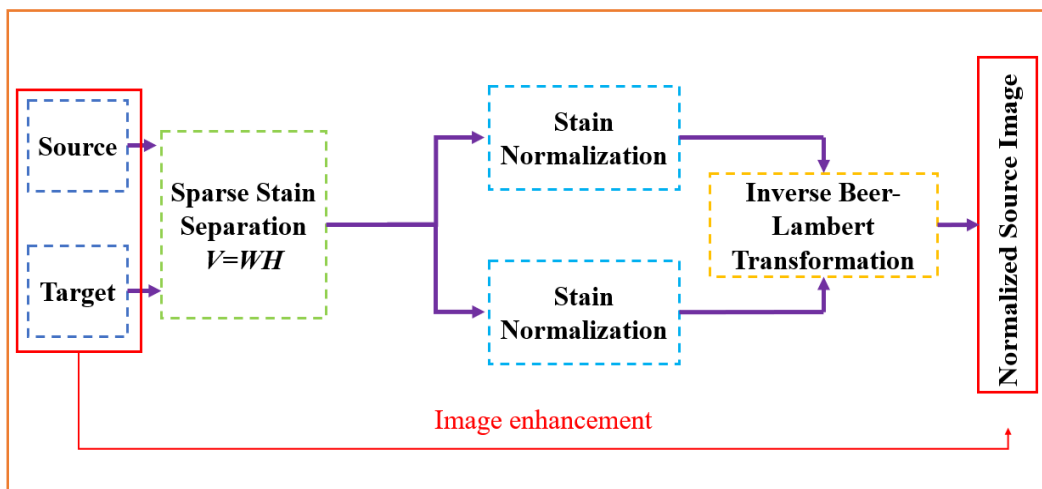


FIGURE 5. Flow chart of structure-preserving color normalization.

were sometimes inconsistent when there were three or more colorants. Khan *et al.* [41] proposed a specific color deconvolution (SCD) method, which reduced image artifacts and improved the stability of tissue image analysis algorithms. Vahadane *et al.* [37] adopted a new structure-preserving color standardization (SPCN) method, as shown in Fig.5. This method not only accurately separated the source image from the target image, but also improved the low-quality tissue. Learn the contrast of the image. Zheng *et al.* [42] applied a new adaptive color deconvolution (ACD) method. This method qualitatively evaluated 500 WSIs, with a normalization failure rate of 0.4%, effectively avoiding structure and color artifacts. It had a good effect under the appearance of various colors. Salvi *et al.* [43] proposed a stain color adaptive normalization (SCAN) method for dyeing color, which could improve the contrast between the tissue and the background without changing the cavity and background color, while preserving the local area.

In summary, the above several color normalization methods are quantitatively compared by evaluating the relative square error (rSE) of a normalized image (global), the nucleus and the matrix (global rSE, nuclear rSE, and matrix rSE). As shown in Fig. 6, the rSE of the stain color adaptive normalization method is lower than those of the other methods, and this method achieves excellent performance in the color normalization process; however, the histogram normalization performance is relatively poor.

C. SUMMARY

Histopathological images are more complex than other images, so the images will also be contrast enhanced while removing noise. Histopathological images rarely use a single denoising or contrast enhancement process. In addition, there are researchers who use threshold segmentation to denoise and obtain good results. In recent years, due to

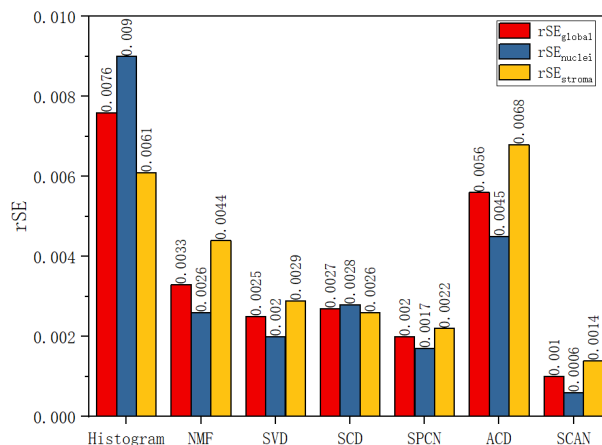


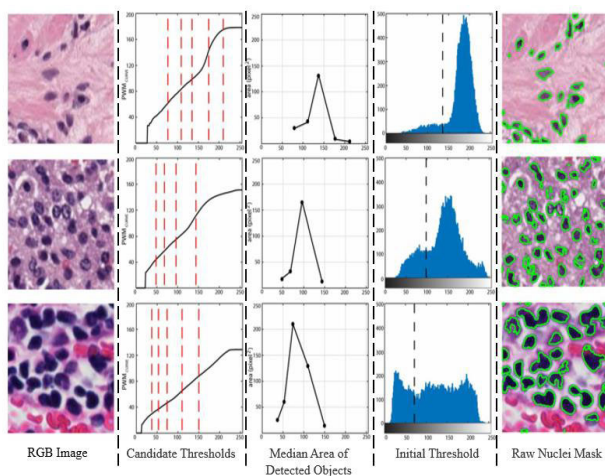
FIGURE 6. Quantitative comparison between seven advanced methods. (rSE_{global} represents the relative square error of the entire image, while rSE_{nuclei} and rSE_{stroma} represent the errors of the nucleus and matrix, respectively).

the flexibility of morphological filtering operation methods, they have been widely used in pathological image enhancement.

Table 2 summarizes the abovementioned color normalization methods. Color normalization is a key part of pathological image preprocessing, and the estimation of dye density maps and color appearance is the core of many color standardization techniques. Although color normalization can reduce color difference changes due to staining and other reasons, during the color normalization process, the tissue structure of an original image may undergo some changes, which affects subsequent image detection. The influence of color normalization on the downstream tasks (segmentation, feature extraction, and classification) of pathological image detection is also worth exploring.

TABLE 2. Color normalization method of histopathological images.

Methods	Authors, year	Advantages	Weaknesses
Histogram	Reinhard et al. (2001) [38], Sertel et al. (2009) [31]	Simple operation, balanced color distribution.	Inability to preserve local structure well.
NMF	Rabinovich et al. (2003) [39]	Color decomposition of tissue samples stained with multiple dyes.	Dyeing leaks and co-localization cause errors.
SVD	Macenko et al. (2009) [40]	Quantitative analysis.	Not suitable for images with more than two dyes, there is noise.
SCD	Khan et al. (2014) [41]	Fewer image artifacts, insensitive to imaging conditions.	Doesn't replicate the less abundant staining well.
SPCN	Vahadane et al. (2016) [37]	Precise dye separation and improve the contrast of low-quality images.	Complicated calculation.
ACD	Zheng et al. (2019) [42]	Effectively avoid structure and color artifacts.	When the color distribution is single, an error will occur.
SCAN	Salvi et al. (2020) [43]	Improve the contrast between the organization and the background while retaining local information.	The input image size is too small, the separation may fail.

**FIGURE 7.** Effectiveness diagram of histopathological image threshold segmentation method [44].

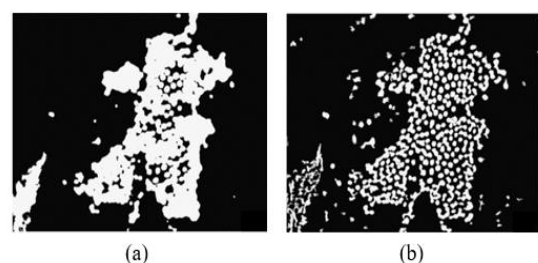
III. IMAGE SEGMENTATION

Image segmentation divides an image into a number of non-interfering areas with different characteristics, extracts the target of interest, and then accurately analyzes the target information. The regions of interest segmented by pathological images include cell bodies, cell nuclei and cytoplasmic contours. Currently, the commonly used segmentation methods of histopathological images include threshold segmentation, ACM segmentation, cluster segmentation, watershed segmentation and neural network.

A. THRESHOLD SEGMENTATION

Threshold segmentation is a method that sets a threshold to separate the target object from the background area according to the difference in the gray value of each position of the image, as shown in Fig. 7. Because threshold segmentation is simple and efficient, it is widely used in image segmentation. At present, the more commonly used methods of histopathological cancer cell segmentation include the global threshold method, adaptive threshold method, local threshold method and others.

Lu and Mandal [26] proposed an effective local threshold method to obtain the complete nuclear area, the main components in the local circular area were preserved. Mouelhi *et al.* [28] extracted all stained nuclei regions and segment overlapping nuclei based on adaptive local thresholds and enhanced morphological methods. Gurcan *et al.* [34] used a combination of morphology and hysteresis threshold to segment cell nuclei, with an average segmentation accuracy of about 90.24%, but some faintly visible cell nuclei could not be detected. Filipczuk *et al.* [45] proposed adaptive threshold segmentation to separate the target from the background, and the result of the separation was a relatively good isolated object. However, this method is difficult to distinguish cells with similar brightness. In addition, the most popular method of global threshold is the Otsu threshold. Vahadane *et al.* [33] applied the Otsu threshold method to avoid artifacts and to improve its segmentation performance for histological images. Kowal *et al.* [29] compared the otsuzu threshold and the adaptive threshold, as shown in Fig.8, the Otsu threshold method cannot be used to detect the fine bright boundaries between cluster nuclei, and this method often produced very large pseudo nuclei composed of several actual nuclei.

**FIGURE 8.** Comparison of threshold segmentation methods [29]. (a)Otsu threshold segmentation method. (b) Adaptive threshold segmentation method.

In summary, threshold segmentation is suitable for the case where there is a strong contrast between the gray level of the target and the background. When the staining of histopathological images is uneven, methods are prone to

oversegmentation and undersegmentation. Therefore, many researchers combine threshold segmentation with other segmentation methods for pathological image segmentation to avoid oversegmentation and undersegmentation problems.

B. ACTIVE CONTOUR MODEL (ACM) SEGMENTATION

The ACM is an object definition framework used to find object boundaries. The ACM finds the region of interest by defining a smooth curve around an image. The model includes GVF, the snake model, geodesic contour model and others.

Malek *et al.* [46] used gradient vector flow (GVF) snake method to segment breast cancer cytological images, and the accuracy of shape description obtained by this method in grayscale images was better than that obtained by traditional Snake method. Yang *et al.* [47] adopted the color ACM to divide the extracellular boundary, which had high computational efficiency. Geodesic active contour is a very popular image segmentation tool, but it is often very sensitive to model initialization. When a color image is converted to a grayscale image, broken edges and weak borders are prone to appear. To solve this problem, Xu *et al.* [48] proposed an ACM based on color gradient to quickly and accurately segment multiple targets on a super large image, which had a more obvious boundary than the traditional gray gradient-based function.

