Journal of Microscopy

Journal of Microscopy, Vol. 258, Issue 3 2015, pp. 233–240 doi: 10.1111/jmi.12237
Received 28 August 2014; accepted 3 February 2015

Automated image analysis of nuclear atypia in high-power field histopathological image

CHENG LU*, MENGYAO JI†, ZHEN MA‡ & MRINAL MANDAL§
*College of Computer Science, Shaanxi Normal University, Xi'an, Shaanxi Province, China
†Department of Gastroenterology, Renmin Hospital of Wuhan University, Wuhan, Hubei, China
‡College of Food Engineering and Nutritional Science, Shaanxi Normal University, Xi'an, Shaanxi Province, China
§Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Alberta, Canada

Key words. Classification, computer-aided image analysis, histopathological image analysis, nuclei atypia.

Summary

Aims

We developed a computer-aided technique to study nuclear atypia classification in high-power field haematoxylin and eosin stained images.

Methods and results

An automated technique for nuclear atypia score (NAS) calculation is proposed. The proposed technique uses sophisticated digital image analysis and machine-learning methods to measure the NAS for haematoxylin and eosin stained images. The proposed technique first segments all nuclei regions. A set of morphology and texture features is extracted from presegmented nuclei regions. The histogram of each feature is then calculated to characterize the statistical information of the nuclei. Finally, a support vector machine classifier is applied to classify a high-power field image into different nuclear atypia classes. A set of 1188 digital images was analysed in the experiment. We successfully differentiated the high-power field image with NAS1 versus non-NAS1, NAS2 versus non-NAS2 and NAS3 versus non-NAS3, with area under receiver-operating characteristic curve of 0.90, 0.86 and 0.87, respectively. In three classes evaluation, the average classification accuracy was 78.79%. We found that texture-based feature provides best performance for the classification.

Conclusion

The automated technique is able to quantify statistical features that may be difficult to be measured by human and demonstrates the future potentials of automated image analysis technique in histopathology analysis.

Introduction

In current practice, microscopic analysis of haematoxylin and eosinstained section is the backbone in the diagnosis of breast cancer and other types of cancers. Traditionally, pathologists examine histological slides under a microscope, and make diagnostic decisions. Such decisions are subjective and often lead to intraobserver and interobserver variability (Ruijter *et al.*, 2000; Allsbrook *et al.*, 2001). Therefore, quantitative analysis that can provide reliable, reproducible and objective results is highly desirable. Recent advances in digital whole slide scanner technology enable the use of quantitative and automatic image analysis methods. Many research works have been conducted to reduce the subjective level and pathologists' workload, and have been shown to have great potential in assisting pathologists for their diagnosis (Gurcan *et al.*, 2009).

Nottingham grading system is a commonly used prognostic tool for breast cancer, in which the morphological measurement of tubule, nuclear atypia and mitotic count are assessed. In the Nottingham grading system, evaluation of cell nuclear atypia plays an important role in the examination and diagnosis of breast cancer (Elston & Ellis, 1991). The level of nuclear atypia, also known as nuclear pleomorphism, indicates evolution stage of the cancer. In Nottingham grading system, nuclear atypia scores (NASs) of 1, 2 and 3 are given depending on the degree of the atypia. A higher score indicates a more serious nuclear atypia stage (Elston & Ellis, 1991). In clinical practice, pathologists examine a tissue slide under 20× or 40× magnification through light microscope and evaluate a visible area, normally high-power field (HPF), with NAS of 1, 2, or 3. Such manual examination is tedious, time consuming and often leads to interobserver variability (Buccoliero, 2008).

C. Lu and M. Ji have contributed equally to this work. Ethics: Ethics approval and/or informed consent were not required for this study.
Correspondence to: C. Lu, College of Computer Science, Shaanxi Normal University, Xi'an, Shaanxi Province, China 710119. Tel/Fax: +086 029-85310161; e-mail: chenglu@snnu.edu.cn