In summary, the ACM has the advantages of simple calculation and the ability to handle topological changes in a curve during the deformation process. However, it lacks robustness in handling acute angles, topological changes and initialization problems. The special advantage of the GVF model is that it is not sensitive to initialization and can move into the boundary pits. The traditional ACM is based on a gray edge gradient, which easily leads to edge fractures and weak boundaries. The color gradient ACM can find the correct edges of cells in a larger range and overcome the problems of edge fractures and weak edges.

C. CLUSTERING SEGMENTATION

Clustering segmentation segments the pixels in the image space into different classes or clusters according to a certain standard and then obtains the segmentation results for an original image. Commonly used clustering algorithms are the K-means clustering algorithm, mean shift clustering algorithm and fuzzy c-means (FCM) clustering algorithm. Clustering algorithms are widely used in histopathological segmentation.

Dimitropoulos *et al.* [49] adopted the K-means method to separate the nucleus from the cytoplasm, but the algorithm relied on prior knowledge of the number of clusters. Sertel *et al.* [50] proposed the mean shift algorithm, which does not require prior knowledge of the number of clusters and does not define the shape of the clusters. Filipczuk *et al.* [51] proposed an improved method of segmentation based on FCM, which could solve the problem of poor representation of nuclei. Even if there are only a

few nuclei in the picture, the correct segmentation can be achieved. Aswathy *et al.* [16] compared three segmentation algorithms: FCM algorithm, K-means clustering algorithm and active contour model, and analyzed their performance to determine effective metrics for breast cancer detection applications. As shown in Fig.9, the K-means clustering algorithm performed better than other methods in segmenting cell nuclei.

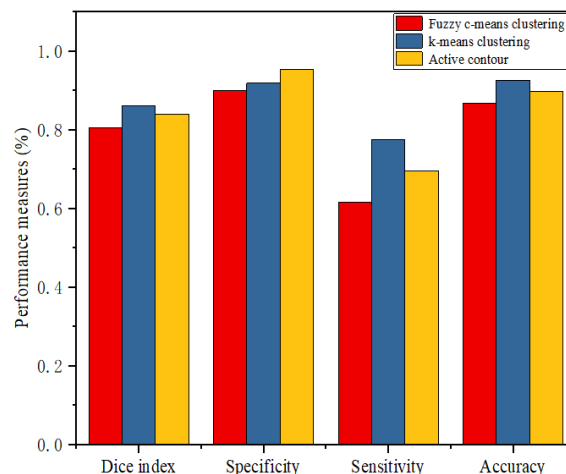


FIGURE 9. Performance comparison of various clustering segmentation algorithms.

In summary, K-means clustering is the most popular clustering algorithm. The algorithm is fast and simple, but the number of clusters needs to be specified, and the selection of K is usually not easy to determine. Compared with K-means clustering, the mean shift method does not need to select the number of clusters, which is a huge advantage. However, the disadvantage of the mean shift method is that the selection of the window size is very important. Compared with K-means hard clustering, FCM provides more flexible clustering results.

D. WATERSHED SEGMENTATION

The watershed segmentation algorithm is a method of connecting the points with similar gray levels around the pixels in the image space into an end-to-end contour. Watershed segmentation is a simple and computationally efficient segmentation technique, and it is a good choice for multiobjective segmentation (such as nuclei in histological images).

Dundar *et al.* [36] adopted watershed segmentation algorithm to separate cell regions of breast cancer histopathological images to identify single cells. Aymen *et al.* [28] separated clustered or overlapping cores using an advanced segmentation method based on watershed algorithm. Because the watershed algorithm is prone to over-segmentation in segmenting histopathological images. Mohammed *et al.* [35] proposed a method combining watershed algorithm and Otsu optimal threshold. This method reduced the over segmentation and under segmentation problems by inhibiting the

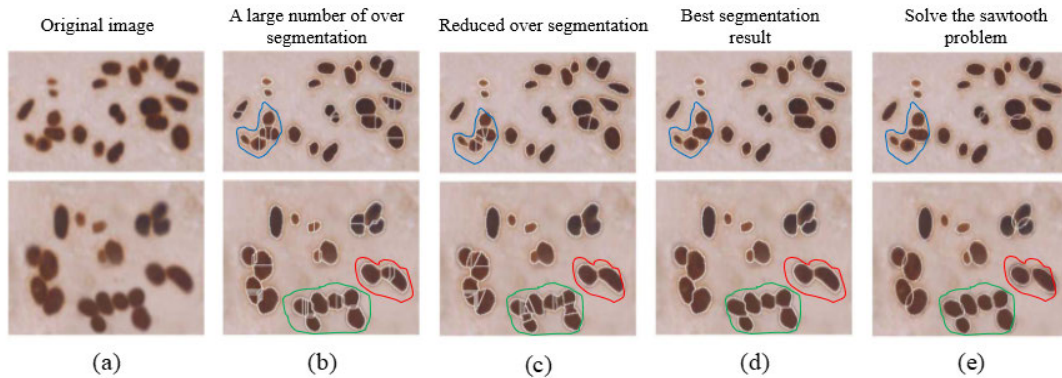


FIGURE 10. Watershed segmentation method [54]. (a) Original image. (b) Classic watershed segmentation method. (c) Shape marking method. (d) Segmentation method based on H-minima transform. (e) Contour adjustment using parameterized method.

local minimum of 1%, and obtained the highest accuracy of 99.92% through the segmentation of lymphocyte nuclei. But it is necessary to use a controllable watershed marker to prevent the loss of lymphocytes. In addition, Schmitt *et al.* [52] found that the iterative voting method based on grayscale was more effective than the watershed method in cell clustering separation. Cheng *et al.* [53] proposed a method for separating aggregated nuclei from fluorescence microscope cell images, which used shape markers and labeling functions in the watershed algorithm, and improved the segmentation accuracy by 6%-7% compared with previous methods. Jung *et al.* [54] proposed a watershed segmentation method based on H-minima transform. This method obtains better results by segmenting the cervical and breast cell images. Figure 10 demonstrated the segmentation results of various watershed methods. The H-Minima transform method was obviously superior to the classical watershed method and the shape labeling method. The H-minima transform method could effectively solve the sawtooth problem after contour adjustment by parameterization method.

In summary, the watershed algorithm is a very effective image segmentation algorithm. The watershed algorithm has a good effect on fuzzy edge segmentation and can obtain complete and continuous edge contours. However, it is very sensitive to some interferences (noise and subtle gray differences) in images, and it easily produces oversegmentation under these interferences. Researchers have proposed two methods to remove the oversegmentation phenomenon. (1) Unnecessary edge information is removed by observational processing. (2) The improved gradient function allows the watershed algorithm to segment only the desired targets.

E. NEURAL NETWORK

With the development of deep learning technology, neural network methods have been applied to pathological image segmentation. Wang *et al.* [55] applied convolutional neural networks (CNNs) to detect mitosis in breast histopathological images. This technology requires very little computing resources to quickly obtain accurate results and can use

a multilayer CNN model to further improve the accuracy. Öztürk *et al.* [56] used a DCNN to semantically segment various cell types in histopathological images. Experiments showed that the training error of this method is 9.2%. Cui *et al.* [57] adopted a FCNN to segment cell nuclei and their boundaries. This method can be applied to other biomedical image segmentation tasks with certain generalizability. Chen *et al.* [58] proposed a novel deep contour-aware network learning framework. This method achieved more accurate detection and segmentation. Alom *et al.* [59] used R2U-Net to segment cell nuclei in public data sets. Experiments showed that the method had accurate segmentation and good robustness, and the segmentation accuracy reached 92.15%. Later, researchers combined the Unet network with other methods and proposed some new methods. Salvi *et al.* [60] introduced a gland segmentation method for prostate histopathological images mixed with deep networks (CNN, ResNet34, Unet) and traditional techniques. This method obtained a dice score of 90.16%. Kang *et al.* [61] proposed a nuclear segmentation method based on the Unets network and Deep Layer Aggregation. Experimental results showed that this method had good generalizability. Kucharski *et al.* [62] investigated a melanocyte segmentation method based on a convolutional autoencoder, which realized the feasibility of segmenting a small ground truth database with a segmentation sensitivity of 0.76 and a specificity of 0.94.

F. SUMMARY

Table 3 shows the comparison of the comprehensive performance of histopathological cancer cell segmentation algorithms. In recent years, a large number of image segmentation methods have appeared, but the segmentation accuracy of these methods is unsatisfactory and needs to be improved. Threshold segmentation is the most widely used image segmentation algorithm, and the key of this method is the selection of the threshold. If some optimal solution algorithms are applied to the selection of the threshold, it is possible to greatly improve the segmentation accuracy. When

TABLE 3. Histopathological image segmentation methods.

Types	Methods	Authors, year	Targets	Comments
Threshold	Local threshold	Lu et al. (2014) [26]	Mitos data set	Obtain a complete nuclear area.
	Adaptive local threshold, Morphology	Mouelhi et al. (2019) [28]	Breast	Segment overlapping nuclei.
	Adaptive threshold, K-means clustering	Filipcuk et al. (2012) [45]	Breast, Fibroadenoma	Cells with similar brightness can be distinguished.
	Otsu threshold	Vahadane et al. (2014) [33]	Breast, Intestine, Prostate	Simple but sensitive to noise.
	GVF-Snake	Malek et al. (2007) [46]	Breast	It is very effective in tracking the angle accuracy of cells.
ACM	Color ACM	Yang et al. (2008) [47]	Lymph	It has high computational efficiency.
	A high-throughput ACM	Xu et al. (2011) [48]	Prostate	Fast and accurate segmentation of multiple objects in an image.
	K-means	Dimitropoulos et al. (2017) [49]	Lymph	Prior knowledge dependent on the number of clusters.
Clustering	Mean shift	Sertel et al. (2010) [50]	Neuroblastoma	No prior knowledge is needed and the shape of clustering is not constrained.
	FCM	Filipcuk et al. (2011) [51]	Breast	Can better represent the nucleus.
Watershed	Watershed	Aymen et al. (2018) [28]	Breast	Separate aggregated or overlapping nuclei.
	Watershed, Otsu threshold	Mohammed et al. (2013) [35]	Lymph	The impact of over-segmentation and under-segmentation is significantly reduced.
	Improved watershed	Cheng et al. (2009) [53]	Mouse neuronal	Improved segmentation accuracy.
	DCNN	Öztürk et al. (2019) [56]	Breast	The blurring problem of patch edge and cell depth has been very successful.
Neural network	FCN	Cui et al. (2019) [57]	Breast, Liver, etc	The seamless prediction of nuclei in large full slide images is realized.
	Deep Contour-Aware Networks	Chen et al. (2017) [58]	Colorectal, Glioblastoma	This method can output accurate probability maps of histological objects and draw a clear outline to separate clustering object instances.
	R2U-Net	Alom et al. (2018) [59]	The dataset	This method has good qualitative performance.
	Unets, Deep Layer Aggregation	Kang et al. (2019) [61]	Bladder, Colorectal, etc	This method can be easily popularized in different cell types of different organs.
	Convolutional Autoencoders	Kucharski et al. (2020) [62]	Melanocytes	This method can be used for histopathological image analysis of small ground truth database.

the segmentation algorithm is applied to practical production, there are still great difficulties in accuracy, real time and operability. In the future, research on high real-time segmentation algorithms will be the focus of efforts. To date, there is no general image segmentation method, which is still the direction of future research work.

With the rapid development of deep learning, many researchers have used neural network algorithms to segment histopathological images. Compared with traditional image segmentation methods (threshold segmentation, Active contour model segmentation, Watershed segmentation, etc.), the neural network segmentation method is faster and more accurate and is also suitable for complex histopathological images. Due to the large amount of computation and high time consumption of deep learning algorithms, their effects need to be further verified.

IV. IMAGE FEATURE EXTRACTION

Feature extraction is a method to reduce the dimensionality of image information. Feature extraction can also be understood

as transforming useful data symbols or information in images into nonimage representations or descriptions, such as values, vectors and, so that a computer can understand the images. Since an image is usually large, the feature subset that needs to be calculated is small, and feature extraction reduces the computational complexity of the classification algorithm [12]. At present, pathological image feature extraction mainly includes shape features, color features, texture features and others.

A. SHAPE FEATURES

Shape features are the most direct way of expressing information about the appearance of an image. Shape features are unrelated to the pixels of images, so they are not sensitive to changes in the grayscale and brightness of images. Shape features include contour features and regional features. The common shape features of histopathological images are regional features, which mainly include geometric features, moment features and topological features.

Petushi *et al.* [63] regarded the average distance between the nearest nuclear centroids in a high-density region as a possible classification feature. Elsalamony *et al.* [64] adopted geometric features such as the area, convex area, perimeter, eccentricity, solidity, and ratio as input variables for classification. Lu *et al.* [65] proposed a method combining regional ellipticity and local pattern features to distinguish melanocytes and candidate nuclear regions, which provided robust parameters for identifying melanocytes. The experimental results on skin histopathological images showed that the sensitivity of this method exceeded 80%. In addition, there are many feature combination methods. Tashk *et al.* [66] proposed an automatic mitotic detection algorithm based on the combination of texture features and stiffness matrix features, which contributed to obtaining more reliable classification.

In histopathological images, because the arrangement and shape of cancer cells and normal cells change greatly, the topological features are not affected by the geometric distortion of the image, and they are a global feature that does not depend on distance changes [67]. Therefore, researchers use graphic technology to describe topological features with the help of changes in cell structure. As shown in Fig.11(a), voronoi tessellation was based on the distance to the points in a specific subset of the plane, and the image was divided into different planes and regions. As shown in Fig.11(b), delaunay triangulation was obtained by connecting point pairs on the plane, so that the triangle formed by connecting three non collinear points with an edge, that is to say, it was surrounded in an outer circle, in which there were no other points. In Fig.11(c), the minimum spanning tree was a connected tree of undirected graphs, which connects all vertices with the edge with the smallest weight.

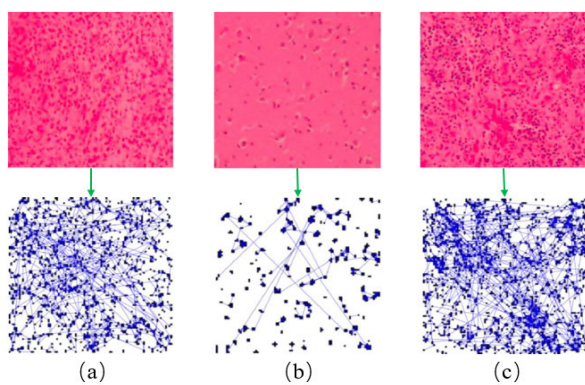


FIGURE 11. Spatial arrangement features of samples extracted from histopathological images [12]. (a) Voronoi mosaic. (b) Delaunay triangulation. (c) Minimum spanning tree.

In summary, the accuracy of shape feature extraction is affected by the segmentation effect. For images with poor segmentation results, the shape parameters cannot even be extracted. When extracting the shapes of irregular objects, the extraction results will be inaccurate. Cancer cells have

irregular shapes, so there are relatively few studies on shape extraction corresponding to cancer cells.

B. COLOR FEATURES

Color features are a basic visual feature and global feature for humans to perceive and distinguish different objects. Color characteristics are very sensitive to the gray information and brightness changes of images. Since pathological images need to be stained during the acquisition process, color normalization is necessary for color feature extraction. The color models mainly include HSV, RGB, HSI and others.

To obtain color information, Rashmi *et al.* [68] dyed cell nuclei into dark purple through H&E, while other tissues were dyed light pink, and then extracted the color features of each pixel in RGB and LAB color spaces. Xu *et al.* [48] transformed RGB color image into HSV color space, eliminated hue and saturation channel, and produced gray image with scalar brightness. The color gradient representation (Fig.12(b)) has a more obvious boundary than the corresponding gray gradient (Fig.12(a)). Bejnordi *et al.* [69] standardized the color distribution of each tissue component by aligning the 2D histogram of the color distribution in the hue-saturation-density model. Kong *et al.* [70] used local fourier transform to extract the color texture at each pixel from the most discriminant color space. This method made the local fourier transform texture features extracted from most discriminant color space achieved the maximum discrimination effect in classification performance. Sertel *et al.* [71] introduced a new method of combining color and texture, which combined model-based intermediate representation with low-level texture features to capture pixel-level tissue features. The combined feature space improved the accuracy of classification.

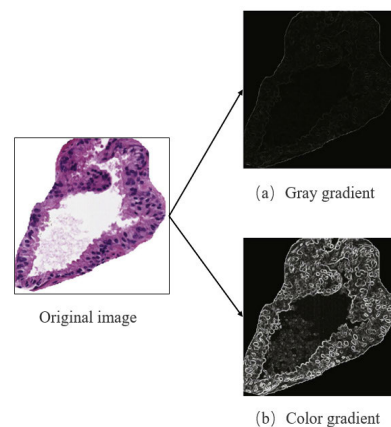


FIGURE 12. Comparison of gray gradient and color gradient of histopathological image [48].

In summary, the diversity of color models provides many choices for the extraction of color features. Histopathological images are usually stained to cause obvious color changes in cancer cells, so the color characteristics of cancer cells are quite important characteristic information. Color images have

more complete image information than grayscale images, but color image processing is complicated and time-consuming, and grayscale images are generally processed.

C. TEXTURE FEATURES

Texture, which is independent of color and brightness, is an important feature to express images. Texture features provide important information on image homogeneity and the organization and arrangement of the surface structure, as well as their connection with the surrounding environment. Texture features are insensitive to image rotation changes and noise and are the most widely used feature extraction method.

In texture feature extraction, gray-level co-occurrence matrix (GLCM) method is widely used. The GLCM method is based on the statistics of two pixels with a certain gray level on the image. The GLCM method is obtained by statistically calculating the gray level status of two pixels kept at a certain distance on the image. As shown in Fig.13, GLCM (0,0) has one group in the original diagram, so $p(0,0) = 1$, while GLCM (1,2) has three groups in the original diagram, so $p(1,2) = 3$. GLCM reflects the spatial information of the image gray level, so it is used to process the texture information of the image primitives and arrangement structure. Albayrak et al. [24] applied Haralick feature descriptors to extract features of spatial dependence and texture relations. Sertel et al. [72] proposed a new method for color texture analysis that used nonlinear color quantization of self-organizing feature maps to modify the GLCM method. The experimental results on the image of follicular lymphoma showed that the new color texture analysis was better than gray texture analysis. Bruno et al. [73] proposed a feature selection method based on curvelet transform, local binary pattern (LBP), and statistical analysis to obtain higher area under the ROC curve and metrics accuracy values: the obtained rates were among 91% and 100%. In addition, Irshad et al. [74] adopted scale invariant feature transform to found strong mitosis detection method. Simsek et al. [75] introduced a set of advanced texture features to represent the prior knowledge of spatial organization, and this method was not sensitive to noise. Öztürk et al. [30] used different feature extraction methods to detect cells in histological images. The

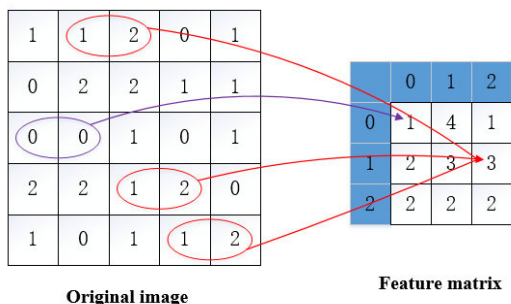


FIGURE 13. Gray-level co-occurrence matrix.

maximally stable extremal region (MSER) algorithm has got better results than other algorithms, and can detect many cells and mark the surroundings. The performance of common feature extraction algorithms is shown in Fig.14.

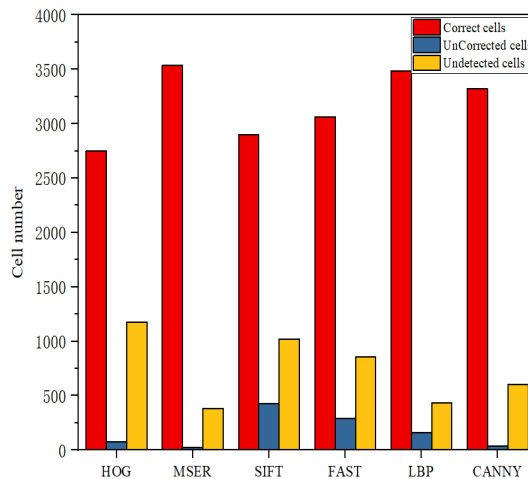


FIGURE 14. Comparison of six commonly used feature extraction algorithms.