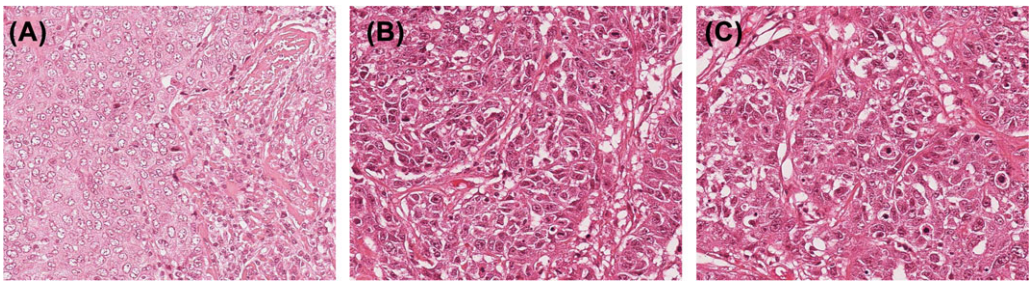

Fig. 1. Examples of breast cancer image with nuclei atypia score (NAS). (A) NAS = 1. (B) NAS = 2. (C) NAS = 3.

Many computerized techniques have been reported which focused mainly on nuclei, tubule and mitosis detection in breast cancer images (Petushi *et al.*, 2006; Fatakdawala *et al.*, 2010; Irshad *et al.*, 2013; Lu & Mandal, 2014). For example, Fatakdawala *et al.* proposed an expectation maximization driven geodesic active contour technique to quantitatively segment lymphocytes in HER2+ breast cancer (Fatakdawala *et al.*, 2010). Lu *et al.* applied morphological and texture features to detect the mitotic cell in multispectral images (Lu & Mandal, 2014). However, nuclear atypia scoring for HPF image is a challenging task and has not been addressed well in the literature (Cosatto *et al.*, 2008; Dalle *et al.*, 2009). Cosatto *et al.* (Cosatto *et al.*, 2008) proposed to classify a region into benign and malignant using median of nuclear area and the number of nuclei found in that region. Dalle *et al.* (Dalle *et al.*, 2009) proposed to extract features from selected nuclear regions and classify NAS 2 and NAS 3. In both of the above-mentioned works, the mean of features obtained from selected nuclei was utilized to characterize certain regions. In practice, in order to achieve good result, a large number of nuclei and their statistical features should be considered in order to measure the nuclear atypia degree in a HPF image. To our best knowledge, there is no fully automated technique for HPF image nuclear atypia scoring.

Digital image processing/analysis techniques are able to enhance the image quality for better object location and to measure repeated patterns that may not be easily captured by human. Machine learning techniques are able to learn features that extracted from existing samples with class labels and to predict class labels for unseen samples. Both of digital image processing/analysis and machine learning techniques have been successfully applied in computer-aided diagnosis in radiology. In this paper, we propose a fully automated framework for HPF image nuclear atypia scoring. We concentrate on quantitative measurement of cellular features and investigate how they contribute to the automated classification among different NASs. The results showed that automated image analysis combined with machine learning techniques have the capability of grading the nuclear atypia for a HPF image. The automated technique is able to mimic human mind to certain extent and is able to quantify statistical features that may be difficult to measure by human.

Materials and methods

Image data

This study was based on a set of 1188 haematoxylin and eosin stained digital images of breast cancer tissue slides. The image data set were obtained from a publicly available Website (MITOS-ATYPIA-14 (Ludovic *et al.*, 2014)). These images were captured under 40× magnification using Aperio Scanscope XT digital scanner (Buffalo Grove, IL). Each image has a size of 1539 × 1376 pixels with a resolution of 0.2455 μm/pixel and corresponds to a physical area of 755.649 μm × 675.616 μm. Two experienced pathologists assigned a NAS of 1, 2 or 3 for each image based on manual examination. If an image has two different scores, a third pathologist examines the image and assigns an independent score. In this case, the final score for the image is the one with majority agreement. A NAS of 1, 2 or 3 is assigned for each image according to the degree of the nuclear atypia. Figure 1 shows examples of images with nuclei atypia score of 1 (A), nuclei atypia score of 2 (B) and nuclei atypia score of 3 (C). It is observed that in the case of NAS 1, there is minor variation in size, shape and chromatin pattern. In the case of NAS 2, the nuclei are distinctly enlarged, and nucleoli are visible. In the case of NAS 3, we can observe markedly enlarged nuclei with prominent nucleoli, and generally variation in terms of nuclei size, shape and chromatin pattern.

The overall schematic of the proposed technique

The schematic for the proposed technique is shown in Figure 2. Stain normalization and colour separation method was first applied on the original haematoxylin and eosin stained RGB image (Niethammer *et al.*, 2010). The grey image of the H-stained image was then selected for further processing since the nuclei regions are well distinguished with background. Using a combination of seed detection, local threshold and morphology operations, the proposed technique segmented all the nuclei regions within the image. Next, a set of morphology and texture features was extracted from presegmented nuclei regions. In order to characterize the statistical information of the nuclei within each HPF image, histogram of each feature

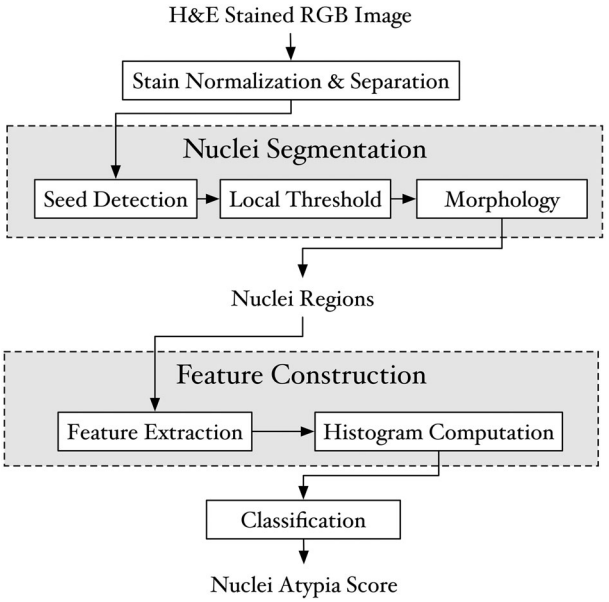

Fig. 2. The overall schematic of the proposed technique.

was then calculated. The feature histograms were then fed into a classifier to obtain the final grading result, i.e. NAS. Details of each step are described in the following sections.

Stain normalization and separation

Before applying the image analysis algorithms, the images were normalized in order to reduce the staining variations due to different storage time, procedures and stains (Niethammer *et al.*, 2010). In the stain normalization, the prior information on the stain vectors (i.e. the hematoxylin stain and eosin stain) is used in the estimation. After obtaining the normalized colour image, the image is separated into hematoxylin stained image (Fig. 3 B) and eosin stained colour image (Fig. 3 C), respectively, based on the estimated stain vectors. Since the nuclei are stained with hematoxylin, the normalized hematoxylin images are chosen for further analysis.

Nuclei segmentation

Once the normalized hematoxylin stain image was obtained, a blue-ratio image $I_{br}(x,y)$ is calculated as follows (Chang *et al.*, 2012):

$$I_{br}(x, y) = \frac{100 \times I_B(x, y)}{1 + I_R(x, y) + I_B(x, y)} \times \frac{256}{1 + I_R(x, y) + I_G(x, y) + I_B(x, y)},$$

where (x, y) represents the 2-D coordinate of the image; $I_R(x, y)$, $I_G(x, y)$ and $I_B(x, y)$ are the red, green and blue channel of the hematoxylin colour image $I(x, y)$, respectively. The blue-ratio image transformation aims to accentuate the nuclei dye in the image so that the nuclei areas are easier to be detected from the whole image (Chang *et al.*, 2012). Considering that the nuclear region has a circle-like shape, and the size is also varied, we apply a multiscale Laplacian of Gaussian (LoG) to detect the seed point for each nuclear in the image. A LoG filter is defined as follows (Lindeberg, 1998):

$$\mathrm{LoG}(x, y; \sigma) = \frac{\partial G(x, y; \sigma)}{\partial x^2} + \frac{\partial G(x, y; \sigma)}{\partial y^2}.$$

In the multiscale LoG framework, $\sigma \in [\sigma_{\min}, \sigma_{\max}]$, where $\sigma_{\min}$ and $\sigma_{\max}$ are the lower and upper limit of the scale. For each predefined scale, $\sigma \in [\sigma_{\min}, \sigma_{\max}]$, we calculate the output of the scale-normalized LoG filter $\mathrm{LoG}_{\mathrm{norm}}(x, y; \sigma) = \sigma^2 \mathrm{LoG}(x, y; \sigma)$ with the blue-ratio image as input as follows:

$$R(x, y) = \underset{\sigma \in [\sigma_{\min}, \sigma_{\max}]}{\arg\max} [\mathrm{LoG}_{\mathrm{norm}}(x, y; \sigma)^* I_{br}(x, y)].$$

Once we obtain the filtered image $R(x, y)$, the mean-shift algorithm (Comaniciu & Meer, 2002; Lu *et al.*, 2013) is applied to find the seed points of nuclei region. Since each seed point location indicates the presence of a nucleus, the nuclei region was calculated by using a local threshold for each seed point location (Lu *et al.*, 2014). Morphological operations are then applied to obtain nuclei regions with smooth boundaries and reasonable size from a HPF image. Figure 3 (D) shows the final nuclei segmentation result for a HPF image shown in Figure 3(A).