In summary, the texture features extracted by the GLCM have good discrimination ability, but the calculation of the GLCM is time-consuming. Fortunately, researchers continue to improve it. The LBP algorithm is simple to calculate and can be used for real-time detection. Some researchers combine the local information expressed by the LBP with other information or algorithms to form a joint feature quantity to obtain a better extraction effect.

D. SUMMARY

Feature extraction is an important step in image processing. Shape features, which are a type of stable information, do not change with the surrounding environment. However, many shape features only describe the local properties of a target while the comprehensive description of a target requires high computational time and storage capacity. Generally, the processing and analysis of colorful images are complex and time-consuming for color features, so color images are usually first converted to grayscale images and then processed using grayscale image processing methods. Texture features are a common method for the feature extraction of histopathological images. This method has strong resistance to noise, but it is sensitive to illumination and reflection. Among the texture feature extraction methods, the GLCM has strong adaptability and robustness, but it lacks global image information. In addition, we discuss the feature extraction methods in Table 4. Feature extraction is very important for the accuracy of cancer cell recognition. The features of the histopathological images obtained by different feature extraction methods will lead to different success rates, so the selection of appropriate features is the key factor to improve the recognition accuracy.

TABLE 4. Feature extraction method of histopathological image.

Type	Method	Authors, year	Comments
Shape features	Geometric features	Elsalamony <i>et al.</i> (2016) [64]	Detect irregular shaped blood cells.
	Double ellipse feature	Lu <i>et al.</i> (2013) [65]	Provides robust parameters.
	Stiffness moment characteristics	Tashk <i>et al.</i> (2015) [66]	Improve classification accuracy.
	RGB and LAB color spaces	Rashmi <i>et al.</i> (2020) [68]	Get color information.
Color features	Color gradient	Xu <i>et al.</i> (2011) [48]	Color gradients have more obvious boundaries than grayscale gradients.
	Color two-dimensional histogram	Bejnordi <i>et al.</i> (2014) [69]	Color standardization.
	Most discriminant color space	Kong <i>et al.</i> (2011) [70]	Improve the accuracy of the system.
	Model-based intermediate representation and low-level textures	Sertel <i>et al.</i> (2009) [71]	Maximize the classification performance.
	GLCM	Albayrak <i>et al.</i> (2013) [24]	Improve cell separation accuracy.
Texture features	Improved GLCM	Sertel <i>et al.</i> (2008) [72]	Better effect on image distortion.
	LBP	Bruno <i>et al.</i> (2016) [73]	Achieved higher area under the ROC curve and metrics accuracy values (the obtained rates were among 91% and 100%).
	Scale-invariant feature transform	Irshad <i>et al.</i> (2013) [74]	Good matches can be filtered out.

V. CLASSIFICATION

Classification realized by classifiers, is the process of constructing a function or designing a model to map feature extracted data to specific categories. The function or model maps the characteristic data to a given category to perform cell identification. Classifiers can be divided into supervised classifiers and unsupervised classifiers. Supervised classifiers include decision trees (DTs), support vector machines (SVMs), the K-nearest neighbors (KNN), sparse representation (SR) and neural network algorithms. Unsupervised classification means that the classification process does not require any prior knowledge, and the similarity is directly searched from the image. Therefore, unsupervised classification is also called clustering. The commonly used unsupervised classification methods in pathology detection include generative adversarial networks (GANs), domain adaptation and principal component analysis (PCA).

A. SUPERVISED CLASSIFIER

1) DTS

DTs, which are a basic classification and regression method, are a type of tree structure classification algorithm. This approach is easy to understand and implement, but when there are too many categories, errors can multiply quickly. Korkmaz *et al.* [76] applied a DT to classify histopathological gastric images, and the highest accuracy result was 86.66%. Rahman *et al.* [77] used a DT classifier to classify oral squamous cells with an accuracy of 99.78%. Tyr *et al.* [78] extracted shape, texture, and color features from histopathological images and achieved 99.4% accuracy using a DT classifier. The random forest is an ensemble classifier composed of many DTs. It has a fast training

speed and is adjustable, and it does not need to adjust many parameters, such as with the SVM. Irshad *et al.* [74] proposed a random forest to select texture features and achieved classification with a higher positive predictive value and F-measure.

2) SVM

The SVM is a binary classification model proposed by Vapnik in 1995 according to statistical learning theory. It maps the input space where the sample points are located to the high-dimensional feature space to achieve the maximum linear separation. The SVM has a good generalization ability and is suitable for small samples, nonlinearity and other problems. The SVM algorithm generally shows good results for highly sparse features [79].

Aswathy *et al.* [16] adopted SVM to classify benign and malignant breast cancer histological images, with an accuracy rate of 91.1%. Masood *et al.* [80] proposed an efficient self-advised SVM to identify skin cancer, with an average diagnosis accuracy rate of 89.1%. Al-Kadi *et al.* [81] used SVM, Bayes and KNN to test the performance of texture fractal features, and the classification accuracy was 94.12%, 92.50% and 79.70%, respectively. SVM has a high accuracy rate. Kuse *et al.* [82] constructed a training data set consisted of 80 lymphocyte patterns and 98 non-lymphoid cell patterns, and classified the data set using SVM, with a classification accuracy of 78%. Yan *et al.* [83] adopted multi-label SVM to achieve multi-classification, and the multi-label method had a strong classification ability in the recognition of multi-label colon pathology images. Krishnan *et al.* [84] defined a set of compact 18 features on the oral tissue image, using SVM classifier, the accuracy rate reached 99.66%.

3) KNN

The KNN is one of the simplest classification algorithms. This method has a simple idea and no need to estimate parameters. It is more suitable for the automatic classification of class domains with large sample sizes. Niwas *et al.* [85] extracted the features of breast tissue images using complex wavelets and used KNN for classification. The classification accuracy reached 93.9%. The disadvantage of this method is that it requires a large amount of calculation because it needs to calculate the distance of known sample points of each classified text in order to obtain the k nearest points. At present, the solution commonly used by researchers is to pre-cut the known sample points to remove samples that are useless for classification.

4) SR

The SR optimization model, which has good performance in pattern recognition, is established from the point of view of signal reconstruction. The best current classification systems often choose SR as their key module. The two main tasks of image SR are dictionary generation and signal sparse decomposition. Han *et al.* [86] proposed an automatic histological classification method based on dictionary learning and sparse coding. The classification accuracy reached 90.05%. Srinivas *et al.* [87] adopted a new sparsity model for multichannel histopathological image representation and classification. Shirale *et al.* [88] applied a class-level dictionary learning method to reduce the workload of pathologists, which achieved good performance on various histopathological image datasets. The highest accuracy of the algorithm reached 96.56%. Li *et al.* [89] proposed a multichannel joint sparse model based on mutual information, which improved the discrimination ability of joint SR coefficients.

5) NEURAL NETWORK

A neural network is a mathematical information processing model similar to the structure of the brain's synaptic connection. A neural network system is more complicated, and it achieves the role of information classification by changing the interconnection relationships between internal nodes. CNNs are often used in the classification of histopathological images.

Zheng *et al.* [90] proposed a CNN to improve the classification of breast cancer. This method effectively reduces the noise and redundancy caused by stroma and achieves the best classification performance. The average classification accuracy of the algorithm was 96.4%. Wahab *et al.* [91] used a two-stage deep convolutional neural network with a simple structure to focus on nonmitotic divisions that are difficult to classify. Saito *et al.* [92] constructed a CNN to automatically detect prominent lesions in WCE small intestine images. The trained CNN was proven to be able to detect prominent lesions in independent test images with a sensitivity of 90.7% and a detection rate of 98.6% for prominent lesions.

In addition, two-dimensional (commonly used in histopathological images) or three-dimensional (commonly used in MRI and CT) convolutional neural networks are commonly used in classification, and one-dimensional CNN is rarely used in classification. Moitra *et al.* [93] adopted one-dimensional convolutional neural network for classification of non-small cell lung cancer, which consumed less time and resources. Panigrahi *et al.* [94] proposed to use capsule network for oral cancer classification. Compared with CNN model, capsule network has better ability in capturing posture information and spatial relationship, and can better distinguish cancerous and non-cancerous images with 97.35% accuracy. Kutlu *et al.* [95] proposed an automatic leukocyte detection method based on deep learning and transfer learning, which was very successful in detecting overlapped or partially visible cells in images. The highest accuracy of the algorithm reached 99.52%.

In summary, there are many types of supervised classification methods, and they all have their own classification characteristics, as shown in Fig. 15. DT has a strong ability to handle large sample data, and the calculation speed is fast; however, the stability of the classification accuracy is not good. The SVM has a higher classification accuracy and better generalization ability than other algorithms on small samples and linear data. However, the prediction rate of the SVM is slightly lower than that of other classifiers, and it is very sensitive to the lack of data. When the value of k in KNN is very small, it is very sensitive to noise; and when the number of samples is unbalanced, the effect will be poor. The SR model is simple, suitable for linear features, and easy to understand and operate. However, the model is no longer accurate when there are changes in posture and no alignment in an image. A neural network is the classifier with the highest accuracy in supervised classification, but it requires a large amount of sample data and expensive computing equipment. Compared with other classifiers, the computational costs of neural networks are more expensive. Supervised classification can control the selection of training areas and training samples, but the selection of the classification system and training sample areas is affected by subjective factors. Therefore, supervised classification can only classify the categories defined by the training samples and cannot identify the categories that are not defined by an analyst, and it is easy to omit of categories.

B. UNSUPERVISED CLASSIFIER

1) GAN

A GAN uses continuous game learning between a generative model and a discriminant model to produce a good output. A GAN is weak in processing discrete data. Man *et al.* [96] proposed an unsupervised anomaly detection screening method based on a GAN. This method classified breast cancer histopathological images, and the highest classification accuracy was 99.13%. Xue *et al.* [97] proposed a new cGAN model, HistoGAN, for high-fidelity histopathological image

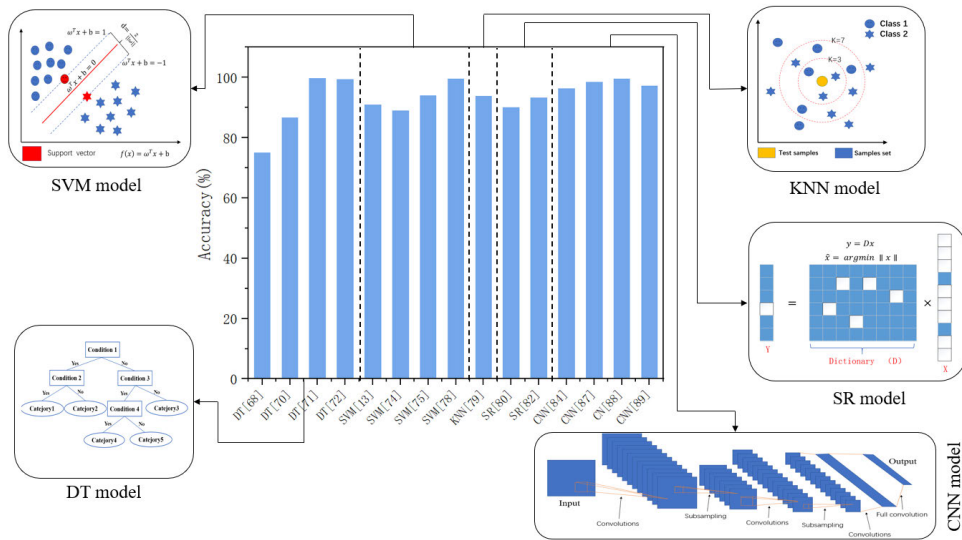


FIGURE 15. Classification accuracy of supervised classifier.

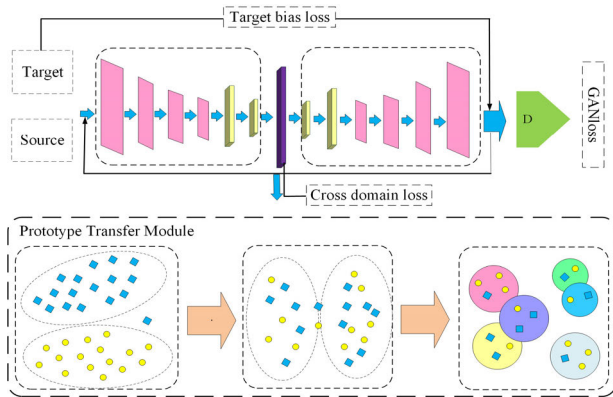


FIGURE 16. Prototype transfer GAN framework diagram.

synthesis. This method could significantly and continuously improve the classification performance of cervical cancer histopathological images and metastatic cancer datasets (by 6.7% and 2.8%, respectively). Wang *et al.* [98] proposed an unsupervised learning method for prototype transfer GANs. As shown in Fig. 16, the accuracy of classifying benign and malignant tissues reached nearly 90%.

2) DOMAIN ADAPTATION

Domain adaptation is a method of transfer learning, which can effectively solve the learning problem of inconsistent probability distribution between training samples and test samples. Because the images of benign tumors are very similar to those of malignant tumors, Pendar *et al.* [99] adopted a new unsupervised domain-adaptive method based on representation learning to overcome these problems. The above achieved an average classification rate of 88.5%, an increase of 5.1% compared with the basic method, and an increase of 1.25% compared with the latest method. Yu *et al.* [100]

proposed a method of introducing unsupervised learning domain adaptation into a typical deep CNN model. As shown in Fig. 17, this method reduced label duplication.

3) PCA

PCA uses the ideas of linear transformation and dimensionality reduction to transform data into a new coordinate system, reduce the dimensionality of a dataset, and make the data intuitively presented in a two-dimensional coordinate system. The PCA method has the characteristics of simple calculation, easy-to-understand results and no parameter limitations. However, it is unable to intervene in the processing process through parameterization methods, and the expected results may not be obtained. Sertel *et al.* [72] used a combination of PCA and linear discriminant analysis and then used a Bayesian classifier to classify lymphoma cells. The overall correct classification rate of this method was 88.9%. Shi *et al.* [101] adopted the quaternion Grassmann averages network algorithm, which performed the best in the classification of color histopathological images. Shi *et al.* [102] proposed a color pattern random binary hashing based PCANet algorithm, which was superior to the original PCANet algorithm and other traditional unsupervised algorithms. Figure 18 is a schematic diagram of the color pattern random binary hashing based PCANet algorithm.

In summary, the unsupervised classifier classification process does not impose any prior knowledge and can directly classify a dataset. An unsupervised classifier can effectively eliminate the subjective factors that people have on the size and shape of cancer cells in the dataset calibration. Furthermore, the detection costs are reduced, and the classification accuracy is improved. The research of unsupervised image classification is still in the development stage, and its research results are less than other directions.

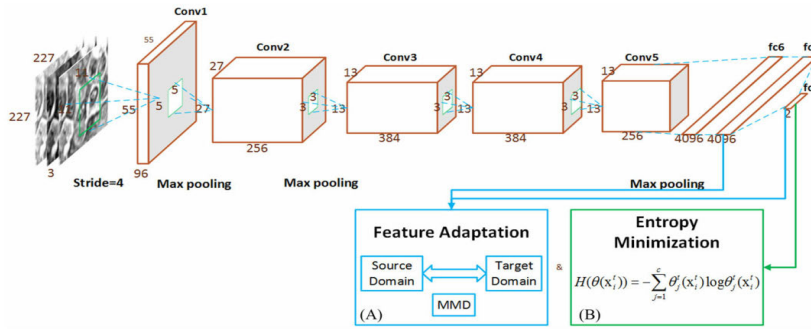


FIGURE 17. Improved CNN model [100]. (a) feature adaptation (b) entropy minimization.

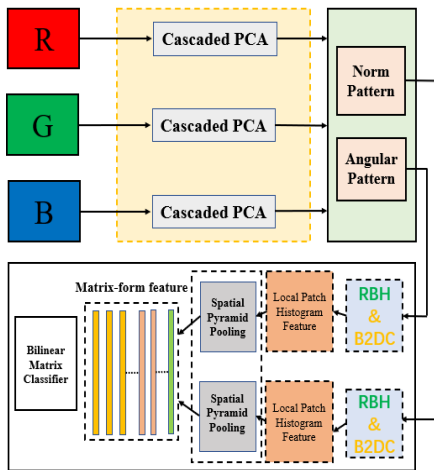


FIGURE 18. The color pattern random binary hashing based PCANet workflow chart and color histopathological image classification framework based on matrix classifier.

C. SUMMARY

Image processing has a crucial impact on the classification speed, stability and accuracy of the model. Image preprocessing can remove useless information or enhance useful information. Image segmentation is to extract regions of interest, further reducing the interference of useless information. Feature extraction can reduce the dimensionality of the data. The data undergoes a series of image processing will significantly improve the classification performance of the model.

At present, the classification accuracy of pathological images is mostly over 85%. As shown in Table 5, this result seems to be very good but unstable. In many studies, the classification results may only be applicable to the research samples, and the classification accuracy may vary greatly for different sample sets, as shown in Fig. 19. The selection of features also has a great impact on the classification accuracy. Korkmaz *et al.* [76] used Fourier transform-mass spectrometry-Fourier transform for feature extraction. When five features were selected, the highest accuracy was 86.66%; and when the number of selected features was increased to 45, the highest accuracy was 68.88%. Therefore, the versatility of classification algorithms and how to choose suitable

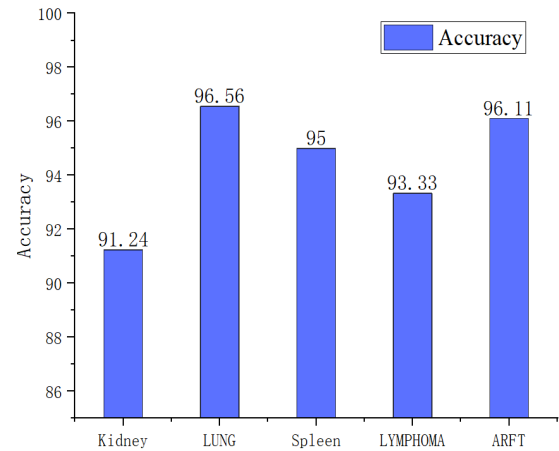


FIGURE 19. Classification accuracy of histopathological image classification method based on SR for different data sets.

features are worth studying. Supervised classifiers are the mainstream image classification methods. Supervised classifiers can obtain the desired classification results according to people’s needs, but they are affected by subjective factors to a certain extent. Unsupervised classification does not require any prior knowledge in the classification process. Analysts only need to set the number of classifications to classify all images and discover some hidden laws. Although supervised classification occupies a dominant position in classification, unsupervised classification can reduce labor and costs and has immeasurable development prospects in the future.

VI. DISCUSSION

Image analysis based on machine vision has become an increasingly important field because of the fast generation of images and the increasing dependence of the medical community on these images. Cancer has become the main cause of death for people. If cancer can be detected in time, the hope of a cure will be greater. Cancer cell detection based on machine vision overcomes the shortcomings of traditional manual detection methods and has become the mainstream cancer cell detection method.

TABLE 5. Classification methods for histopathological images.

Type	Methods	Authors, year	Targets	Accuracy
Supervised classifier	DT	Korkmaz et al. (2018) [76]	Gastric	86.66%
	DT	Rahman et al. (2020) [77]	Oral	99.78%
	SVM	Aswathy et al. (2020) [16]	Breast	91.1%
	SVM	Masood et al. (2015) [80]	Skin	89.1%
	SVM	Al-Kadiet al. (2015) [81]	Brain	94.12%
	SVM	Krishnan et al. (2012) [84]	Oral	99.66%
	KNN	Niwas et al. (2013) [85]	Breast	93.9%
	SR	Han et al. (2011) [86]	Brain	90.05%
	SR	Shirale et al. (2018) [88]	ADL dataset, Lymphoma dataset, ARFT	96.11%
	SR	Li et al. (2020) [89]	ADL dataset, BreaKHis datasets	Average accuracy over 90%
Unsupervised classifier	CNN	Zheng et al. (2017) [90]	Breast	96.4%
	CNN	Saito et al. (2020) [92]	Intestine	98.6%
	Capsule network	Panigrahi et al. (2020) [94]	Oral	97.35%
	GAN	Man et al. (2020) [96]	Breast	99.13%
	GAN	Wang et al. (2021) [98]	Breast	90%
	Domain adaptation	Pendar et al. (2018) [99]	BreaKHis dataset	88.5%
	PCA	Shi et al. (2016) [101]	ADL dataset	81.8%
PCA	Shi et al. (2019) [102]	HCC dataset	94.58%	

There are many types of cancer, including breast cancer, lung cancer, liver cancer, stomach cancer, colon cancer and others. At present, breast cancer is the most researched cancer based on machine vision cancer cell detection. As shown in Tables 1 and 3, there are relatively few other cancers that are researched. At present, the detection of various cancer cells is still in the exploratory stage; and the detection of other cancers should be increased, which is conducive to the establishment of a complete cancer cell detection system based on machine vision. However, many researchers' methods can only detect fixed types of cancer cells. Even if these methods can detect other types of cancer cells, the detection accuracies are relatively low and unstable. These problems must be resolved through further research.