Feature construction

In order to classify the nuclei atypia for each HPF image, we extracted a set of 142 morphological and statistical texture features, shown in Table 1, from each presegmented nuclear region. In other words, for each presegmented nuclear region, we have 142 feature values. Denote the feature vector for the ith nuclear region in a HPF image as F_i.

It is natural to utilize the mean of feature vectors extracted from all nuclei regions to characterize each HPF image and differentiate/distinguish image with different NAS. However, in this study, we considered the histogram of all features and we will demonstrate, in section 'Results and discussions', that such higher level of statistical information is better than just using the mean of all features. Therefore, to characterize each HPF image using extracted features, we compute histogram of each feature and denote it as F^k, where k represents the index of feature, $k \in [1...142]$. For a HPF image, we have 142 histograms for different features. If the number of bin for each histogram is set to 50, a vector dimension with 7100 variables vector is used to describe a HPF image.

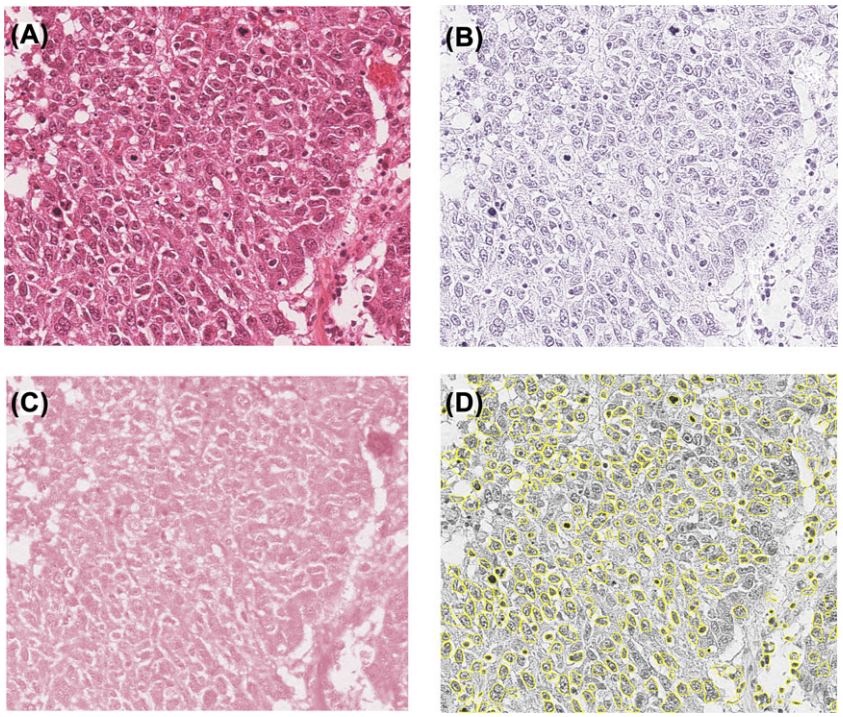

Fig. 3. (A) Original haematoxylin and eosin (H&E) colour image. (B) Normalized hematoxylin stain image. (C) Normalized eosin stain image. (D) Automated segment nuclei regions (yellow contours) overlaid onto the grey level hematoxylin stain image.

Table 1. Features extracted for image classification. Each feature is extracted from presegmented nuclear region, and a histogram is computed for each feature in a HPF image

Feature name	Dimension
Nuclear size	1
Stain intensity features: mean of stain, SD of stain	2
Gradient features: sum, mean, entropy of gradient magnitude (Gonzalez & Woods 2002)	3
Tamura texture features: coarseness, contrast direction (Tamura *et al.*, 1978)	3
Grey level run-length matrix-based texture features (Tang, 1998)	44
Co-occurrence matrix-based texture features in four directions (Haralick *et al.*, 1973)	88
In total	142

SD, standard deviation.

Classification/nuclei atypia grading

In this paper, we use machine-learning method to classify images with different nuclei atypia scores. Specifically, we utilize the histogram of extracted features to represent all HPF images and divide all data into two subsets. We first utilize one subset as the training set to train a classifier, and then use the other subset as unseen testing subset to test the classifier. Basically, a classifier is a mathematical model that is able to differentiate different classes by tuning its parameters based on the training set samples [Duda2001]. A simplest classifier would be a single threshold t for one dimension feature like size of nuclei (denoted as S). For a nuclei region R_i, one can set $R_i \in C_1$ if $S_i > t$, otherwise $R_i \in C_2$, where C_1 and C_2 represent two different classes of nuclei. However, in this paper, due to the high dimension features and nonlinearity of the data, we use a sophisticated classifier model named support vector machine (SVM) to perform the classification task.

Classification performance analysis

During the evaluation process, we initially analyse 2-class classification performance, i.e. we separate all three NAS into three pairs of two classes: NAS 1 versus non-NAS1 (aggregate of NAS 2 and NAS 3), NAS 2 versus non-NAS2 (aggregate of NAS 1 and NAS 3), and NAS 3 versus non-NAS3 (aggregate of NAS 1 and NAS 2). In these three pairs 2-class classification, a SVM classifier model with radial basis function is used.

We then evaluate the classification performance of three classes, i.e. NAS 1 versus NAS 2 versus NAS 3. In the 3-class classification, a SVM classifier with radial basis function using the one-against-one strategy is used (Hsu & Lin, 2002). We also compare the contribution of the histogram of features and the mean value of features in the classification.

The importance of each type of features (see Table 2) is also analysed. In other words, the contribution of each type of features that lead to the classification is evaluated. All of six types of features are iteratively excluded from the whole feature set and six different incomplete feature sets are generated. All of these six incomplete feature sets were then used for the classification and the classification performances were sorted ascendant.

Parameter selections

There are several parameters used in the proposed technique. In nuclei segmentation module, parameters $\sigma_{\min}$ and $\sigma_{\max}$ are the lower and upper limit of the scale used to detect nuclei in image. Because size of nuclei should be within a reasonable range even if they are malignant in certain magnification, these two parameters are easy to be determined by measuring a few manual selected samples. In this study, we set $\sigma_{\min} = 6$ and $\sigma_{\max} = 20$ empirically (this depends on magnification of image, the above-mentioned values are set under magnification of $40\times$).

In classification module, the SVM classifier with radial basis function has two parameters: C and δ. These two parameters can adjust the separation/classification plane to get different classification performance. In this study, to reduce the computational complexity, we select these two parameters using a grid search method. That is, we create a 2-D grid where each grid point contains one possible parameters combination. In order to show overall classification performance of the proposed technique in 2-class classification with different parameter settings, we selected the SVM radial basis function parameters using grid search method. The receiver-operating characteristic curve is shown in Figure 4. In the grid search method, the value of parameter C ranges from 1 to 18, with a step of 1; the value of parameter δ range from 0.1 to 2 with a step of 0.2.

Evaluation metrics

The main objective of the evaluation is to determine whether the automated grading results obtained by the proposed technique are consistent with the manually assigned ones. Several evaluation metrics that were used in this study are described as follows.

Leave-one-out validation. In order to better utilize all available data and to evaluate the proposed technique, we use "leave-one-out" validation method to evaluate the classification performance. In other words, we first divide the whole data set into two subsets, in which one subset contains only one image data (denote this as testing set) and the other subset contains remaining image data (denote this as training set). The training set (i.e. bigger subset) is used as the experienced samples to train a classifier, whereas the testing set (i.e. smaller subset) is used as an unseen sample to test the trained classifier. Denote the total number of image as N, we then repeat such procedure by N times with different training and testing sets. Finally, the average performance is calculated and used as the final classification results.

Point-based metric. In order to evaluate the classification performance of three classes, we define a point-based metric. Denote an automated computed score, obtained from the proposed technique, for image i as A_i; the manually assigned score as M_i, a point-based metric is calculated as follows:

$$\mathrm{P} = \sum_{i=1}^{N} 1(A_i = M_i) - \sum_{i=1}^{N} 1(|A_i = M_i| = 2),$$

where N is the total number of images involved in the evaluation, and function $\mathbf{1}$(expression) is an indicator function, which is defined as follows:

$$1(\mathrm{exp\,ression}) = \begin{cases} 1 & \text{if expression is true} \\ 0 & \text{if expression is false.} \end{cases}$$