The machine vision cancer detection process includes image preprocessing, target region segmentation, feature extraction and selection, and cell recognition and classification. A large number of algorithms appear in each processing flow, and these algorithms have their own advantages, disadvantages and scope of adaptation. How to improve the accuracy, execution efficiency, real-time performance and robustness of algorithms has always been the focus of researchers. In addition, with the rapid development of deep learning, neural networks have appeared in various visual detection processes. neural networks have the advantages of high speed and high precision and is deeply loved by researchers; however, neural networks have some problems, such as a large amount of calculation and high time consumption. The effects of neural networks needs to be further verified. The deep learning model is essentially a black box and does not provide the interpretability of the decision-making process, which in turn makes debugging difficult when needed. Poor interpretability can lead to distrust of clinicians trained to make interpretable clinical inferences. The future trend may be to create a general interpretable structure for medical image computing.

VII. OUTLOOK

At present, machine vision is used in various industries, such as manufacturing, agriculture, medicine, military, aerospace, and scientific research. Machine vision technology has the characteristics of high speed, high precision and multiple functions, which greatly promote the development of society. Machine vision has also developed rapidly in the field of medicine. Computer-aided diagnosis and result prediction models based on machine learning are helpful for clinical decision-making. In many cases, machine learning models can provide excellent accuracy in cancer cell detection, but there are still some problems in practical application. Therefore, cancer cell detection methods based on machine vision still need continuous improvement, and future research should focus on the following aspects.

- 1) In terms of image preprocessing, affected by external factors such as light, signal transmission, and acquisition environment, it is difficult to distinguish some weak signals from noise, which affects the subsequent detection accuracy. How to build a stable and reliable detection system to overcome the interference of the external environment is one of the problems to be solved. Although color normalization can reduce the influence of color differences in subsequent images, some changes in the organizational structure may also affect the subsequent image detection, which also needs to be explored. At present, the color normalization algorithm can be integrated into the deep learning framework to improve the performance for segmentation and classification [103]–[106]. But all of them use antagonistic neural network to normalize the color. In the future, deep learning will be a problem worth studying in color normalization.
- 2) For histopathological image segmentation, traditional segmentation methods will have various problems, such as undersegmentation and oversegmentation.

Most of the methods used for tissue segmentation are based on the modification and adjustment of existing image processing techniques to adapt to new applications, which may not be the best choice for tissue image processing needs. With the development of artificial intelligence and deep learning, neural networks have the characteristics of fast speed and accuracy in image segmentation. In the future, neural networks will become the focus of image segmentation research.

- 3) Feature extraction is a key step in cell detection based on machine vision, which has a significant impact on the subsequent recognition accuracy, computational complexity, and stability. At present, multifeature fusion technology is the mainstream feature extraction method. The deep learning method can learn additional feature libraries that cannot be represented by any handmade features, thereby further improving the classification accuracy. In the future, deep learning needs further research in feature extraction.
- 4) In terms of classification, traditional classifiers have the problems of long calculation times, poor accuracy, and instability. A classifier based on a neural network is obviously better than a traditional classifier. With the improvement of histopathological databases, neural networks will become the mainstream classification methods. In recent years, many neural network algorithms have appeared, including SPPNet, Faster RCNN, Mask RCNN, YOLO series and other target detection methods, which provide the possibility for the detection of cancer cells. Because unsupervised classification does not need prior knowledge, the computational costs are reduced. In the future, unsupervised classifiers will become the main research direction of classification.
- 5) Although a series of excellent algorithms continue to appear, there are relatively few open source codes (algorithm implementation) for histopathological image processing, and the algorithm verification lacks standards and ground truth as a reference. Future work should address the issue of openness in the field. This requires the close cooperation of experts in various fields, such as computer scientists, clinicians and pathologists.

REFERENCES

- [1] P. Vineis and C. P. Wild, "Global cancer patterns: Causes and prevention," *Lancet*, vol. 383, no. 9916, pp. 549–557, Feb. 2014.
- [2] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA, Cancer J. Clin.*, vol. 68, no. 6, pp. 394–424, 2018.
- [3] W. Cao, H. Chen, Y. Yu, N. Li, and W. Chen, "Changing profiles of cancer burden worldwide and in China: A secondary analysis of the global cancer statistics 2020," *Chin. Med. J.*, vol. 134, no. 7, pp. 783–791, 2021.
- [4] G. Logambal and V. Saravanan, "Cancer diagnosis using automatic mitotic cell detection and segmentation in histopathological images," in *Proc. Global Conf. Commun. Technol. (GCCT)*, Apr. 2015, pp. 128–132.
- [5] K. M. Fong, Y. Sekido, and J. D. Minna, "Molecular pathogenesis of lung cancer," *J. Thoracic Cardiovascular Surg.*, vol. 118, no. 6, pp. 1136–1152, 1999.
- [6] M. Iranifam, "Analytical applications of chemiluminescence methods for cancer detection and therapy," *TrAC Trends Anal. Chem.*, vol. 59, pp. 156–183, Jul. 2014.
- [7] Y. Xue and N. Ray, "Cell detection in microscopy images with deep convolutional neural network and compressed sensing," 2017, *arXiv:1708.03307*.
- [8] L. He, L. R. Long, S. Antani, and G. R. Thoma, "Histology image analysis for carcinoma detection and grading," *Comput. Methods Programs Biomed.*, vol. 107, no. 3, pp. 538–556, 2012.
- [9] T. A. A. Tosta, L. A. Neves, and M. Z. D. Nascimento, "Segmentation methods of H&E-stained histological images of lymphoma: A review," *Inform. Med. Unlocked*, vol. 9, pp. 35–43, Jan. 2017.
- [10] P. Kleczek, J. Jaworek-Korjakowska, and M. Gorgon, "A novel method for tissue segmentation in high-resolution H&E-stained histopathological whole-slide images," *Computerized Med. Imag. Graph.*, vol. 79, Jan. 2020, Art. no. 101686.
- [11] R. Bidart and A. Wong, "TriResNet: A deep triple-stream residual network for histopathology grading," in *Proc. Int. Conf. Image Anal. Recognit.*, 2019, pp. 369–382.
- [12] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological image analysis: A review," *IEEE Rev. Biomed. Eng.*, vol. 2, pp. 147–171, 2009.
- [13] W. Ming, F. Shen, X. Li, Z. Zhang, J. Du, Z. Chen, and Y. Cao, "A comprehensive review of defect detection in 3C glass components," *Measurement*, vol. 158, Jul. 2020, Art. no. 107722.
- [14] J. Jose, A. Chacko, and D. A. S. Dhas, "Comparative study of different image denoising filters for mammogram preprocessing," in *Proc. Int. Conf. Inventive Syst. Control (ICISC)*, Jan. 2017, pp. 1–6.
- [15] S. Tewary, I. Arun, R. Ahmed, S. Chatterjee, and C. Chakraborty, "AutoIHC-scoring: A machine learning framework for automated Allred scoring of molecular expression in ER- and PR-stained breast cancer tissue," *J. Microsc.*, vol. 268, no. 2, pp. 172–185, Nov. 2017.
- [16] M. A. Aswathy and M. Jagannath, "Performance analysis of segmentation algorithms for the detection of breast cancer," *Proc. Comput. Sci.*, vol. 167, pp. 666–676, Jan. 2020.
- [17] B. Tang, J. Kong, and S. Wu, "Summary of machine vision surface defect detection," *J. Imag. Graph.*, vol. 22, no. 12, pp. 1640–1663, 2017.
- [18] P. Wang, X. Hu, Y. Li, Q. Liu, and X. Zhu, "Automatic cell nuclei segmentation and classification of breast cancer histopathology images," *Signal Process.*, vol. 122, pp. 1–13, May 2016.
- [19] M. de Bruijne, "Machine learning approaches in medical image analysis: From detection to diagnosis," *Med. Image Anal.*, vol. 33, pp. 94–97, Oct. 2016.
- [20] D. R. Snead, Y. W. Tsang, A. Meskiri, P. K. Kimani, R. Crossman, N. M. Rajpoot, E. Blessing, K. Chen, K. Gopalakrishnan, and P. Matthews, "Validation of digital pathology for primary histopathological diagnosis of routine, inflammatory dermatopathology cases," *Amer. J. Dermatopathol.*, vol. 68, no. 7, pp. 1063–1072, 2016.
- [21] F.-T. Johra and M. M. H. Shuvo, "Detection of breast cancer from histopathology image and classifying benign and malignant state using fuzzy logic," in *Proc. 3rd Int. Conf. Electr. Eng. Inf. Commun. Technol. (ICEEICT)*, Sep. 2016, pp. 1–5.
- [22] Y. Mao, Z. Yin, and J. Schober, "A deep convolutional neural network trained on representative samples for circulating tumor cell detection," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2016, pp. 1–6.
- [23] H. Sharma, N. Zerbe, I. Klempert, O. Hellwich, and P. Hufnagl, "Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology," *Comput. Med. Imag. Graph.*, vol. 61, pp. 2–13, Nov. 2017.
- [24] A. Albayrak and G. Bilgin, "Breast cancer mitosis detection in histopathological images with spatial feature extraction," *Proc. SPIE*, vol. 9067, pp. 103–108, Dec. 2013.
- [25] H. Irshad, "Automated mitosis detection in histopathology using morphological and multi-channel statistics features," *J. Pathol. Informat.*, vol. 4, no. 1, p. 10, 2013.
- [26] C. Lu and M. Mandal, "Toward automatic mitotic cell detection and segmentation in multispectral histopathological images," *IEEE J. Biomed. Health Informat.*, vol. 18, no. 2, pp. 594–605, Mar. 2014.
- [27] P. Wang, L. Wang, Y. Li, Q. Song, S. Lv, and X. Hu, "Automatic cell nuclei segmentation and classification of cervical Pap smear images," *Biomed. Signal Process. Control*, vol. 48, pp. 93–103, Feb. 2019.