Basically, if an automated computed score is identical to the manually assigned one, the proposed technique will obtain one point; if an automated computed score is different by two units from the manually assigned one, the proposed technique will obtain a negative point, i.e. penalty by one point. The higher the P is, the better the proposed technique performs. In order to better present the metric regardless the size of a data set, a percentage version of this point-based metric is calculated as follows:

$$\overset{*}{P} = \frac{P}{N} \times 100\%.$$

Average classification accuracy. Another metric to evaluate the three classes classification is the average classification accuracy. The image is graded into three classes, with NAS 1, NAS 2 and NAS 3, respectively. For each class, we evaluate the classification/grading accuracy for each class as follows:

$$C^k = \frac{\sum_{i=1}^{N^k} 1(A_i^k = M_i^k)}{N^k} \times 100\%,$$

where the superscript k indicates the class, $k \in [1, 2, 3]$; N^k is the total number of images belongs to class k; A_i^k and M_i^k represent the automatically obtained score by the computer-aided technique and manually assigned score for the ith image in class k, respectively. To calculate the average classification

Table 2. 3-Class classification performances

Techniques	Point-based metric	Average accuracy
Single feature based	694/1188 (58.42%)	63.38%
Feature mean based	544/1188 (45.79%)	50.87%
The proposed (histogram of features)	936/1188 (78.79%)	78.79%

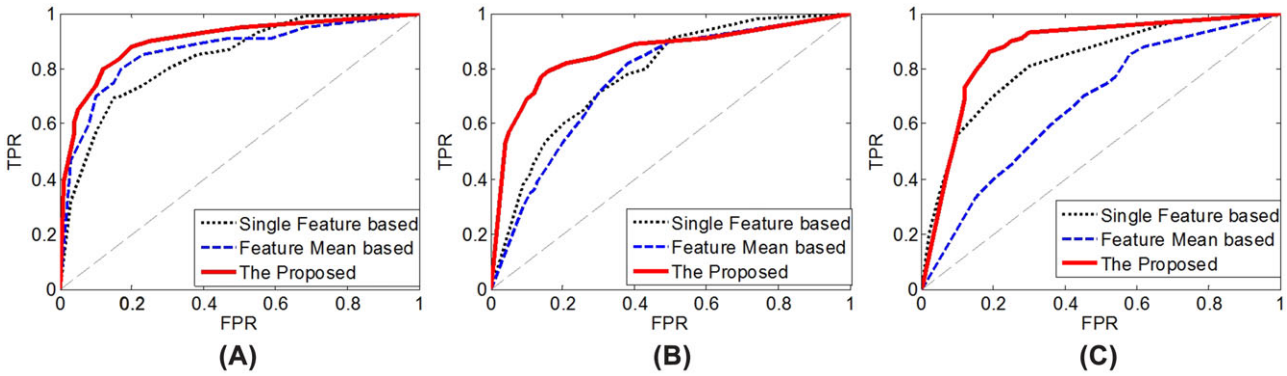

Fig. 4. (A) The receiver-operating characteristic (ROC) curves of NAS1 versus NAS2 and NAS3. Area under the curve (AUC) is 0.84, 0.86 and 0.90 for single feature based, feature mean based and the proposed technique, respectively ($p < 0.0001$). A random guess that corresponds to the dash line in ROC curve with a AUC value of 0.5. (B) The ROC curves of NAS2 versus NAS1 and NAS3. The AUC is 0.78, 0.76 and 0.86 for single feature based, feature mean based and the proposed technique, respectively ($p < 0.0001$). (C) The ROC curves of NAS3 versus NAS1 and NAS2. The AUC is 0.82, 0.67 and 0.87 for single feature based, feature mean based and the proposed technique, respectively ($p < 0.0001$).

accuracy in three classes classification, the following equation is applied.

$$C_{\text{ave}} = \frac{1}{3}\sum_{k=1}^{3} C^k.$$

Results and discussions

2-Class classification comparison

During the evaluation process, we firstly analysed 2-class classification. In order to compare with the existing techniques, we also implemented a similar technique described in (Cosatto *et al.*, 2008), where a single feature – median nuclear area over a region is used for two aptypia nuclei classification (we refer this technique as single feature based technique hence forth). In another technique (Dalle *et al.*, 2009), the mean of all features are generally utilized for classification (we refer this technique as feature mean based technique hence forth). In this study, we also compare the performance with the feature mean based technique. In the comparison, we used the SVM as the base classifier. The overall performance is evaluated using receiver-operating characteristic curve by tuning parameters with fine steps. The receiver-operating characteristic curves obtained by the single feature based, feature mean based and the proposed techniques are shown in Figure 4. It is observed that by using all the histogram of all features, the automated technique is able to separate the different classes, NAS1 versus non-NAS1, NAS2 versus non-NAS2 and NAS3 versus non-NAS3, with AUC of 0.90, 0.86 and 0.87, respectively. It is clear that when single feature and the feature mean are used in the classification, the classification power is worse than that by using the histogram of features.

3-Class classification comparison

The final goal of the proposed technique in this study is to assign a NAS automatically with an unseen HPF image. The 3-class classification results are shown in Table 3. Two kinds of evaluation metrics are calculated, i.e. point-based metric and average accuracy as described previously in section 'Parameter selections'. The results shown in Table 3 are the best result in terms of the classification accuracy. For 1188 HPF image, the proposed technique using histogram of all features has a score of 936, which corresponds to 78.79% accuracy. Similar to the result of 2-class classification as shown in section '2-class classification comparison', the classification performance of single feature based technique and feature mean based technique are worse than the proposed technique.

Importance of features

In the analysis, we also evaluated the importance of features that contribute to the classification. In other words, we evaluated which type of features contributed the most for classification. Table 2 presents the importance of all six types of features

Table 3. Importance of features for classification

Rank	NAS1 versus NAS2 and NAS3	NAS2 versus NAS1 and NAS3	NAS3 versus NAS1 and NAS2	NAS1 versus NAS2 versus NAS3
1	Grey level run-length matrix-based features	Grey level run-length matrix-based features	Grey level run-length matrix-based features	Grey level run-length matrix-based features
2	Co-occurrence matrix based-features	Co-occurrence matrix based-features	Co-occurrence matrix based-features	Co-occurrence matrix based-features
3	Gradient features	Tamura texture features	Gradient features	Nuclear size
4	Nuclear size	Nuclear size	Nuclear size	Gradient features
5	Stain intensity features	Stain Intensity features	Stain Intensity features	Stain Intensity features
6	Tamura texture features	Gradient features	Tamura texture features	Tamura texture features

that were used for classification. The higher the ranking, the more important the feature is.

From Table 3, it is clear that for all classifications, the texture features are the most important ones in separating different classes. More specifically, the grey level of the run-length matrix-based features and co-occurrence matrix-based features are most important during the classification. Besides the texture features, the size of nuclear and gradient features are also important for the classification. The observed results are in accordance with the subjective experience of the pathologists in grading NAS where the chromatin pattern and nuclear size are the primary factors during observation. With the higher score of NAS, the chromatin density is higher and the size of nuclear is larger. The texture features are aimed to quantitatively measure the chromatin pattern within the nuclear area and the experimental result shows their contribution for classifying the NAS.

Overall, the proposed technique is able to classify HPF image with different NAS. High-level statistical information, i.e. histogram of features within a HPF image, has showed superior classification ability than the mean of features. During the process of classifying 2 classes and 3 classes of HPF images, the proposed technique achieved good result and showed great potential in providing an objective method to assist the pathologists in the clinic diagnosis. One may seek for an ideal one-dimensional feature that would be able to classify different classes with no mistakes. However, to the best of our knowledge, this does not exist. Combining different statistical and quantitative features is a reasonable way to provide the best result. Within all the extracted features, the texture features contributed the most in classification of HPF images with different NAS.

This study is unique in combining efficient image analysis and machine learning methods to study the NAS in the diagnosis of cancer histology and provides a foundation for future research. The proposed framework is generic and can be extended to histopathological image of different cancer types. In future, we intend to include more features, such as contour-based features, to improve the classification performance. However, a high accuracy segmentation technique for nuclei region is required to obtain more accurate nuclei contour.

Acknowledgement

This research is partially supported by the National Natural Science Foundation of China (Grant No. 61401263) and Fundamental Research Funds for the Central Universities of China (Grant No.GK201402037). In this study, Cheng Lu designed the automated computerized technique, Mengyao Ji analysed the image data and result, Zhen Ma wrote the paper and Mrinal Mandal designed the research study.

References

Allsbrook, W., Mangold, K., Johnson, M.H., *et al.* (2001) Inter-observer reproducibility of Gleason grading of prostatic carcinoma: general pathologist. *Hum. Pathol.* **32**, 81–88.

Buccoliero, A.M. (2008) The problems and promise of central pathology review. *Pediatr. Dev. Pathol.* **11**, 327–328.