- [28] A. Mouelhi, H. Rmili, J. B. Ali, M. Sayadi, R. Doghri, and K. Mrad, "Fast unsupervised nuclear segmentation and classification scheme for automatic allred cancer scoring in immunohistochemical breast tissue images," *Comput. Methods Programs Biomed.*, vol. 165, pp. 37–51, Oct. 2018.
- [29] M. Kowal, P. Filipczuk, A. Obuchowicz, J. Korbicz, and R. Monczak, "Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images," *Comput. Biol. Med.*, vol. 43, no. 10, pp. 1563–1572, 2013.
- [30] Ş. Öztürk and A. Bayram, "Comparison of HOG, MSER, SIFT, FAST, LBP and Canny features for cell detection in histopathological images," *Helix*, vol. 8, no. 3, pp. 3321–3325, Apr. 2018.
- [31] O. Sertel, U. V. Catalyurek, H. Shimada, and M. N. Guican, "Computer-aided prognosis of neuroblastoma: Detection of mitosis and karyorrhexis cells in digitized histological images," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Sep. 2009, pp. 1433–1436.
- [32] K. Belkacem-Boussaid, J. Prescott, G. Lozanski, and M. N. Gurcan, "Segmentation of follicular regions on H&E slides using a matching filter and active contour model," *Proc. SPIE*, vol. 7624, pp. 902–912, Mar. 2010.
- [33] A. Vahadane and A. Sethi, "Towards generalized nuclear segmentation in histological images," in *Proc. 13th IEEE Int. Conf. Bioinf. BioEng.*, Nov. 2013, pp. 1–4.
- [34] M. N. Gurcan, T. Pan, H. Shimada, and J. Saltz, "Image analysis for neuroblastoma classification: Segmentation of cell nuclei," in *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2006, pp. 4844–4847.
- [35] E. A. Mohammed, M. M. A. Mohamed, C. Naugler, and B. H. Far, "Chronic lymphocytic leukemia cell segmentation from microscopic blood images using watershed algorithm and optimal thresholding," in *Proc. 26th IEEE Can. Conf. Electr. Comput. Eng. (CCECE)*, May 2013, pp. 1–5.
- [36] M. M. Dundar, S. Badve, G. Bilgin, V. Raykar, R. Jain, O. Sertel, and M. N. Gurcan, "Computerized classification of intraductal breast lesions using histopathological images," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 7, pp. 1977–1984, Jul. 2011.
- [37] A. Vahadane, T. Peng, A. Sethi, S. Albarqouni, L. Wang, M. Baust, K. Steiger, A. M. Schlitter, I. Esposito, and N. Navab, "Structure-preserving color normalization and sparse stain separation for histological images," *IEEE Trans. Med. Imag.*, vol. 35, no. 8, pp. 1962–1971, Aug. 2016.
- [38] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley, "Color transfer between images," *IEEE Comput. Graph. Appl.*, vol. 21, no. 5, pp. 34–41, Sep./Oct. 2001.
- [39] A. Rabinovich, S. Agarwal, C. Laris, J. Price, and S. Belongie, "Unsupervised color decomposition of histologically stained tissue samples," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 16, 2003, pp. 667–674.
- [40] M. Macenko, M. Niethammer, J. S. Marron, D. Borland, J. T. Woosley, X. Guan, C. Schmitt, and N. E. Thomas, "A method for normalizing histology slides for quantitative analysis," in *Proc. IEEE Int. Symp. Biomed. Imag., Nano Macro*, Jun. 2009, pp. 1107–1110.
- [41] A. M. Khan, N. Rajpoot, D. Treanor, and D. Magee, "A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 6, pp. 1729–1738, Jun. 2014.
- [42] Y. Zheng, Z. Jiang, H. Zhang, F. Xie, J. Shi, and C. Xue, "Adaptive color deconvolution for histological WSI normalization," *Comput. Methods Programs Biomed.*, vol. 170, pp. 107–120, Mar. 2019.
- [43] M. Salvi, N. Michielli, and F. Molinari, "Stain color adaptive normalization (SCAN) algorithm: Separation and standardization of histological stains in digital pathology," *Comput. Methods Programs Biomed.*, vol. 193, Sep. 2020, Art. no. 105506.
- [44] M. Salvi and F. Molinari, "Multi-tissue and multi-scale approach for nuclei segmentation in H&E stained images," *Biomed. Eng. OnLine*, vol. 17, no. 1, pp. 1–13, Dec. 2018.
- [45] P. Filipczuk, M. Kowal, and A. Obuchowicz, "Breast fibroadenoma automatic detection using k-means based hybrid segmentation method," in *Proc. 9th IEEE Int. Symp. Biomed. Imag. (ISBI)*, May 2012, pp. 1623–1626.
- [46] J. Malek, A. Sebri, S. Mabrouk, K. Torki, and R. Tourki, "Automated breast cancer diagnosis based on GVF-snake segmentation, wavelet features extraction and fuzzy classification," *J. Signal Process. Syst.*, vol. 55, nos. 1–3, pp. 49–66, Apr. 2009.
- [47] L. Yang, O. Tuzel, P. Meer, and D. J. Foran, "Automatic image analysis of histopathology specimens using concave vertex graph," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2008, pp. 833–841.
- [48] J. Xu, A. Janowczyk, S. Chandran, and A. Madabhushi, "A high-throughput active contour scheme for segmentation of histopathological imagery," *Med. Image Anal.*, vol. 15, no. 6, pp. 851–862, Dec. 2011.
- [49] K. Dimitropoulos, P. Barmpoutis, T. Koletsis, I. Kostopoulos, and N. Grammalidis, "Automated detection and classification of nuclei in PAX5 and H&E-stained tissue sections of follicular lymphoma," *Signal, Image Video Process.*, vol. 11, no. 1, pp. 145–153, Jan. 2017.
- [50] O. Sertel, U. V. Catalyurek, G. Lozanski, A. Shanaah, and M. N. Gurcan, "An image analysis approach for detecting malignant cells in digitized H&E-stained histology images of follicular lymphoma," in *Proc. 20th Int. Conf. Pattern Recognit.*, Aug. 2010, pp. 273–276.
- [51] P. Filipczuk, M. Kowal, and A. Obuchowicz, "Fuzzy clustering and adaptive thresholding based segmentation method for breast cancer diagnosis," in *Computer Recognition Systems*, vol. 4, 2011.
- [52] O. Schmitt and M. Hasse, "Radial symmetries based decomposition of cell clusters in binary and gray level images," *Pattern Recognit.*, vol. 41, no. 6, pp. 1905–1923, Jun. 2008.
- [53] J. Cheng and J. C. Rajapakse, "Segmentation of clustered nuclei with shape markers and marking function," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 3, pp. 741–748, Mar. 2009.
- [54] C. Jung and C. Kim, "Segmenting clustered nuclei using H -minima transform-based marker extraction and contour parameterization," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 10, pp. 2600–2604, Oct. 2010.
- [55] H. Wang, A. C. Roa, A. N. Basavanahally, H. L. Gilmore, N. Shih, M. Feldman, J. Tomaszewski, F. Gonzalez, and A. Madabhushi, "Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features," *J. Med. Imag.*, vol. 1, no. 3, Oct. 2014, Art. no. 034003.
- [56] Ş. Öztürk and B. Akdemir, "Cell-type based semantic segmentation of histopathological images using deep convolutional neural networks," *Int. J. Imag. Syst. Technol.*, vol. 29, no. 3, pp. 234–246, Sep. 2019.
- [57] Y. Cui, G. Zhang, Z. Liu, Z. Xiong, and J. Hu, "A deep learning algorithm for one-step contour aware nuclei segmentation of histopathology images," *Med. Biol. Eng. Comput.*, vol. 57, no. 9, pp. 2027–2043, Sep. 2019.
- [58] H. Chen, X. Qi, L. Yu, Q. Dou, J. Qin, and P.-A. Heng, "DCAN: Deep contour-aware networks for object instance segmentation from histology images," *Med. Image Anal.*, vol. 36, pp. 135–146, Feb. 2017.
- [59] M. Z. Alom, C. Yakopcic, T. M. Taha, and V. K. Asari, "Nuclei segmentation with recurrent residual convolutional neural networks based U-Net (R2U-Net)," in *Proc. IEEE Nat. Aerosp. Electron. Conf. (NAECON)*, Jul. 2018, pp. 228–233.
- [60] M. Salvi, M. Bosco, L. Molinaro, A. Gambella, M. Papotti, U. R. Acharya, and F. Molinari, "A hybrid deep learning approach for gland segmentation in prostate histopathological images," *Artif. Intell. Med.*, vol. 115, May 2021, Art. no. 102076.
- [61] Q. Kang, Q. Lao, and T. Fevens, "Nuclei segmentation in histopathological images using two-stage learning," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2019, pp. 703–711.
- [62] D. Kucharski, P. Kleczek, J. Jaworek-Korjakowska, G. Dyduch, and M. Gorgon, "Semi-supervised nests of melanocytes segmentation method using convolutional autoencoders," *Sensors*, vol. 20, no. 6, p. 1546, Mar. 2020.
- [63] S. Petushi, F. U. Garcia, M. M. Haber, C. Katsinis, and A. Tozeran, "Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer," *BMC Med. Imag.*, vol. 6, no. 1, pp. 1–11, Apr. 2006.
- [64] H. A. Elsalamony, "Healthy and unhealthy red blood cell detection in human blood smears using neural networks," *Micron*, vol. 83, pp. 32–41, Apr. 2016.
- [65] C. Lu, M. Mahmood, N. Jha, and M. Mandal, "Automated segmentation of the melanocytes in skin histopathological images," *IEEE J. Biomed. Health Informat.*, vol. 17, no. 2, pp. 284–296, Mar. 2013.
- [66] A. Tashk, M. S. Helfroush, H. Danyali, and M. Akbarzadeh-jahromi, "Automatic detection of breast cancer mitotic cells based on the combination of textural, statistical and innovative mathematical features," *Appl. Math. Model.*, vol. 39, no. 20, pp. 6165–6182, Oct. 2015.
- [67] K. Fukuma, V. B. S. Prasath, H. Kawanaka, B. J. Aronov, and H. Takase, "A study on nuclei segmentation, feature extraction and disease stage classification for human brain histopathological images," *Proc. Comput. Sci.*, vol. 96, pp. 1202–1210, Jan. 2016.
- [68] R. Rashmi, K. Prasad, C. B. K. Udupa, and V. Shwetha, "A comparative evaluation of texture features for semantic segmentation of breast histopathological images," *IEEE Access*, vol. 8, pp. 64331–64346, 2020.

- [69] B. E. Bejnordi, N. Timofeeva, I. Otte-Höller, N. Karssemeijer, and J. A. van der Laak, "Quantitative analysis of stain variability in histology slides and an algorithm for standardization," *Proc. SPIE*, vol. 9041, Mar. 2014, Art. no. 904108.
- [70] H. Kong, M. Gurcan, and K. Belkacem-Boussaid, "Partitioning histopathological images: An integrated framework for supervised color-texture segmentation and cell splitting," *IEEE Trans. Med. Imag.*, vol. 30, no. 9, pp. 1661–1677, Sep. 2011.
- [71] O. Sertel, J. Kong, U. V. Catalyurek, G. Lozanski, J. H. Saltz, and M. N. Neves, "Histopathological image analysis using model-based intermediate representations and color texture: Follicular lymphoma grading," *J. Signal Process. Syst.*, vol. 55, nos. 1–3, pp. 169–183, Apr. 2009.
- [72] O. Sertel, J. Kong, G. Lozanski, A. Shana'ah, U. Catalyurek, J. Saltz, and M. Gurcan, "Texture classification using nonlinear color quantization: Application to histopathological image analysis," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2008, pp. 597–600.
- [73] D. O. T. Bruno, M. Z. D. Nascimento, R. P. Ramos, V. R. Batista, L. A. N. Neves, and A. S. Martins, "LBP operators on curvelet coefficients as an algorithm to describe texture in breast cancer tissues," *Expert Syst. Appl.*, vol. 55, pp. 329–340, Aug. 2016.
- [74] H. Irshad, S. Jalali, L. Roux, D. Racoceanu, G. Naour, L. Hwee, and F. Capron, "Automated mitosis detection using texture, SIFT features and HMAX biologically inspired approach," *J. Pathol. Informat.*, vol. 4, no. 2, p. 12, 2013.
- [75] A. C. Simsek, A. B. Tosun, C. Aykanat, C. Sokmensuer, and C. Gunduz-Demir, "Multilevel segmentation of histopathological images using cooccurrence of tissue objects," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 6, pp. 1681–1690, Jun. 2012.
- [76] S. A. Korkmaz, "Recognition of the gastric molecular image based on decision tree and discriminant analysis classifiers by using discrete Fourier transform and features," *Appl. Artif. Intell.*, vol. 32, nos. 7–8, pp. 629–643, Sep. 2018.
- [77] T. Y. Rahman, L. B. Mahanta, H. Choudhury, A. K. Das, and J. D. Sarma, "Study of morphological and textural features for classification of oral squamous cell carcinoma by traditional machine learning techniques," *Cancer Rep.*, vol. 3, no. 6, p. e1293, Dec. 2020.
- [78] T. Y. Rahman, L. B. Mahanta, A. K. Das, and J. D. Sarma, "Automated oral squamous cell carcinoma identification using shape, texture and color features of whole image strips," *Tissue Cell*, vol. 63, Apr. 2020, Art. no. 101322.
- [79] T. J. Alhindi, S. Kalra, K. H. Ng, A. Afrin, and H. R. Tizhoosh, "Comparing LBP, HOG and deep features for classification of histopathology images," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2018, pp. 1–7.
- [80] A. Masood and A. Al-Jumaily, "Differential evolution based advised SVM for histopathological image analysis for skin cancer detection," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 781–784.
- [81] O. S. Al-Kadi, "A multiresolution clinical decision support system based on fractal model design for classification of histological brain tumours," *Computerized Med. Imag. Graph.*, vol. 41, pp. 67–79, Apr. 2015.
- [82] M. Kuse, T. Sharma, and S. Gupta, "A classification scheme for lymphocyte segmentation in H&E stained histology images," in *Proc. Int. Conf. Pattern Recognit.*, 2010, pp. 235–243.
- [83] Y. Xu, L. Jiao, S. Wang, J. Wei, Y. Fan, M. Lai, and E. I.-C. Chang, "Multi-label classification for colon cancer using histopathological images," *Microsc. Res. Technique*, vol. 76, no. 12, pp. 1266–1277, Dec. 2013.
- [84] M. M. R. Krishnan, C. Chakraborty, R. R. Paul, and A. K. Ray, "Hybrid segmentation, characterization and classification of basal cell nuclei from histopathological images of normal oral mucosa and oral submucous fibrosis," *Expert Syst. Appl.*, vol. 39, no. 1, pp. 1062–1077, Jan. 2012.
- [85] S. I. Niwas, P. Palanisamy, K. Sujathan, and E. Bengtsson, "Analysis of nuclei textures of fine needle aspirated cytology images for breast cancer diagnosis using complex Daubechies wavelets," *Signal Process.*, vol. 93, no. 10, pp. 2828–2837, Oct. 2013.
- [86] J. Han, H. Chang, L. Loss, K. Zhang, F. L. Baehner, J. W. Gray, P. Spellman, and B. Parvin, "Comparison of sparse coding and kernel methods for histopathological classification of Glioblastoma multiforme," in *Proc. IEEE Int. Symp. Biomed. Imag., Nano Macro*, Mar. 2011, pp. 711–714.
- [87] U. Srinivas, H. S. Mousavi, V. Monga, A. Hattel, and B. Jayarao, "Simultaneous sparsity model for histopathological image representation and classification," *IEEE Trans. Med. Imag.*, vol. 33, no. 5, pp. 1163–1179, May 2014.
- [88] N. N. Shirale, "Sparse representation based class level dictionary learning approach for histopathology image classification," in *Proc. 4th Int. Conf. Conver. Technol. (I2CT)*, Oct. 2018, pp. 1–5.
- [89] X. Li, H. Tang, D. Zhang, T. Liu, L. Mao, and T. Chen, "Histopathological image classification through discriminative feature learning and mutual information-based multi-channel joint sparse representation," *J. Vis. Commun. Image Represent.*, vol. 70, Jul. 2020, Art. no. 102799.
- [90] Y. Zheng, Z. Jiang, F. Xie, H. Zhang, Y. Ma, H. Shi, and Y. Zhao, "Feature extraction from histopathological images based on nucleus-guided convolutional neural network for breast lesion classification," *Pattern Recognit.*, vol. 71, pp. 14–25, Nov. 2017.
- [91] N. Wahab, A. Khan, and Y. S. Lee, "Two-phase deep convolutional neural network for reducing class skewness in histopathological images based breast cancer detection," *Comput. Biol. Med.*, vol. 85, pp. 86–97, Jun. 2017.
- [92] H. Saito, T. Aoki, K. Aoyama, Y. Kato, A. Tsuboi, A. Yamada, M. Fujishiro, S. Oka, S. Ishihara, T. Matsuda, M. Nakahori, and S. Tanaka, "Automatic detection and classification of protruding lesions in wireless capsule endoscopy images based on a deep convolutional neural network," *Gastrointestinal Endoscopy*, vol. 92, no. 1, pp. 144–151, 2020.
- [93] D. Moitra and R. K. Mandal, "Classification of non-small cell lung cancer using one-dimensional convolutional neural network," *Expert Syst. Appl.*, vol. 159, Nov. 2020, Art. no. 113564.
- [94] S. Panigrahi, J. Das, and T. Swarnkar, "Capsule network based analysis of histopathological images of oral squamous cell carcinoma," *J. King Saud Univ.-Comput. Inf. Sci.*, Nov. 2020.
- [95] H. Kutlu, E. Avci, and F. Özyurt, "White blood cells detection and classification based on regional convolutional neural networks," *Med. Hypotheses*, vol. 135, Feb. 2020, Art. no. 109472.
- [96] R. Man, P. Yang, and B. Xu, "Classification of breast cancer histopathological images using discriminative patches screened by generative adversarial networks," *IEEE Access*, vol. 8, pp. 155362–155377, 2020.
- [97] Y. Xue, J. Ye, Q. Zhou, L. R. Long, S. Antani, Z. Xue, C. Cornwell, R. Zaino, K. C. Cheng, and X. Huang, "Selective synthetic augmentation with HistoGAN for improved histopathology image classification," *Med. Image Anal.*, vol. 67, Jan. 2021, Art. no. 101816.
- [98] D. Wang, Z. Chen, and H. Zhao, "Prototype transfer generative adversarial network for unsupervised breast cancer histology image classification," *Biomed. Signal Process. Control*, vol. 68, Jul. 2021, Art. no. 102713.
- [99] P. Alirezazadeh, B. Hejrati, A. Monsef-Esfahani, and A. Fathi, "Representation learning-based unsupervised domain adaptation for classification of breast cancer histopathology images," *Biocybern. Biomed. Eng.*, vol. 38, no. 3, pp. 671–683, 2018.
- [100] X. Yu, H. Zheng, C. Liu, Y. Huang, and X. Ding, "Classify epithelium-stroma in histopathological images based on deep transferable network," *J. Microsc.*, vol. 271, no. 2, pp. 164–173, Aug. 2018.
- [101] J. Shi, X. Zheng, J. Wu, B. Gong, Q. Zhang, and S. Ying, "Quaternion Grassmann average network for learning representation of histopathological image," *Pattern Recognit.*, vol. 89, pp. 67–76, May 2019.
- [102] J. Shi, J. Wu, Y. Li, Q. Zhang, and S. Ying, "Histopathological image classification with color pattern random binary hashing-based PCANet and matrix-form classifier," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 5, pp. 1327–1337, Sep. 2017.
- [103] F. G. Zanjani, S. Zinger, B. E. Bejnordi, J. A. W. M. van der Laak, and P. H. N. de With, "Stain normalization of histopathology images using generative adversarial networks," in *Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2018, pp. 573–577.
- [104] M. T. Shaban, C. Baur, N. Navab, and S. Albarqouni, "StainGAN: Stain style transfer for digital histological images," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 953–956.
- [105] P. Salehi and A. Chalechale, "Pix2Pix-based stain-to-stain translation: A solution for robust stain normalization in histopathology images analysis," in *Proc. Int. Conf. Mach. Vis. Image Process. (MVIP)*, Feb. 2020, pp. 1–7.
- [106] D. Mahapatra, B. Bozorgtabar, J. P. Thiran, and L. Shao, "Structure preserving stain normalization of histopathology images using self supervised semantic guidance," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2020, pp. 309–319.

...