Cosatto, E., Miller, M., Graf, H.P. & Meyer, J.S. (2008) Grading nuclear pleomorphism on histological micrographs. In *Proceedings of the 19th International Conference on Pattern Recognition*. Tampa, Folorida. doi:10.1109/ICPR.2008.4761112.

Chang, H., Loss, L. & Parvin, B. (2012) Nuclear segmentation in H and E sections via multi-reference graphcut (MRGC). In *Proceedings of the International Symposium Biomed*. Barcelona, Spain 1 June, 2014. Available from: http://www.cris.ucr.edu/IGERT/papers/ISBI2012.pdf.

Comaniciu, D. & Meer, P. (2002) Mean shift: a robust approach toward feature space analysis. In *Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence*. 603–619.

Dalle, J., Racoceanu, D. & Putti, T.C. (2009) Nuclear pleomorphism scoring by selective cell nuclei detection. In *Proceedings of the IEEE Workshop on Applications of Computer Vision*. 7–8. Snowbird, Utah.

Elston, C.W. & Ellis, I.O. (1991) Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathol.* **19**, 403–410.

Fatakdawala, H., Basavanhally, A., Xu, J., Bhanot, G., Ganesan, S., Feldman, M., Tomaszewski, J.E. & Madabhushi, A. (2010) Expectation–maximization-driven geodesic active contour with overlap resolution: application to lymphocyte segmentation on breast cancer. *IEEE Trans. Biomed. Eng.* **57**, 1676–1689.

Gonzalez, R. & Woods, R. (2002) *Digital Image Processing*. Wiley, New York.

Gurcan, M.N., Boucheron, L.E. Can, A., Madabhushi, A., Rajpoot, N.M. & Yener, B. (2009) Histopathological image analysis: a review. *IEEE Rev. Biomed. Eng.* **2**, 147–171.

Haralick, R., Shanmugam, K. & Dinstein, I. (1973) Textural features for image classification. *IEEE Trans. Syst., Man, Cybern* **3**, 610–621.

Hsu, C.W. & Lin, C.J. (2002) A comparison of methods for multiclass support vector machines. *IEEE Trans. Neural Net.* **13**, 415–425.

Irshad, H., Jalali, S. & Roux, L. (2013) Automated mitosis detection using texture, SIFT features and HMAX biologically inspired approach. *J. Pathol. Info*. **4**, S12. doi:10.4103/2153-3539.109870.

Lindeberg, T. (1998) Edge detection and ridge detection with automatic scale selection. *Int. J. Comput. Vis.* **30**, 117–156.

Lu, C., Mahmood, M., Jha, N. & Mandal, M. (2013) Automated segmentation of the melanocytes in skin histopathological images. *IEEE J. Biomed. Health Info*. **17**, 284–296.

Lu, C. & Mandal, M. (2014) Toward automatic mitotic cell detection and segmentation in multispectral histopathological images. *IEEE J. Biomed. Health Info.* **18**, 594–605.

Ludovic, R., Daniel, R., Capron, F., Calvo, J., Attieh, E., Naour, G.L. & Gloaguen, A. (2014) [Intenet]. MITOS&ATYPIA-detection of mitosis and evaluation of nuclear atypia score in breast cancer histological images. Available from: MITOA-ATYPIA-14 Grand Challenge http://mitos-atypia-14.grand-challenge.org/home/.

Niethammer, M., Borland, D., Marron, J.S., Woolsey, J. & Thomas, N.E. (2010) Appearance normalization of histology slides. In *Proceedings of the International Workshop Machine Learning in Medical Imaging*. **1**, 58–66. Beijing, China.

Petushi, S., Garcia, F.U., Haber, M.M., Katsinis, C. & Tozeren, A. (2006) Large-scale computations on histology images reveal graded differentiating parameters for breast cancer. *BMC Med. Imaging.* **6**, 1–11.

Ruijter, E., van Leenders, G., Miller, G., Debruyne, F. & van de Kaa, C. (2000) Errors in histological grading by prostatic needle biopsy specimens: frequency and predisposing factors. *J. Pathol*. **192**, 229–233.

Tamura, H., Mori, S. & Yamawaki, T. (1978) Textural features corresponding to visual perception. *IEEE Trans. Syst., Man, Cybern.* **8**, 460–473.

Tang, X. (1998) Texture information in run-length matrices. *IEEE Trans. Image Processing.* **7**, 1602–1609.